AI Model Performance Tiers
(TrueSkill ELO Forest Plot by Performance Level)

Elite (ELO ≥ 30)

MOSES — 33.9

O3 — 32.1

Advanced (25 ≤ ELO < 30)

GPT-4.1 — 26.1

LightRAG — 25.4

LightRAG-nano — 25.3

Intermediate (20 ≤ ELO < 25)

MOSES-nano — 22.9

GPT-4.1-nano — 22.6

GPT-4o — 20.1

Basic (ELO < 20)

GPT-4o-mini — 17.0

TrueSkill ELO Rating ($\mu - 3\sigma$)