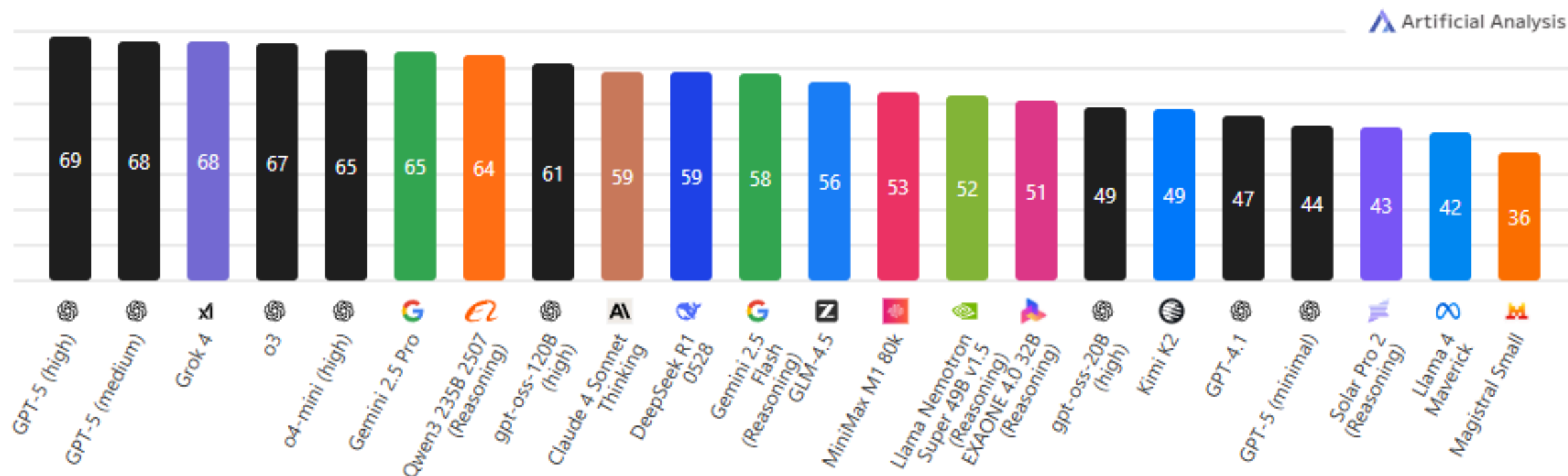


Språkmodeller för utvecklare

- **Syfte:** Förstå skillnader mellan olika språkmodeller och när man ska välja vilken typ av modell.

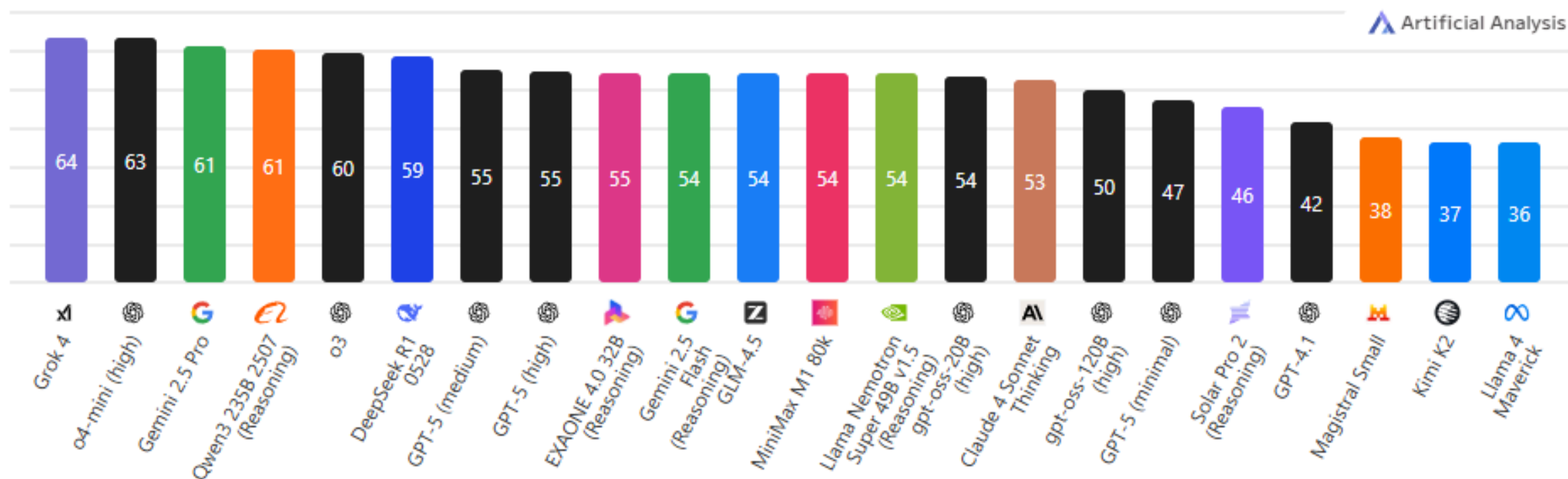
Artificial Analysis Intelligence Index

Artificial Analysis Intelligence Index v2.2 incorporates 8 evaluations: MMLU-Pro, GPQA Diamond, Humanity's Last Exam, LiveCodeBench, SciCode, AIME, IFBench, AA-LCR



Artificial Analysis Coding Index

Represents the average of coding benchmarks in the Artificial Analysis Intelligence Index (LiveCodeBench & SciCode)



Att välja modell

- Benchmarks säger någonting, men långt ifrån allt.
- Det finns ingen modell som är bäst på allt.
- Svarskvalitet är inte det enda att ta hänsyn till, även hastighet och kostnad spelar roll.

Andra faktorer

- **Context window:** hur många tokens (*1 token \approx 3–5 tecken*) modellen kan hantera samtidigt.
- **Säkerhet:** var skickas din data?
- **Function/tool calling:** behöver du att modellen klarar av att använda externa verktyg?

Tool calling

- **Vad är tool calling?** Modellen anropar definierade verktyg med strukturerade argument för att t.ex. söka/läsa/skriva filer, köra kommandon eller skicka ett meddelande i Slack.
- Stor skillnad mellan modeller.

Vad är en reasoning-modell?

- Många av de nya modellerna idag är reasoning-modeller.
- Är tränade på att bryta ned problem i mindre delar och tänka stegvis.
- Resonerandet blir en del av modellens svar.
- Starkare på oklara, flerstegs- och öppna problem.
- Långsammare och dyrare.

När ska man inte välja en reasoning-modell?

- **Snabba ändringar nära markören:** små kodfixar, kommentarer, enklare frågor.
- **Kostnadskänsliga uppgifter:** summeringar, översättningar, extrahering, loggsökningar.

Sammanfattning

- **Välj rätt modell för jobbet:** Ingen modell är bäst på allt. Valet beror på uppgiften – kodning, analys, snabba ändringar, etc.
- **Benchmarks ofta opålitliga:** Titta på relevanta tester för ditt användningsområde, testa och få en känsla för nya modeller.
- **Balansera kvalitet, hastighet och kostnad:** En "smartare" modell är ofta långsammare och dyrare. Använd dem för komplexa problem. För enklare uppgifter räcker ofta en snabbare och billigare modell.