

5. Stanford AI Index

El grup Human-Centered Artificial Intelligence de la Universitat de Stanford publica anualment el seu **AI Index Report**.

<https://hai.stanford.edu/research/ai-index-report>

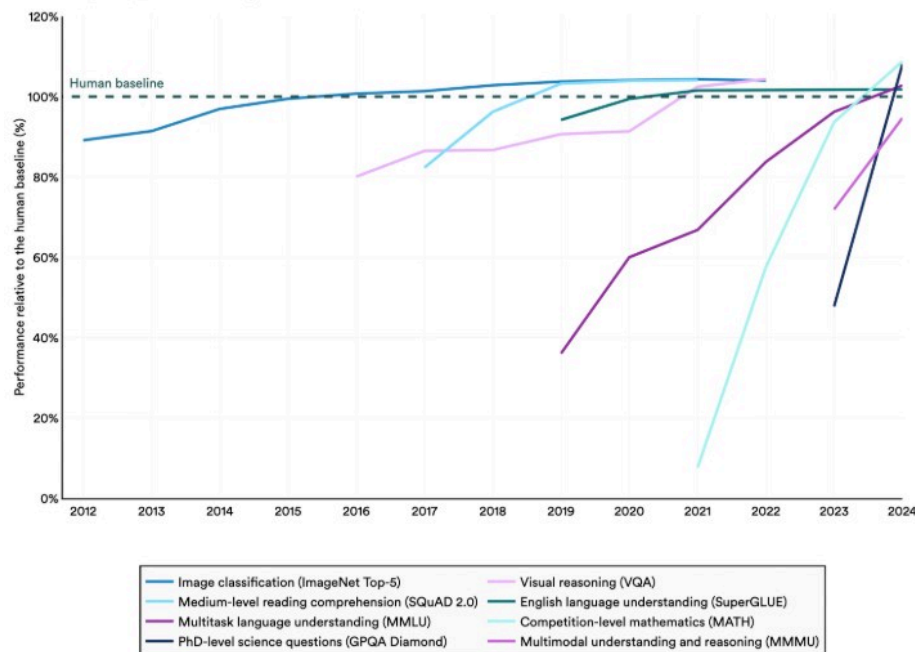
En els apartats següents detallam una dotzena de punts clau que hi destaquen en aquesta vuitena edició del 2025.

5.1. Millora del rendiment de la IA

El 2023, els investigadors van introduir nous punts de referència (MMMUI, GPQA i SWE-bench) per provar els límits dels sistemes d'IA avançats. Tot just un any després, el rendiment va augmentar bruscament: les puntuacions van augmentar en 18,8, 48,9 i 67,3 punts percentuals a MMMU, GPQA i SWE-bench, respectivament. Més enllà dels *benchmarks*, els sistemes d'IA van fer grans avenços en la generació de vídeo d'alta qualitat i, en alguns entorns, els agents de models de llenguatge fins i tot van superar els humans en tasques de programació amb temps limitat.

Select AI Index technical performance benchmarks vs. human performance

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

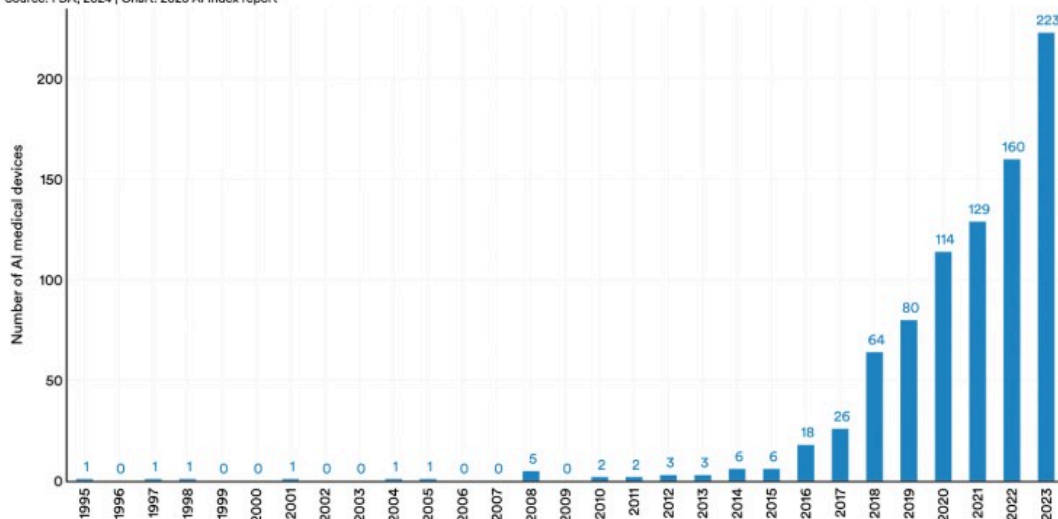


5.2. IA cada pic més integrada

Des de l'assistència sanitària fins al transport, la IA s'està movent ràpidament del laboratori a la vida diària. El 2023, la FDA va aprovar 223 dispositius mèdics compatibles amb IA, en comparació amb només sis el 2015. A les carreteres, els cotxes autònoms ja no són experimentals: Waymo, un dels majors operadors dels Estats Units, ofereix més de 150.000 viatges autònoms cada setmana, mentre que els robotaxis assequibles Apollo Go de Baidu ara són en servei a nombroses ciutats de la Xina.

Number of AI medical devices approved by the FDA, 1995–2023

Source: FDA, 2024 | Chart: 2025 AI Index report

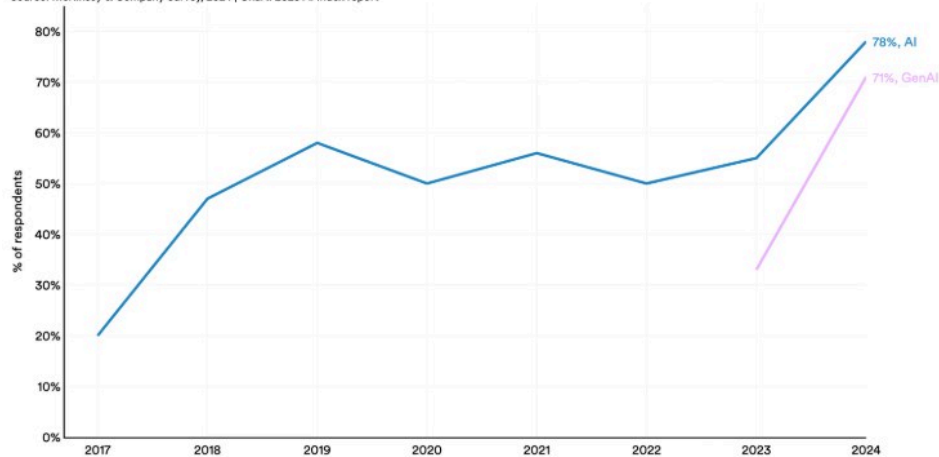


5.3. Millora de la productivitat

El 2024, la inversió privada en IA dels EUA va créixer fins als 109.000 milions de dòlars, gairebé 12 vegades els 9.300 milions de dòlars de la Xina i 24 vegades els 4.500 milions del Regne Unit. La [IA generativa](#) va experimentar un impuls especialment fort, atraient 33.900 milions de dòlars a nivell mundial en inversió privada, un augment del 18,7% respecte al 2023. L'ús de les empreses d'IA també s'està accelerant: el 78% de les organitzacions van declarar que utilitzen IA el 2024, en comparació amb el 55% de l'any anterior. Mentrestant, un nombre creixent d'investigacions confirma que la IA augmenta la productivitat i, en la majoria dels casos, ajuda a reduir les bretxes d'habilitats entre la força de treball.

Share of respondents who say their organization uses AI in at least one function, 2017–24

Source: McKinsey & Company Survey, 2024 | Chart: 2025 AI Index report

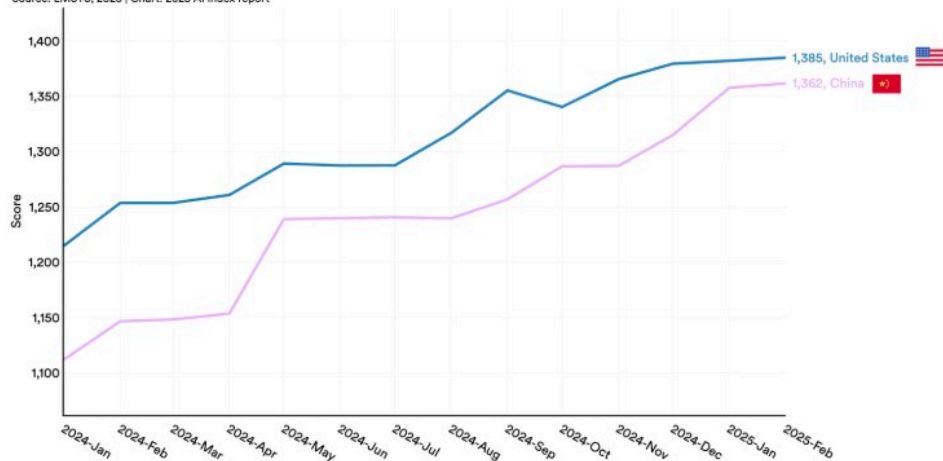


5.4. La Xina escurça distància amb els EUA

El 2024, les institucions amb seu als Estats Units van produir 40 models d'IA notables, superant significativament els 15 de la Xina i els tres d'Europa. Tot i que els EUA mantenen el seu lideratge en quantitat, els models xinesos han tancat ràpidament la bretxa de qualitat: les diferències de rendiment en els principals punts de referència com MMLU i HumanEval es van reduir de dos dígit el 2023 a gairebé la paritat el 2024. Mentrestant, la Xina continua liderant les publicacions i patents d'IA. Al mateix temps, el desenvolupament de models és cada cop més global, amb llançaments notables des de regions com l'Orient Mitjà, Amèrica Llatina i el sud-est asiàtic.

Performance of top United States vs. Chinese models on LMSYS Chatbot Arena

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



5.5. IA responsable

Els incidents relacionats amb la intel·ligència artificial estan augmentant molt, però les avaluacions estandarditzades de RAI (*Responsible AI*) segueixen sent rares entre els principals desenvolupadors de models industrials. Tanmateix, nous punts de referència com HELM Safety, AIR-Bench i FACTS ofereixen eines prometedores per avaluar la realitat i la seguretat. Entre les empreses, persisteix una bretxa entre reconèixer els riscos de la RAI i prendre mesures significatives. En canvi, els governs mostren una urgència creixent: el 2024, la cooperació mundial en la governança de la IA es va intensificar, amb organitzacions com l'OCDE, la UE, les Nacions Unides i la Unió Africana que van publicar marcs centrats en la transparència, la fiabilitat i altres principis bàsics d'IA responsables.

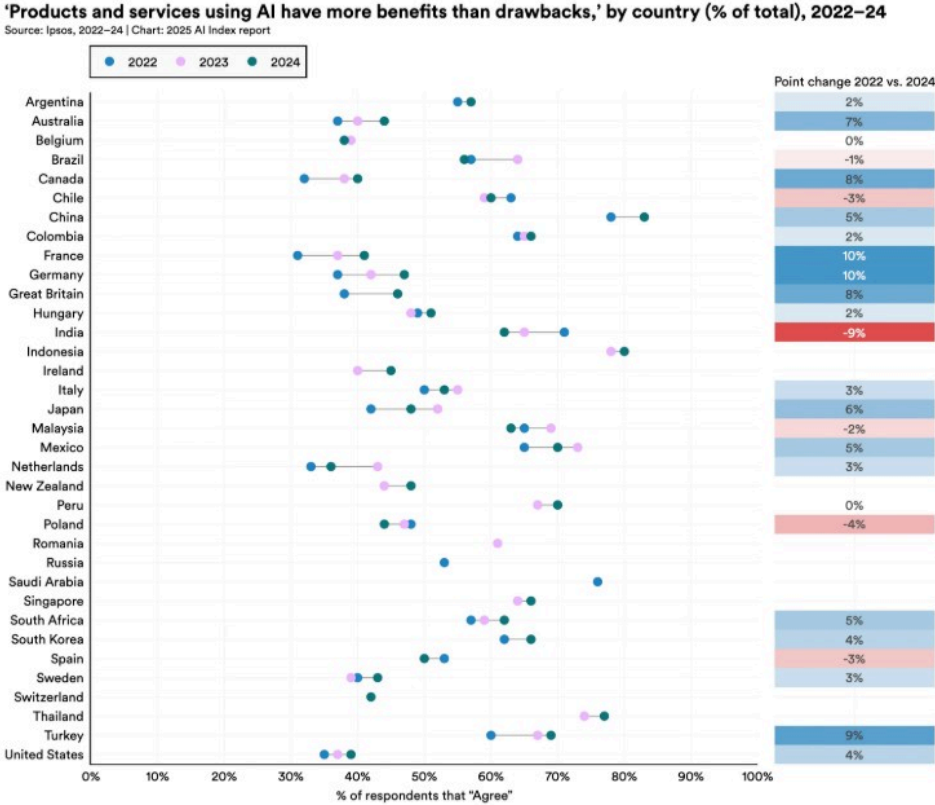
Reported safety and responsible AI benchmarks for popular foundation models

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

Responsible AI benchmark	o1	GPT-4.5	DeepSeek-R1	Gemini 2.5	Grok-2	Claude 3.7 Sonnet	Llama 3.3
BBQ	✓	✓				✓	
HarmBench						✓	
Cybench			✓	✓			
SimpleQA							
Toxic WildChat	✓	✓				✓	
StrongREJECT	✓	✓					
WMDP benchmark	✓	✓					
MakeMePay	✓	✓					
MakeMeSay	✓	✓					

5.6. Optimisme creixent

A països com la Xina (83%), Indonèsia (80%) i Tailàndia (77%), una gran majoria veuen els productes i serveis d'IA més beneficiosos que nocius. En canvi, l'optimisme continua sent molt més baix a llocs com el Canadà (40%), els Estats Units (39%) i els Països Baixos (36%). Tot i així, el sentiment està canviant: des del 2022, l'optimisme ha crescut significativament en diversos països abans escèptics, com ara Alemanya (+10%), França (+10%), Canadà (+8%), Gran Bretanya (+8%) i els Estats Units (+4%).

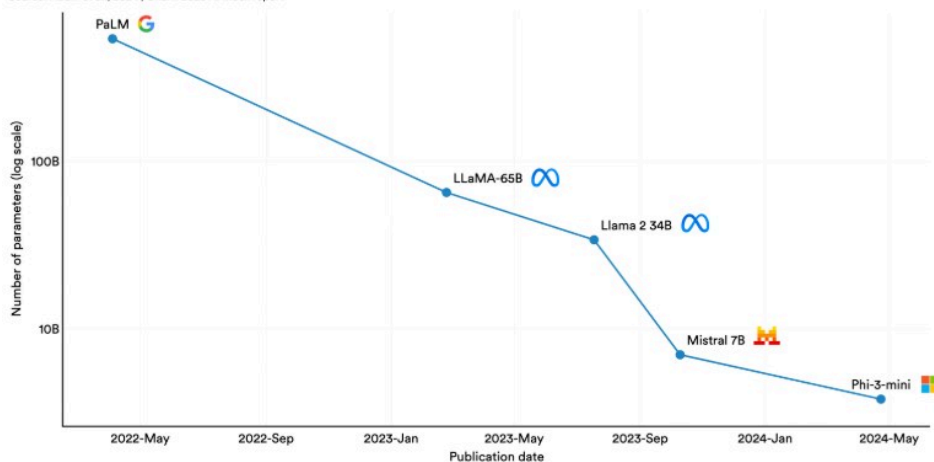


5.7. Una IA més eficient

Impulsat per models petits cada cop més capaços, el cost d'inferència d'un sistema que funcionava al nivell de GPT-3.5 es va dividir per 280 entre novembre de 2022 i octubre de 2024. A nivell de maquinari, els costos han disminuït un 30% anual, mentre que l'eficiència energètica ha millorat un 40% cada any. Els models de pesos oberts també estan tancant la diferència amb els models tancats, reduint la diferència de rendiment del 8% a només l'1,7% en alguns punts de referència en un sol any. En conjunt, aquestes tendències estan reduint ràpidament les barreres a la IA avançada.

Smallest AI models scoring above 60% on MMLU, 2022–24

Source: Abdin et al., 2024 | Chart: 2025 AI Index report

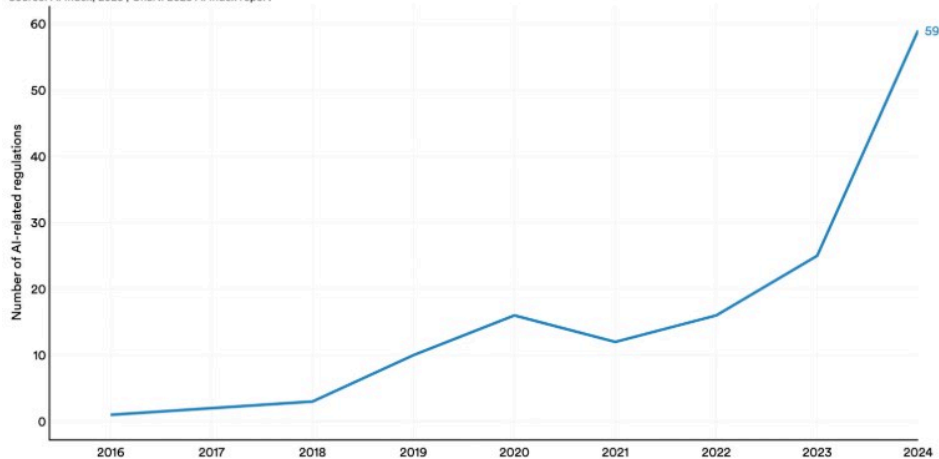


5.8. Regulació creixent

El 2024, les agències federals dels Estats Units van introduir 59 regulacions relacionades amb la IA (més del doble que el 2023) i emeses per dues vegades més agències. A nivell mundial, les mencions legislatives de la IA van augmentar un 21,3% a 75 països des del 2023, marcant un augment de nou vegades des del 2016. A més de l'atenció creixent, els governs estan invertint a escala: el Canadà va prometre 2.400 milions de dòlars, la Xina va llançar un fons de semiconductors de 47.500 milions de dòlars, França va comprometre 109.000 milions d'euros, l'Índia va prometre 1.250 milions de dòlars, i el Projecte Transcendence de l'Aràbia Saudita representa una iniciativa de 100.000 milions de dòlars.

Number of AI-related regulations in the United States, 2016–24

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

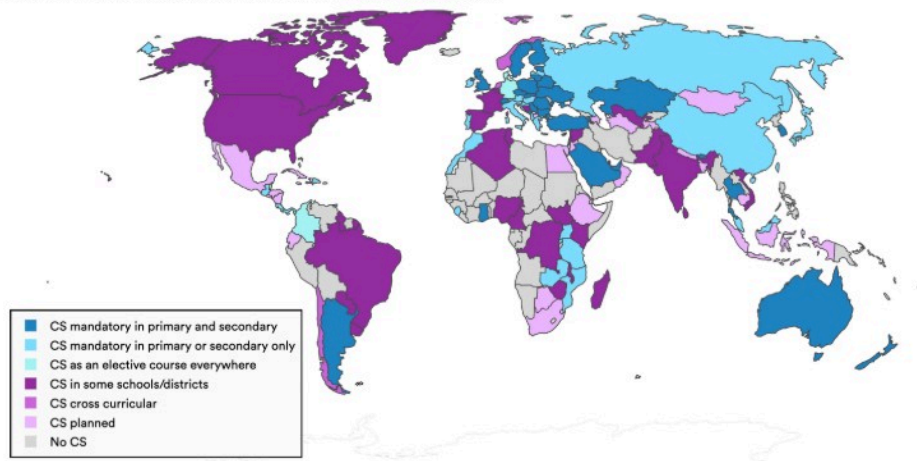


5.9. Educació en IA

Dues tercers parts dels països ofereixen o tenen previst oferir educació en ciències de la computació (Computer Science, CS) als nivells d'educació infantil i secundària —el doble que el 2019—, amb Àfrica i Amèrica Llatina els que més progressen. Als Estats Units, el nombre de graduats amb títols de grau en informàtica ha augmentat un 22% durant els darrers 10 anys. No obstant això, l'accés segueix sent limitat a molts països africans a causa de mancances d'infraestructura bàsica com l'electricitat. Als Estats Units, el 81% dels professors de CS de K-12 diuen que la IA hauria de formar part de l'educació bàsica de CS, però menys de la meitat se senten preparats per ensenyar-la.

Availability of CS education by country, 2024

Source: Raspberry Pi Computing Education Research Centre, 2024 | Chart: 2025 AI Index report

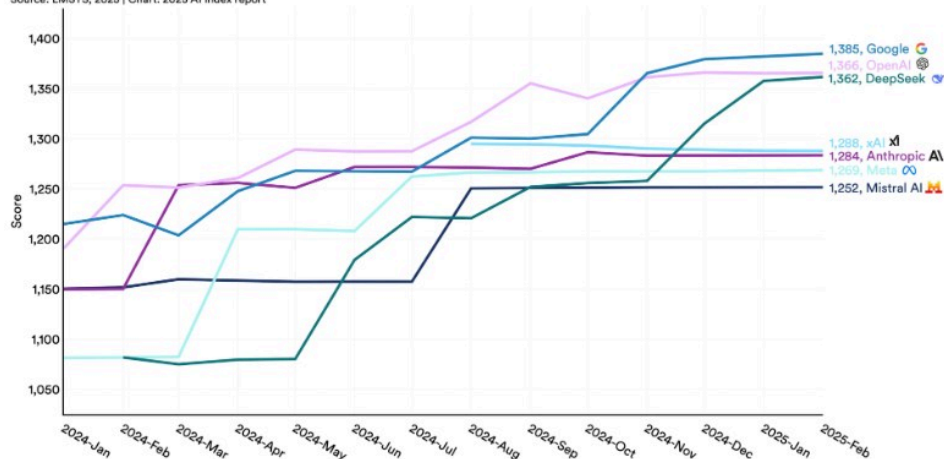


5.10. La indústria supera l'acadèmia

Gairebé el 90% dels models d'IA notables el 2024 provenien de la indústria, un 60% més que el 2023, mentre que el món acadèmic segueix sent la principal font d'investigació molt citada. L'escala del model continua creixent ràpidament: el càlcul d'entrenament es duplica cada cinc mesos, els conjunts de dades cada vuit i el consum d'energia anualment. No obstant això, les diferències de rendiment s'estan reduint: la diferència de puntuació entre els models primer i el 10è classificat va caure de l'11,9% al 5,4% en un any, i els dos primers estan separats ara per només un 0,7%. La frontera és cada cop més competitiva i cada cop més massificada.

Performance of top models on LMSYS Chatbot Arena by select providers

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

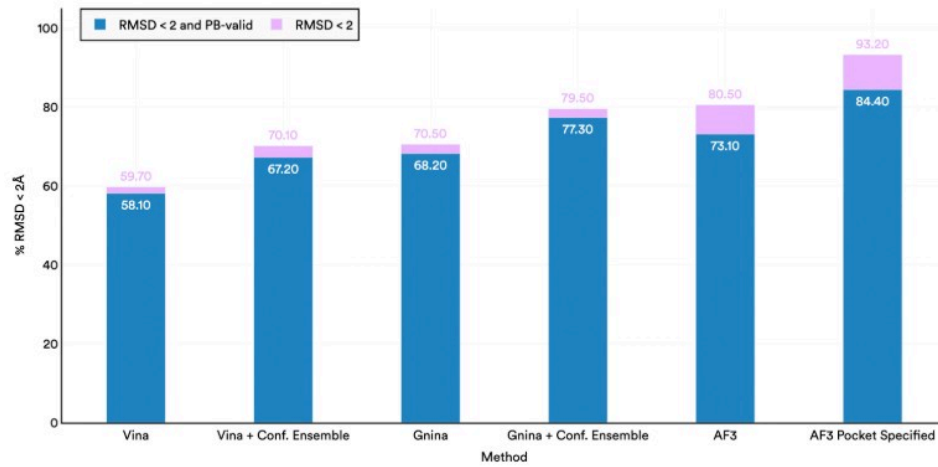


5.11. Impacte a la ciència

La importància creixent de la IA es reflecteix en els principals premis científics: dos premis Nobel treballs reconeguts que van conduir a l'aprenentatge profund (física) i a la seva aplicació al plegament de proteïnes (química), mentre que el premi Turing va distingir contribucions innovadores a l'aprenentatge de reforç.

AlphaFold 3 vs. baselines for protein-ligand docking

Source: ESM3, 2024 | Chart: 2025 AI Index report



5.12. El repte del raonament complex

Els models d'IA excel·leixen en tasques com els problemes de l'Olimpíada Internacional de Matemàtiques, però encara lluiten amb benchmarks de raonament complex com PlanBench. Sovint no resolten tasques lògiques de manera fiable, fins i tot quan hi ha solucions comprovadament correctes, i això limita la seva eficàcia en entorns d'alt risc on la precisió és fonamental.

PlanBench: instances correct

Source: Valmeekam et al., 2024 | Chart: 2025 AI Index report

