Preparing for an AI-First and Agent-Driven Era

Cyrill Glockner

Today: Tokens to Subscriptions

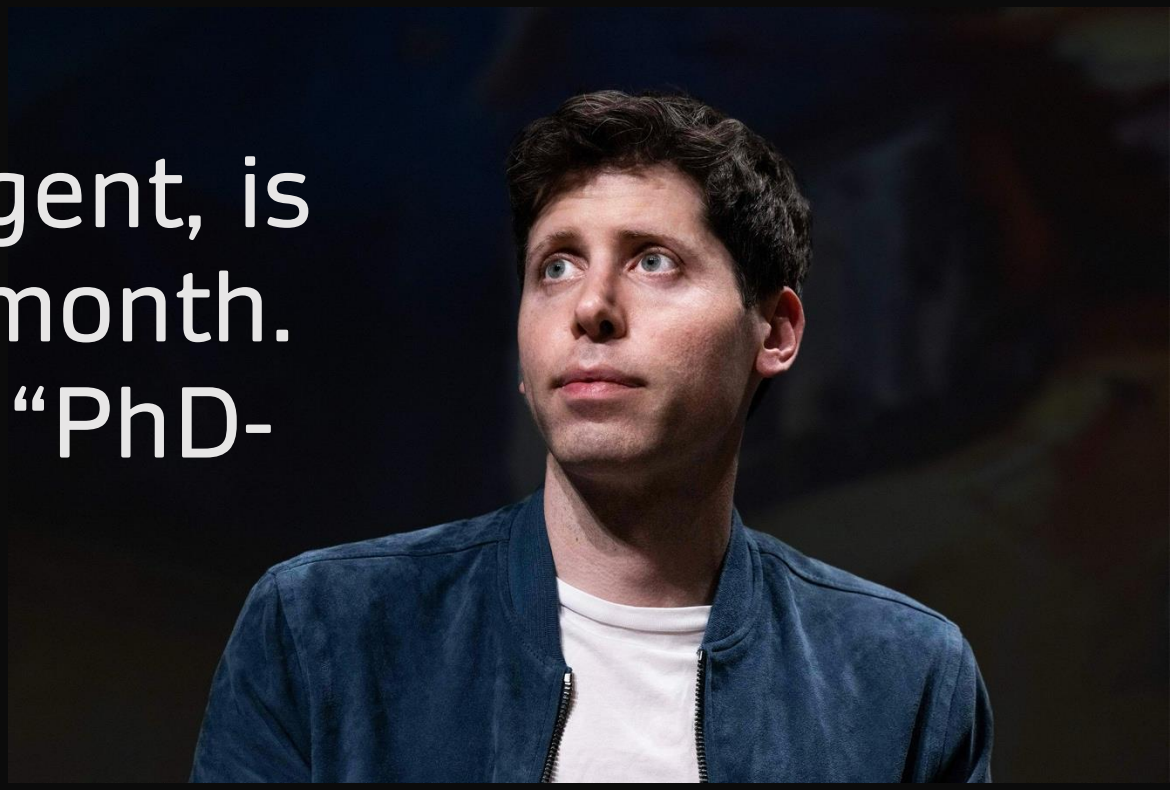
AI services cost based on token

- Example: Paying $0.01 per 1,000 tokens processed by an AI summarization service.
- SKU Translation: Based on the average monthly token usage across customers (e.g., 1 million tokens at $10 cost per user), bundle compute, support, and service value into a $99/month subscription plan.
- Monitor usage to ensure margin sustainability, with fair use thresholds if necessary.
- OpenAI is an example
- Experiment and change as needed

Businesses must package underlying token costs into predictable subscription SKUs for scalability and margin control.

Open AI: Agent for Sale Rumor

1. A “high-income knowledge worker” agent, priced at $2,000 a month.
2. A software developer agent, is said to cost $10,000 a month.
3. $20,000-per-month for “PhD-level” researcher.



Tomorrow: Value-Based Pricing

- Pricing based on expertise and capabilities, not compute cycles.
- Token pricing survives for foundational tasks.

Hiring an AI agent will mirror hiring human experts

Companies that rely on offering lower-level toolchains instead might struggle.

Example: Rufous

- Rufous connects human and digital workforces into a unified platform.
- Their model embraces agent-based collaboration: digital agents manage tasks, payments, and workflow orchestration.
- Talent and work coordination becomes increasingly autonomous.

Rufous is an early example of an agent-first company, blending AI and human workforces into seamless operations.

https://rufous.com/

What is a Job?

What is a Job? (Modern Definition)
A **Job** = a **bundle of tasks** that must be completed,
using a specific **set of skills** and **knowledge**
to deliver a **desired outcome** within a **given context**.

- Formal work frameworks like O*NET and McKinsey list ~800–1,000 base task types.
- Real-world economic activity spans **5,000–20,000 distinct task types** globally.
- New tasks emerge constantly as technology evolves.

The task universe is vast but semi-structured, perfect for agents to map, learn, and automate over time.

OpenAI SWE-Lancer Benchmark

- Evaluates AI on 1,400 real freelance coding tasks worth $1M.
- **Top result:** Claude 3.5 Sonnet completed 26.2% of coding tasks (~$400K earned).
- **Managerial tasks:** Claude 3.5 Sonnet had a 44.9% success rate in selecting proposals.
- AI still struggles with a lot of tasks, highlighting current limitations.

URL: https://openai.com/index/swe-lancer/

PSL's Vision for AI-Native Startups

Phase 1: AI assist tools layered onto human-led processes.

- Example: 50%-80% of coding, testing, and integration handled by AI agents.

Phase 2: Full AI-native process redesign.

- Example: Small human leadership layer orchestrates agent workforces.

Phase 3: AI-native teams operate with exponential advantage.

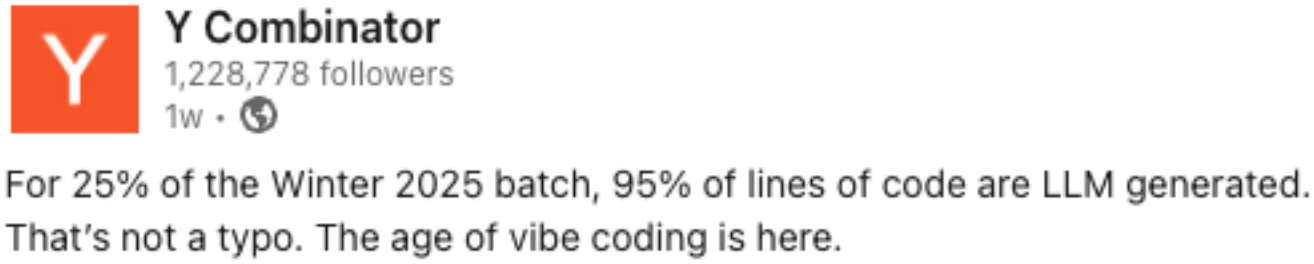
- Example: Engineering shifts from "building software" to "designing agent systems." achieving instant scalability.

https://www.psl.com/feed-posts/three-years-out-my-forecasts-for-ais-effects-on-product-dev-teams

URL: https://www.swebench.com/

SWE-bench

- Evaluates AI models' ability to resolve real-world GitHub issues,
- **Dataset:** Comprises 2,294 issue-pull request pairs from 12 popular Python repositories.
- **Evaluation:** Uses unit test verification, comparing model-generated patches against original pull requests to assess correctness.
- **Performance:** As of February 2025, leading reasoning model achieved a 64.6% success rate on the SWE-bench Lite subset.



Seed-Strapping

“Achieve scale with a single round of venture funding”

Top Lean AI Native Companies Leaderboard OFFICIAL

View Source Submit a Startup

.	Company	Description	Location	Annual Revenue	# of Employe	Revenue/Empl	Profitable ?	Total Funding	Valuation	Valuation/Employee	Founded
	Telegram	Messaging	Dubai	$1,000,000,000	30	$33,333,333	Yes	$3,200,000,000	$30,000,000,000	$1,000,000,000	2013
	Midjourney	Image Generation	San Francisco	$500,000,000	40	$12,500,000	Yes	$0	$10,000,000,000	$250,000,000	2022
	Anysphere (Cursor)	AI Code Editor	San Francisco	$100,000,000	20	$5,000,000		$173,400,000	$2,600,000,000	$130,000,000	2022
	Cal AI	AI Calorie Tracker	New York	$12,000,000	4	$3,000,000	Yes	$0			2024
	Mercor	Talent Marketplace + Data	San Francisco	$75,000,000	30	$2,500,000		$135,600,000	$2,000,000,000	$66,666,667	2023
	Eleven Labs	AI Voice, Text-to-Speech	New York	$100,000,000	50	$2,000,000		208000000	$3,300,000,000	$66,000,000	2022
	Fal.ai	Generative media platform	San Francisco	$40,000,000	25	$1,600,000	Yes	$72,000,000			2021
	OpenArt	AI for Image and Video	San Francisco	$12,000,000	8	$1,500,000		$5,000,000			2022
	Solvely.ai	AI for Homework	San Francisco	$6,000,000	4	$1,500,000					2023
	Oleve	AI Consumer Portolio	New York	$6,000,000	4	$1,500,000	Yes	$600,000	$10,000,000	$2,500,000	2024
	Stackblitz (Bolt.new)	AI Code Editor	San Francisco	$20,000,000	15	$1,333,333		$87,900,000	$700,000,000	$46,666,667	2017
	Aragon AI	AI Avatar Generation	San Francisco	$10,000,000	9	$1,111,111	Yes	$900,000	$10,000,000	$1,111,111	2022
	Photoroom	Image Editing	Paris	$50,000,000	50	$1,000,000	Yes	$62,000,000	$500,000,000	$10,000,000	2019
	Lovable	AI Code Editor	Stockholm	$17,000,000	20	$850,000		$22,500,000			2023
	AKOOL	AI Video Generation	Santa Clara	$40,000,000	50	$800,000	Yes				2022
	Gamma	AI for Presentations	San Fransisco	$20,000,000	28	$714,286	Yes	$19,000,000			2020
	Praktika.ai	AI Avatars for Learning	Remote	$20,000,000	30	$666,667		$36,500,000			2021
	Chatbase	AI Customer Support	Remote	$5,000,000	11	$454,545	Yes	$0			2023
	Jenni.ai	AI for Writing	Remote	$10,000,000	23	$434,783	Yes	$850,000			2021
	Vapi	AI Voice API	San Francisco	$8,000,000	20	$400,000		$20,125,000	$130,000,000	$6,500,000	2023
	Recall.ai	API for Video Conferences	San Francisco	$10,000,000	25	$400,000	Yes	$12,700,000			2022

https://leanaileaderboard.com/

“Reason before responding”

OpenAI Deep Research

- An AI agent designed to autonomously browse the web and generate comprehensive, cited reports on user-specified topics within 5 to 30 minutes.
- **Performance:** Achieved a 26.6% accuracy rate on the "Humanity's Last Exam" (HLE) benchmark, surpassing competitors like DeepSeek's R1 (9.4%) and GPT-4o (3.3%).
- **Use Cases:** Ideal for professionals in finance, science, and engineering, as well as for consumers seeking detailed analyses, such as product comparisons or financial assessments.
- May occasionally produce inaccuracies or rely on less authoritative sources.

URL: https://openai.com/index/introducing-deep-research/

FUTURE: Intelligence Explosion (2027)

- Next-gen AI agents become AI researchers.
- AI starts to invent new AI architectures, optimizations, and training techniques.
- Acceleration of AI capabilities becomes exponential, not linear.

AI not just doing work better, but it’s making itself better at doing work, leading to significant leaps forward.

AI Race, Companies and Countries

Escape Velocity in AI Competition

- First organization to scale self-improving agents achieves escape velocity.
- Compounding gains become unreachable for slower competitors.
- Strategic inflection point creates market dominance.

Once escape velocity is reached, no traditional organization will catch up.

We are getting closer to AGI

“AI systems that are as capable as humans across a wide range of cognitive tasks.”

The Narrative has changed

Dario Amodei

Machines of Loving Grace[1]

How AI Could Transform the World for the Better

https://darioamodei.com/machines-of-loving-grace

Leading AI researchers focused on AI risk throughout their panel discussion.

WORLD ECONOMIC FORUM

Global Race towards AGI is Accelerating

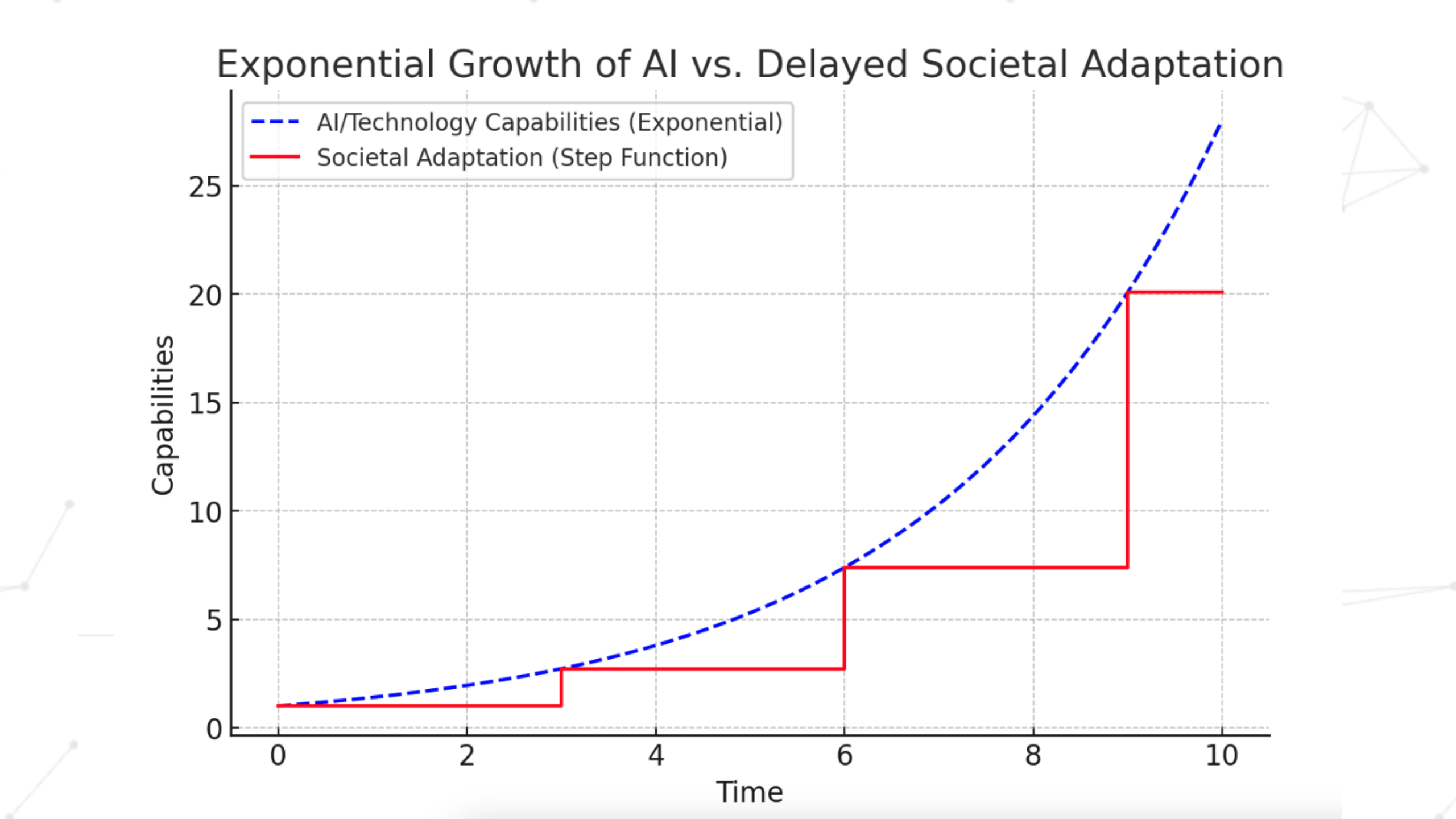
Dario Amodei (Anthropic CEO):
“AI systems can eventually help make even smarter AI systems; a temporary lead could be parlayed into a durable advantage”

Sam Altman (OpenAI CEO)
‘My guess is we will hit AGI sooner than most people in the world think and it will matter much less,’ says OpenAI’s CEO”

Demis Hassabis (Google Deepmind CEO)
“AGI being a system that’s capable of exhibiting all the cognitive capabilities humans can. I think we’re getting closer and closer, but we’re still probably a handful of years away.”

Yoshua Bengio (AI Researcher)
My current estimate places a 95% confidence interval for the time horizon of superhuman intelligence at 5 to 20 years. 50% for 5, what is the date assuming 10% or 20%?

But why are things changing slowly?


Exponential Growth of AI vs. Delayed Societal Adaptation
AI/Technology Capabilities (Exponential)
Societal Adaptation (Step Function)
Capabilities
0
5
10
15
20
25
0
2
4
6
8
10
Time

“The Future is already here, it’s just not evenly distributed”

William Gibson

In Closing