**Best Model Performance by Prompt** Best: OpenChat 3.5506516Nous Hermes Mistral 7B DPGest: Zephyr 7B **Best: Mistral 7B** 0.99 Privacy Score 1.0 0.97 0.95 Quality Score 0.89 0.80 0.79 8.0 Normalized Score (higher is better) 0.60 0.60 0.4 0.2 0.0

prompt2

prompt1

**Prompt** 

prompt3

prompt4