Vishal Kamalakrishnan - April 25, 2025 - STS 4600

Sociotechnical Synthesis: Bridging Actors And Accountability In AI Failures

My technical project and STS research paper explore socio-technical systems in artificial intelligence (AI) products. My technical project develops a framework to improve the accuracy of Large Language Models (LLM) in generating Kusto Query Language (KQL) queries. At the same time, my STS research paper employs Actor-Network Theory (ANT) to analyze systemic failures in AI deployment. Both projects look at AI systems from a holistic view and emphasize that AI systems can only succeed with a combination of technical solutions and an understanding of the socio-technical systems shaping their implementation.

My technical project[1] addresses the issue of AI hallucinations in LLMs, where LLMs respond with factually inaccurate or misleading information to their given query. It proposes a three-step framework to enhance the reliability of LLMs for KQL generation. First, a ground-truth dataset of KQL queries was compiled from public repositories to benchmark open-source LLMs such as Gemini and Llama. Second, the STRUCTCHEM framework—originally designed for chemistry reasoning—was adapted to KQL generation. STRUCTCHEM works through structured prompting, formula generation, and iterative confidence-based reviews that help refine AI-generated results by providing AI tools with additional context. Finally, this system was refined by employing several AI learning techniques and was tested using tools like Microsoft Sentinel and Microsoft Defender.

My STS research paper examines Air Canada's 2024 chatbot failure through ANT and argues that the chatbot's failure was a result of its actor-network rather than an isolated technical error. It examined several factors within Air Canada's network that contributed to this failure

[1] Note: technical project referenced is LLM Threat Query that I have been working on during fall 2024 - spr 2025 which is unrelated to my capstone research - Earthquake Analysis And Prediction, that I completed during spr 2024.

including corporate strategy, industry trends, cheaper GPU access, and legal ambiguities and accountability. By tracing ANT's translation phases—problematization, interessement, enrollment, and mobilization—my STS research paper reveals how institutional priorities and actor networks shape the outcomes of AI products.

Working on both projects simultaneously enhanced my understanding of AI products and failures in their implementation. The technical projects allowed me to focus on combating AI hallucinations from a technical perspective, enriched by insights from STS research, which shows how socio-technical dynamics shape technical design. This dual approach allowed me to develop a more robust technical solution that aligns better with the complex actor networks governing real-world deployment and increases tolerance for failures arising from a misaligned actor network. Ultimately, working on both my STS research paper and my technical project together this past year has allowed me to explore AI failures from multiple angles, and each work contributed to improving the quality of the other.

412 words