

Cognitive MMLU Methodology

- Confidence-aware benchmarking
- â¢ 27 scenario cohort expanding to 300+
- â¢ Dual reviewer workflow + rubric
- â¢ Outputs linked to evaluation notes