GPT-4o -50 functions - Error vs. Incoherence Scatter Plot (log₂ scale)

HumanEval