Short post for LinkedIn

Headline: Reproducible Audit of Semantic-Ambiguity Failures in LLMs

Body: Over the past weeks my team reproduced semantic-ambiguity failures in LLMs (e.g., homophone confusion and first-person self-harm phrasing) and built an auditable evidence bundle documenting the cases, checksums, and signing instructions.

We also implemented a canonical "Sixteen Frequencies" mapping and exposed it via an API for reproducible research.

Key artifacts: `audit/evidence_bundle.zip`, `audit/manifest.csv`, `data/frequencies/SIXTEEN_FREQUENCIES_MAPPING.merged.json`.

Link to full write-up and reproducible artifacts: (link to Substack post or repo PR)

AILab #ResponsibleAI #Safety #Research