

GPQA on the diamond set: accuracy

Source: AI Index, 2025 | Chart: 2025 AI Index report

