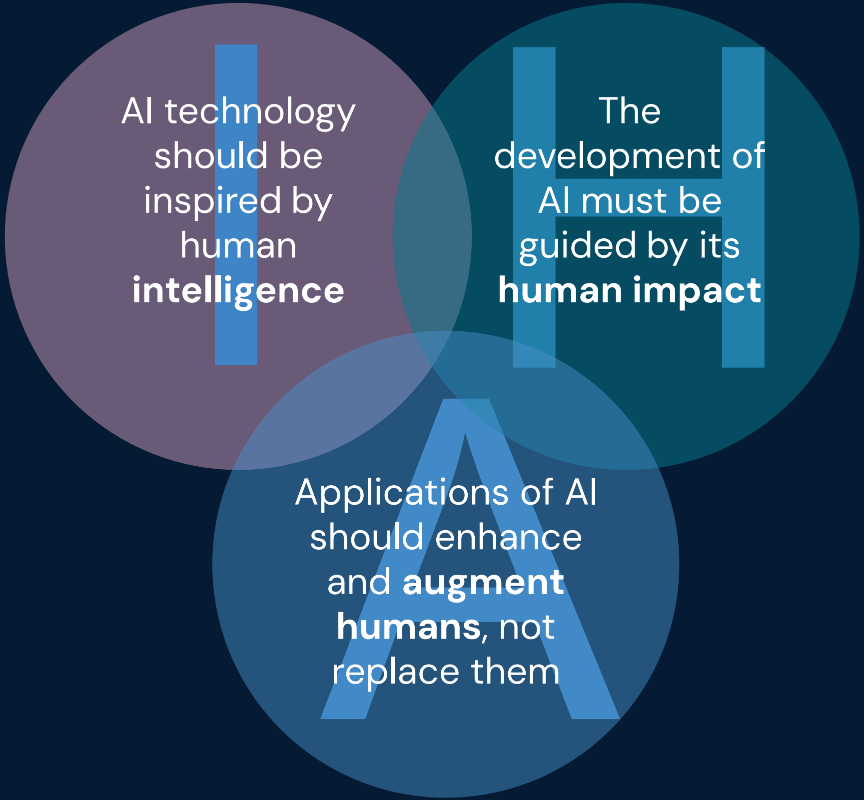
Stanford University
Human-Centered
Artificial Intelligence

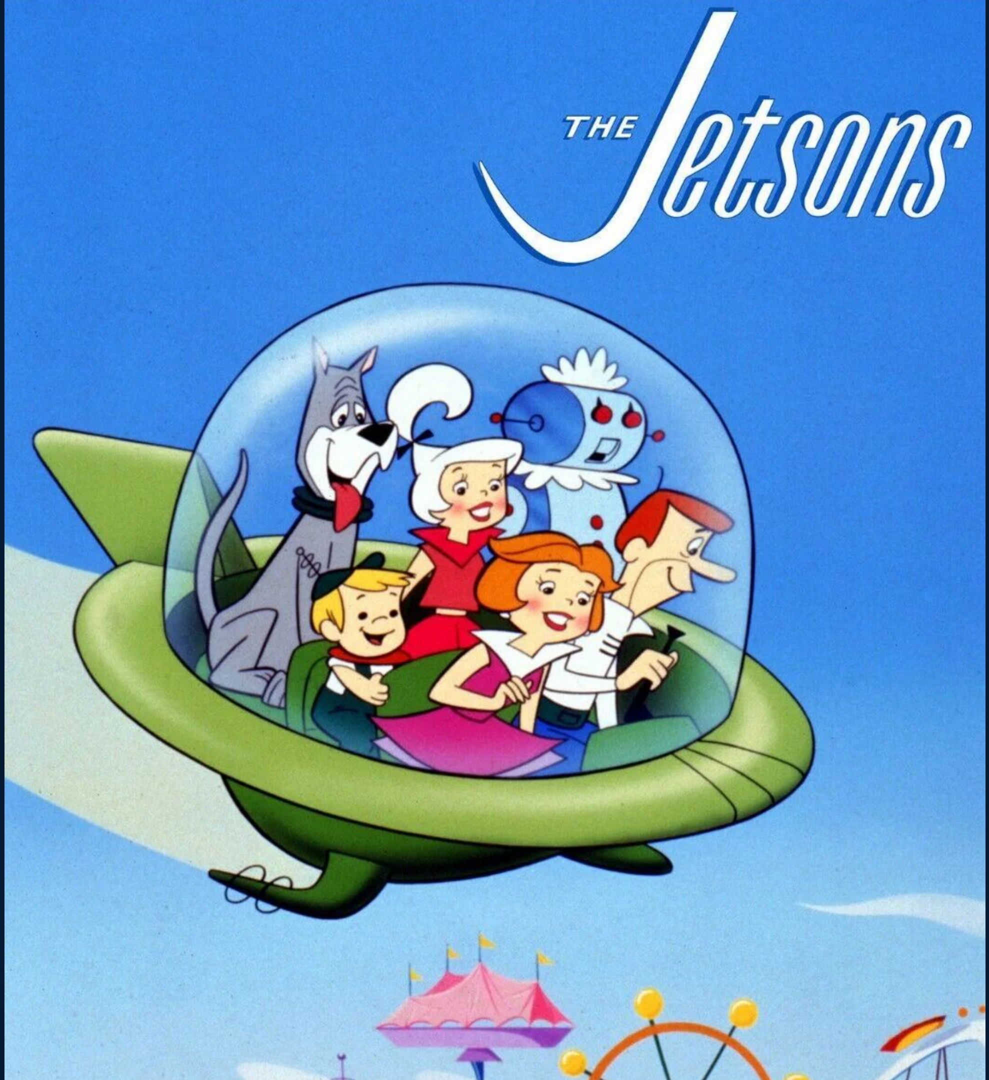
World Summit AI 2025

Vanessa Parli, Director of Research Programs

Our Goal

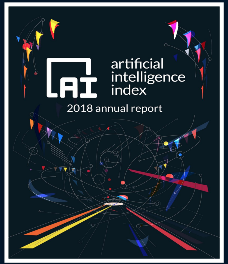
To advance AI **research**, **education**, **policy** and **practice** to improve the human condition

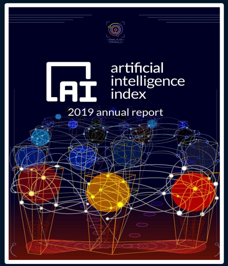


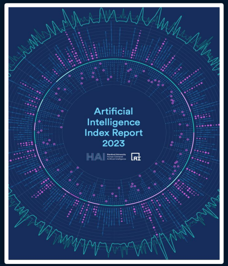

PLUG BACK IN
A NETFLIX SERIES
BLACK MIRROR

THE Jetsons

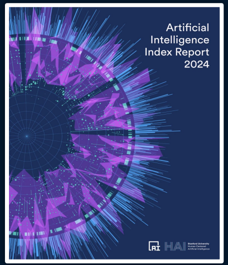
2025

8th Edition

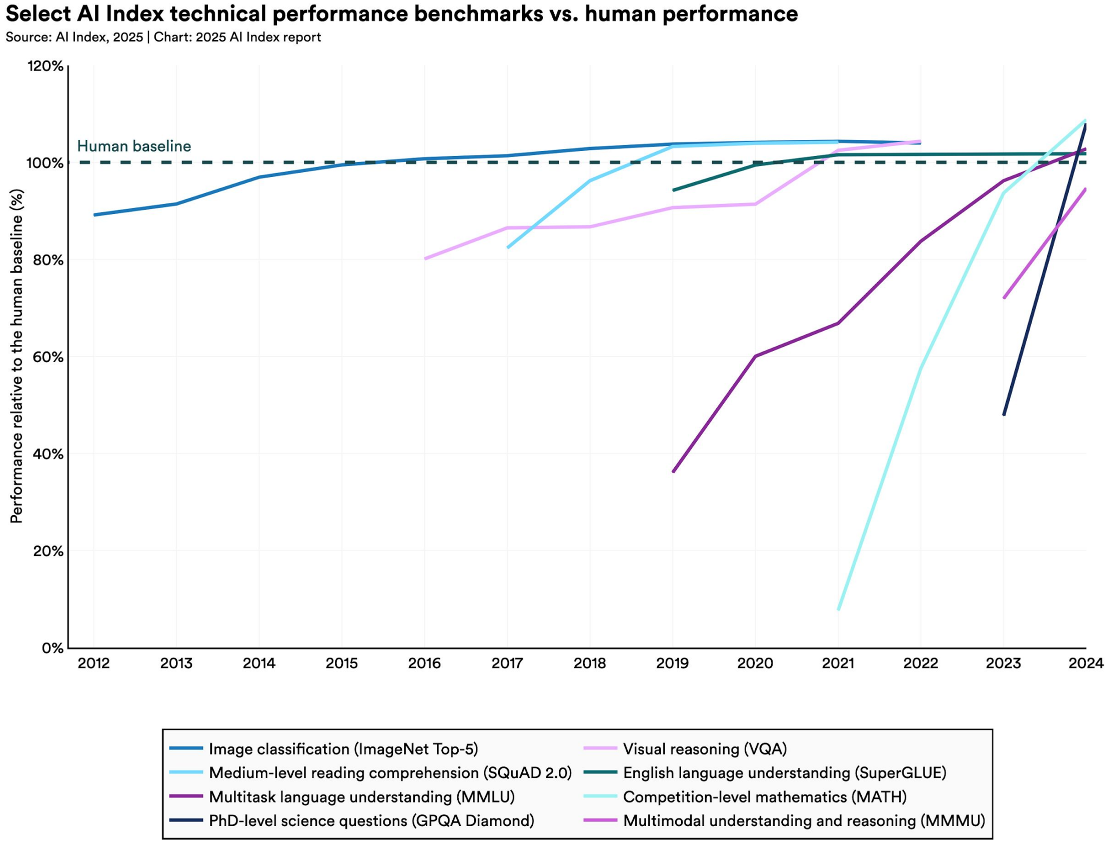








AI continues to exceed humans on benchmarks

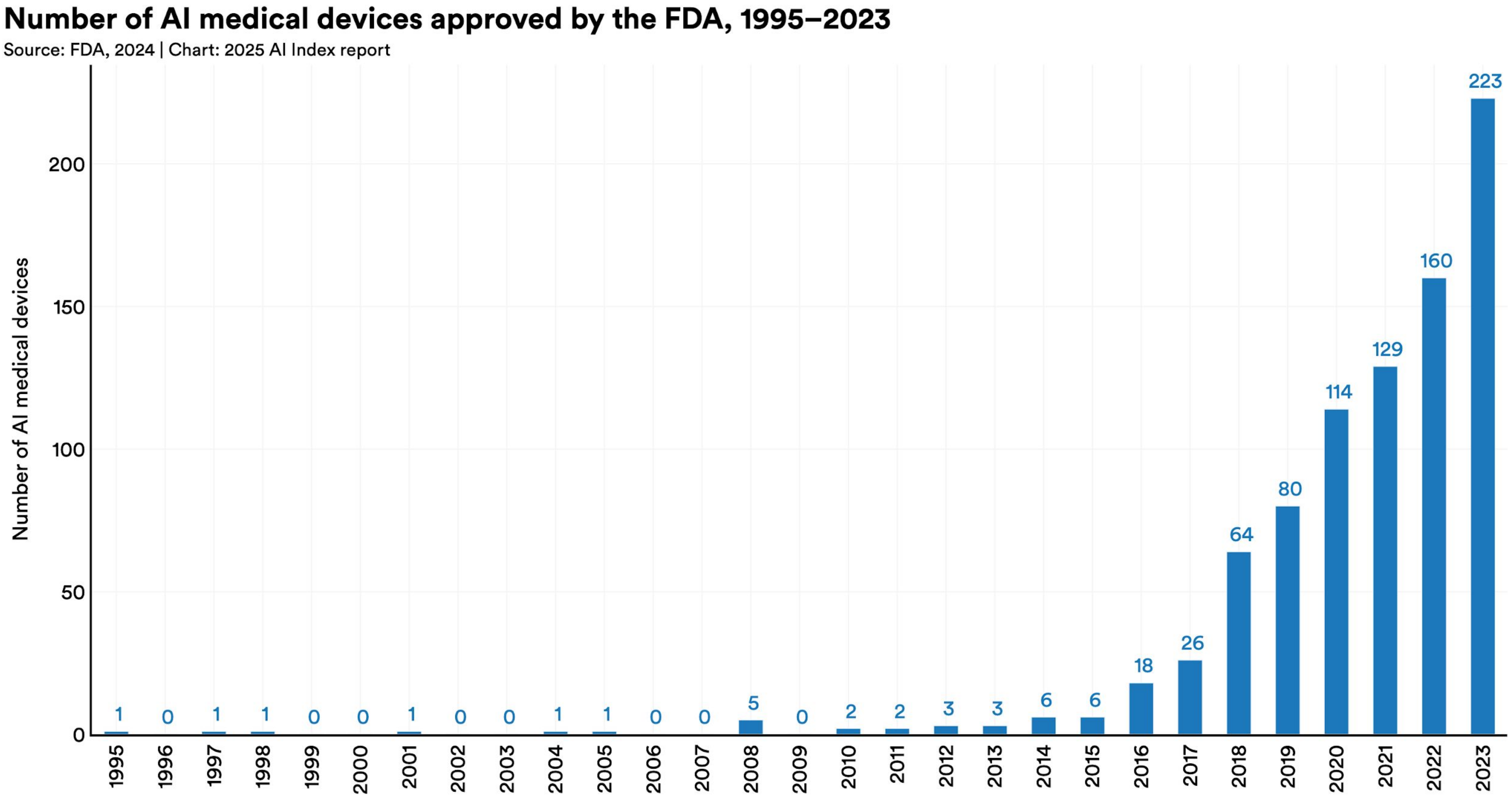


GPQA, + 49%

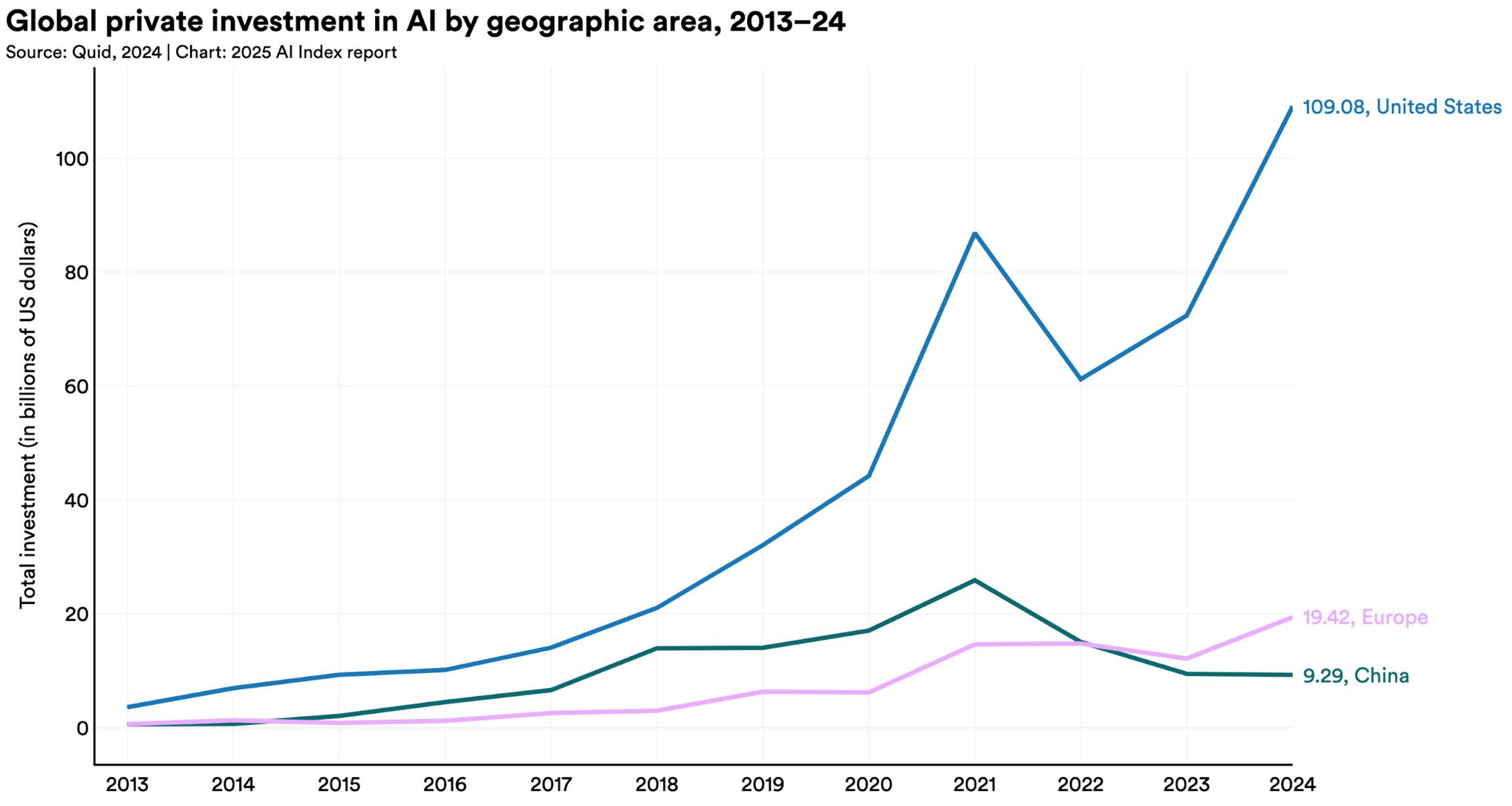
MMMU, +19%

SWE-bench, +67%

AI is increasingly embedded in everyday life



Business is all in on AI



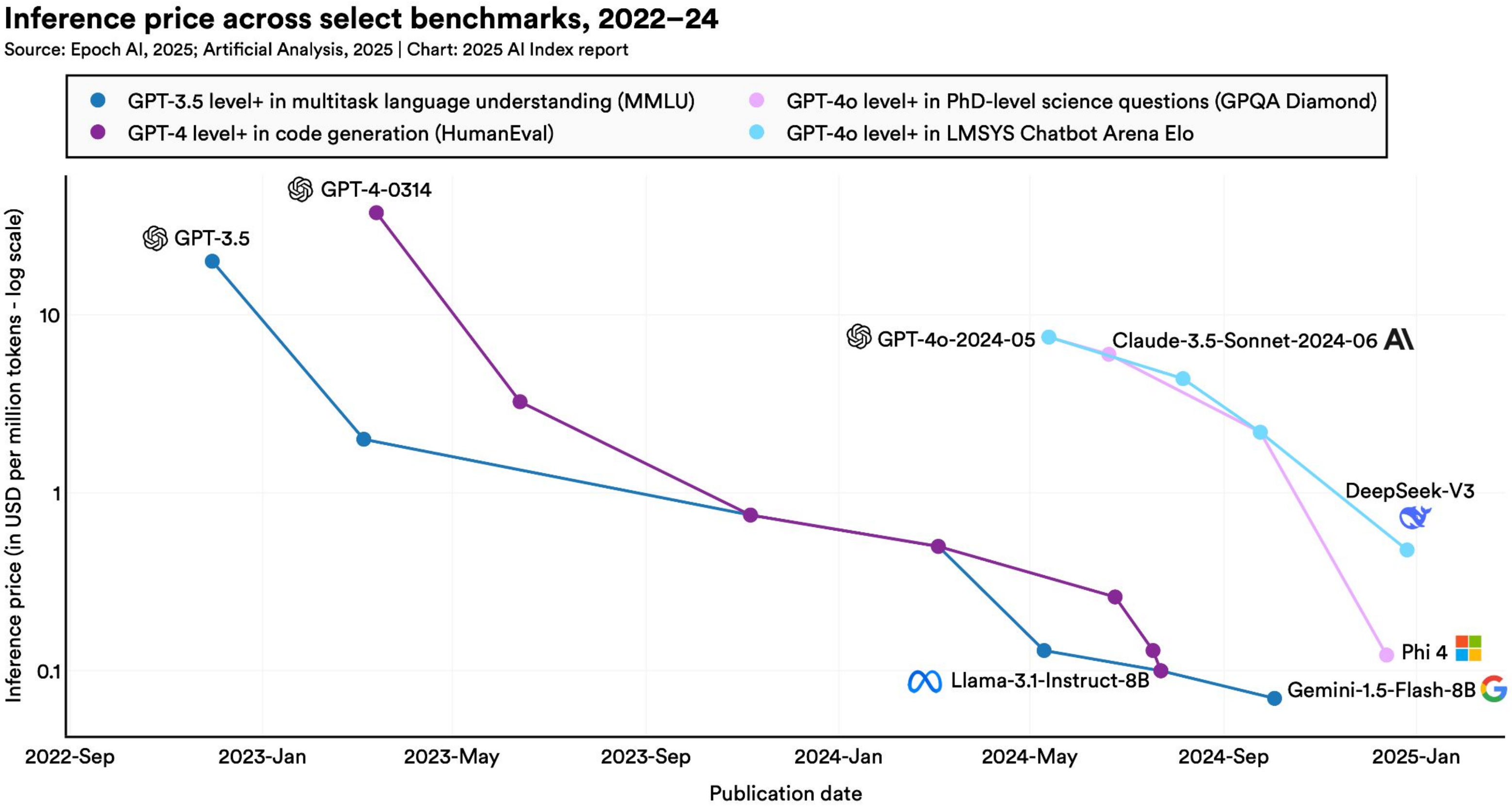
And we are *beginning* to understand impact

AI's productivity equalizing effects

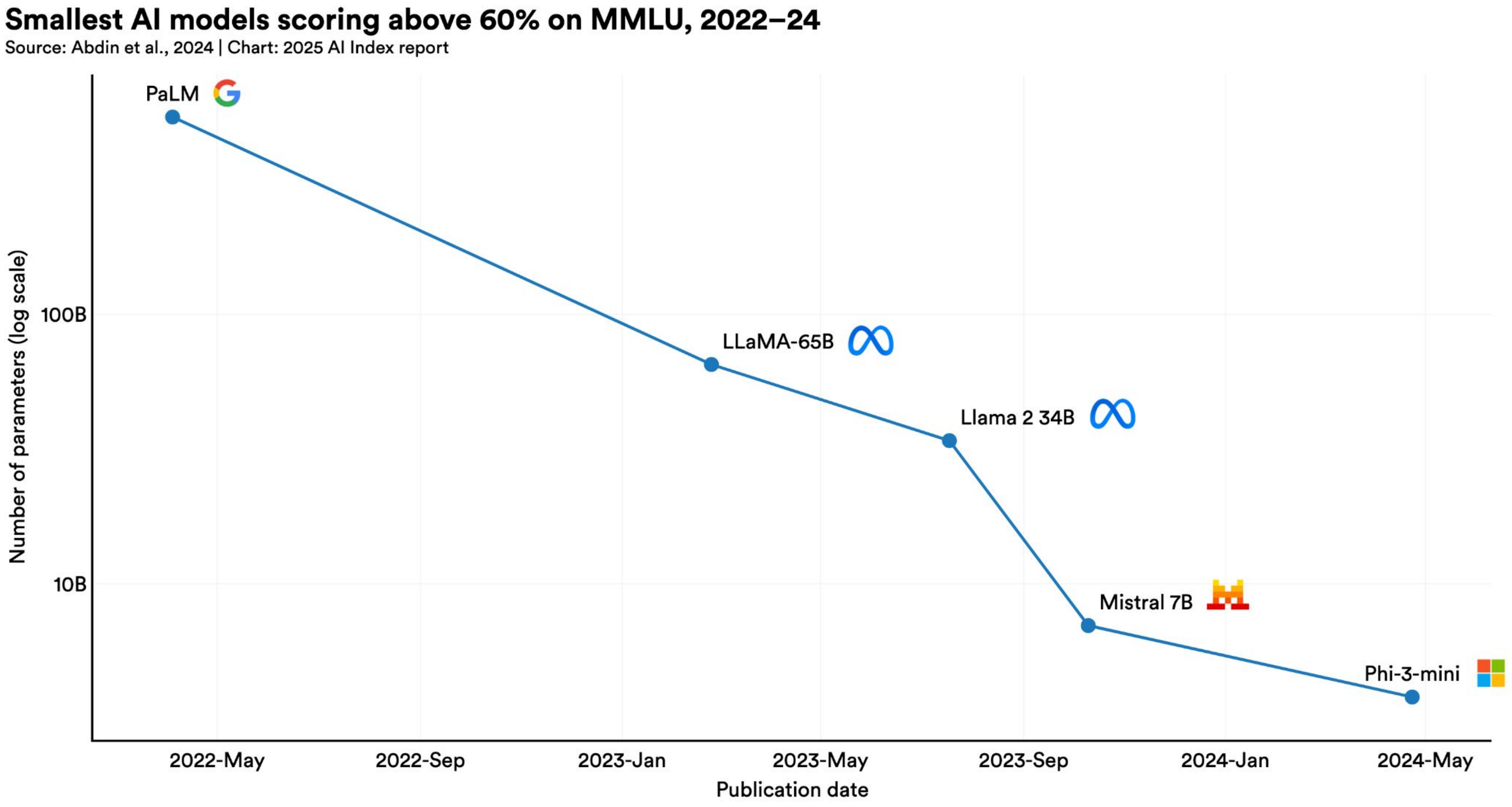
Study	Task	Low-skill worker productivity gain	High-skill worker productivity gain
Brynjolfsson et al., 2023	Customer support	34%	Indistinguishable from zero
Dell'Acqua et al., 2023	Consulting	42.96%	16.5%
Cui et al., 2024	Software engineering	21–40%	7–16%
Hoffman et al., 2024	Software engineering	12–27%	5–10%

Figure 4.4.10

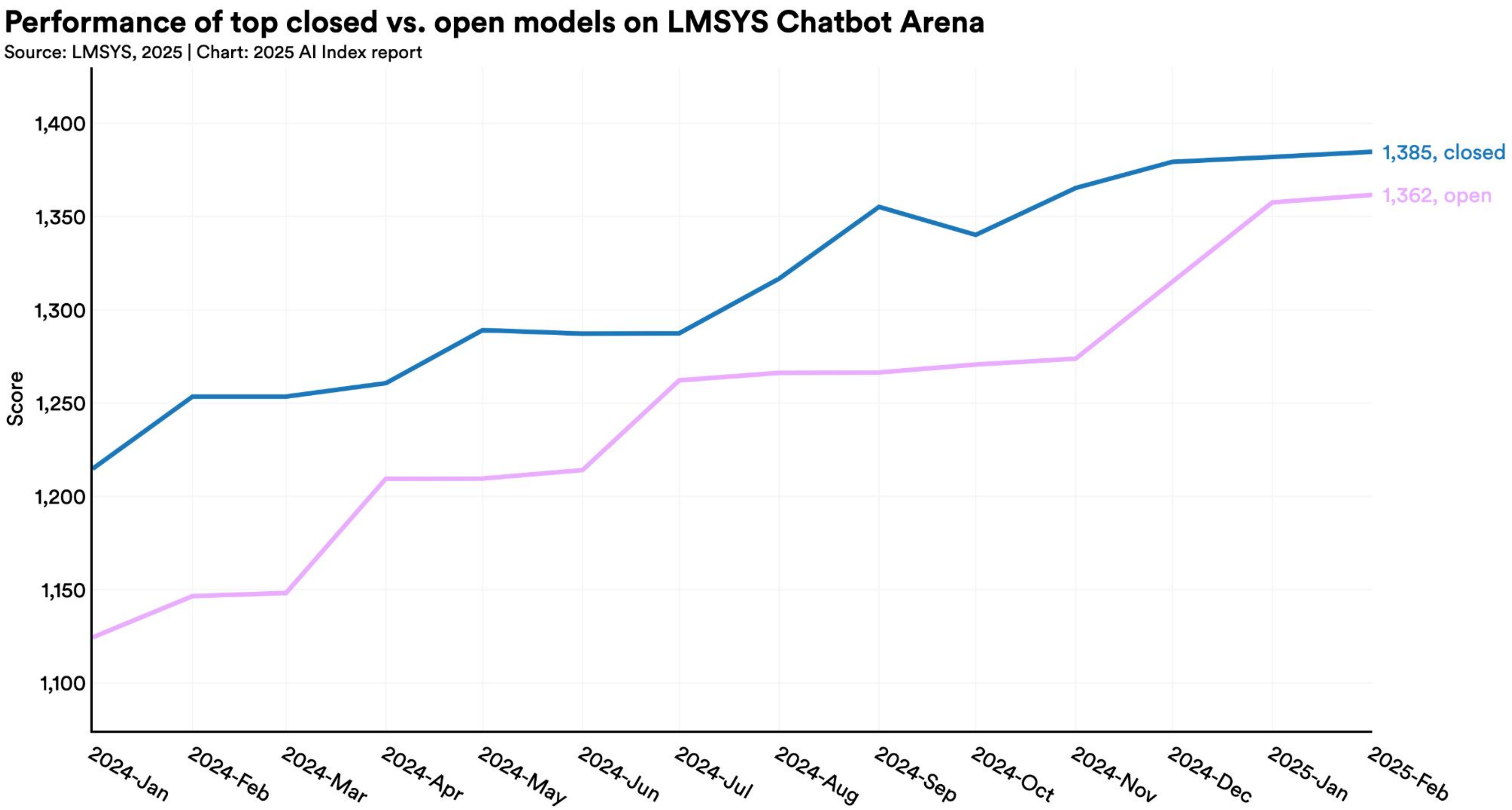
AI becomes more efficient, affordable and accessible



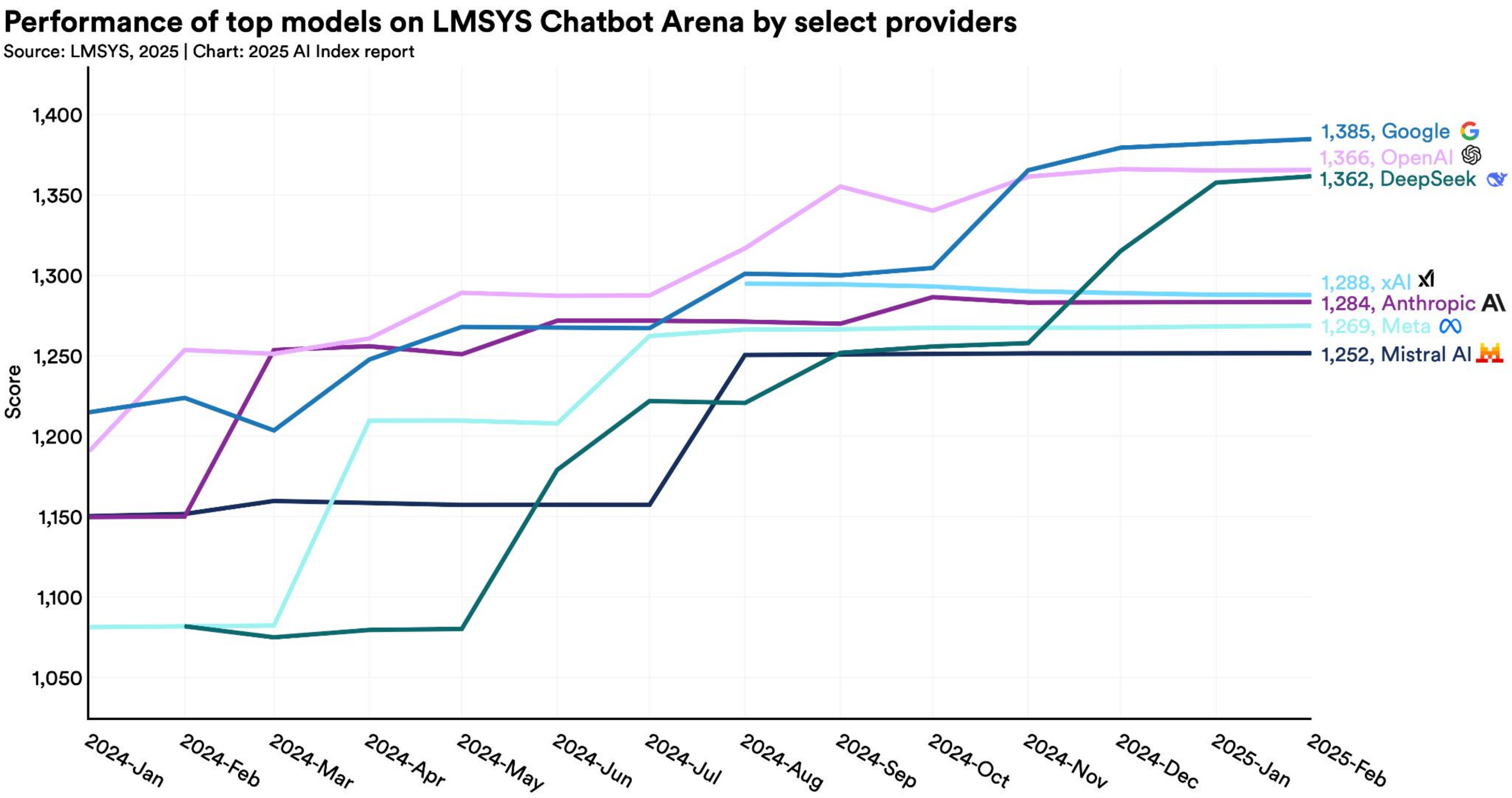
AI becomes more efficient, affordable and accessible



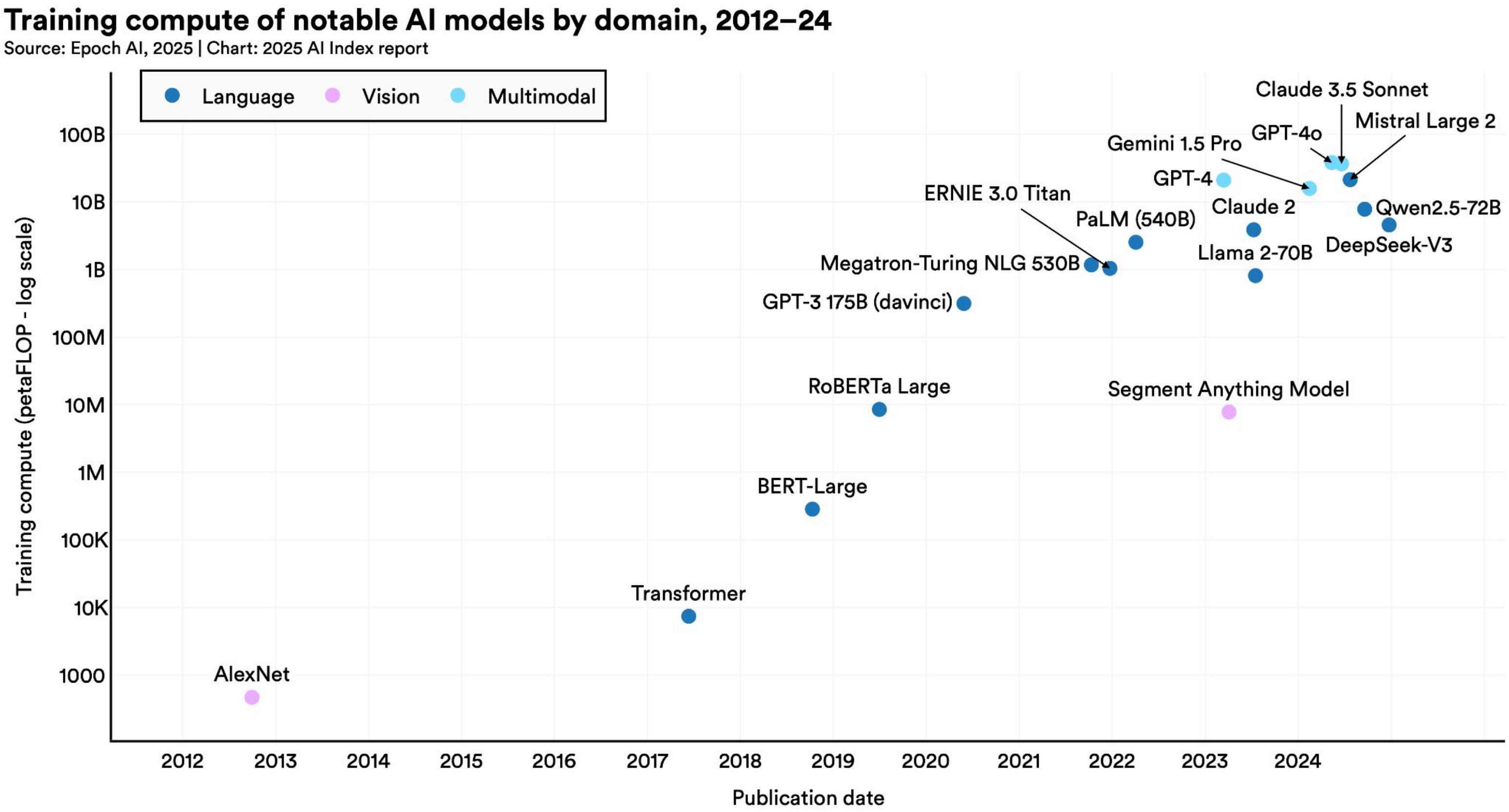
AI becomes more efficient, affordable and accessible



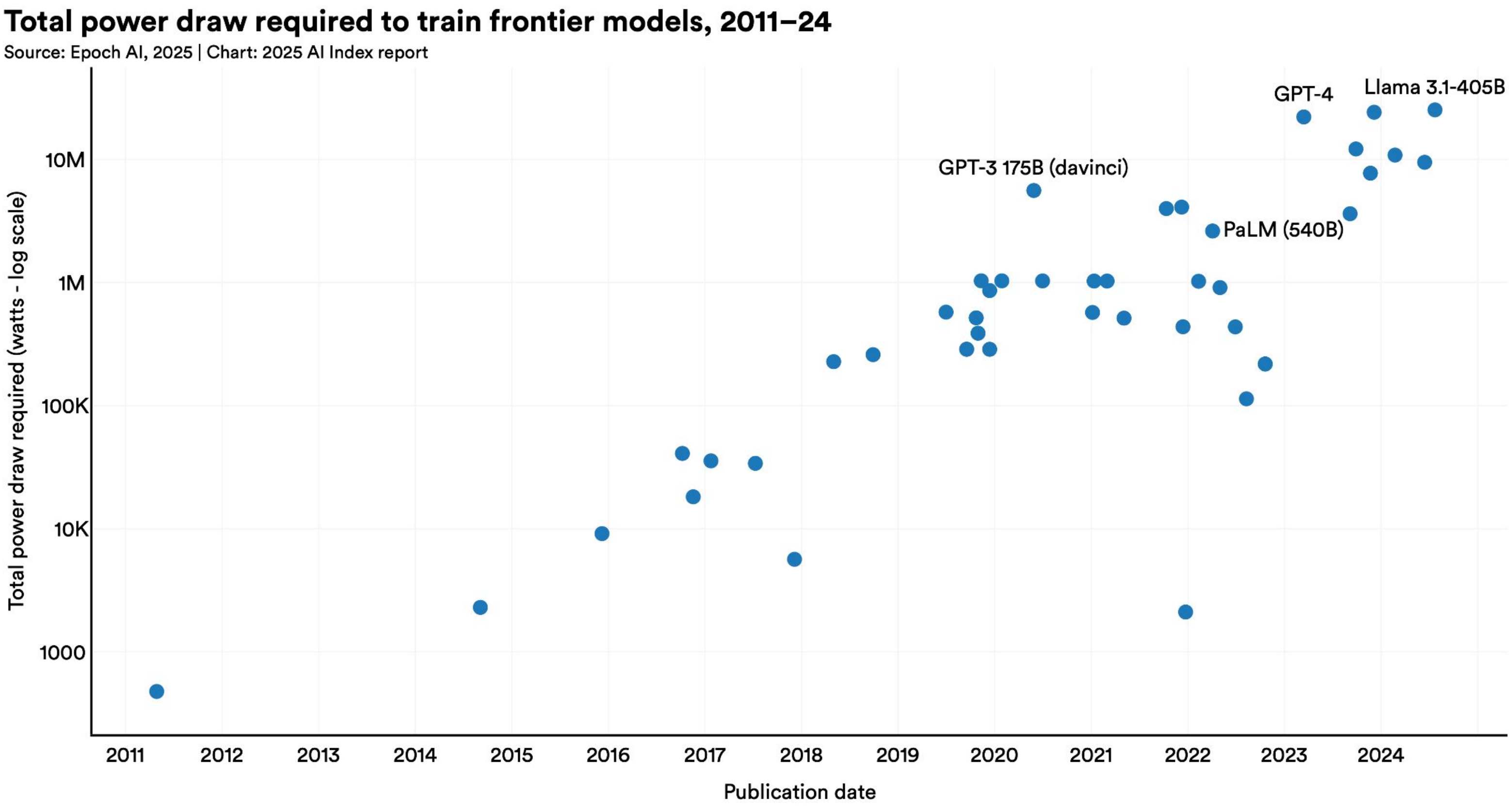
The frontier is tightening



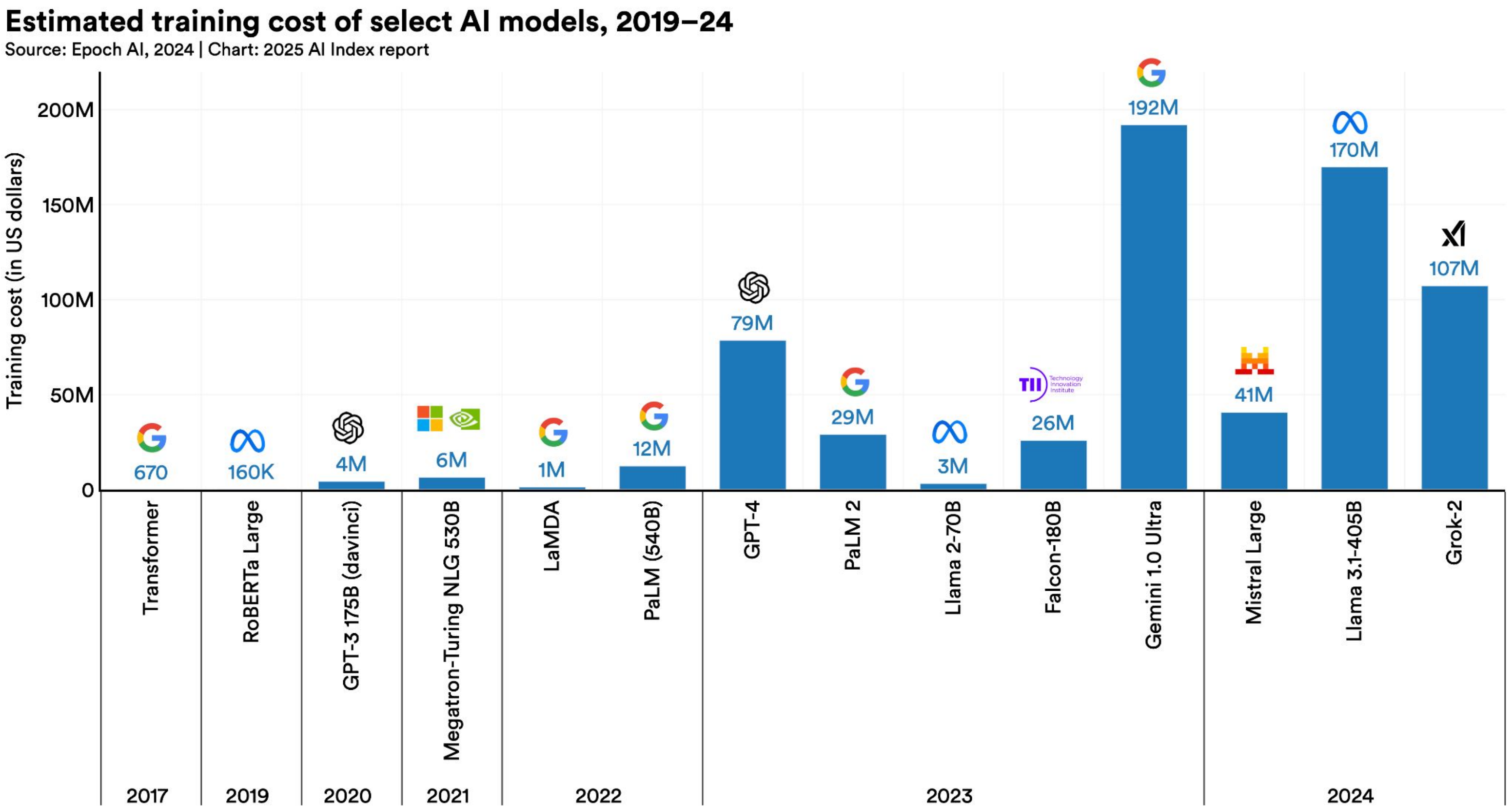
But, resource requirements remain significant



But, resource requirements remain significant



But, resource requirements remain significant



Responsible AI benchmarks evolve

Reported general capability benchmarks for popular foundation models

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

Capability benchmark	o1	GPT-4.5	DeepSeek-R1	Gemini 2.5	Grok-2	Claude 3.7 Sonnet	Llama 3.3
MMLU, MMLU-Pro or MMMLU	✓	✓	✓	✓	✓	✓	✓
GPQA or GPQA-Diamond	✓	✓	✓	✓	✓	✓	✓
MATH-500	✓		✓		✓	✓	✓
AIME 2024	✓	✓	✓	✓		✓	
SWE-bench verified	✓	✓	✓	✓		✓	
MMMU	✓	✓		✓	✓	✓	

Responsible AI benchmarks evolve

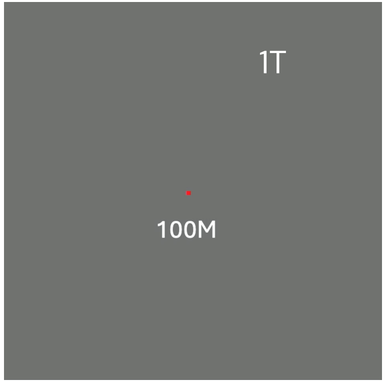
Reported safety and responsible AI benchmarks for popular foundation models
Source: AI Index, 2025 | Table: 2025 AI Index report

RAI 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	✓	✓					

Research –
Inspiring Action

The science of intelligence

- Data efficiency
- Energy efficiency
- Beyond evolution
- Explainability
- Melding mind and machine



Surya Ganguli
Associate Professor of Applied Physics, Senior Fellow at the Stanford Institute for Human-Centered AI and Associate Professor, by courtesy, of Neurobiology and of Electrical Engineering

Identifying culture in language models

Question

Do you personally believe that drinking alcohol is morally acceptable, morally unacceptable, or is it not a moral issue?

A. Morally acceptable

B. Morally unacceptable

C. Not a moral issue

D. Depends on the situation

Existing language models were less aligned on ‘everyday’ judgments of value, compared to the more high-stakes political value judgments.


Tatsu Hashimoto
Assistant Professor of Computer Science

Jeanne Tsai
Dunlevie Family Professor, Professor of Psychology

Human-Centered AI

Proliferation of AI for Good Institutes

HAI
Stanford University
Human-Centered
Artificial Intelligence

AJL

BROWN
HCRI
HUMANITY CENTERED
ROBOTICS INITIATIVE
RHODE ISLAND

UNIVERSITY OF
CANBERRA
HUMAN CENTRED
TECHNOLOGY
RESEARCH CLUSTER

Future
of Humanity
Institute
UNIVERSITY OF OXFORD

NUS National University of Singapore
NUS BUSINESS SCHOOL
Centre on AI Technology
for Humankind

Chicago Human+AI

VUB
AI
XPERIENCE
CENTRE

Australian
National
University
NATURAM PRIMUM COGNOSCERE RERUM
HMI

Human-AI Interaction @ UBC

r/ai
center
for
responsible
ai

UNIVERSITY OF AMSTERDAM
RPA
Humane AI

CHAI
Center for
Human-Compatible
Artificial
Intelligence

JOHNS HOPKINS
INSTITUTE *for*
ASSURED AUTONOMY

Human-centered AI

SOCIETY-CENTERED

↕

COMMUNITY-CENTERED

↕

USER-CENTERED

James Landay
Denning Co-Director of Stanford HAI, Anand Rajaraman and Venky Harinarayan Professor of Computer Science

What does Human-Centered AI mean in your field?

RESOURCES

- Sign up for the HAI newsletter
- Attend free public HAI events
- Explore professional development opportunities

vparli@stanford.edu