Joshua Claassen (DZHW)

Web surveys under attack: Novel strategies for detecting LLM-driven bots

There is an ongoing discussion on the threat of bots – programs designed to interact with web-based systems – to the data quality and integrity of self-administered web surveys. Most recently, t-online had to shut down a web survey on the car manufacturer Tesla because of suspiciously high completion rates and sudden shifts in survey outcomes pointing to bot infiltration. In this talk, I take a close look at the capabilities of bots that are linked to Large Language Models (LLMs), such as Google's Gemini Pro. LLM-driven bots cannot only conduct complex tasks that go far beyond the capabilities of their rule-driven counterparts but also understand questions and provide meaningful responses. Overall, I present three case studies that 1) describe LLM-driven bot behavior, 2) analyze the performance of prompt injections, and 3) predict LLM-based text in open narrative responses.