



ChatGPT and the Adoption of AI Tools

January 2023

Whitney K. Rothe



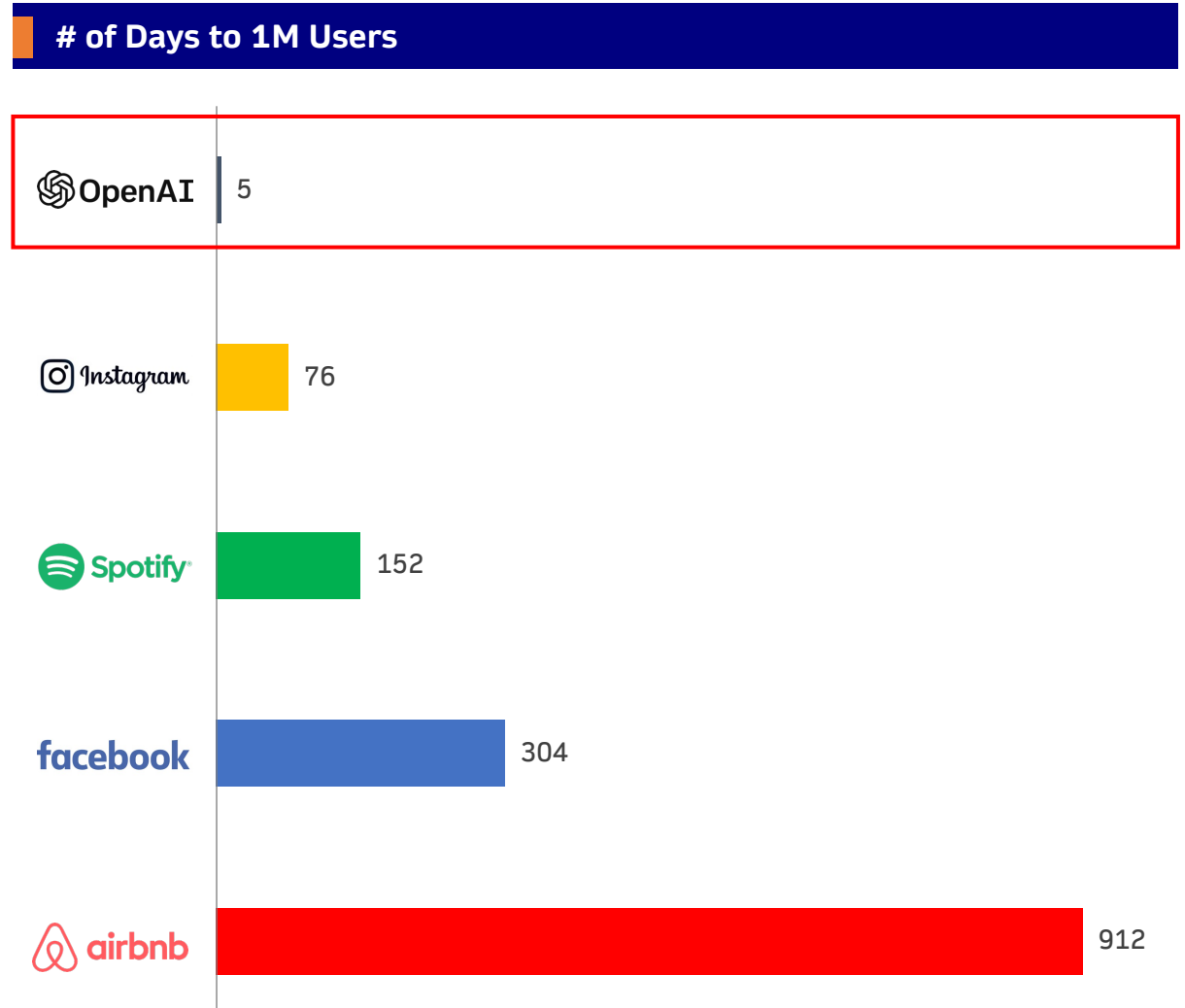
Today's Discussion

- How to think about ChatGPT
- Evaluating the commercial viability of AI
- Where the technology is headed

How to think about ChatGPT

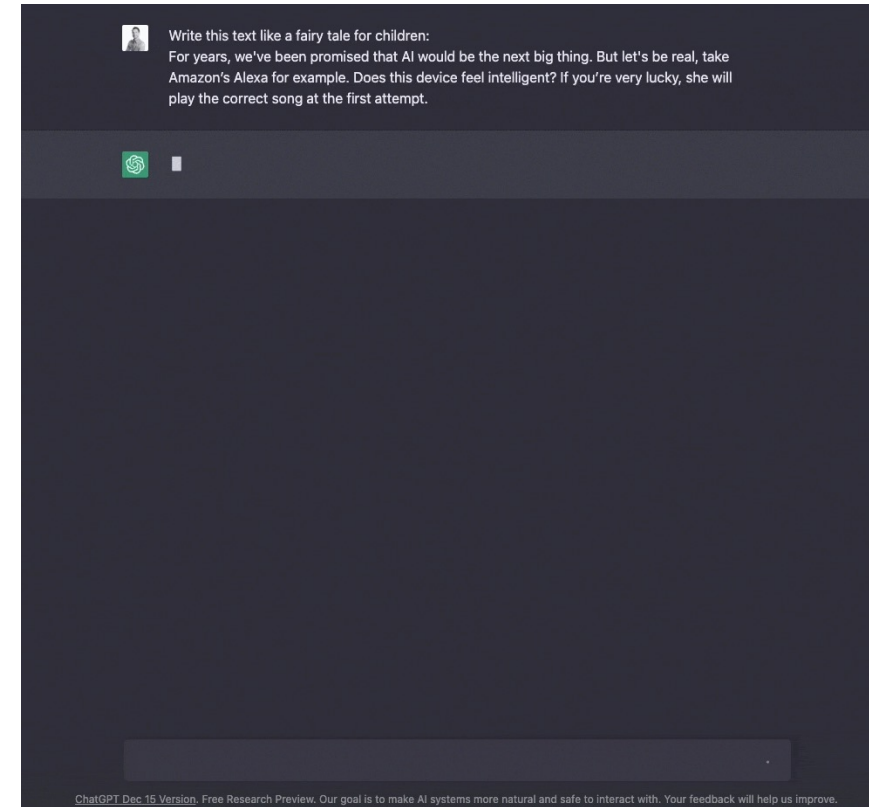
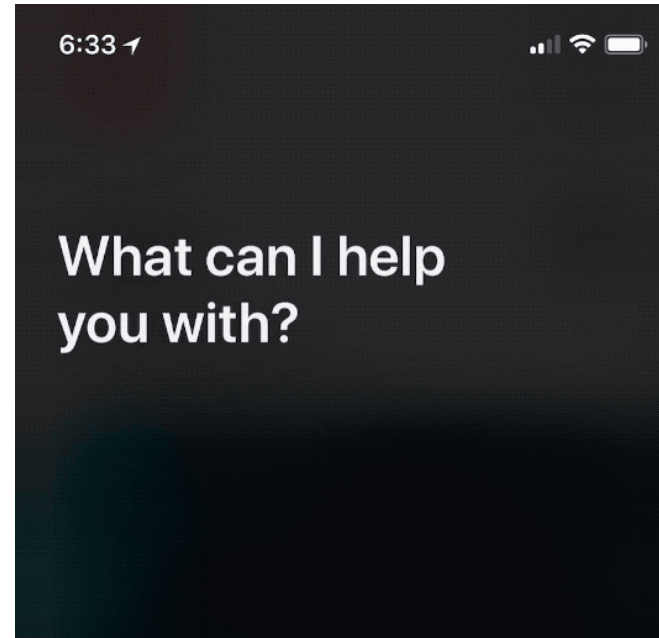
About ChatGPT

- Launched in November 2022
- Built on GPT-3 language model
- Developed by OpenAI (2015)
- Backed by Silicon Valley insiders
- Strategic partnership with Microsoft
- \$1B committed to date
- Projecting \$200M in revenue by end of 2023



Why has ChatGPT created such buzz?

Natural language processing capabilities have reframed the established expectations of end users.



How ChatGPT works

- ChatGPT's key breakthrough is in its usage of **Reinforcement Learning from Human Feedback (RLHF)**

- Human feedback is embedded within the training loop to minimize harmful, untruthful, and/or biased outputs

AI is evaluated across two key dimensions:

- **Capability** how well it can **optimize its objective function** (the goal defined by the mathematical expression)
- **Alignment** what we want the **model to do** (versus the data on which it is trained)

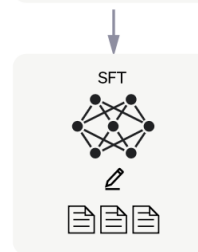
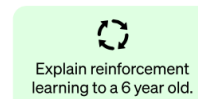
1

A list of like prompts are selected and a group of human labelers create the expected output response.

Step 1

Collect demonstration data and train a supervised policy.

A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3.5 with supervised learning.

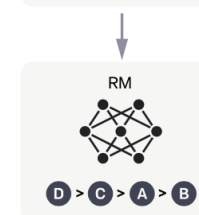
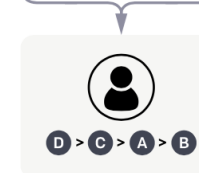
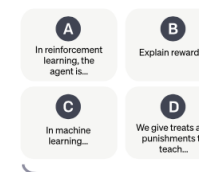
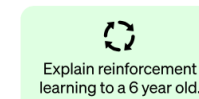
2

Based on the expected response, multiple outputs are then generated. Human labelers evaluate and rank the best.

Step 2

Collect comparison data and train a reward model.

A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.

This data is used to train our reward model.

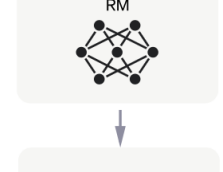
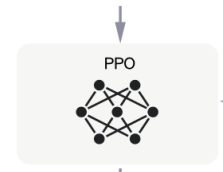
3

This ranking used to train the model to solve for a Reward Score that is aligned to human preferences.

Step 3

Optimize a policy against the reward model using the PPO reinforcement learning algorithm.

A new prompt is sampled from the dataset.



The PPO model is initialized from the supervised policy.

The policy generates an output.

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.

The pros and cons of ChatGPT



- Human-like natural language processing
- Low friction trial
- Scale of internet audience fuels virality
- API creates a path forward for monetization
- Microsoft \$10B investment is a shortcut to enterprise adoption

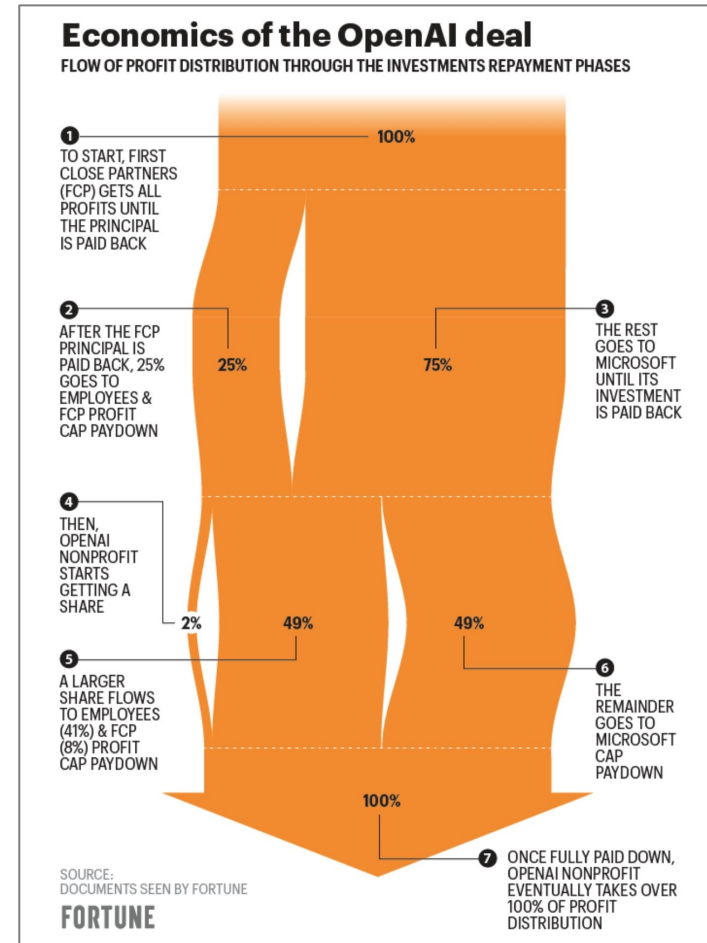
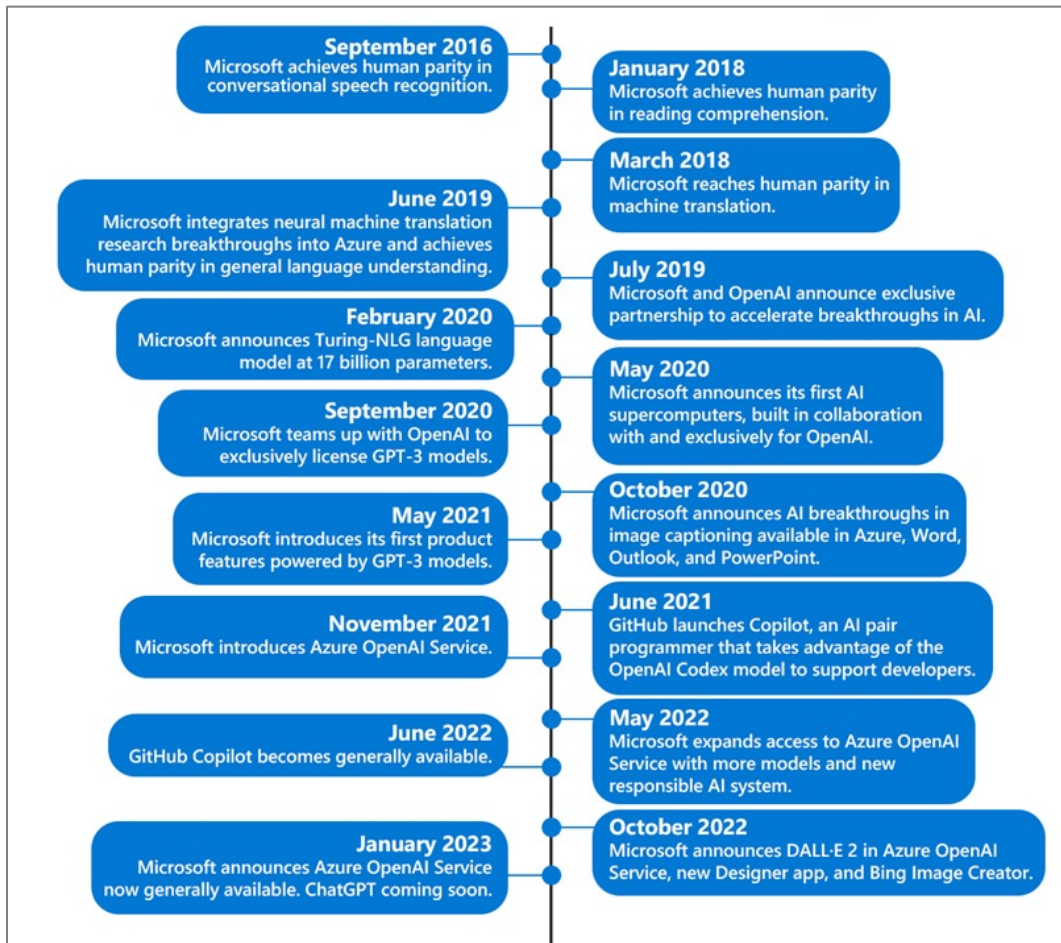


- Public API = Copy cats
- Free -> “\$3M in burn per month”
- Human curation -> slow & \$\$\$
- Compute capacity constrained
- Trained on 2021 data
- Capability \neq Alignment
- Information labeled is not vetted

Why Microsoft is investing \$10B

Exclusive access to the model. OpenAI will be integrated to bolster legacy offerings (Windows, Office, Bing) and accelerate developer ecosystems (Azure, GitHub)

After initial partners are paid back on initial investment round, MSOFT receives 75% of proceeds until its investment is paid back



Evaluating the commercial
viability of AI tools

Signals that an AI tool has commercial viability

- Application Programming Interface (API) first
- Integrates within larger software ecosystems
- Low friction, self-serve end user adoption
- Community-led go-to-market motion

Commercial signals: API first

AI engines are only as valuable as their ability to integrate within existing technologies.

- APIs act as the communication bridge between applications to pass through data
- An API first approach to software development means that the API is a primary product value proposition
- Generative AI scales by being embedded within the context of application layers
- Usage based pricing models are the defacto for AI programs

Image models

Build DALL·E directly into your apps to generate and edit novel images and art. Our image models offer three tiers of resolution for flexibility.

[LEARN MORE ↗](#)

| RESOLUTION | PRICE |
|------------|-----------------|
| 1024×1024 | \$0.020 / image |
| 512×512 | \$0.018 / image |
| 256×256 | \$0.016 / image |

Language models

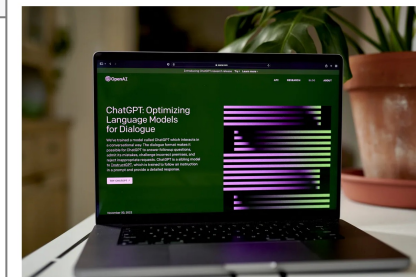
Base models

| | | | |
|---------------------|---------------------|---------------------|------------------------------|
| Ada <i>Fastest</i> | Babbage | Curie | Davinci <i>Most powerful</i> |
| \$0.0004 /1K tokens | \$0.0005 /1K tokens | \$0.0020 /1K tokens | \$0.0200 /1K tokens |

Multiple models, each with different capabilities and price points. **Ada** is the fastest model, while **Davinci** is the most powerful.

Prices are per 1,000 tokens. You can think of tokens as pieces of words, where 1,000 tokens is about 750 words. This paragraph is 35 tokens.

ChatGPT users report \$42 a month pricing for 'pro' access but no official announcement yet



/ OpenAI said earlier this month it would be testing a faster, paid version of the AI chatbot, but it's not clear if the \$42 pricing will stick around.

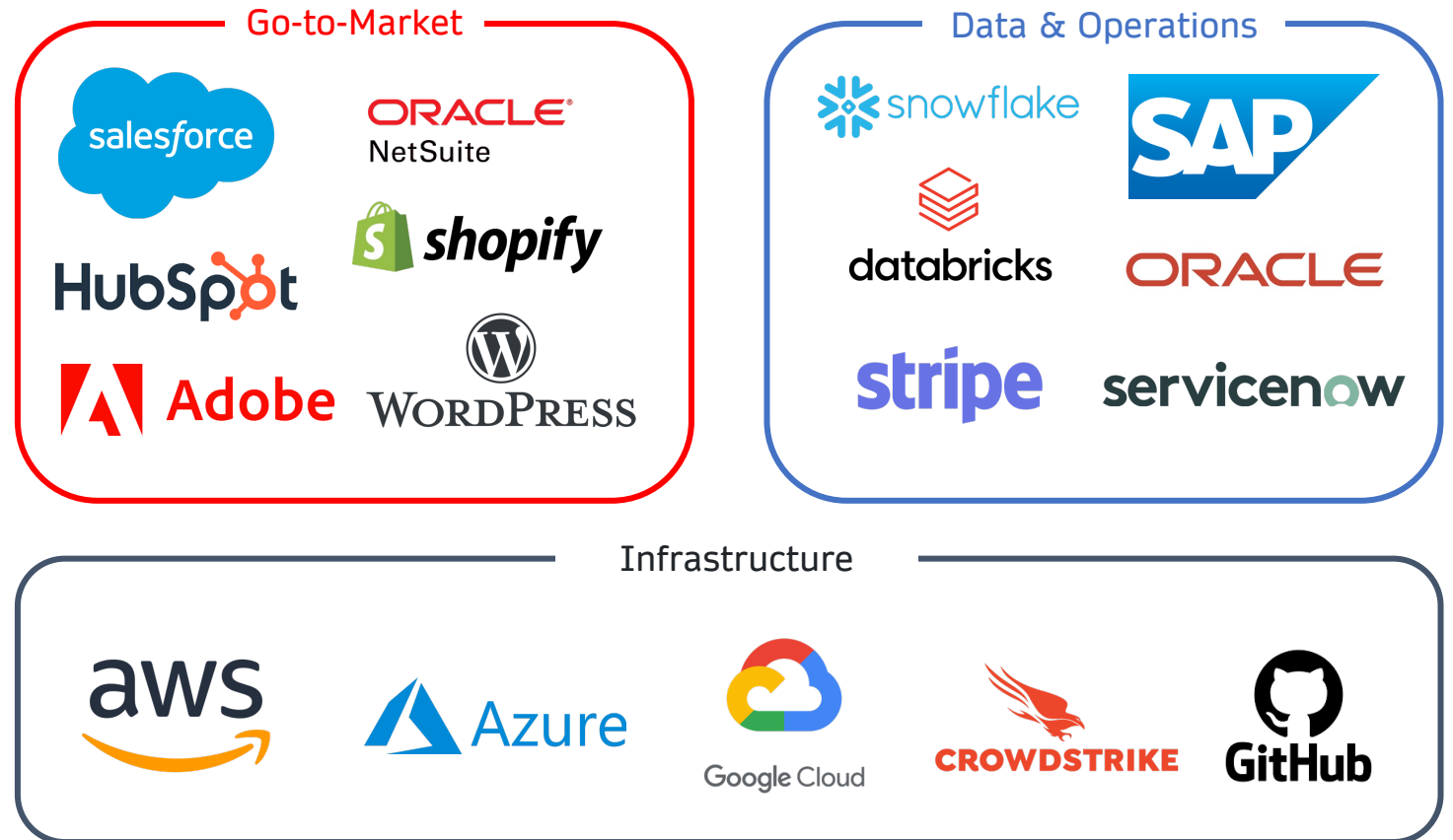
By **JAMES VINCENT**
Jan 23, 2023 at 9:09 AM EST | [1 Comment](#) / [1 New](#)



Commercial signals: Enterprise software ecosystem

Artificial intelligence must be value accretive to the relevant major players within the tech stack.

- In current market environment, IT buyers have a lower willingness to rip and replace proven tech for new feature functionality
- Mission critical software ecosystems have established marketplaces and app stores with in-market customers
- Winners in this environment will **integrate seamlessly** and improve tech stack efficiency

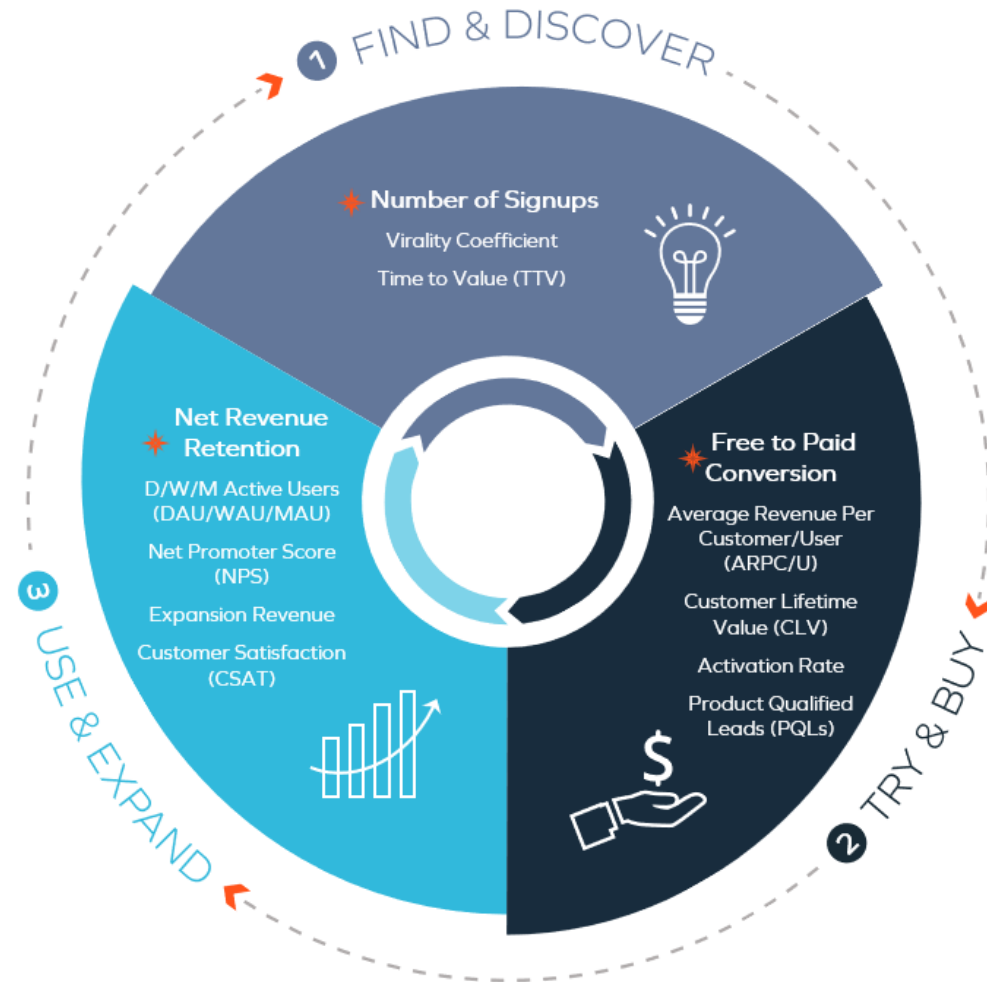


Commercial signals: Low friction, self-serve adoption

Artificial intelligence products adopt PLG (Product Led Growth) to scale efficiently.

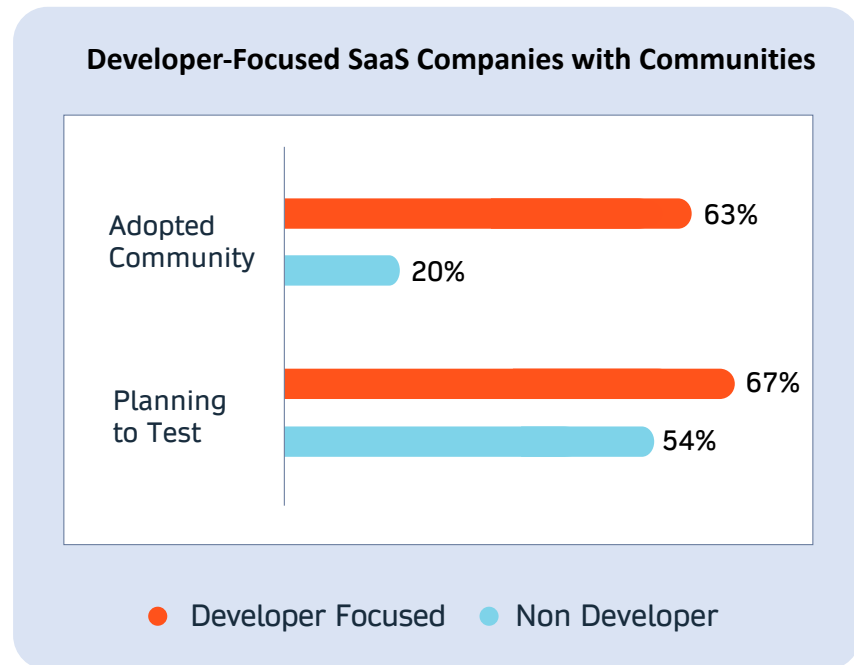
Companies with product-led motions grow efficiently by driving the flywheel of demand

- **Drive users to the website:** Build awareness in the user ecosystem that you solve their problem
- **Onboard users successfully:** Ensure that they start seeing value as quickly as possible
- **Generate a usage habit:** Keep users engaged and active
- **Retain users:** Win back customers who drop off
- **Expand customer use cases:** Help customers to uncover new areas of value and product features



Commercial signals: Community-led GTM

Technical end user communities are the driving force of product credibility and word-of-mouth referrals.



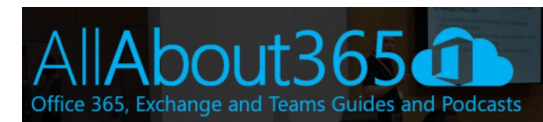
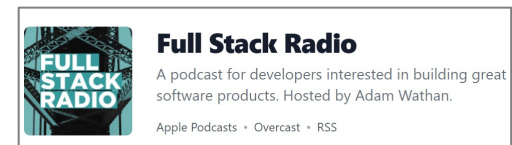
Communities

Destinations for community end users collaborate on projects related software solutions and learn best practices from peers

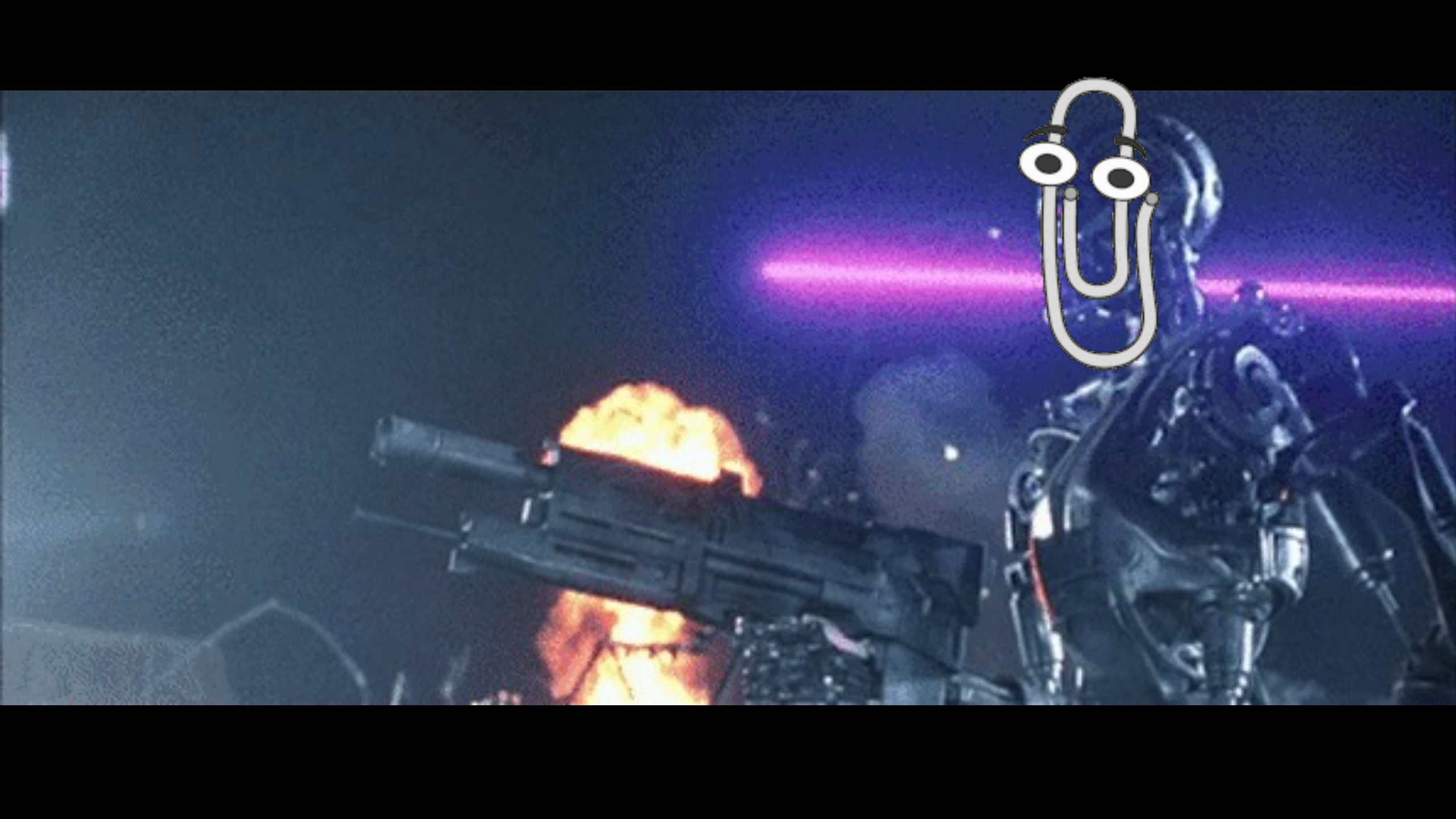


Forums/Blogs/Podcasts

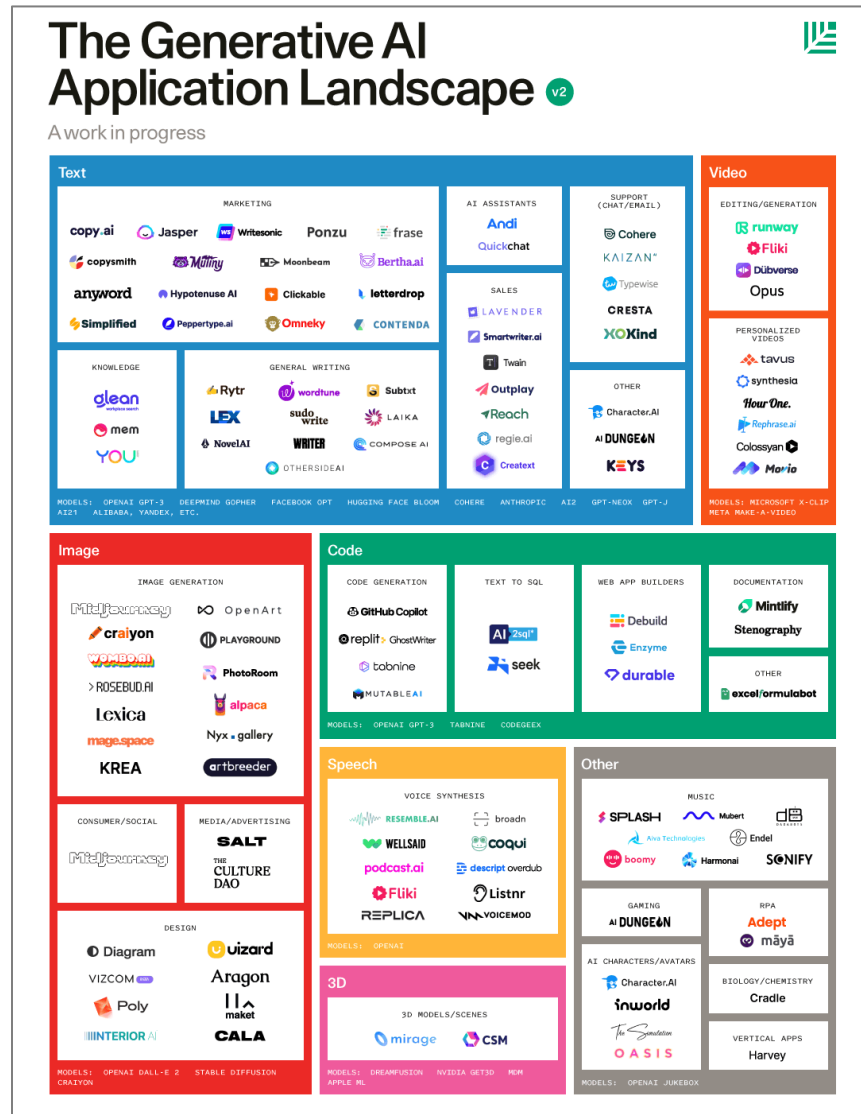
Trusted sources where end user personas go to learn about broader industry trends and from the technical experts in their field










Where the technology is
headed



AI engines will be embedded in our digital life

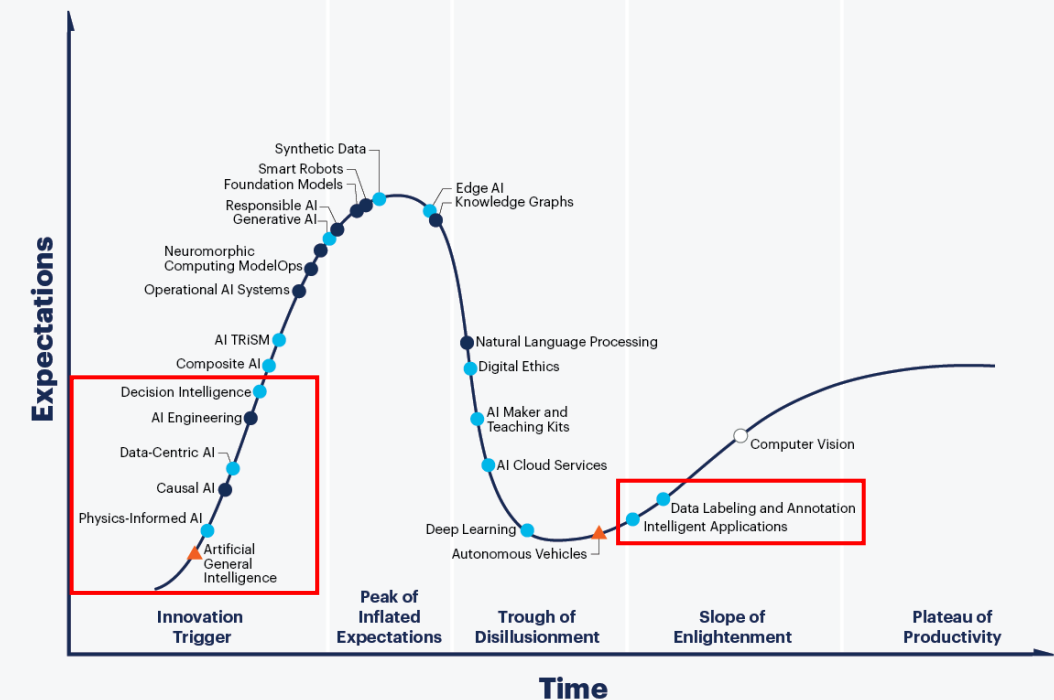


| | Funding to Date | Value Proposition |
|---|-----------------|--|
|  Jasper | \$131M | Content generator optimized for marketing use cases such as social media, ads, blog posts and web copy |
|  | N/A | Independent AI art generation engine that turns text prompts into high resolution images |
|  GitHub Copilot | \$1B | Trained on billions of lines of code, Copilot turns natural language prompts into coding suggestions across dozens of languages |
|  Hour One. | \$25M | Automatically turn text into videos with computer generated presenters. These lifelike virtual characters are based on real humans and can be animated with human expressions just from text |
|  Grain | \$20M | Record, transcribe, and clip moments from research interviews, sales calls, and customer meetings |
|  MEM | \$29M | AI-powered workspace that's personalized to you; Mem makes knowledge storage, sharing, and generation easy and intuitive. |
|  wand | TBD | Build AI engines in minutes modeled on your business data to power your organization with predictive insights to guide your org's outcomes. |

AI innovations fall into four major categories

- **Data-centric (< 2 years)**
 - Data that is AI generated rather than obtained from direct observations of the real world. Data can be generated using different methods, such as statistically rigorous sampling from real data
- **Model-centric (2-5 years)**
 - AI models still need manual training. AI techniques will be applied to remove human curation to train models without larger data sets and improve quality
- **Application-centric (2-5 years)**
 - Rapidly improving decision intelligence to yield better outcomes. These technologies will reduce technical debt, increase data capacity, lengthen model sustainability to produce better predictability
- **Human-centric (5-10 years)**
 - AI that replaces human decisions; it can take into account for good and bad outcomes alike. Responsible AI enables the right outcomes by resolving dilemmas rooted in delivering value versus tolerating risks

Hype Cycle for Artificial Intelligence, 2022



Plateau will be reached:

○ less than 2 years

● 2 to 5 years

● 5 to 10 years

▲ more than 10 years

⊗ obsolete before plateau

As of July 2022

gartner.com

Source: Gartner
© 2022 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner and Hype Cycle are registered trademarks of Gartner, Inc. and its affiliates in the U.S. 1957302

Gartner

Key points to take with you

Conversation points for your next family dinner.

- ChatGPT represents the watershed moment of “consumer” AI tech
- Natural language processing has made a substantial leap forward and will become table stakes
- Proprietary data sets increase in value: AI is only as good as the data on which its being trained
- Within 5 years, AI engines will be woven through the fabric of our digital lives
- Human-like decision making capabilities are 10+ years out

Questions? Feedback?

Please reach out!

Whitney K. Rothe
wrothe@gmail.com
[linkedin.com/in/wkrothe](https://www.linkedin.com/in/wkrothe)

