

Benchmark: MMLU Global Facts  
Performance Metric: Brier Score  
Correlation Metric: Spearman

