

The AI Index: A Compass for Navigating AI's Future



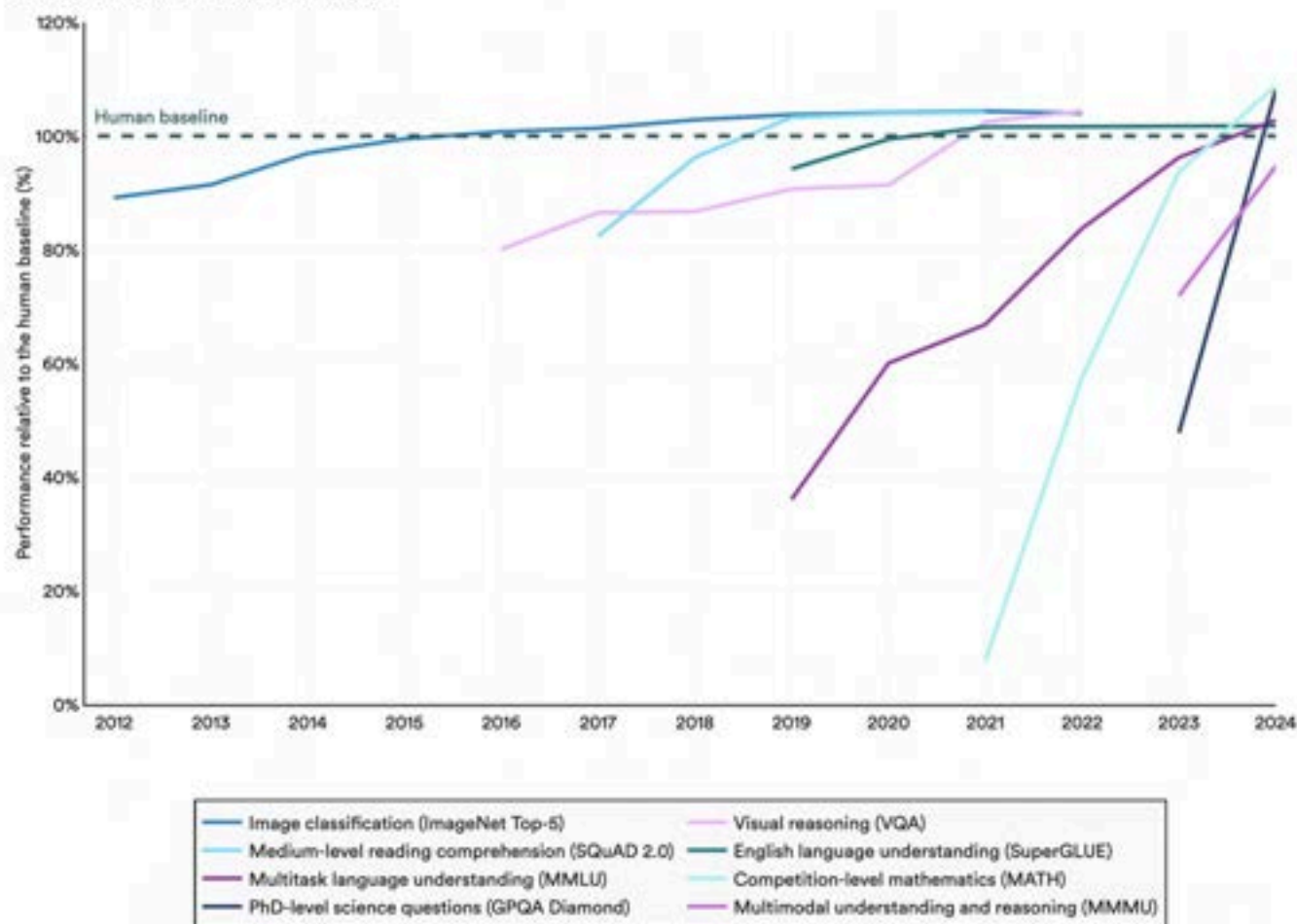
Top Takeaways

1. AI performance on demanding benchmarks continues to improve.

In 2023, researchers introduced new benchmarks—MMMU, GPQA, and SWE-bench—to test the limits of advanced AI systems. Just a year later, performance sharply increased: scores rose by 18.8, 48.9, and 67.3 percentage points on MMMU, GPQA, and SWE-bench, respectively. Beyond benchmarks, AI systems made major strides in generating high-quality video, and in some settings, language model agents even outperformed humans in programming tasks with limited time budgets.

Select AI Index technical performance benchmarks vs. human performance

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

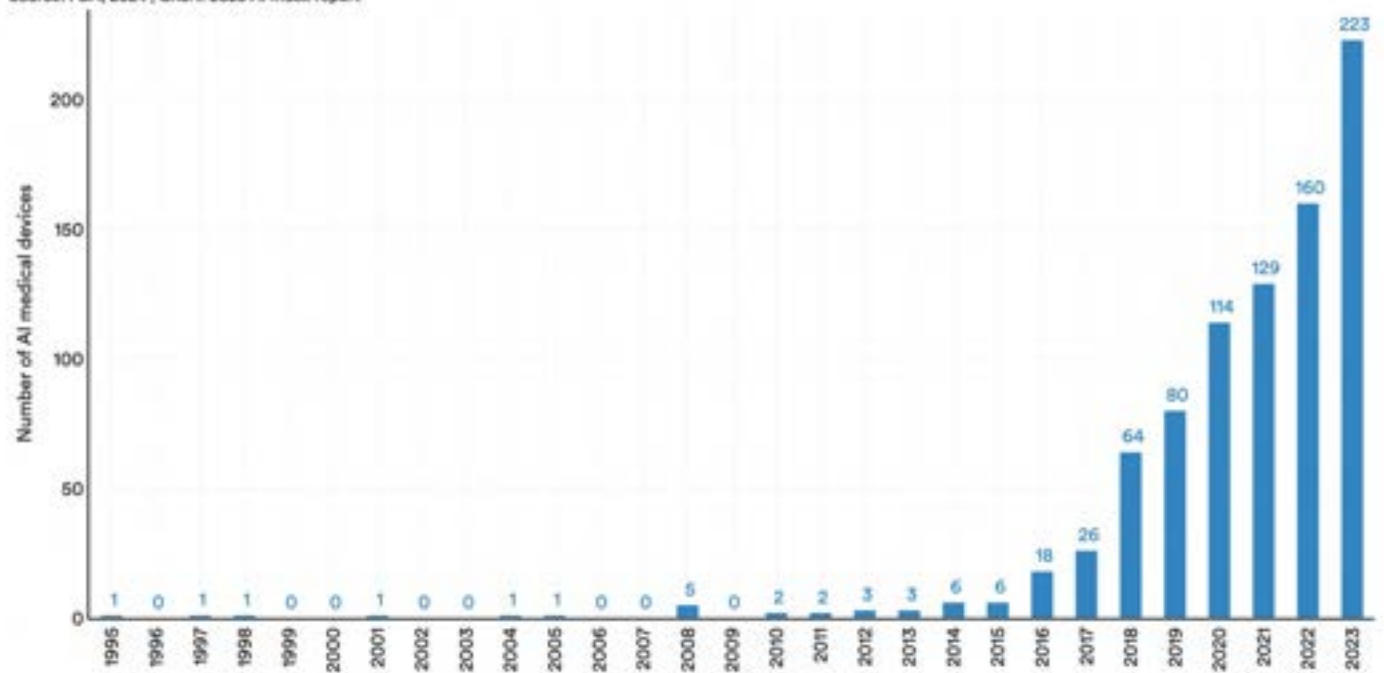


2. AI is increasingly embedded in everyday life.

From healthcare to transportation, AI is rapidly moving from the lab to daily life. In 2023, the FDA approved 223 AI-enabled medical devices, up from just six in 2015. On the roads, self-driving cars are no longer experimental: Waymo, one of the largest U.S. operators, provides over 150,000 autonomous rides each week, while Baidu's affordable Apollo Go robotaxi fleet now serves numerous cities across China.

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

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

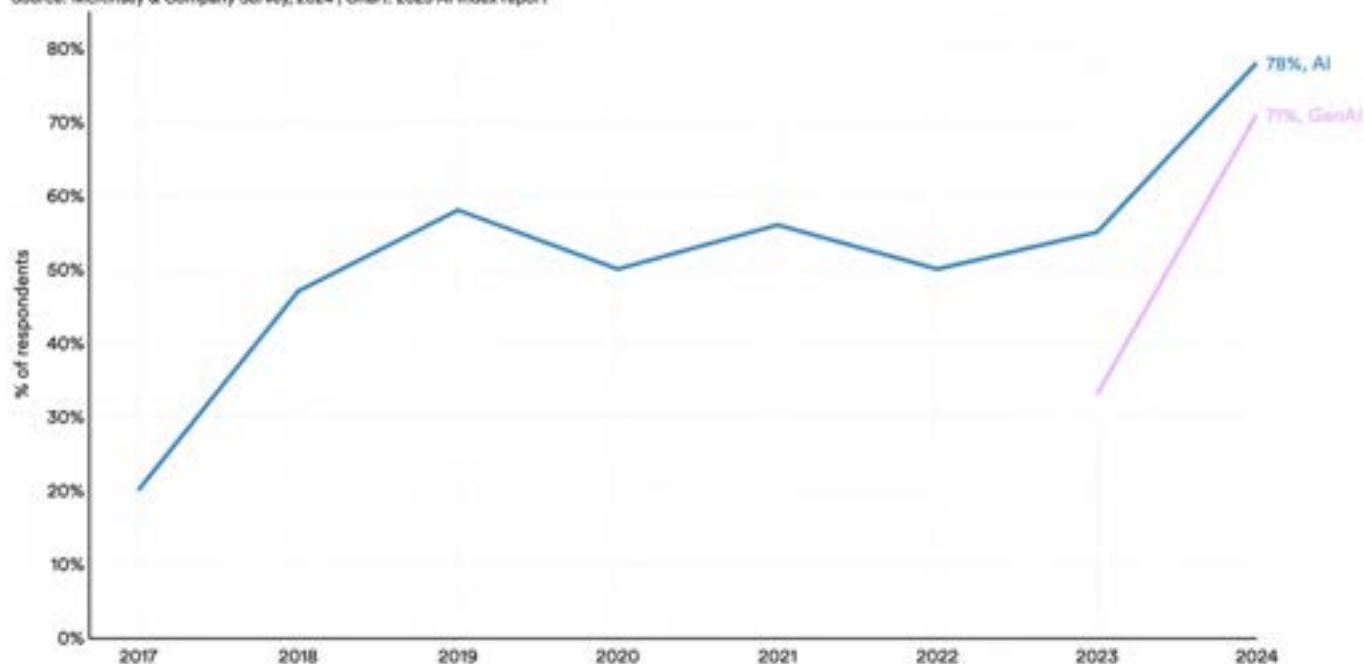


3. Business is all in on AI, fueling record investment and usage, as research continues to show strong productivity impacts.

In 2024, U.S. private AI investment grew to \$109.1 billion—nearly 12 times China’s \$9.3 billion and 24 times the U.K.’s \$4.5 billion. Generative AI saw particularly strong momentum, attracting \$33.9 billion globally in private investment—an 18.7% increase from 2023. AI business usage is also accelerating: 78% of organizations reported using AI in 2024, up from 55% the year before. Meanwhile, a growing body of research confirms that AI boosts productivity and, in most cases, helps narrow skill gaps across the workforce.

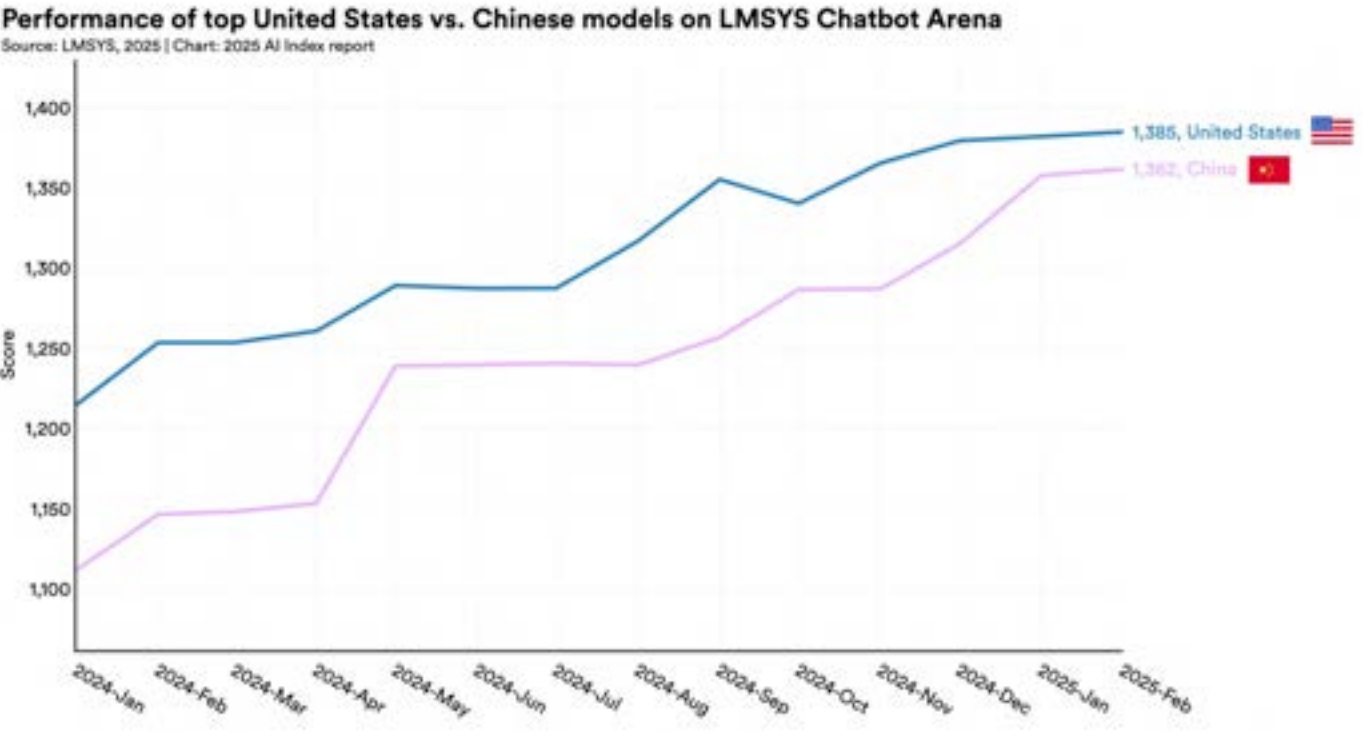
Share of respondents who say their organization uses AI in at least one function, 2017–24

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



4. The U.S. still leads in producing top AI models—but China is closing the performance gap.

In 2024, U.S.-based institutions produced 40 notable AI models, significantly outpacing China's 15 and Europe's three. While the U.S. maintains its lead in quantity, Chinese models have rapidly closed the quality gap: performance differences on major benchmarks such as MMLU and HumanEval shrank from double digits in 2023 to near parity in 2024. Meanwhile, China continues to lead in AI publications and patents. At the same time, model development is increasingly global, with notable launches from regions such as the Middle East, Latin America, and Southeast Asia.



5. The responsible AI ecosystem evolves—unevenly.

AI-related incidents are rising sharply, yet standardized RAI evaluations remain rare among major industrial model developers. However, new benchmarks like HELM Safety, AIR-Bench, and FACTS offer promising tools for assessing factuality and safety. Among companies, a gap persists between recognizing RAI risks and taking meaningful action. In contrast, governments are showing increased urgency: In 2024, global cooperation on AI governance intensified, with organizations including the OECD, EU, U.N., and African Union releasing frameworks focused on transparency, trustworthiness, and other core responsible AI principles.

Reported safety and responsible AI benchmarks for popular foundation models

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

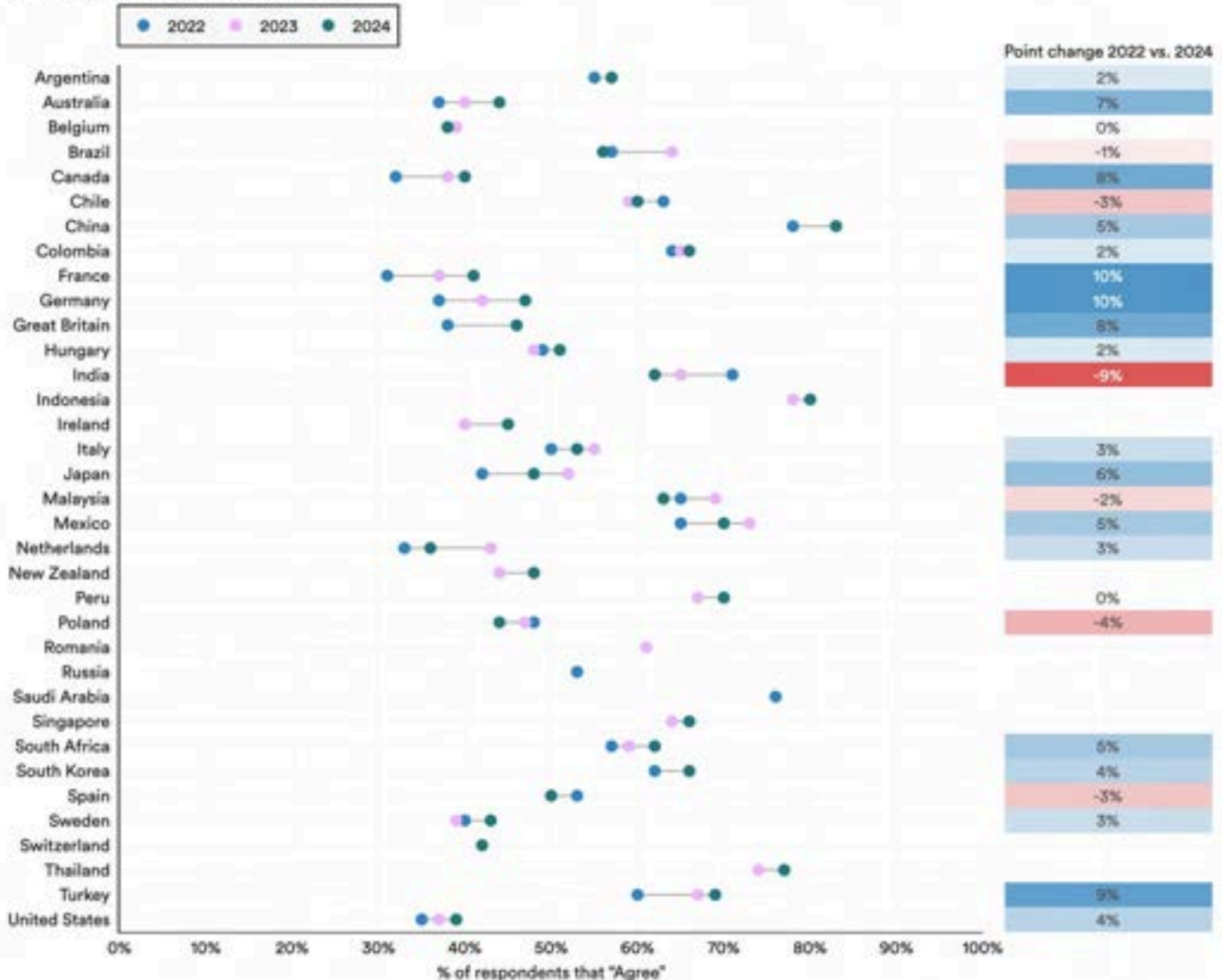
Responsible AI benchmark	o1	GPT-4.5	DeepSeek-R1	Gemini 2.5	Grok-2	Claude 3.7 Sonnet	Llama 3.3
BBQ	✓	✓				✓	
HarmBench							
Cybench						✓	
SimpleQA			✓	✓			
Toxic WildChat	✓	✓				✓	
StrongREJECT	✓	✓					
WMDP benchmark	✓	✓					
MakeMePay	✓	✓					
MakeMeSay	✓	✓					

6. Global AI optimism is rising—but deep regional divides remain.

In countries like China (83%), Indonesia (80%), and Thailand (77%), strong majorities see AI products and services as more beneficial than harmful. In contrast, optimism remains far lower in places like Canada (40%), the United States (39%), and the Netherlands (36%). Still, sentiment is shifting: since 2022, optimism has grown significantly in several previously skeptical countries—including Germany (+10%), France (+10%), Canada (+8%), Great Britain (+8%), and the United States (+4%).

'Products and services using AI have more benefits than drawbacks,' by country (% of total), 2022–24

Source: Ipsos, 2022–24 | Chart: 2025 AI Index report

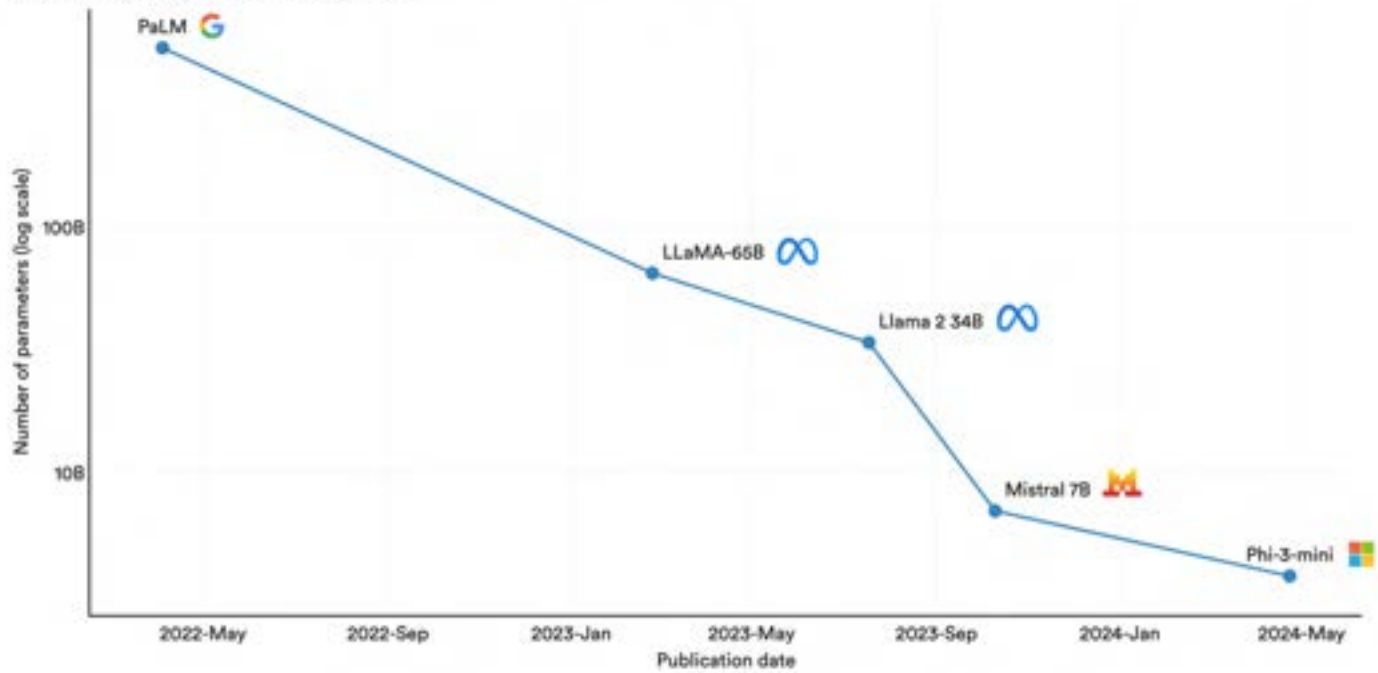


7. AI becomes more efficient, affordable and accessible.

Driven by increasingly capable small models, the inference cost for a system performing at the level of GPT-3.5 dropped over 280-fold between November 2022 and October 2024. At the hardware level, costs have declined by 30% annually, while energy efficiency has improved by 40% each year. Open-weight models are also closing the gap with closed models, reducing the performance difference from 8% to just 1.7% on some benchmarks in a single year. Together, these trends are rapidly lowering the barriers to advanced AI.

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

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

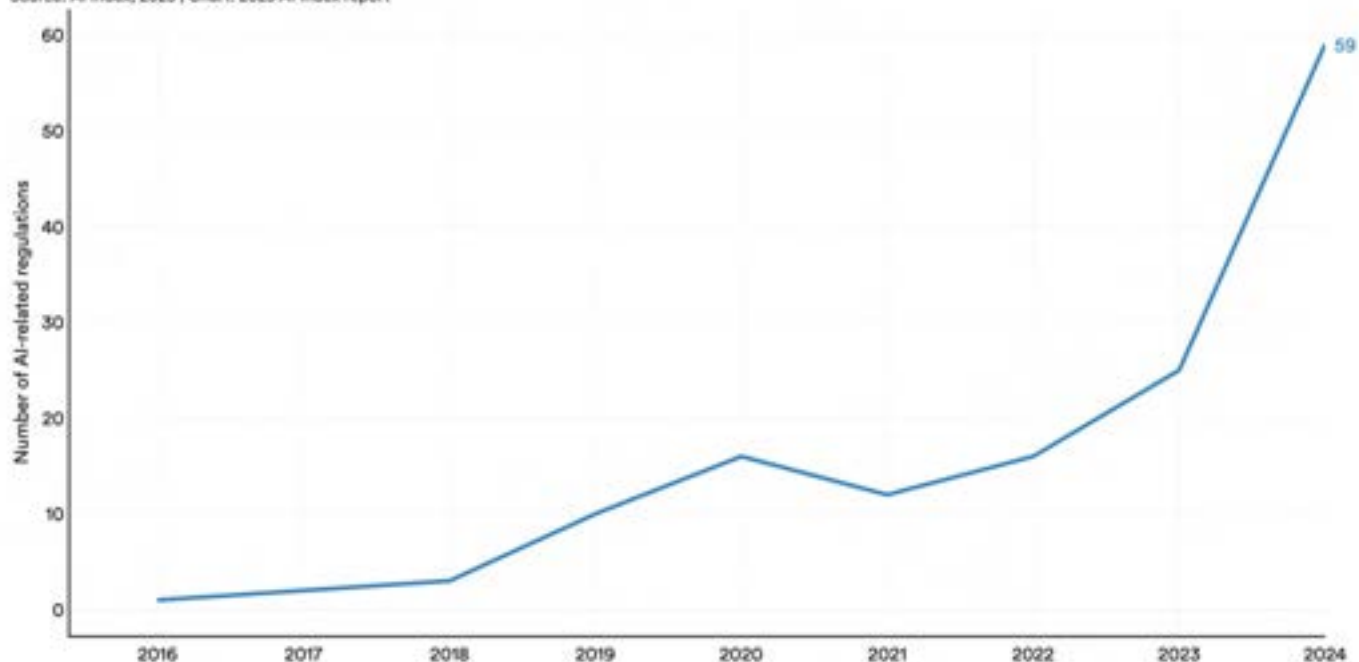


8. Governments are stepping up on AI— with regulation and investment.

In 2024, U.S. federal agencies introduced 59 AI-related regulations—more than double the number in 2023—and issued by twice as many agencies. Globally, legislative mentions of AI rose 21.3% across 75 countries since 2023, marking a ninefold increase since 2016. Alongside growing attention, governments are investing at scale: Canada pledged \$2.4 billion, China launched a \$47.5 billion semiconductor fund, France committed €109 billion, India pledged \$1.25 billion, and Saudi Arabia's Project Transcendence represents a \$100 billion initiative.

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

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

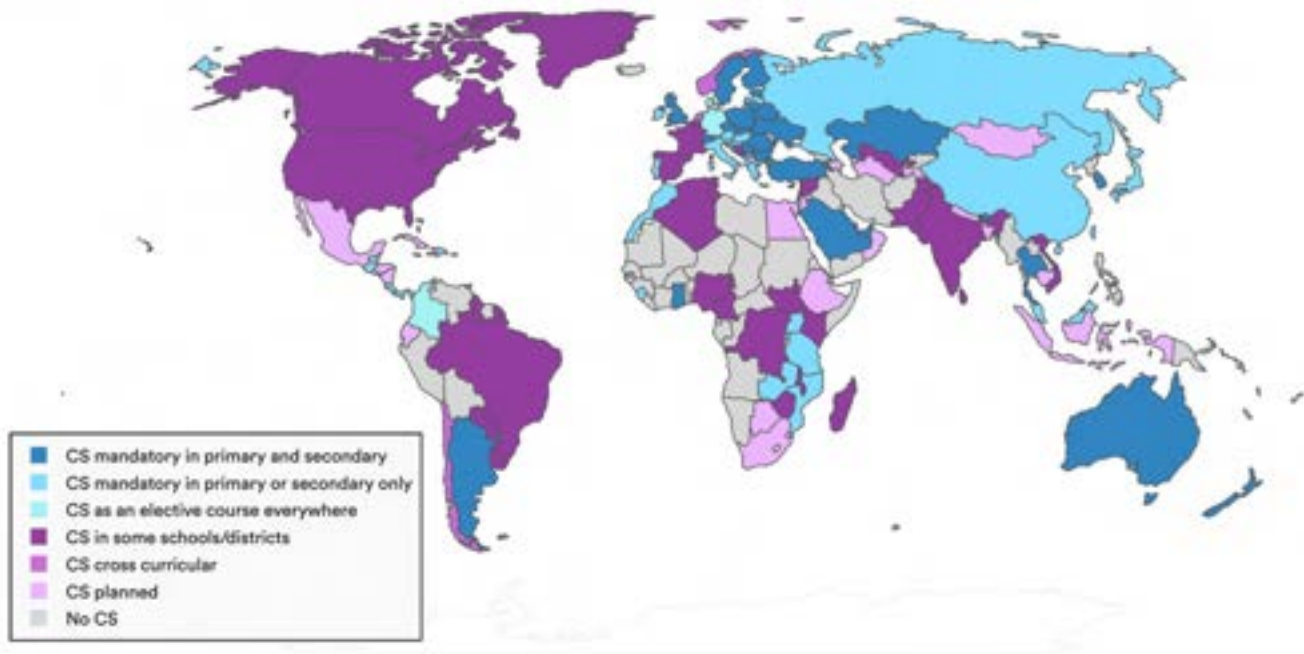


9. AI and computer science education is expanding—but gaps in access and readiness persist.

Two-thirds of countries now offer or plan to offer K–12 CS education—twice as many as in 2019—with Africa and Latin America making the most progress. In the U.S., the number of graduates with bachelor’s degrees in computing has increased 22% over the last 10 years. Yet access remains limited in many African countries due to basic infrastructure gaps like electricity. In the U.S., 81% of K–12 CS teachers say AI should be part of foundational CS education, but less than half feel equipped to teach it.

Availability of CS education by country, 2024

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

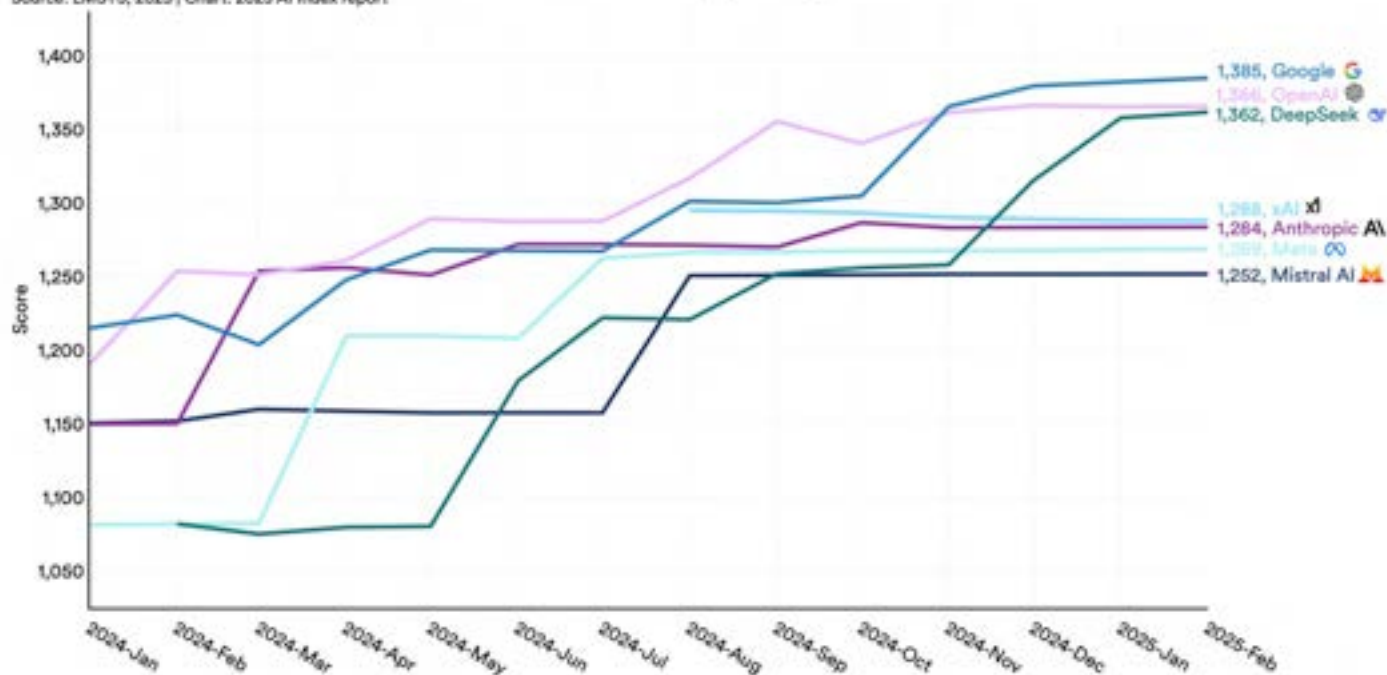


10. Industry is racing ahead in AI—but the frontier is tightening.

Nearly 90% of notable AI models in 2024 came from industry, up from 60% in 2023, while academia remains the top source of highly cited research. Model scale continues to grow rapidly—training compute doubles every five months, datasets every eight, and power use annually. Yet performance gaps are shrinking: the score difference between the top and 10th-ranked models fell from 11.9% to 5.4% in a year, and the top two are now separated by just 0.7%. The frontier is increasingly competitive—and increasingly crowded.

Performance of top models on LMSYS Chatbot Arena by select providers

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

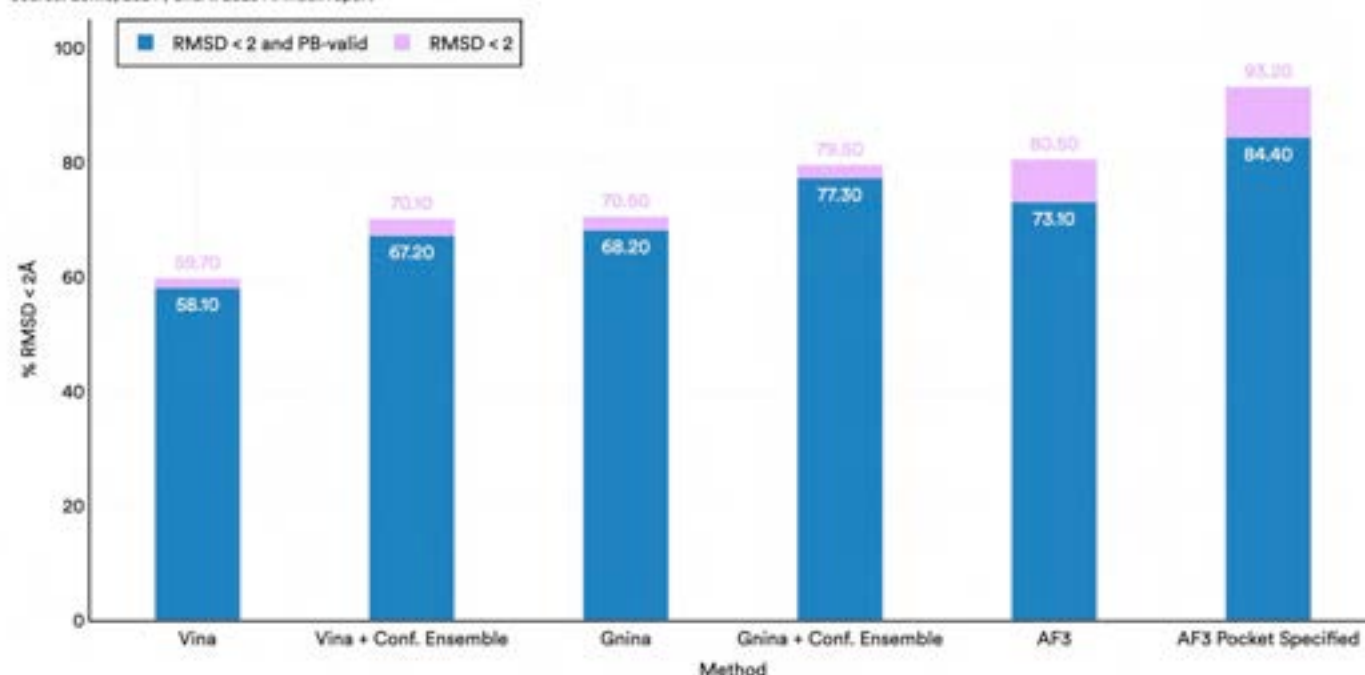


11. AI earns top honors for its impact on science.

AI's growing importance is reflected in major scientific awards: two Nobel Prizes recognized work that led to deep learning (physics), and to its application to protein folding (chemistry), while the Turing Award honored groundbreaking contributions to reinforcement learning.

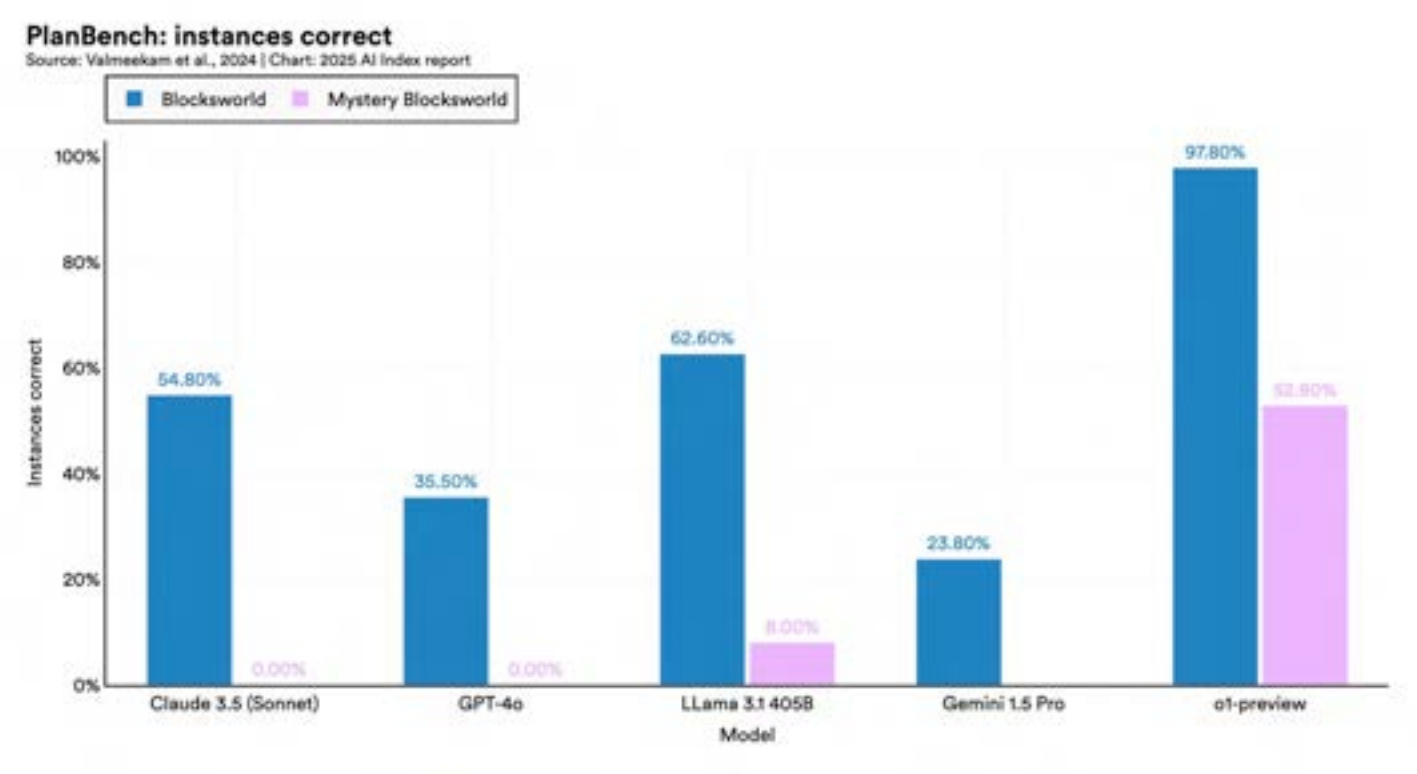
AlphaFold 3 vs. baselines for protein-ligand docking

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



12. Complex reasoning remains a challenge.

AI models excel at tasks like International Mathematical Olympiad problems but still struggle with complex reasoning benchmarks like PlanBench. They often fail to reliably solve logic tasks even when provably correct solutions exist, limiting their effectiveness in high-stakes settings where precision is critical.



Chapter Lineup

Chapter 1: Research and Development

This chapter explores trends in AI research and development, beginning with an analysis of AI publications, patents, and notable AI systems.



Measuring trends in Intelligence

The AI Index report tracks, collates, distills, and visualizes data related to artificial intelligence (AI). Our mission is to provide unbiased, rigorously vetted, broadly sourced data in order for policymakers, researchers, executives, journalists, and the general