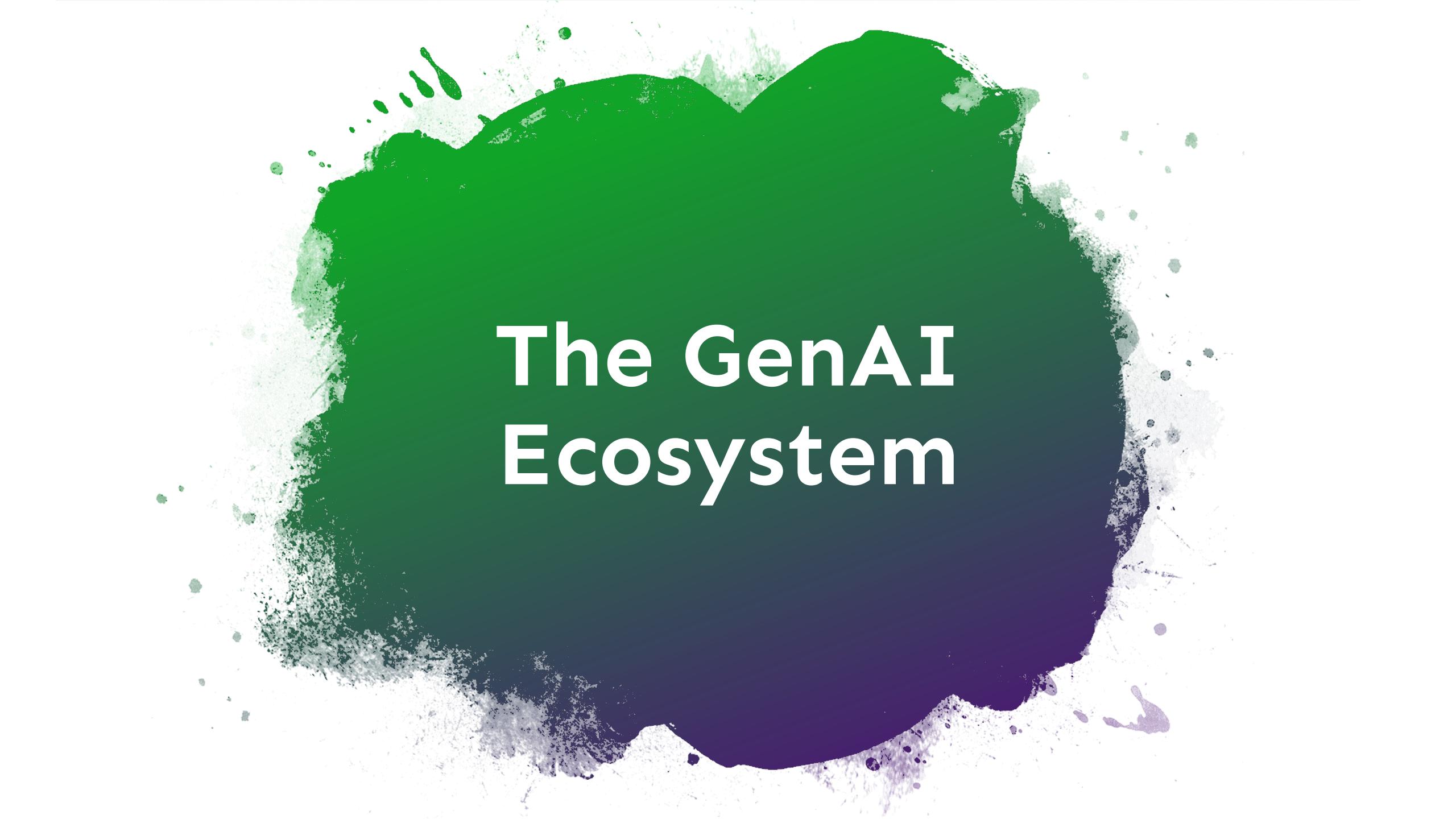


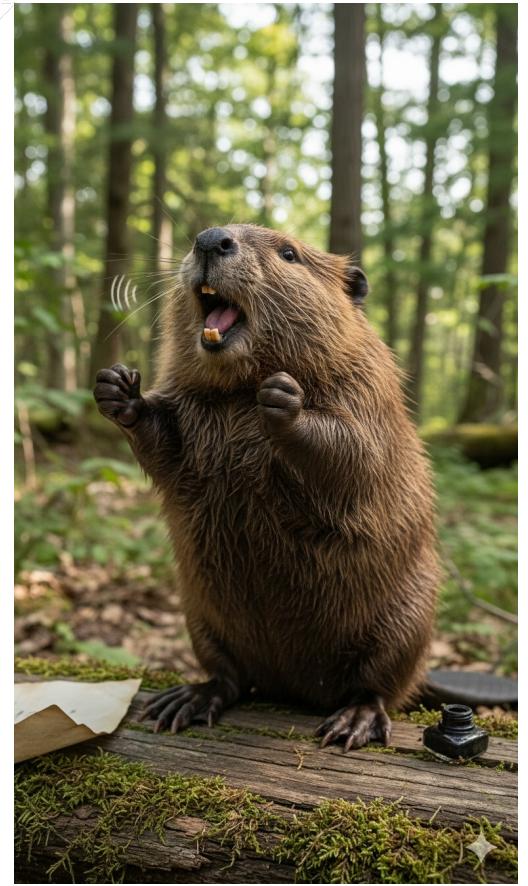


Practical AI 08 – GenAI & Business

Emanuele Fabbiani

The background features a large, irregularly shaped central circle filled with a dark green gradient. This circle is surrounded by a textured, white and light gray pattern resembling a cracked or peeling surface. Small, scattered colored splatters in shades of green, blue, and purple are visible around the perimeter of the central circle.

The GenAI Ecosystem



- [11Labs](#)
- [Suno AI](#)
- [OpenAI GPT-4o Transcribe](#)
- [Google Gemini](#)

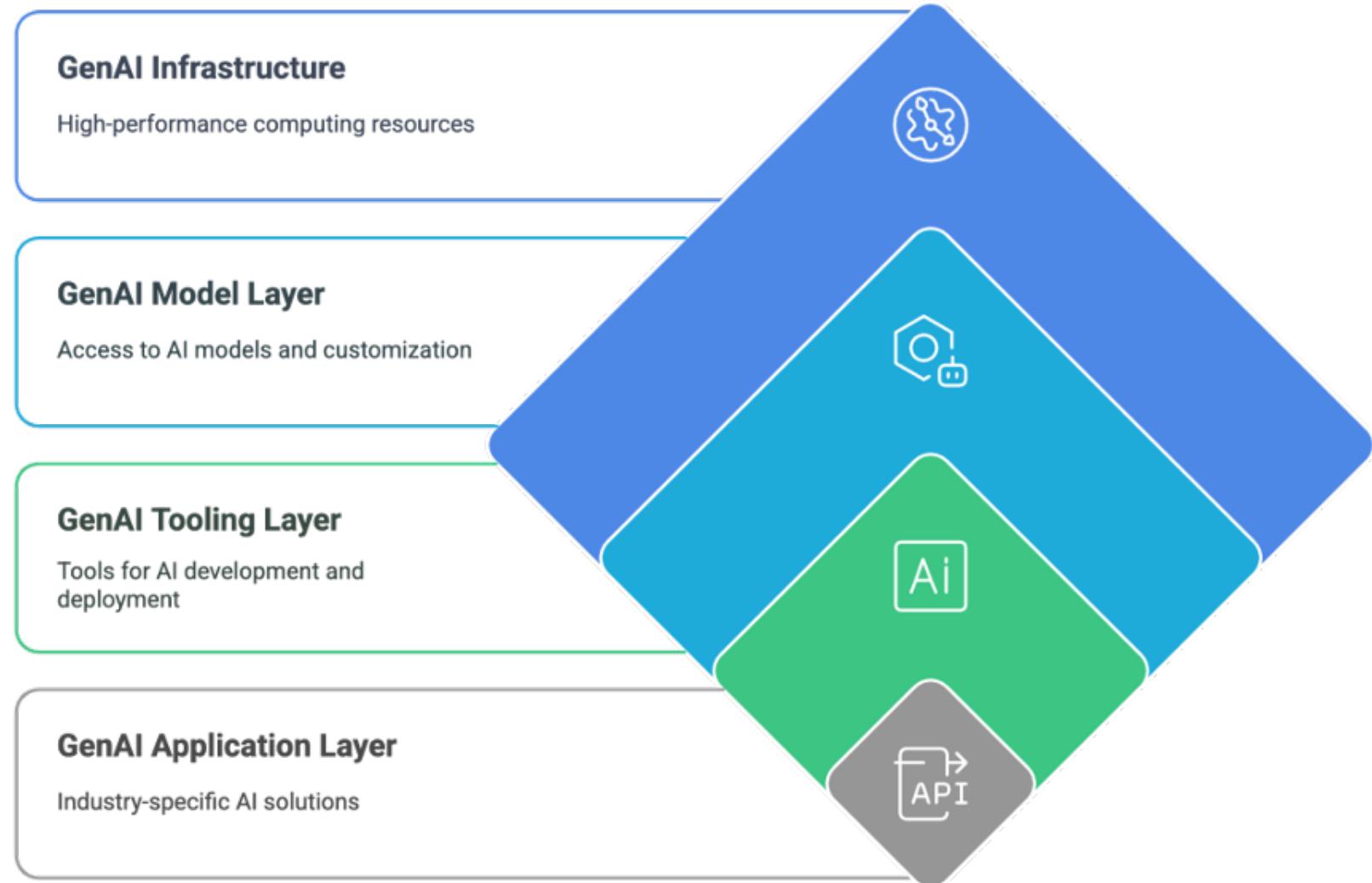
- [Midjourney](#)
- [Flux](#)
- [OpenAI GPT-Image](#)
- [Google Gemini Image](#)

- [Google Veo](#)
- [OpenAI Sora](#)
- [Runway](#)
- [Midjourney](#)

- [OpenAI ChatGPT](#)
- [Anthropic Claude](#)
- [Google Gemini](#)
- [Meta LLaMA](#)



	Product	Company	Main Sponsor	
	ChatGPT	OpenAI	Microsoft	
	Claude	Anthropic	Amazon	
	Gemini	DeepMind	Google	
	Llama	Meta	Meta	

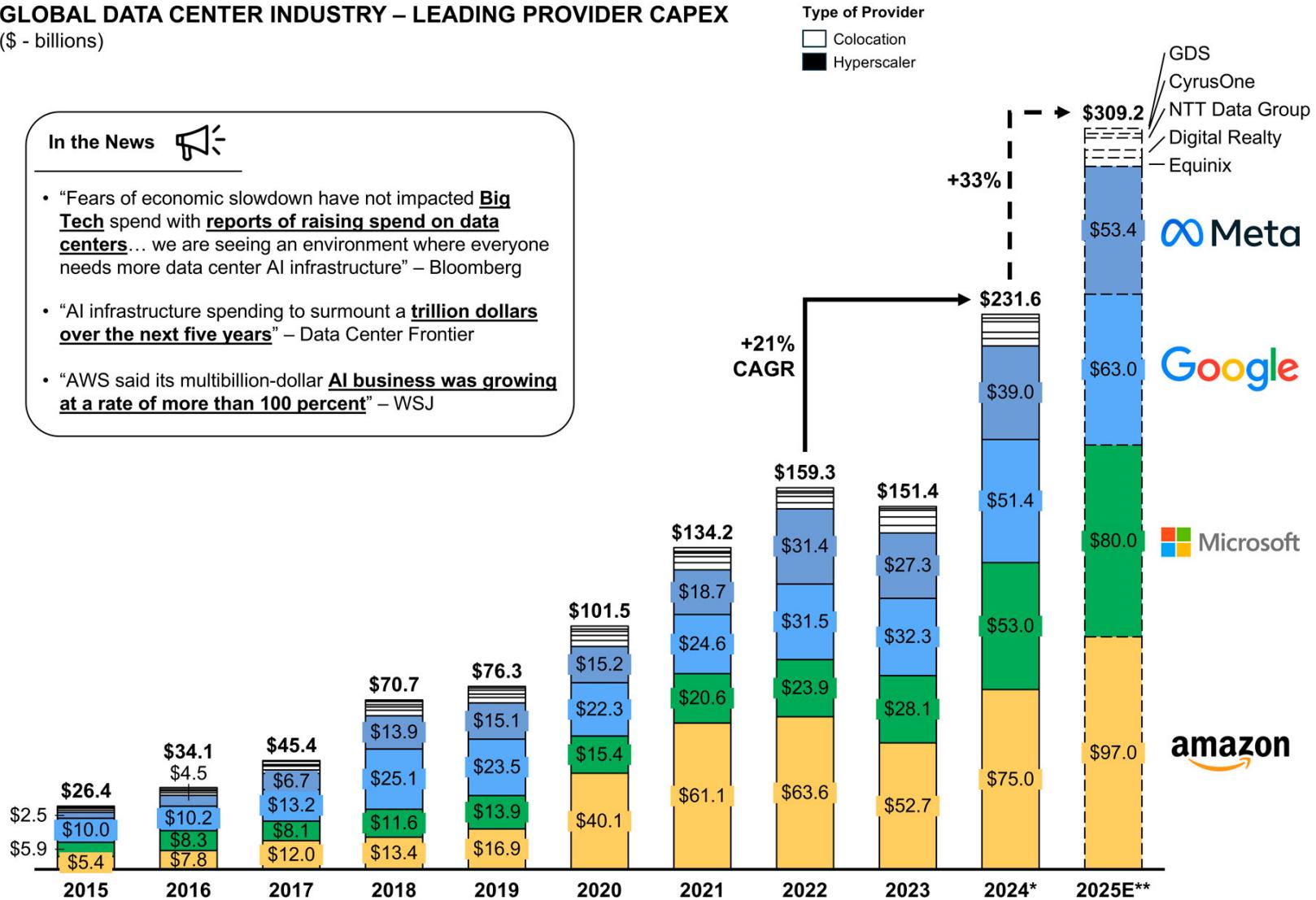




GLOBAL DATA CENTER INDUSTRY – LEADING PROVIDER CAPEX (\$ - billions)

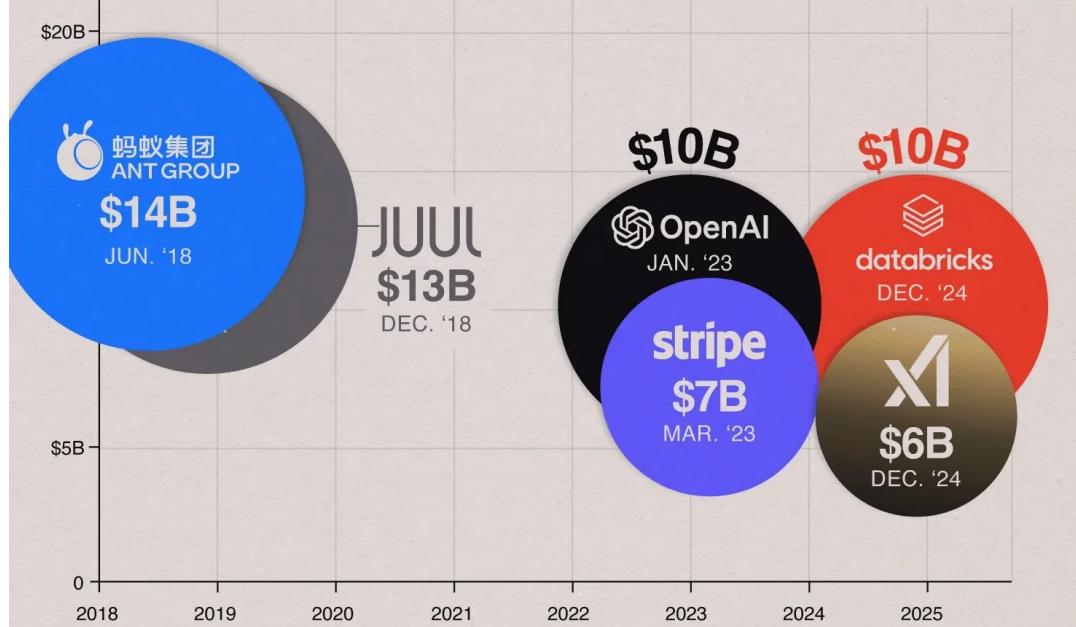
In the News 🎤

- “Fears of economic slowdown have not impacted **Big Tech** spend with reports of raising spend on data centers... we are seeing an environment where everyone needs more data center AI infrastructure” – Bloomberg
- “AI infrastructure spending to surmount a trillion dollars over the next five years” – Data Center Frontier
- “AWS said its multibillion-dollar AI business was growing at a rate of more than 100 percent” – WSJ



Biggest Private Tech Deals

SINCE 2018



As of April 1, 2025
Source: Pitchbook, CNBC

VISUAL CAPITALIST

ANTHROPIC'S VALUATION NEARLY TRIPLES TO \$183 BILLION AFTER LATEST FUNDING ROUND

POST-MONEY VALUATION

\$200B

\$150B

\$100B

\$50B

\$0B

SERIES F: \$183B
+198% FROM LAST ROUND

SERIES E-1: \$61.5B

SOURCE: YAHOO FINANCE

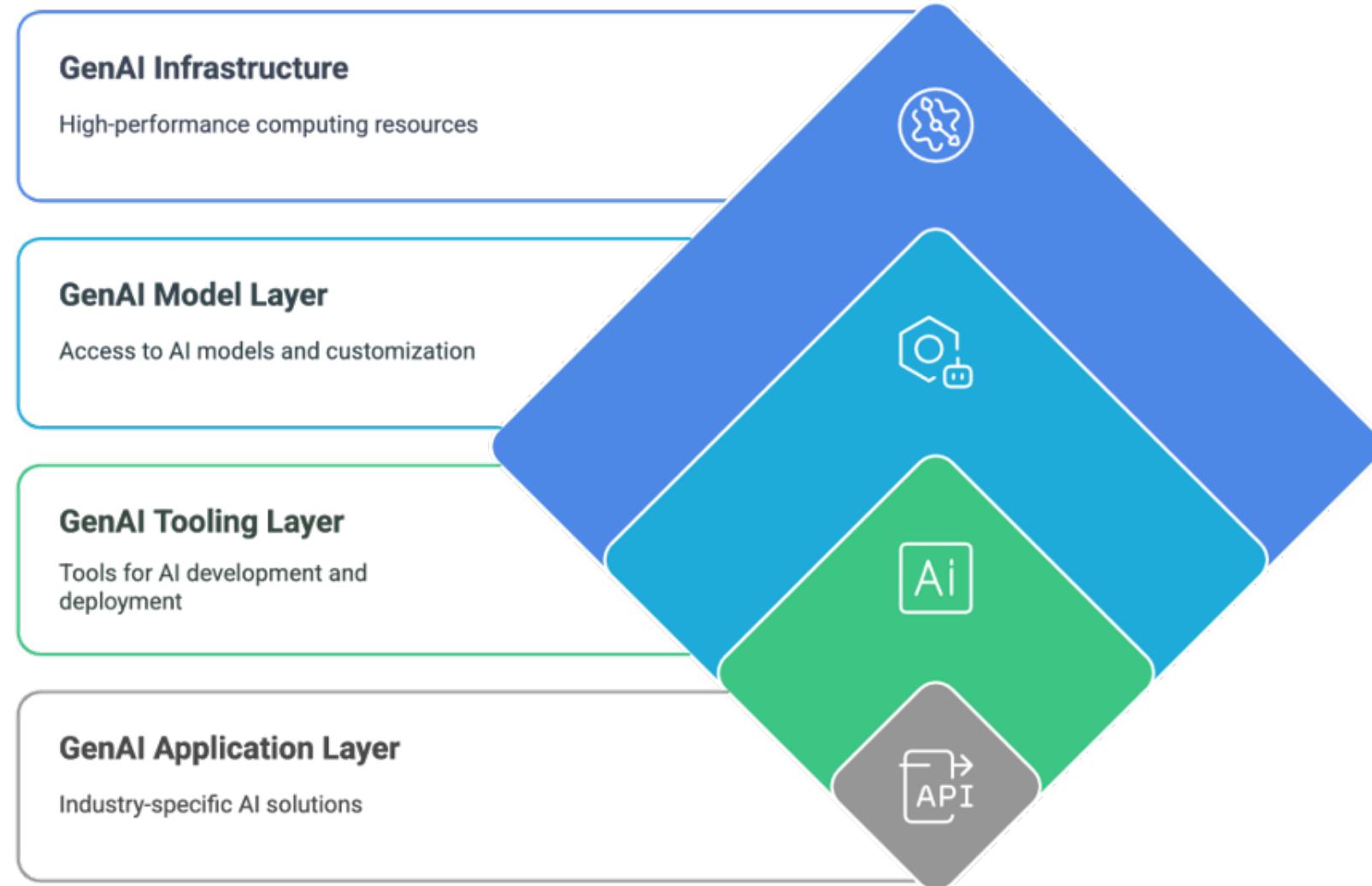
yahoo!finance



Very high entry barriers, natural oligopoly, mostly US and China

By nerds, for nerds

Here's where the most value will be created





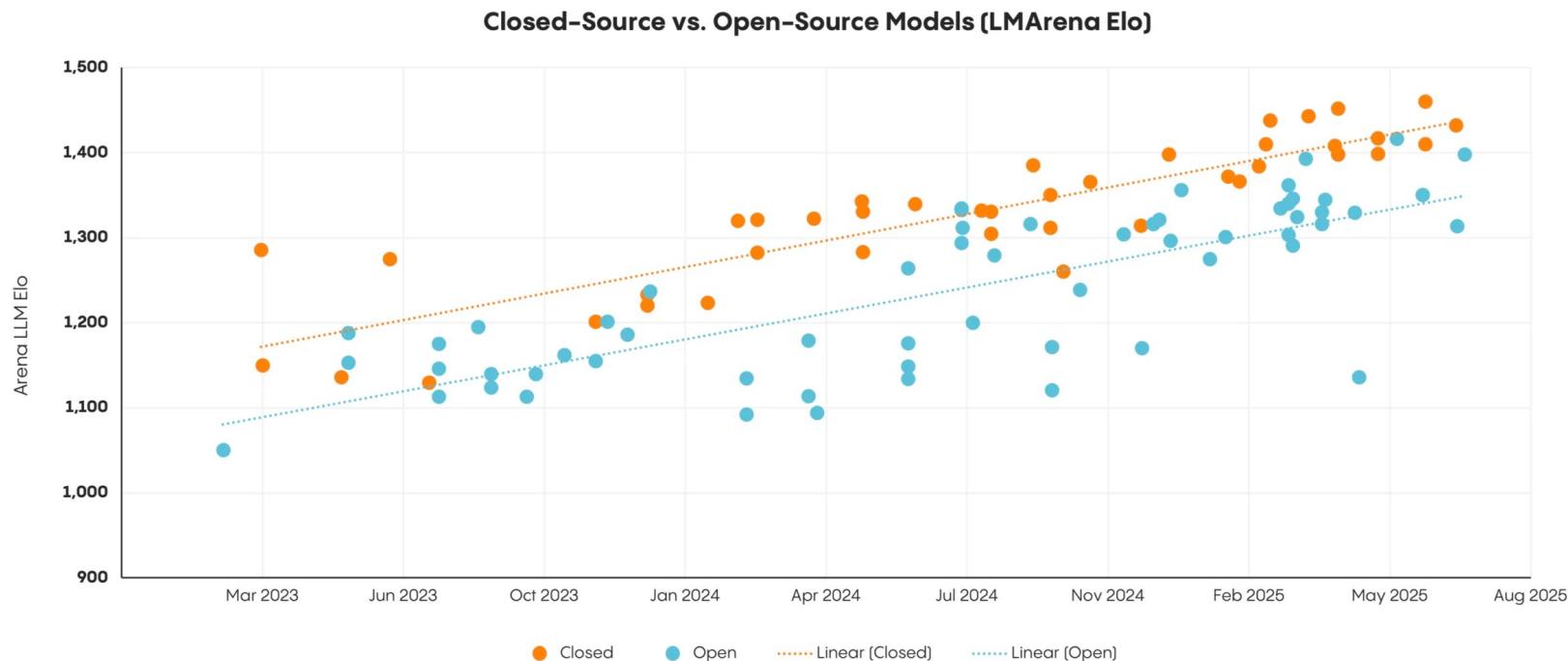
Open-Source vs Closed Source

Features	Open-Source LLMs	Close-Source LLMs
Cost	Lower initial cost but possibly higher maintenance.	Higher initial cost but lower maintenance.
Innovation	Collaborative and fast innovation from a global community.	Slower innovation directed by a single entity.
Support	Reliant on community-based support, which can be extensive but inconsistent.	Consistent and quality support offered through dedicated channels.
Customization	Highly customizable with access to source code.	Customization possible within the platform's limits.
IP	Shared IP under open-source licenses.	IP is closely held, providing competitive advantage.



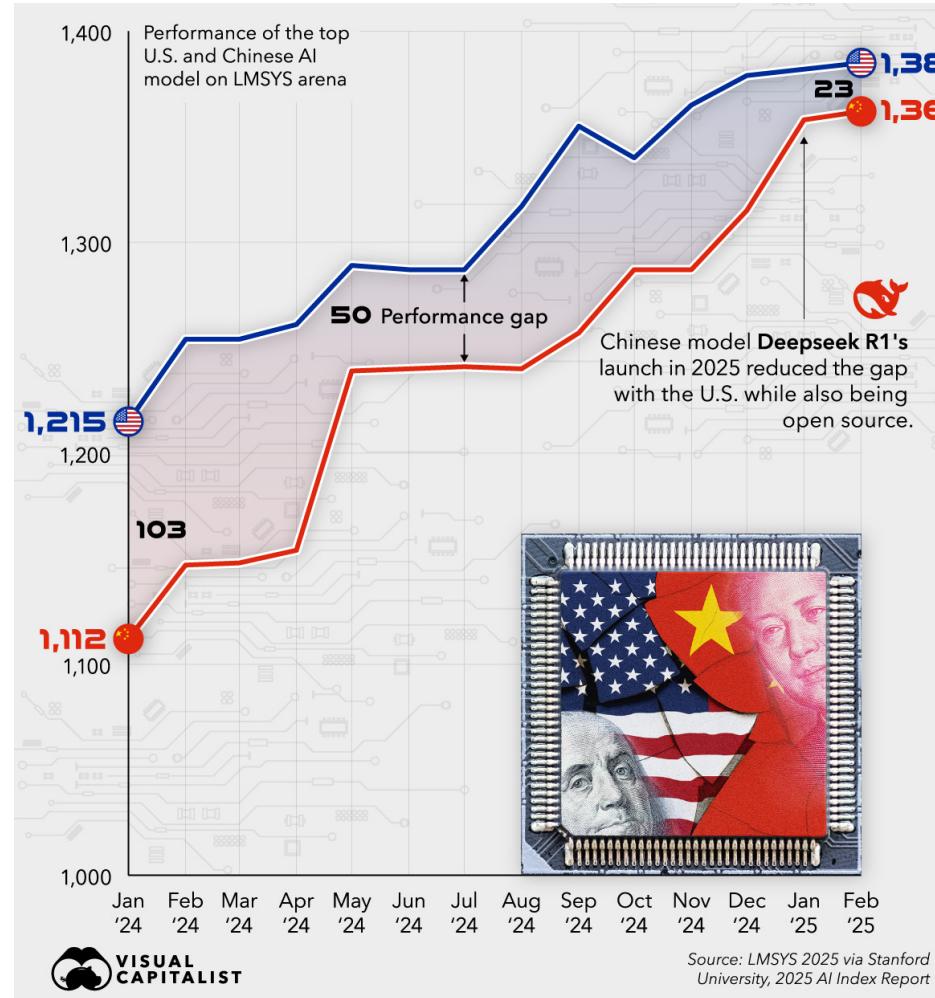
Open-Source Models are Good

Closed-Source vs. Open-Source Models





And Are More and More Chinese



Discussion

Why does Meta offer its models for free, open-sourcing its code?



AI Startups

Discussion

What AI startups do you know?





Cursor

The screenshot shows the Cursor IDE interface. On the left, there is a code editor with a dark theme containing Rust code for a prime number checker. The code defines a function `is_prime` that takes an `u32` and returns a `bool`. It uses a loop to check divisibility up to the square root of the number. The `main` function reads a number from standard input and prints whether it is prime or not. In the center, a large text box displays an AI-generated suggestion: "Based on the code, I would suggest naming this file `prime_checker.rs`. This name accurately reflects the purpose of the code, which is to check whether a given number is prime or not." At the bottom of the code editor, there is a message from the AI asking for a follow-up question. On the right side of the interface, there is a "CHAT" section with a message from the AI asking for a file name.

```

JS todo.js • ⇝ Cursor Settings ⌂ Untitled-1 •
1 | %%L to chat, %%K to generate
2
3 ````Untitled-1
4 // Start Generation Here
5
6
7 fn is_prime(n: u32) -> bool {
8     if n <= 1 {
9         return false;
10    }
11    for i in 2..(n as f64).sqrt() as u32 + 1 {
12        if n % i == 0 {
13            return false;
14        }
15    }
16    true
17 }

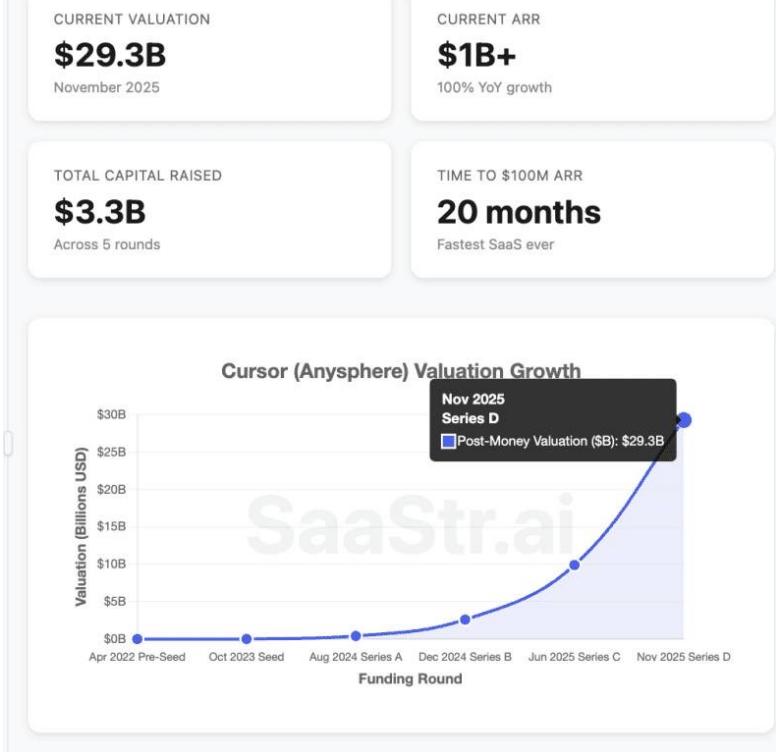
18 fn main() {
19     let mut input = String::new();
20     println!("Enter a number: ");
21     std::io::stdin().read_line(&mut input).expect("Failed to read line");
22     let num: u32 = match input.trim().parse() {
23         Ok(num) => num,
24         Err(_) => panic!("Please type a number!"),
25     };
26     if is_prime(num) {
27         println!("{} is a prime number", num);
28     } else {
29         println!("{} is not a prime number", num);
30     }
31 }
32 }

33
34
35
36
37

```

Cursor's Valuation Journey

From \$0 to \$29.3B in 43 Months: The Fastest SaaS Growth Story Ever





Lovable

The screenshot shows the Lovable application interface. On the left, there is a code editor window with some React component code. In the center, there is a preview of a 'Pricing' page with two plans: 'Plus' (\$19) and 'Pro' (\$49). On the right, there is a dark sidebar with a message from the team.

```

const Label = React.forwardRef<
  React.ElementType,
  React.PropsWithRef<LabelPrimitive.Root>
>((props, ref) => {
  <LabelPrimitive.Root
    ref={ref}
    className={cn(labelVariants, className)}
    {...props}
  >
)
Label.displayName = LabelPrimitive.Root.displayName

export { Label }

...

```

Pricing

Check out our affordable pricing plans

Monthly Yearly

Plus	Pro
For personal use	For professionals
\$19	\$49
Billed \$228 annually	Billed \$588 annually
Up to 5 team members	Everything in Plus, and:
<input type="radio"/> Basic components library	<input type="radio"/> Unlimited team members
<input type="radio"/> Community support	<input type="radio"/> Advanced components
<input type="radio"/> 10GB storage space	<input type="radio"/> Priority support
<input type="radio"/> Unlimited storage	<input type="radio"/> Unlimited storage

Remember: Do not change the component's code unless it's required to integrate or the user asks you to.
IMPORTANT: Create all mentioned files in full, without abbreviations. Do not use placeholders like "Insert the rest of the code here" – output every line of code exactly as it is, so it can be copied and pasted directly into the project.

Lovable

I'll help you integrate the Pricing2 component into the project. Since most of these files already exist in your project like button.tsx, switch.tsx, separator.tsx, and card.tsx, we only need to create the pricing2.tsx file and update the index.tsx to include it.

Edit #98 Add pricing component Preview

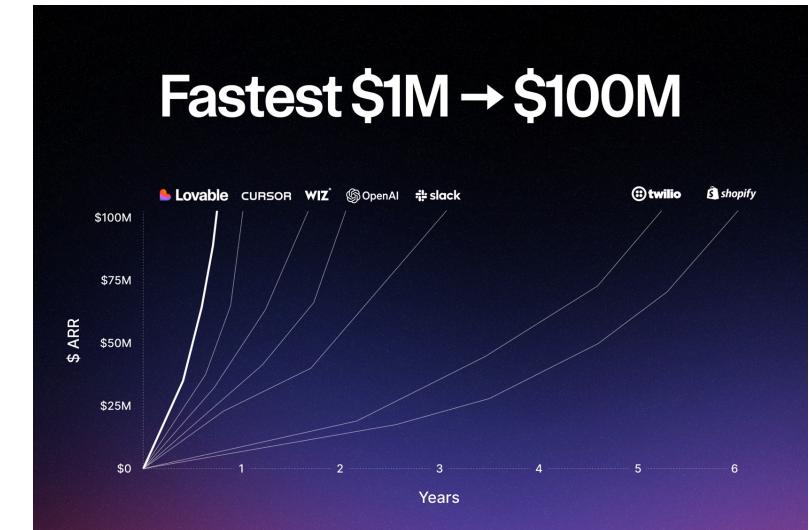
Added the Pricing2 component with a toggle between monthly and yearly pricing plans, featuring responsive design and integration with the existing UI components.

From You

Stay up to date with the news.

Get notified about the latest features and updates, straight from the team.

Juk Lovable...





Harvey

Harvey

All General Transactional Litigation Financial services

Recommended for You

- Translate into Another Language** ★
Upload a document and Harvey will translate it into the language of your selection.
Output · 2 steps
- Proofread for Spelling and Grammar**
Upload a document and Harvey will return an updated version without spelling or grammar...
Draft · 1 step

For Litigation Work

- Analyze a Deposition Transcript for Key Topics**
Upload a deposition and Harvey will produce a summary of the key topics discussed in the...
Output · 2 steps
- Analyze a Trial Transcript for Key Topics** ★
Upload a trial transcript and Harvey will produce a summary of the key topics discussed in the...
Output · 2 steps

For Transactional Work

- Summarize Material Changes from Redlines**
- Draft an Item 1.01 Disclosure**





Abridge

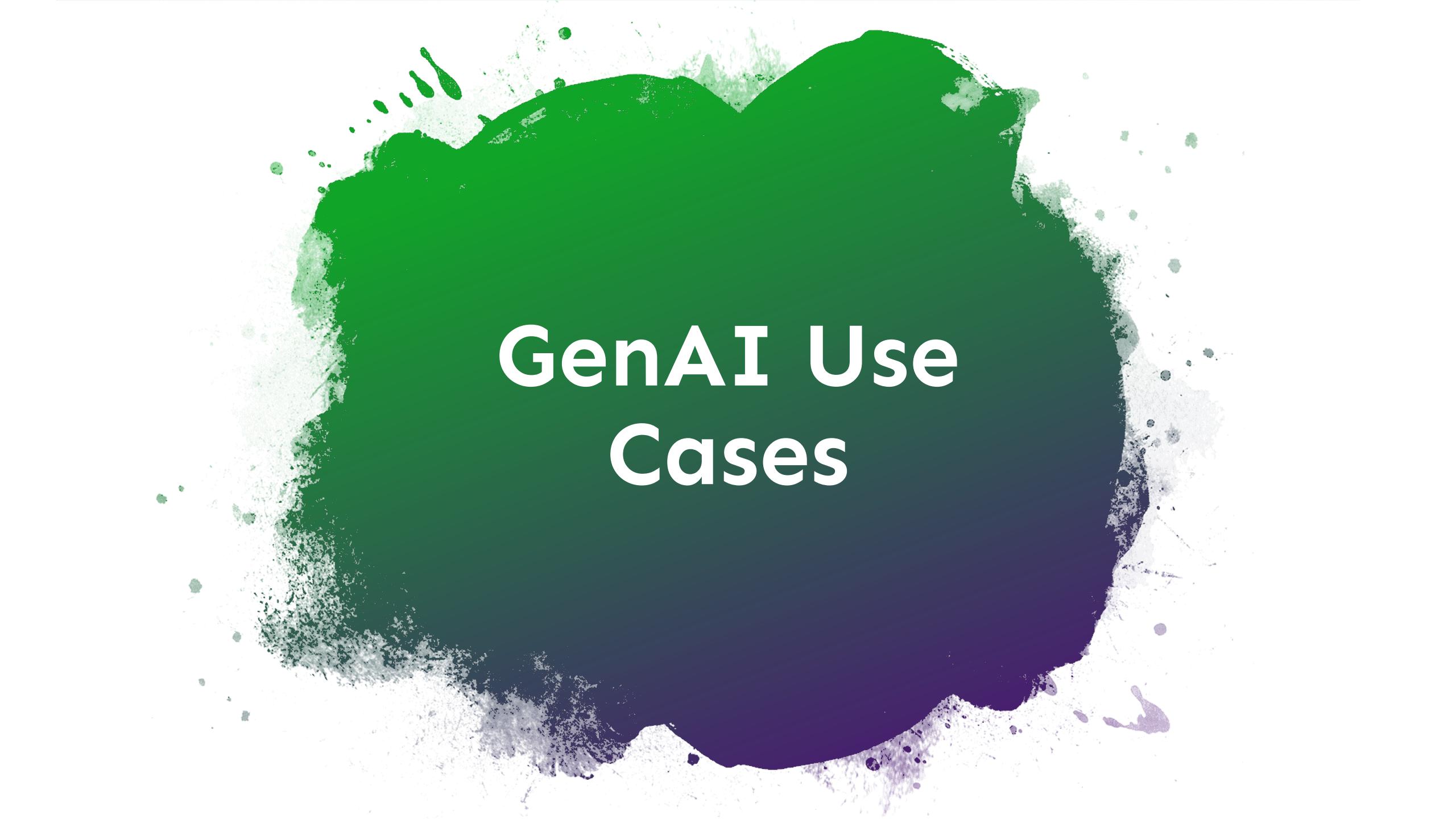
The image shows the Abridge mobile application running on an iPhone and its integration with the Epic electronic medical record system on a laptop. The mobile app features a red background with the word "Abridge" in white, a patient summary, and a "CREATE NOTE" button. The laptop screen displays the Epic interface for a patient named Cobb, Jessica, showing clinical notes, a transcript, and a detailed assessment & plan section. The Abridge logo is visible in the top right corner of the laptop screen.

© 2024 Epic Systems Corporation

A ANNOUNCEMENTS
Posted June 24, 2025
by Abridge

ABRIDGE SECURES \$300M SERIES E LED BY A16Z TO PIONEER A NEW PARADIGM OF CARE INTELLIGENCE





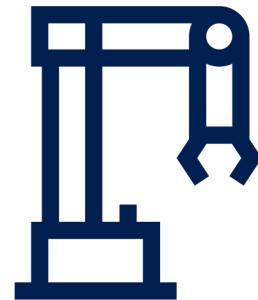
GenAI Use Cases



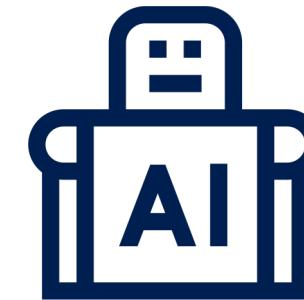
Use Cases



Content
generation &
transformation



Text mining &
Information
extraction



Agentic
Automation



Conversational
Interaction



Internal Operations



**Writing emails
and documents**



**Summarize
meetings**



**Help analyzing
data with Excel**



Copilot

Your everyday AI companion





Marketing

Content
Generation



Automated
Market Research



Synthetic*users*





Marketing



World ▾ Business ▾ Markets ▾ Sustainability ▾ Legal ▾ Commentary ▾ Technology ▾ Investigations ▾ Mc

Klarna using GenAI to cut marketing costs by \$10 mln annually

By Supantha Mukherjee

May 28, 2024 12:30 PM GMT+2 · Updated May 28, 2024





Marketing

"Using GenAI tools like Midjourney, DALL-E, and Firefly for image generation, Klarna said it has **reduced image production costs by \$6 million**.

It is using GenAI to update images on its app and website on a weekly basis, reflecting key retail events such as Valentine's Day, Mother's Day and summer sales.

A further **\$4 million in savings** come from cutting spending on external marketing suppliers for translation, production, and social agencies."

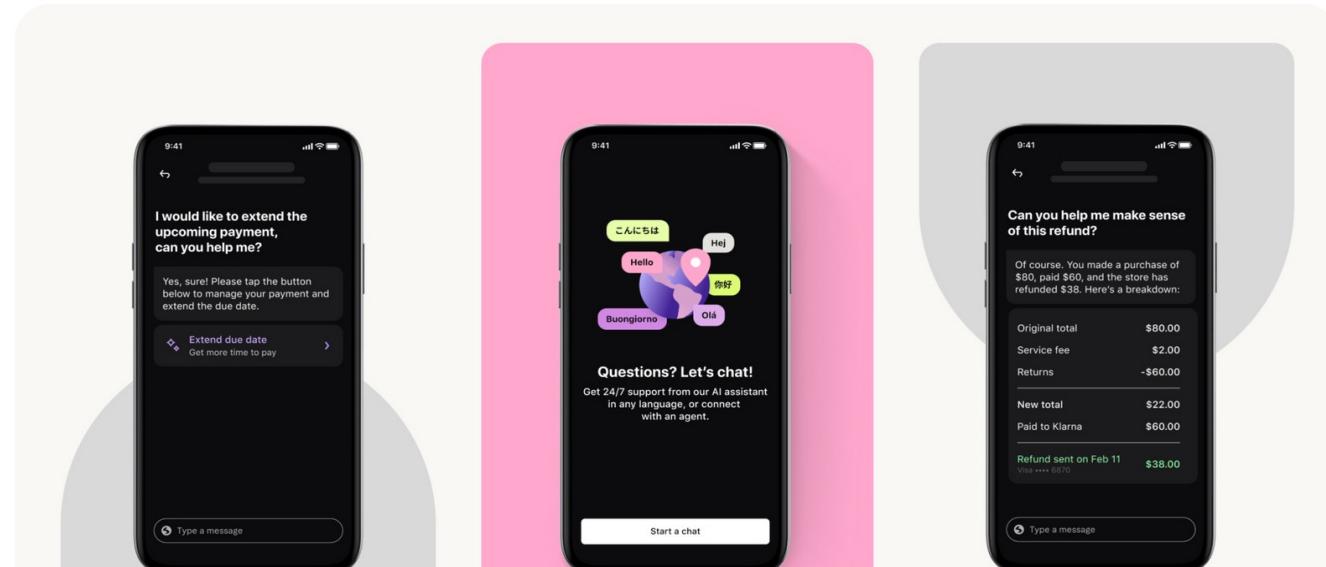


Customer Support

Klarna[Investor relations](#) [Corporate governance](#) [ESG](#) [News](#) [About us](#)

Klarna AI assistant handles two-thirds of customer service chats in its first month

February 27, 2024

Klarna
Klarna



Customer Support

Events Video Special Issues Jobs

VentureBeat

Subscribe GamesBeat

Artificial Intelligence ▾ Security ▾ Data Infrastructure ▾ Automation ▾ Enterprise Analytics ▾ More ▾

Feature

Wells Fargo's AI assistant just crossed 245 million interactions – no human handoffs, no sensitive data exposed

Matt Marshall
@mmarshall

April 8, 2025 12:01 PM

f X in





Compliance

≡ SEARCH

FORTUNE

SUBSCRIBE NOW

SIGN IN

HOME NEWS TECH FINANCE LEADERSHIP WELL EDUCATION FORTUNE 500

FINANCE · A.I.

Goldman Sachs CEO says that AI can draft 95% of an IPO prospectus in minutes

BY PAOLO CONFINO

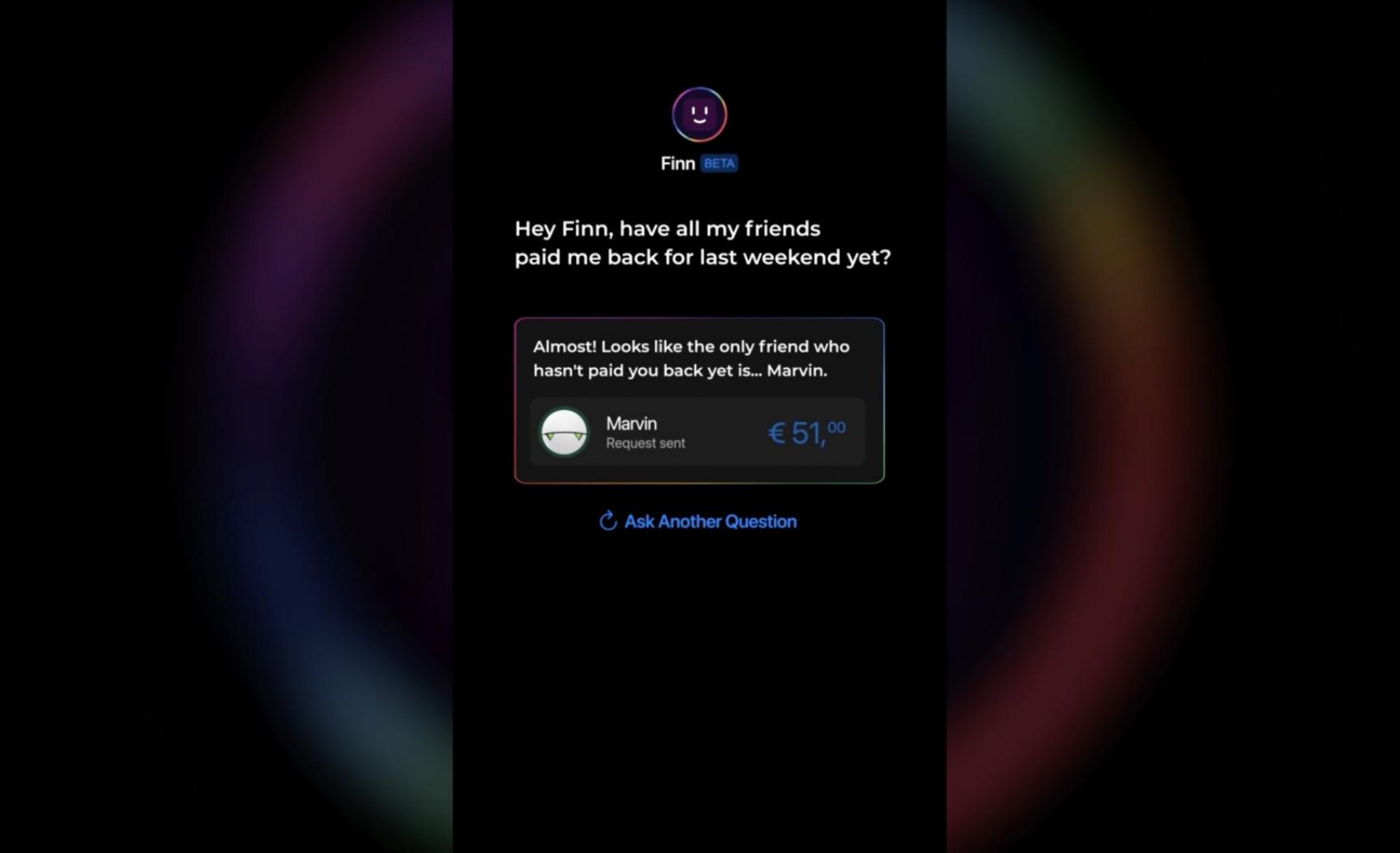
January 17, 2025 at 8:28 PM GMT+1



Goldman Sachs CEO David Solomon
HOLLIE ADAMS/BLOOMBERG



Product





Product



DOVE MANGIARE

INTERVISTE

CURIOSITÀ

RISTORAZIONE



PRENOTA ORA

Clicca qui per scaricare l'app di TheFork!

TheFork abbraccia l'intelligenza artificiale con Operator di OpenAI

12/03/2025

2 minuti



Paolo



TheFork lancia l'integrazione con Operator di OpenAI per la prenotazione automatica dei ristoranti grazie all'intelligenza artificiale

Inseguire l'innovazione è da sempre uno dei punti di forza di TheFork, che ancora una volta si conferma pioniera nell'avanguardia tecnologica. Ecco infatti che, con orgoglio, si posiziona ora tra le prime aziende a **integrare Operator di OpenAI** nella propria piattaforma. Si tratta di un avanzato assistente IA progettato per navigare sul web ed eseguire in modo indipendente una varietà di attività per gli utenti, tra cui la prenotazione di ristoranti.

*"Grazie a OpenAI, stiamo sfruttando il meglio dell'efficienza e della personalizzazione offerte dall'intelligenza artificiale per capire e migliorare ogni aspetto dell'esperienza culinari", ha dichiarato **Almir Ambeskovic**, CEO di TheFork. "L'IA ha già dimostrato un fattore strategico per migliorare l'esperienza offerta da TheFork, e questo è solo uno dei tanti passi avanti per realizzare la nostra visione innovativa, rendendo TheFork ancora più prezioso per i nostri utenti."*

Come funziona Operator

Da **venerdì 21 febbraio** i clienti del Regno Unito possono avvalersi dell'integrazione di TheFork con Operator di OpenAI per **automatizzare le prenotazioni, ricevere consigli personalizzati sui ristoranti e gestire le prenotazioni in modo semplice ed efficace**. Inoltre, le offerte speciali disponibili nei ristoranti saranno applicate automaticamente da Operator di OpenAI, garantendo agli utenti le migliori offerte possibili.

GenAI Pricing Models

Discussion

How would you bill a
GenAI product like
Cursor or Harvey?





Tiered Subscription

Pricing

Monthly Yearly

Individual Plans

Hobby Free

Includes:

- ✓ One-week Pro trial
- ✓ Limited Agent requests
- ✓ Limited Tab completions

Download

Pro

\$20/mo.

Everything in Hobby, plus:

- ✓ Extended limits on Agent
- ✓ Unlimited Tab completions
- ✓ Background Agents
- ✓ Maximum context windows

Get Pro

Pro+ Recommended

\$60/mo.

Everything in Pro, plus:

- ✓ 3x usage on all OpenAI, Claude, Gemini models

Get Pro+

Ultra

\$200/mo.

Everything in Pro, plus:

- ✓ 20x usage on all OpenAI, Claude, Gemini models
- ✓ Priority access to new features

Get Ultra



Usage-based

Flagship models

Our frontier models are designed to spend more time thinking before producing a response, making them ideal for complex, multi-step problems.

GPT-5.2

The best model for coding and agentic tasks across industries

Price

Input:
\$1.750 / 1M tokens

Cached input:
\$0.175 / 1M tokens

Output:
\$14.000 / 1M tokens

GPT-5.2 pro

The smartest and most precise model

Price

Input:
\$21.00 / 1M tokens

Cached input:
-

Output:
\$168.00 / 1M tokens

GPT-5 mini

A faster, cheaper version of GPT-5 for well-defined tasks

Price

Input:
\$0.250 / 1M tokens

Cached input:
\$0.025 / 1M tokens

Output:
\$2.000 / 1M tokens



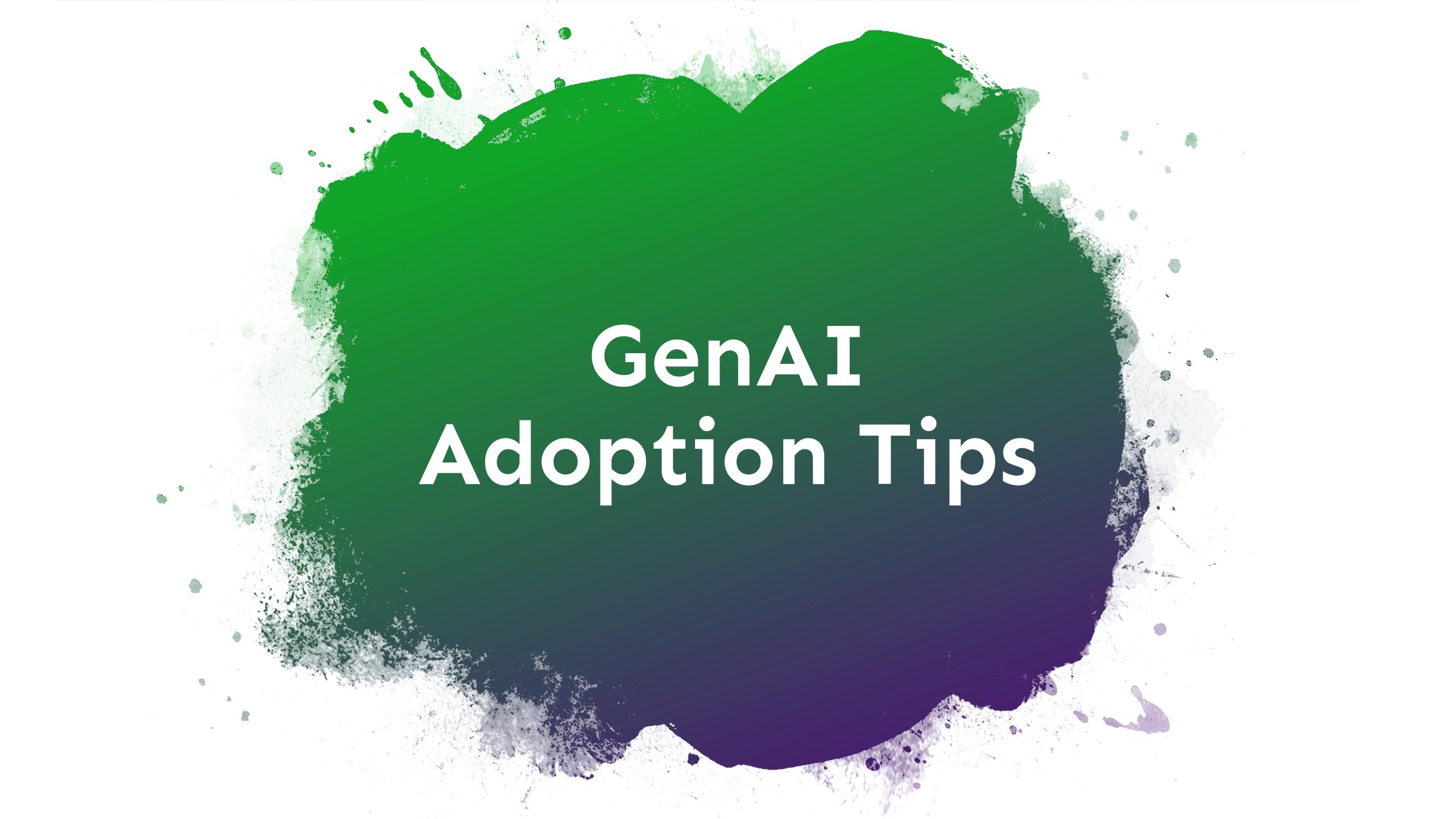
Value-Based

INTERCOM CUSTOMER SERVICE SUITE

Get the #1 AI Agent on one seamless platform.

Billed annually Billed monthly

Essential	Pro	Team	Enterprise
<p>Includes Fin AI Agent</p> <p>The customer support plan for individuals, startups, and small businesses.</p> <p>\$0.99 per Fin resolution ⓘ \$29 per seat/mo, billed annually</p> <p>Free 14 day trial</p> <p>KEY FEATURES INCLUDE</p> <ul style="list-style-type: none">✓ Fin AI Agent • \$0.99 per resolution✓ Multiple team Inboxes✓ Workflow automation builder	<p>Resolutions pricing</p> <p>Fin AI Agent is included in all Intercom plans and can also be used with your current helpdesk.</p> <p>Fin is priced at \$0.99 per resolution — meaning you only pay for the value it delivers.</p> <p>A resolution is counted when either:</p> <ul style="list-style-type: none">• The customer confirms the AI Answer resolved their issue, or• The customer doesn't request more help after the last AI Answer. <p>You'll only be charged once per conversation, even if Fin resolves multiple questions. Learn more</p> <p>For Fin Voice, high-volume pricing or specialized needs, please talk to Sales</p> <p>\$85 per seat/mo, billed annually</p> <p>Free 14 day trial Get a demo</p> <p>EVERY ESSENTIAL FEATURE, PLUS</p> <ul style="list-style-type: none">✓ SSO & identity management✓ GDPR & CCPA support	<p>\$132 per seat/mo, billed annually</p> <p>Free 14 day trial Get a demo</p> <p>EVERY ADVANCED FEATURE, PLUS</p> <ul style="list-style-type: none">✓ Set up in under an hour on your current helpdesk	<p>Already have a helpdesk?</p> <p>Fin AI Agent</p> <p>Use Fin with your current helpdesk including Zendesk, Salesforce, and more. Learn more</p> <p>\$0.99 per Fin resolution ⓘ Minimum commitments apply</p> <p>Free 14 day trial Get a demo</p> <p>FEATURES INCLUDE</p> <ul style="list-style-type: none">✓ Set up in under an hour on your current helpdesk



GenAI Adoption Tips



These are **informed suggestions** — feel free to disagree.

Questions and comments are welcome.



Don't Build Your Model

Developing an AI model from scratch is **complex**, costly, and requires specialised expertise.

For most use cases, general-purpose models are more than sufficient.

Let them handle the heavy lifting—your time is better spent focusing on what truly sets you apart: your **unique data**, insights, and domain knowledge.



BloombergGPT: A Large Language Model for Finance

Shijie Wu^{1,*}, Ozan İrsoy^{1,*}, Steven Lu^{1,*}, Vadim Dabrowski¹, Mark Dredze^{1,3},
Sebastian Gehrmann¹, Prabhanjan Kambadur¹, David Rosenberg², Gideon Mann¹

¹ Bloomberg, New York, NY USA

² Bloomberg, Toronto, ON Canada

³ Computer Science, Johns Hopkins University, Baltimore, MD USA

Abstract

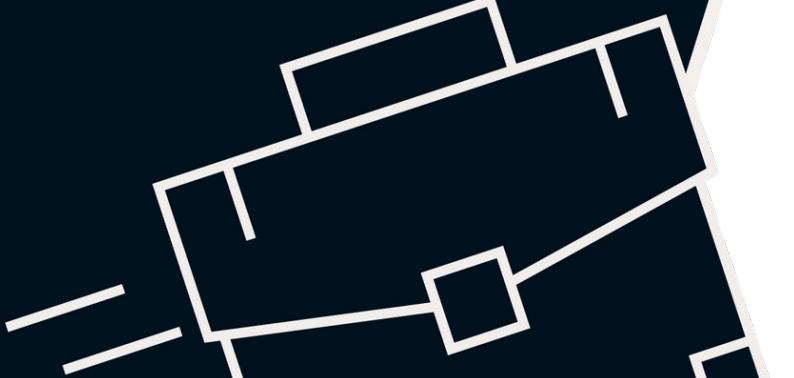
The use of NLP in the realm of financial technology is broad and complex, with applications ranging from sentiment analysis and named entity recognition to question answering. Large Language Models (LLMs) have been shown to be effective on a variety of tasks; however, no LLM specialized for the financial domain has been reported in literature. In this work, we present BLOOMBERGGPT, a 50 billion parameter language model that is trained on a wide range of financial data. We construct a 363 billion token dataset based on Bloomberg's extensive data sources, perhaps the largest domain-specific dataset yet, augmented with 345 billion tokens from general purpose datasets. We validate BLOOMBERGGPT on standard LLM benchmarks, open financial benchmarks, and a suite of internal benchmarks that most accurately reflect our intended usage. Our mixed dataset training leads to a model that outperforms existing models on financial tasks by significant margins without sacrificing performance on general LLM benchmarks. Additionally, we explain our modeling choices, training process, and evaluation methodology. We release Training Chronicles ([Appendix C](#)) detailing our experience in training BLOOMBERGGPT.

Contents

1	Introduction	3
1.1	BLOOMBERGGPT	3
1.2	Broader Contributions	4
2	Dataset	5
2.1	Financial Datasets (363B tokens – 51.27% of training)	7
2.1.1	Web (298B tokens – 42.01% of training)	7
2.1.2	News (38B tokens – 5.31% of training)	7
2.1.3	Filings (14B tokens – 2.04% of training)	7
2.1.4	Press (9B tokens – 1.21% of training)	8
2.1.5	Bloomberg (5B tokens – 0.70% of training)	8
2.2	Public Datasets (345B tokens – 48.73% of training)	9
2.2.1	The Pile (184B tokens – 25.9% of training)	9
2.2.2	C4 (138B tokens – 19.48% of training)	9
2.2.3	Wikipedia (24B tokens – 3.35% of training)	9
2.3	Tokenization	9



Don't Build Generic Apps



Generative AI is still a new technology, but the ecosystem of GenAI-powered applications is **growing fast**.

Big Tech and specialized companies are rapidly developing solutions for common use cases—like meeting summarization.

Don't waste time building what others can do better and faster. Focus on what sets you apart.



ChatGPT Record

Capture and summarize meetings & voice notes

Updated this week

Overview

Note: Currently, ChatGPT record is available for Enterprise, Edu, Team, and Pro workspaces, and only available for the macOS desktop app

ChatGPT may make mistakes, including in its transcriptions, so please check important information.

With record mode, ChatGPT can transcribe and summarize audio recordings like meetings, brainstorms, or voice notes.

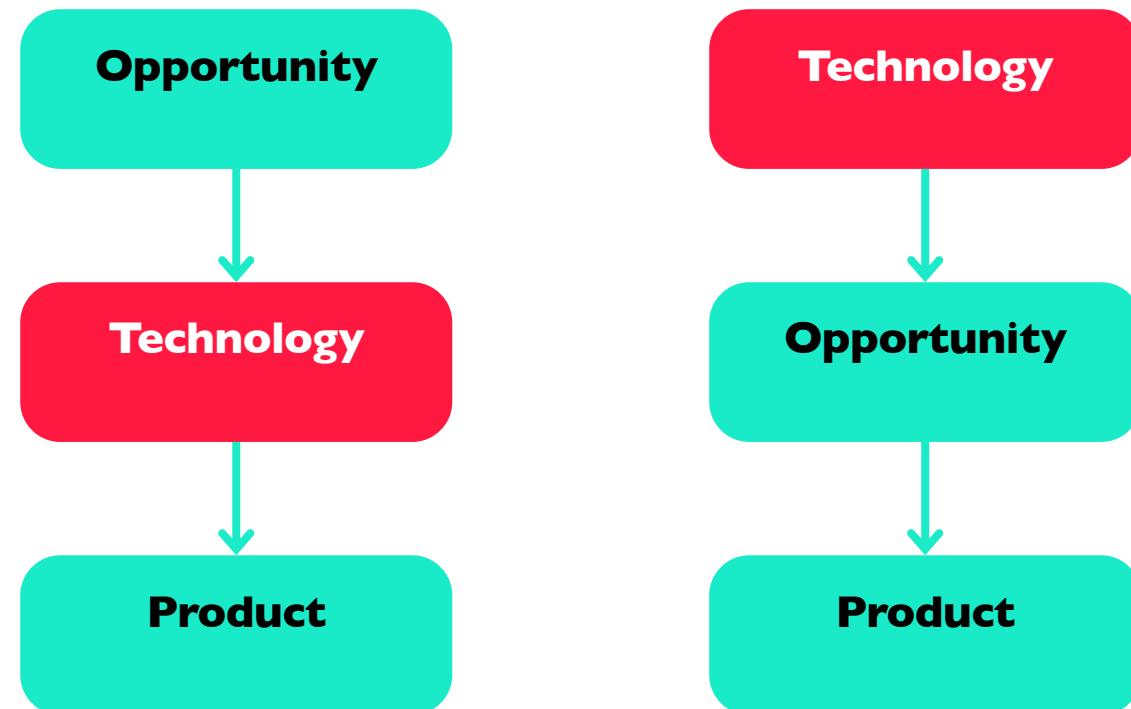
These summaries are saved as canvases in your chat history and can be turned into helpful outputs like project plans, emails, or even code. ChatGPT can also reference canvases and transcripts from your past recordings, in order to provide more helpful responses across conversations.

Note: Use responsibly when recording others.

Please make sure you check local laws and always get the right consents before recording others. You're responsible for making sure that your use of record mode follows applicable laws, which may vary depending on where you and the people you're recording are located.



Balance Tech Pull and Tech Push



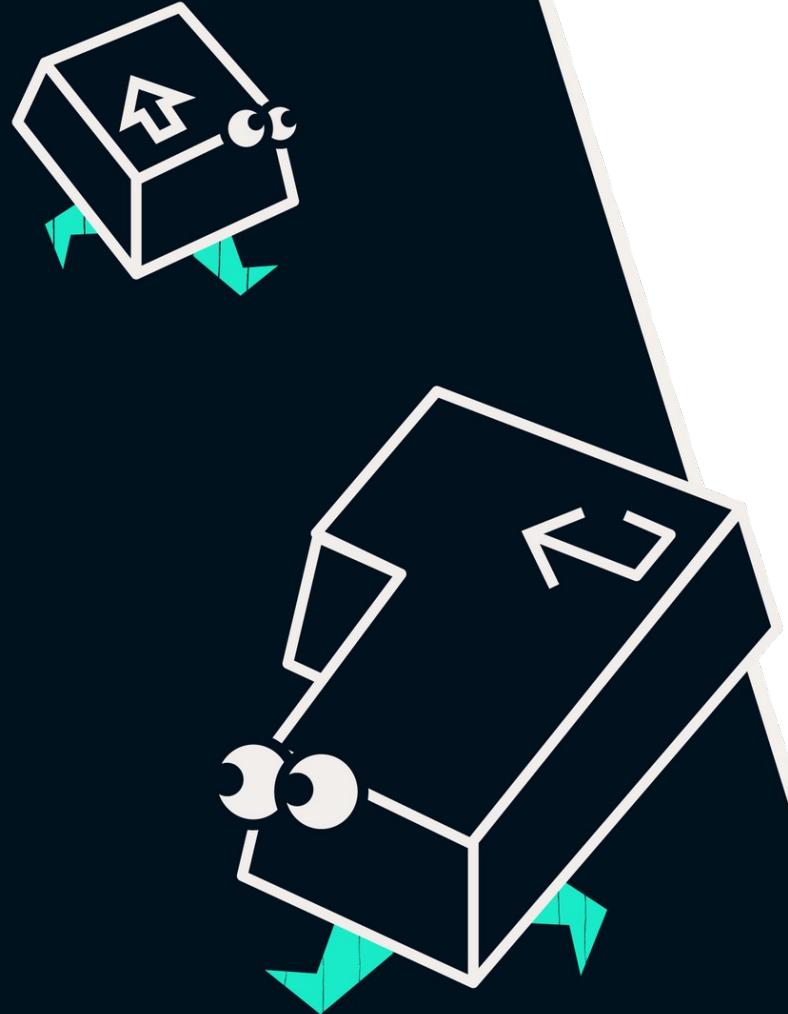


Prototype is Easy, Production is Hard

Tools like v0 and Cursor make it incredibly easy to build prototypes—great for fast iteration and experimentation.

This can significantly impact how teams approach **sales and R&D**.

However, turning a prototype into a production-ready product still requires substantial time and investment. Make sure to plan for that.



GenAI Is NOT The Product

GenAI enables features and products that weren't possible before—but it won't solve your customers' problems on its own.

Be intentional about **where you integrate GenAI**; don't just add AI for the sake of it.

And remember: you don't always need a **chatbot**.



People Need Education and Tools

GenAI offers a strong **productivity boost** for internal processes—and more people are starting to realize it.

They'll use it, whether the company supports it or not.

That's why it's better to equip them with the **right skills** and **tools** from the start.



68%

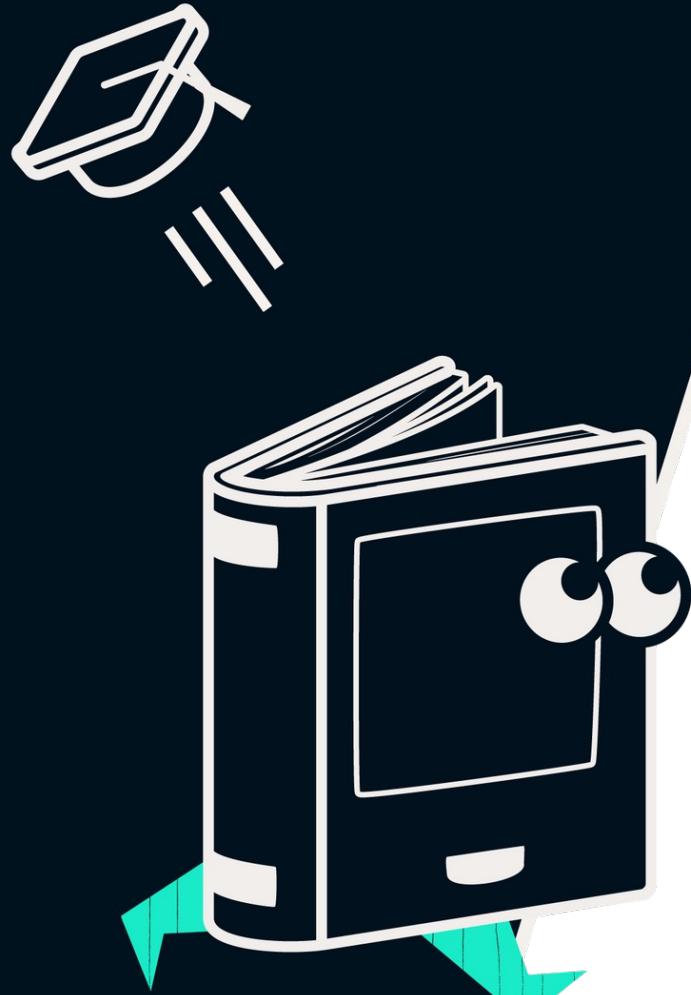
Of managers recommended a GenAI tool to solve a team member's problem last month

22%

Of corporates are realising gains from AI adoption.

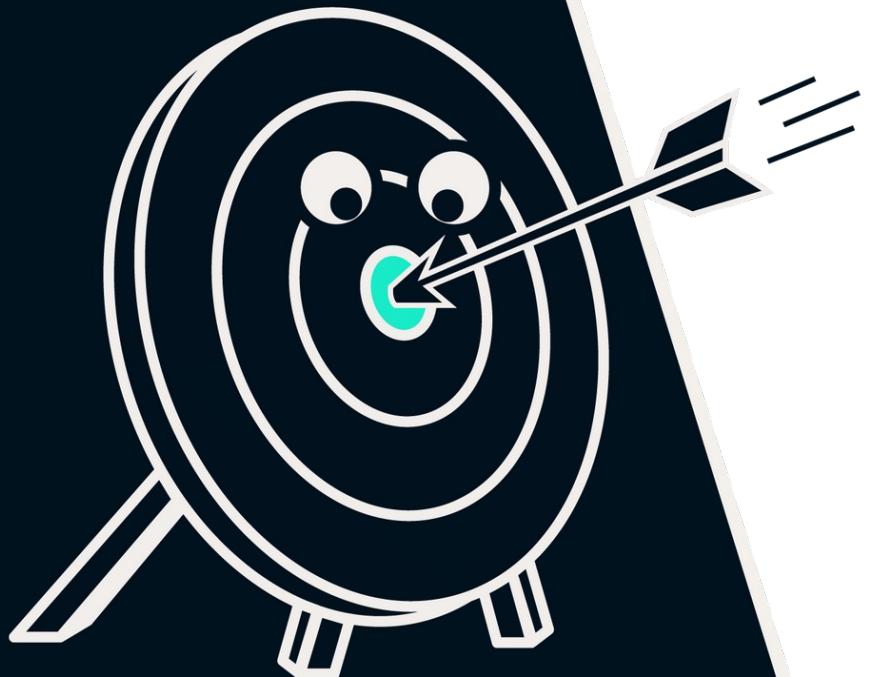
<8%

Of workers have participated into upskilling and reskilling programs to foster GenAI adoption.



EU AI Act - Art. 4

“Providers and deployers of AI systems shall take measures to ensure, to their best extent, a **sufficient level of AI literacy of their staff** and other persons dealing with the operation and use of AI systems on their behalf. [omissis]”



Not All AI is GenAI

Predictive AI isn't dead—far from it. In fact, it can now be powered by higher-quality data generated through GenAI-driven document processing.

The key is to use the **right tool for the right task**: traditional AI models remain the best choice for problems like real estate valuation, credit scoring, and similar applications.



The End