gpt5 - Phase 1: Comprehensive Metrics (2 experiments average) 1.0 0.692 0.718 0.697 Type 8.0 Party 0.6 Metric Score 0.4 0.480 0.692 0.710 Action 0.2 0.722 Asset 0.733 0.0 Accuracy Macro-F1 Weighted-F1