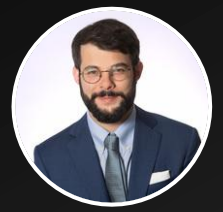
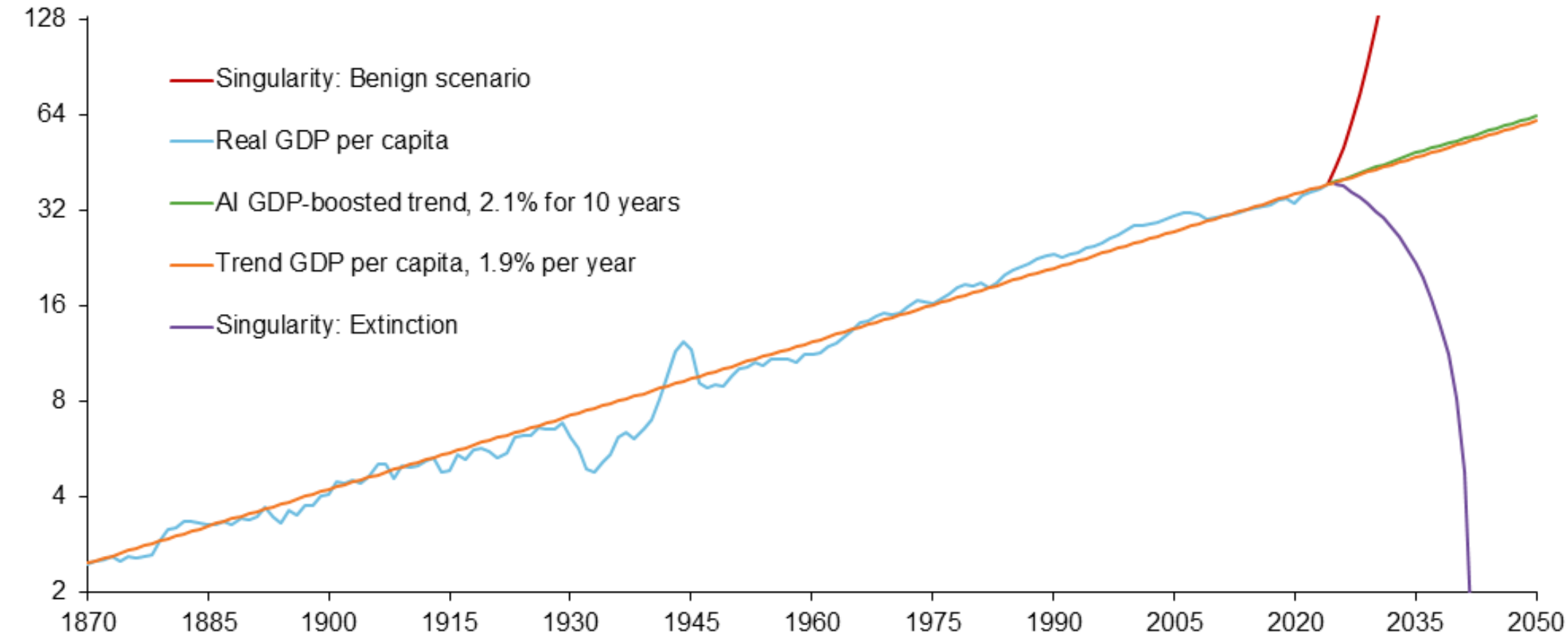
What's Next for AI in 2026:

From LLM Limits to Image Models



Anthony Schiavo
Principal Analyst

Chart 1
AI scenarios



NOTES: The blue line is real gross domestic product (GDP) per capita in 1990 dollars. The orange line is a trend line fitted to the data for 1870–2024 with a trend growth rate of 1.9 percent per year. The red, green and purple lines are hypothetical paths for per capita GDP based on different scenarios.
SOURCES: Bureau of Economic Analysis; Haver Analytics; Macrohistory.net; United Nations; authors' calculations.

Federal Reserve Bank of Dallas

Source: https://www.dallasfed.org/-/media/Images/research/economics/2025/0624/dfe0624c1.png

What To Expect

01 Looking back and ahead on model development

02 Unpacking data center investment dynamics

03 Four future scenarios for AI adoption

AI is a probabilistic tool, trained on large amounts of data, which can replicate the outputs of complex tasks.

Looking back: 2024’s AI predictions

1. AI models are becoming commodified.
2. AI companies are going to have to change their business models to make money.
3. A crash is coming in the next two years (by November 2026).

Model commodification has arrived

We have many models available on both the cutting edge…

Model	Organization	Global Average	Reasoning Average	Coding Average	Agentic Coding Average	Mathematics Average	Data Analysis Average	Language Average	IF Average
Claude 4.5 Opus Thinking High Effort	Anthropic	75.58	84.12	79.65	55.00	94.52	71.98	81.26	62.55
GPT-5.1 Codex Max	OpenAI	75.18	91.43	81.38	46.67	92.87	71.41	75.39	67.13
Claude 4.5 Opus Thinking Medium Effort	Anthropic	74.87	84.94	77.48	58.33	92.99	69.51	80.35	60.52
Gemini 3 Pro Preview High	Google	74.14	79.89	74.60	45.00	94.12	74.91	84.62	65.85

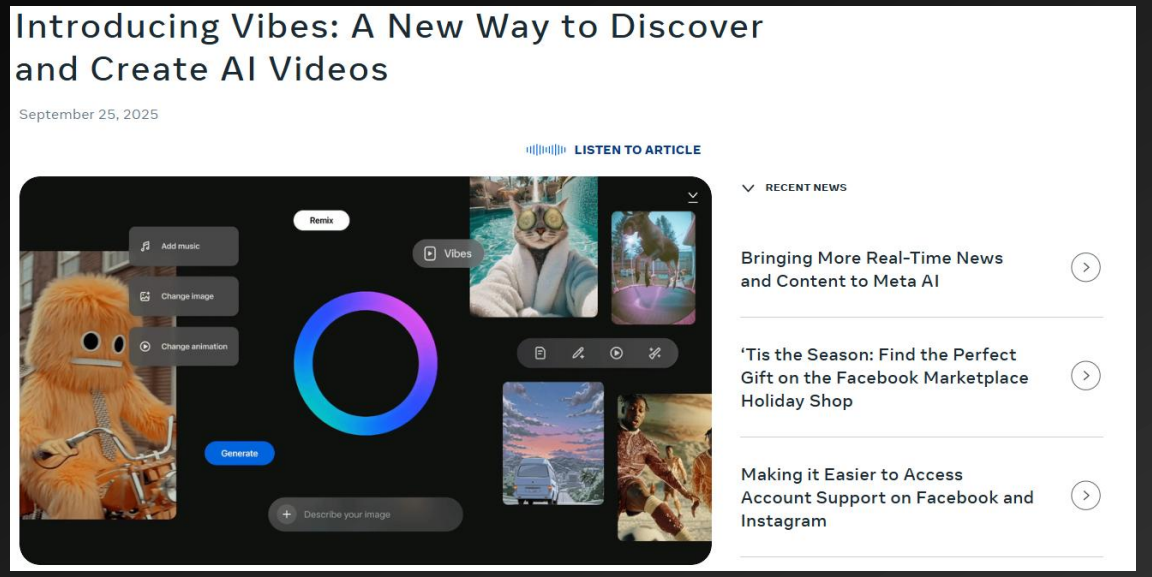
And the middle tier of performance

GPT-5 Mini Low	OpenAI	58.52	47.87	69.55	33.33	80.32	67.48	60.41	50.71
Claude 4.1 Opus	Anthropic	58.33	40.53	76.07	41.67	80.44	66.95	76.75	25.92
Qwen 3 235B A22B Thinking 2507	Alibaba	57.74	59.20	68.97	6.67	84.53	74.65	69.52	40.64
DeepSeek V3.2	DeepSeek	57.38	45.00	75.69	45.00	81.94	66.71	64.24	23.06

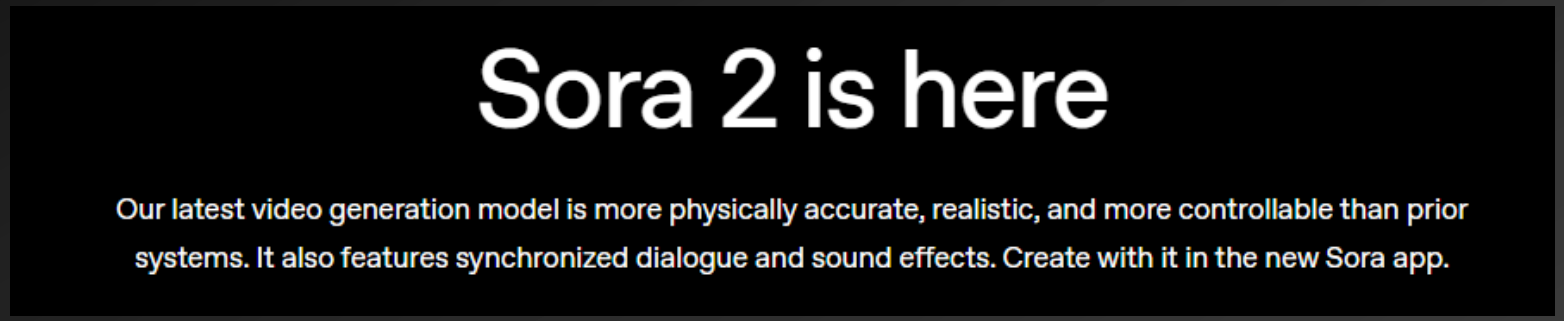
Data from https://livebench.ai/#/

OpenAI and Meta have begun to pivot to consumer business models

Meta



OpenAI



A crash? Not yet…

Why Fears of a Trillion-Dollar AI Bubble Are Growing

Wall Street Races to Cut Its Risk From AI's Borrowing Binge

OpenAI Goes From Stock Market Savior to Burden as AI Risks Mount

Image source: Bloomberg.com

Looking ahead: Predictions for 2026

1. Miniature models become much more powerful.

2. R&D focuses on world models.

3. Image model applications proliferate, especially in medical and industrial markets.

Miniature models become much more powerful

Miniature models, capable of running on edge devices like regular desktop computers, are already roughly as good as ChatGPT 3.5 Turbo.

GPT-3.5 Turbo	OpenAI	34.54	26.67	27.74	26.93	41.21	24.22	60.47
Gemma 2 9B	Google	31.34	17.33	22.46	23.98	35.06	27.64	61.55

Does Model Size Matter? A Comparison of Small and Large Language Models for Requirements Classification

In many simple, real-world business tasks, miniature models already perform similarly to large LLMs.

Table 2. Results across all datasets using CoT ∪ Few-shot.

	PROMISE			PROMISE Reclass			SecReq		
	P	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**
Qwen2-7B	0.86	0.80	0.83	0.55	**0.96**	0.70	0.83	0.90	0.86
Falcon3-7B	0.86	0.80	0.83	0.58	**0.96**	0.72	0.79	0.90	0.84
Granite-3.2	0.84	0.82	0.83	0.62	0.87	0.72	0.84	0.90	0.87
Ministral-8B	0.84	0.77	0.80	0.64	0.89	0.75	0.83	0.88	0.85
Llama-3-8B	0.84	0.77	0.80	0.70	0.87	0.78	0.86	0.91	0.88
SLMs (Mean)	0.85	0.79	0.82	0.62	0.91	0.73	0.83	0.90	0.86
Grok-4	**0.88**	**0.83**	**0.85**	0.60	0.82	0.69	0.84	0.89	0.86
GPT-5o	0.85	0.81	0.83	0.68	0.88	0.77	0.85	0.90	0.87
Claude-4	0.85	0.80	0.82	**0.72**	0.90	**0.80**	**0.87**	**0.92**	**0.89**
LLMs (Mean)	0.86	0.81	0.83	0.67	0.87	0.75	0.85	0.90	0.88

Image models drive hype with new applications

Creator Behind Billboard-Charting AI 'Artist' Xania Monet Defends Her Music Against Backlash From Kehlani And More

Xania Monet's vocals are generated by Suno, an AI platform that was sued by major record labels and the Recording Industry Association of America last year for using copyrighted material to train its AI tools.

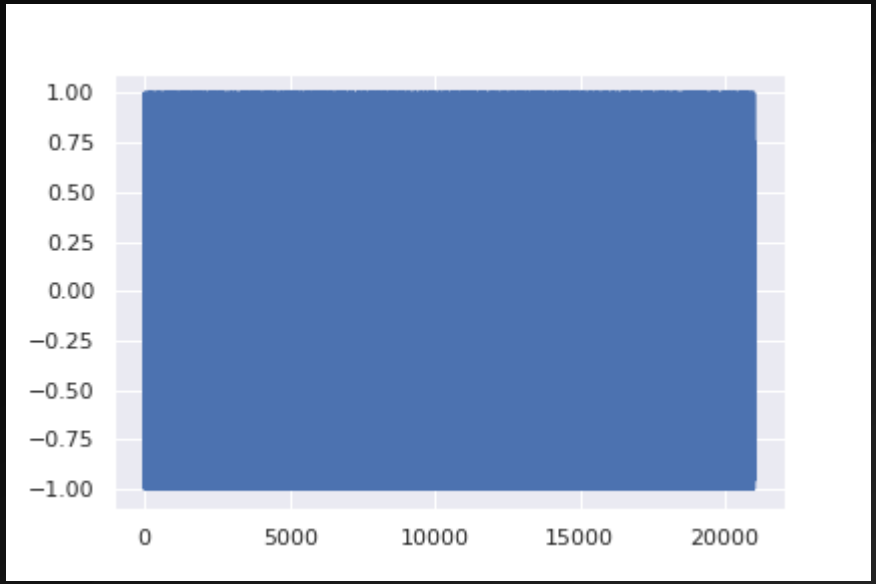


Image diffusion models are the "secret sauce" behind audio generation.

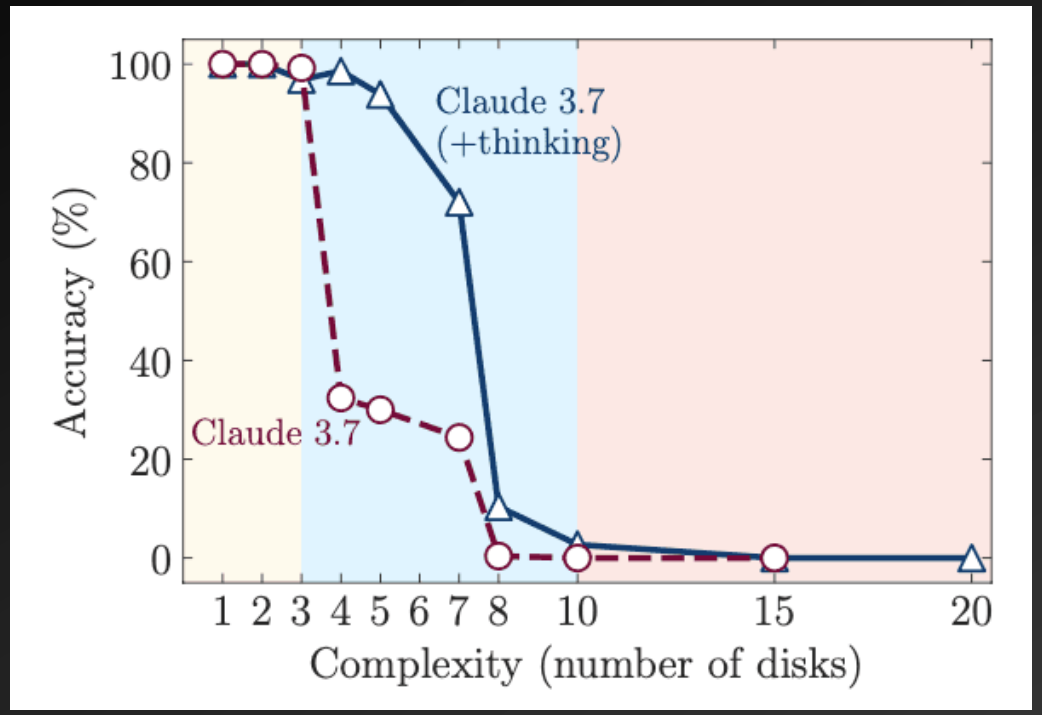
What else can be done with image models?

Image sources – Forbes, https://github.com/simonrouard/CRASH/blob/master/assets/gif_snare.gif

World models get more attention as we reach the limits of LLMs

Machine Learning Research

The Illusion of Thinking: Understanding the Strengths and Limitations of Reasoning Models via the Lens of Problem Complexity



After leaving Meta, French AI pioneer Yann LeCun will launch start-up in Paris

What To Expect

01 | Looking back and ahead on model development

02 | Unpacking data center investment dynamics

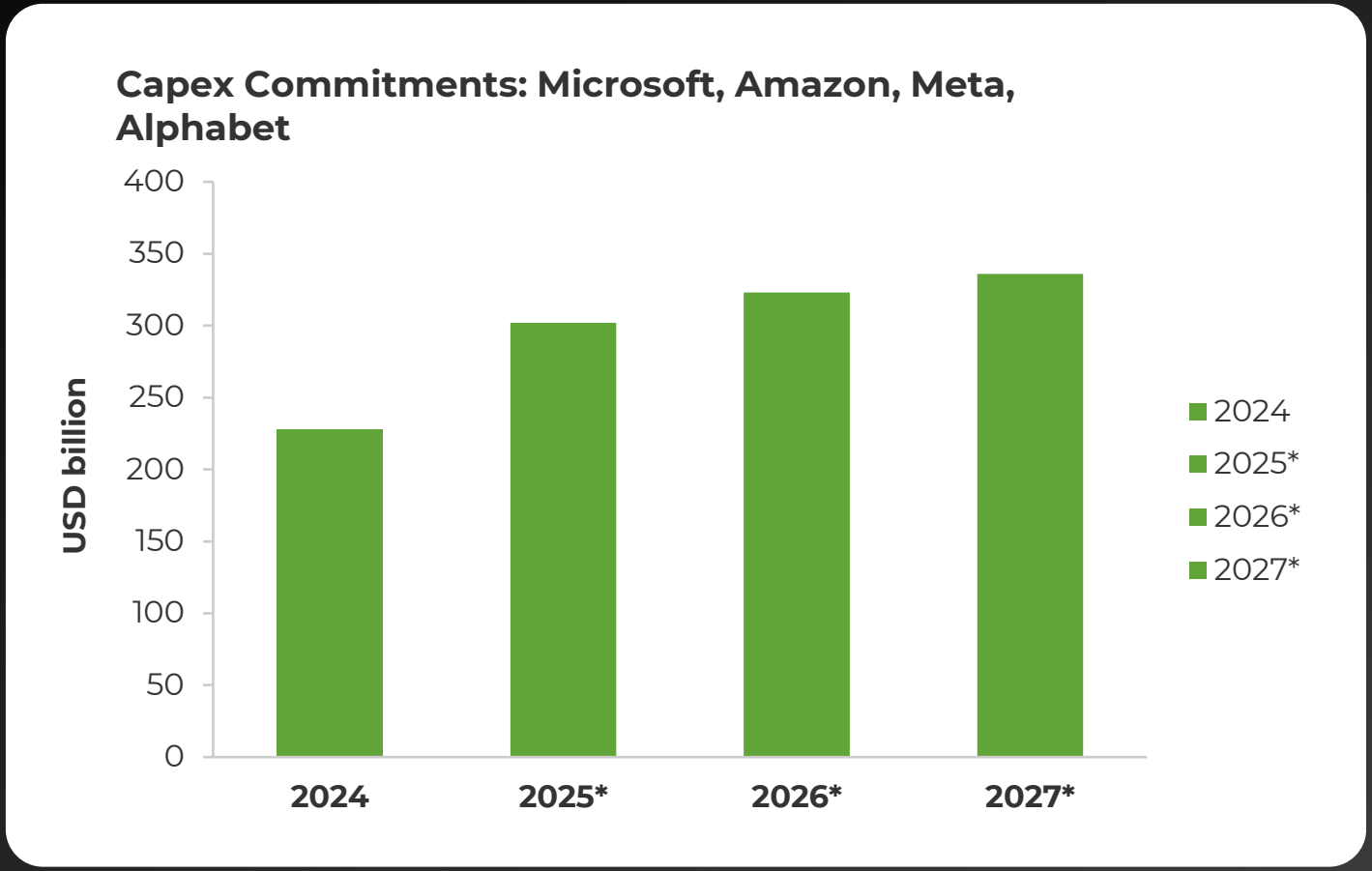
03 | Four future scenarios for AI adoption

Can data center investments really be profitable?

The major hyperscalers are committed to spending around USD 1.1 trillion on capex by 2027.

OpenAI is committed to USD 1.4 trillion in capex spending.

Can this ever work out?

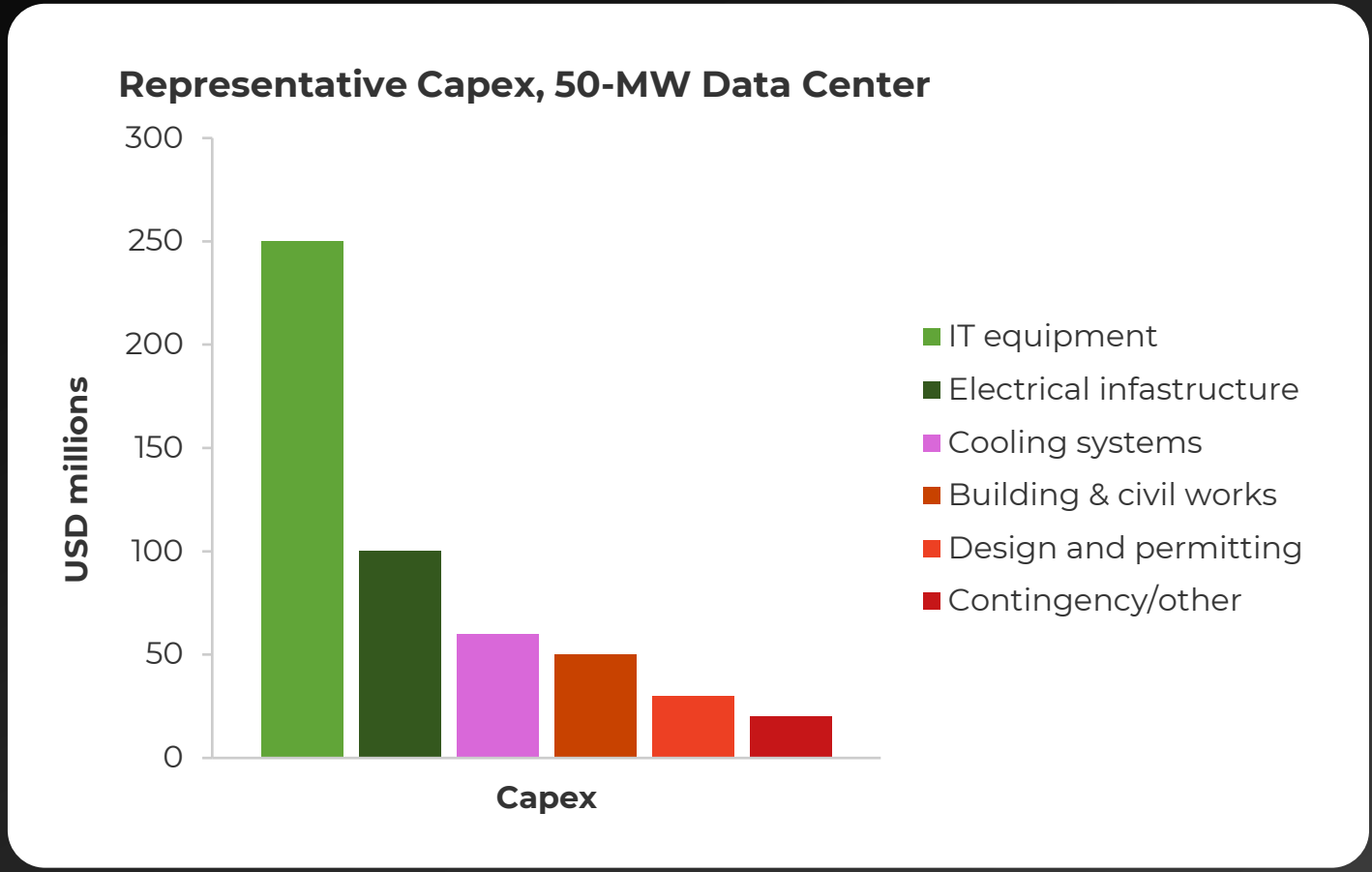


Data source: JP Morgan, Bloomberg

Capex is critical to overall returns

Data centers cost around USD 10 million/MW of IT load.

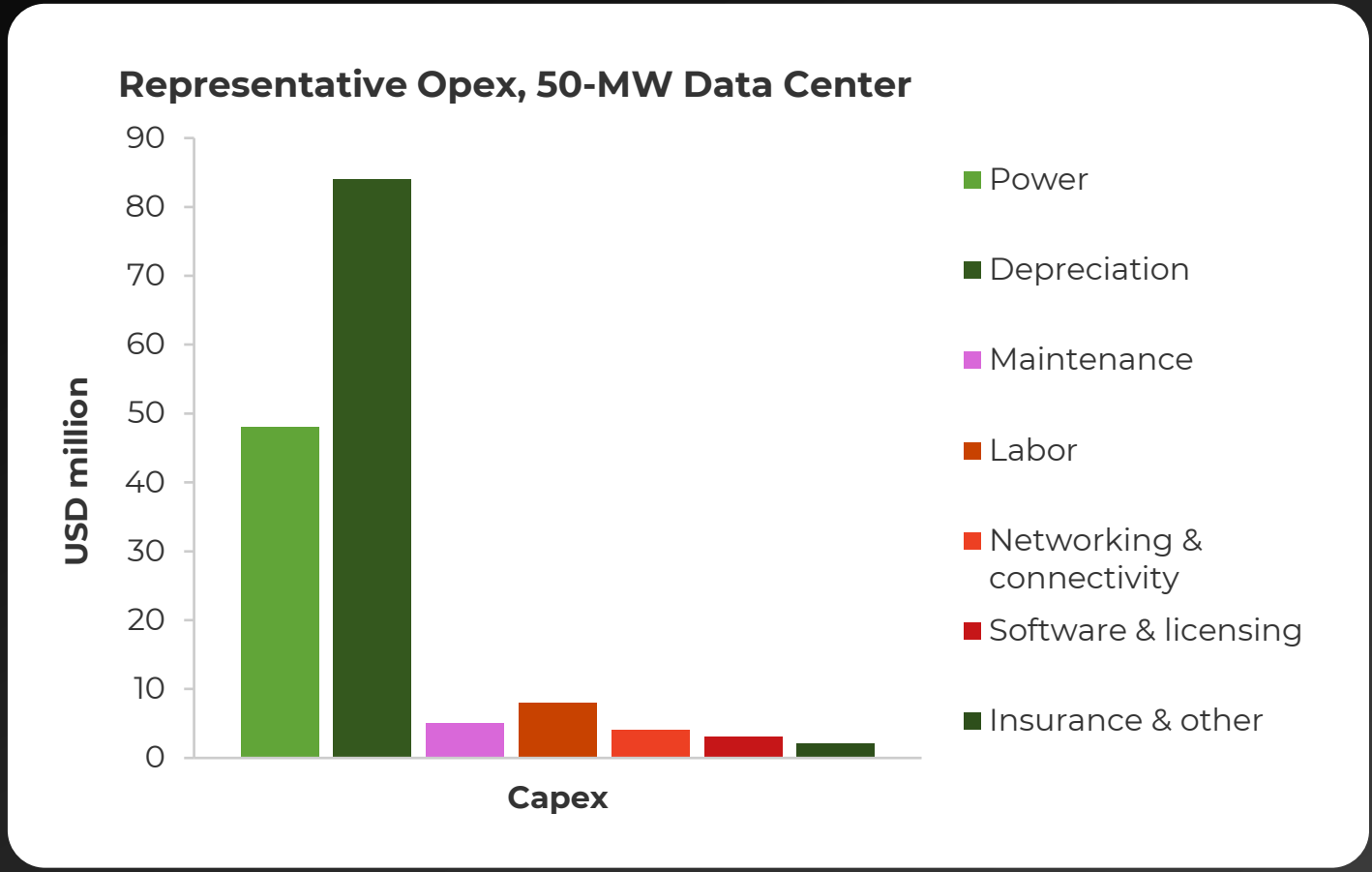
The chart on the right has representative estimates of the capex breakdown of a 50-MW data center.



Opex is dominated by depreciation and energy costs

Power and depreciation are the two biggest ongoing costs for data centers.

The chart on the right has representative estimates of the opex breakdown of a 50-MW data center.



Data center returns are highly sensitive to uptime, GPU life, and rental revenue

We conducted a basic cost modeling exercise to determine data center profitability. The key factors are:

GPU life	Usage/uptime	Rental price
4 Years	99%	USD 0.4/h
2 Years	80%	USD 0.35/h

If two of these factors are in the red, the data center isn't likely to be profitable — if it's funded by debt.

Two major risks to data center returns:

Shrinking demand for compute for more efficient models.

Oversupply of GPUs due to overaggressive capex spend.

What To Expect

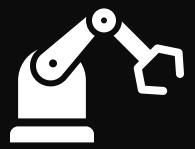
01 | Looking back and ahead on model development

02 | Unpacking data center investment dynamics

03 | Four future scenarios for AI adoption

Four possible futures for AI

Slop future: The best way to extract profit from AI turns out to be advertising and social media.



Work future: AI provides broad-based improvements to economic efficiency.

Tech future: AI leads to significant breakthroughs in high knowledge tasks, including drug development, medical diagnostics, and technology creation.

Trash future: AI remains largely unprofitable, even if it is capable of automating tasks or successfully solving business challenges.

Slop future misses the promise of AI

Outlook: In this future, the only real way to make money with AI is from consumers — business just doesn't have hundreds of billions of additional spend in them.

Lux Take: This future primarily benefits the existing social media platforms — it'll be tough for new apps like Sora to breakthrough their stranglehold. This outcome is relatively unlikely; there's no real shortage of online content, and the social media model mostly works because content is free.

Winners:

Meta

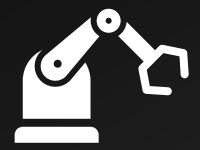
OpenAI

Losers:

amazon

Google

Work future most closely resembles the AI promised land



Outlook: In this future, business uses dominate the AI landscape — and those uses are both profitable and generate meaningful revenue. In this scenario, much of the frontier of economic productivity gains come from AI, across both knowledge and physical industries.

Lux Take: The existing business service companies will benefit the most, as they are best positioned to capture these flows. A lot has to go right: The technology is one thing but perhaps more importantly are value creation and capture.

Winners:

amazon

ANTHROP\C

Microsoft

Losers:

Meta

OpenAI

Science future will take a long time to manifest

Outlook: In this future, the benefits of AI are not in business efficiency but very high knowledge tasks — the creation of drugs, the diagnosis of illnesses, or the development of new materials and chips. These benefits will be good for the economy and society but hard for most current AI companies to capture.

Lux Take: This is perhaps the most likely future, as it relies mostly on the technology improving (which it has done consistently) and less on specific value capture scenarios. Google and Nvidia — which are investing in these longer-term research areas — stand to benefit above and beyond the others.

Winners:

Google

NVIDIA

Losers:

Meta

OpenAI

Microsoft

Failure is always an option

Outlook: AI can fail for a lot of reasons; poor economics and value capture are the most likely culprits.

Lux Take: This is the other likely outcome, as any number of market failures could push us in this direction. The best move may be not to play: Apple is not making investments like its tech peers, but it will still take a cut of any AI success through its walled garden.

Winners:

Apple

Losers:

Meta

Google

OpenAI

NVIDIA

Microsoft

Key Takeaways

1

LLM development is slowing, and there are major cracks in the economics of data center buildout.

There's still a lot of development left for AI in areas like image and world models, but the need for huge cash flows very soon makes AI risky.

2

After the boom and bust, AI is likely to find uses — but AI companies are going to struggle to capture value.

AI may dramatically improve business operations or accelerate science — but can AI companies capture that value?

3

Don't overinvest in AI solutions or a data center boom.

The cost and size of LLMs are likely to keep falling, making in-house adoption more likely. Take advantage of these drops by developing the abilities to run models at the edge and continue data center tech development without overcommitting.

Thank You

VISIT
www.luxresearchinc.com

FOLLOW
@LuxResearch

CONNECT
LuxResearch

READ
http://www.luxresearchinc.com/blog/

EMAIL
questions@luxresearchinc.com

About Lux

Lux Research fuels innovators to not only imagine what's possible in the future but also operationalize innovation success in the near term. We deliver research and advisory services to inspire, illuminate, and ignite innovative thinking that reshapes and grows businesses. Using quality data derived from primary research, fact-based analysis, and opinions that challenge traditional thinking, our experts focus on finding truly disruptive innovations that are also realistic and make good business sense.

The "Lux Take" is trusted by innovation leaders around the world, many of whom seek our advice directly before placing a bet on a startup or partner — our clients rely on Lux insights to make decisions that generate fantastic business outcomes. We pride ourselves on taking a rigorous, scientific approach to avoid the hype and generate unique perspectives and insights that innovation leaders can't live without.

VISIT

www.luxresearchinc.com

FOLLOW

@LuxResearch

CONNECT

LuxResearch

READ

http://www.luxresearchinc.com/blog/

EMAIL

questions@luxresearchinc.com