Model Performance on Gold Standard Dataset

Model	Security Understanding Index (SUI)
Llama 3.1 405B	14.3%
Grok 4	15.0%
DeepSeek V3.2	16.8%
Claude Opus 4.5	24.3%
GPT-5.2	26.6%
Gemini 3 Pro	38.3%