

AI, and everything else

Benedict Evans

December 2023

www.ben-evans.com

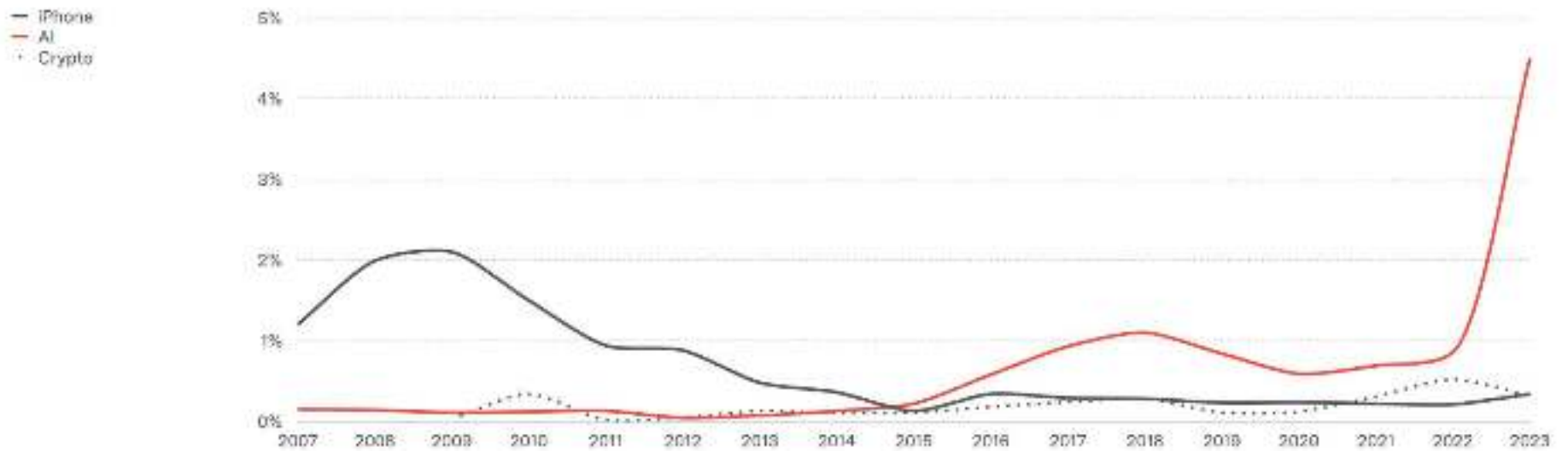
“In my lifetime, I’ve seen two demonstrations of technology that struck me as revolutionary... the GUI and ChatGPT”

Bill Gates, March 2023

Huge excitement within tech

AI has taken over the tech industry hype cycle

Hacker News trends (% front page articles)

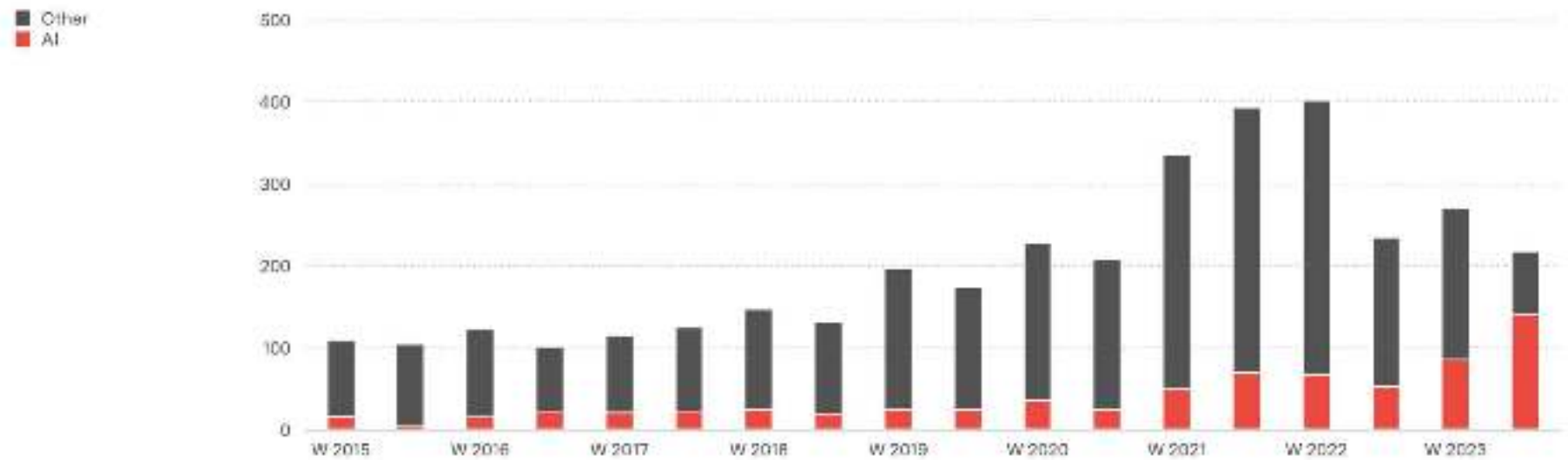


Source: toddwschneider.com

The startup surge

If your startup isn't working on generative AI... what's wrong with you?

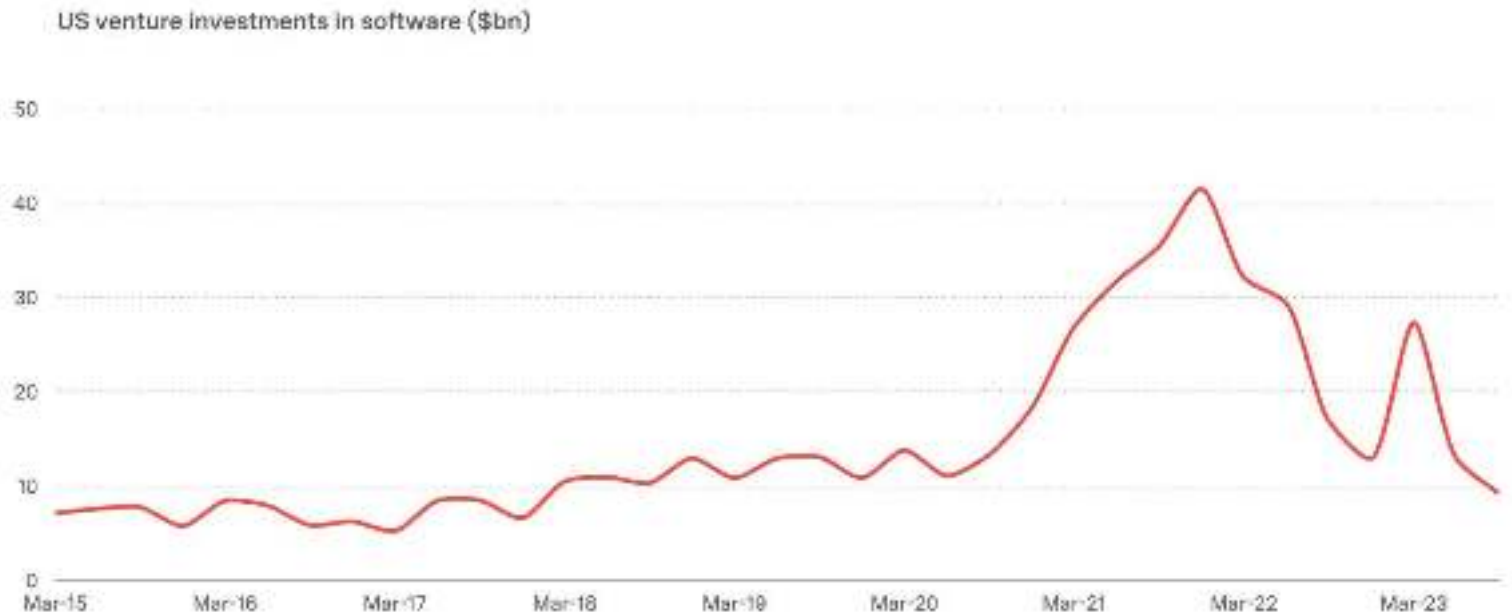
Y Combinator startups by field



Source: Y Combinator

Even as the broader industry slows down

As the 'Covid Rotation' rotates, tech venture slows down - except in AI

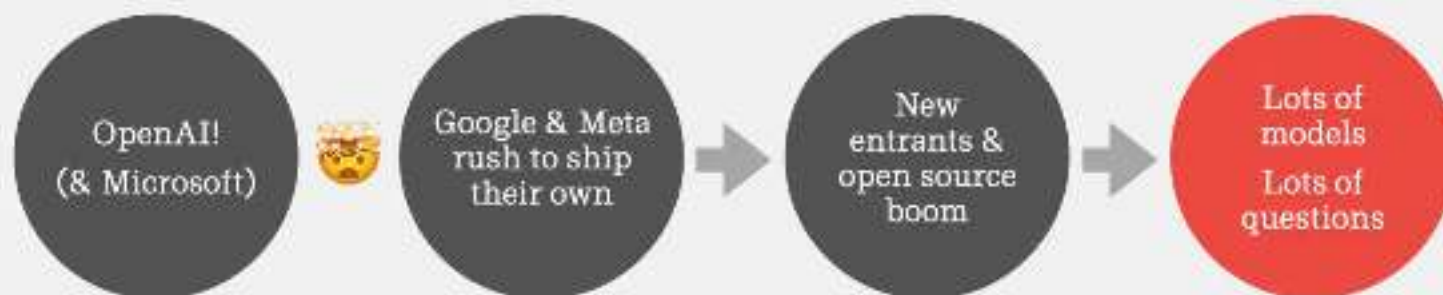


Source: NVCA

Benedict Evans — December 2023 5

The great model boom of 2023

OpenAI showed the way, but there are dozens of projects scrambling to catch up, of all shapes and sizes



Lots of science questions...

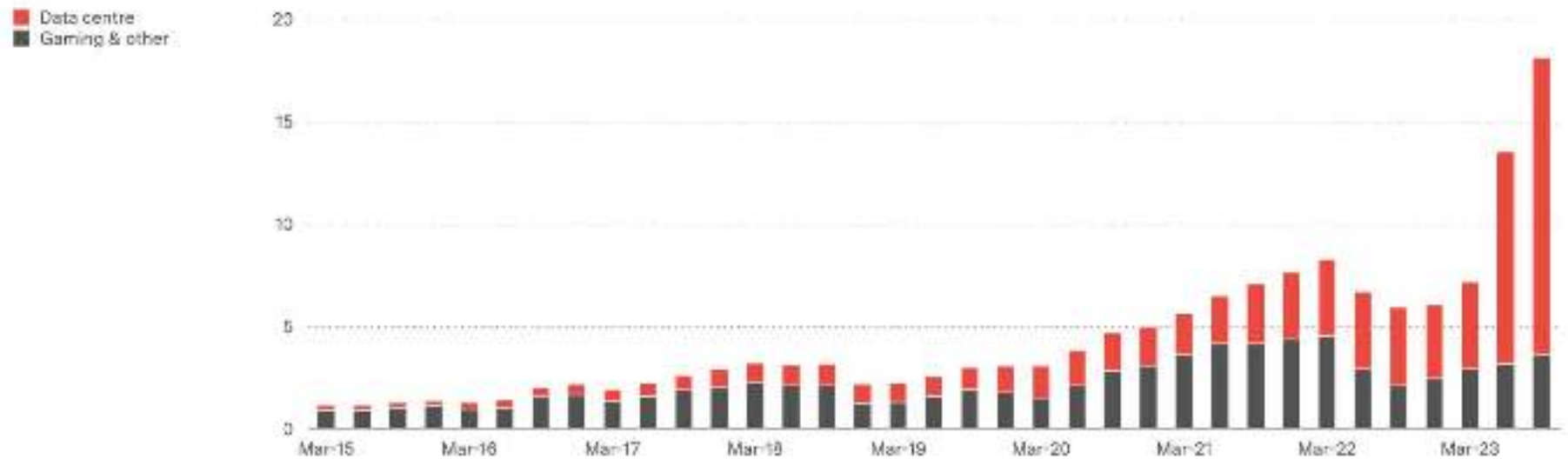
OpenAI broke through by making a giant model, so now, do models get still-bigger, or smaller?



Nvidia can't keep up...

The rush to buy GPU compute capacity (at 40% operating profit)

Nvidia quarterly revenue by segment (\$bn)



Source: Nvidia

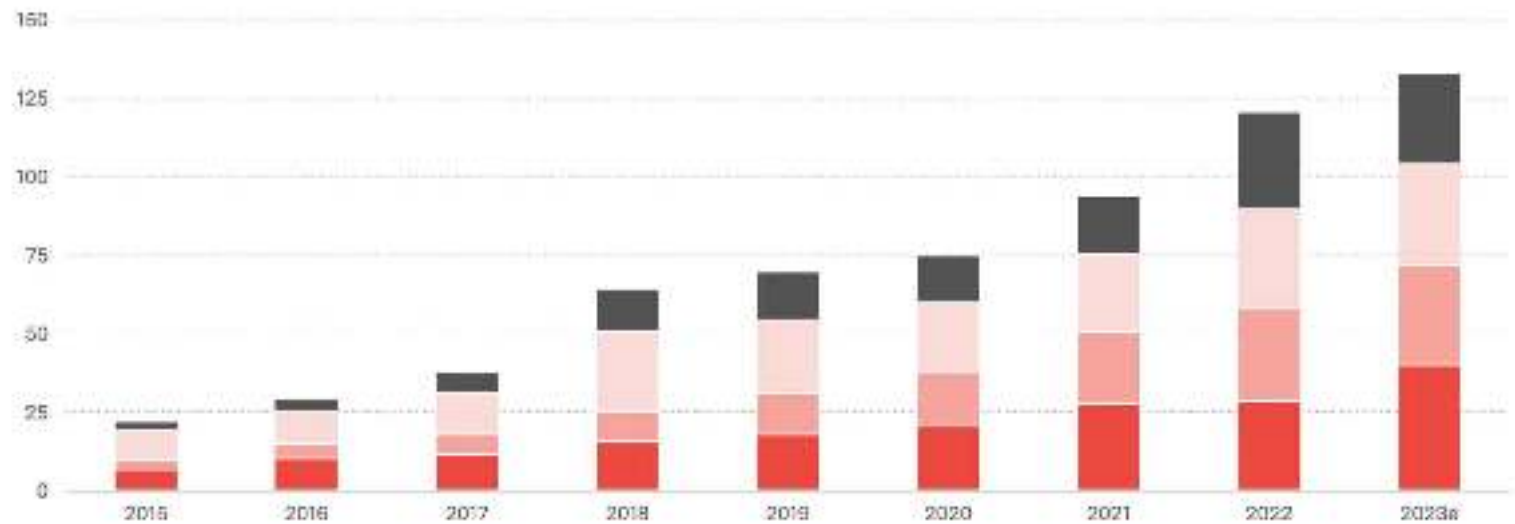
Benedict Evans — December 2023 5

And the hyperscalers... scale

The 'big three' cloud platforms may spend \$100bn on infrastructure this year

Platform capex* (\$bn)

- Meta
- Alphabet
- AWS
- Microsoft

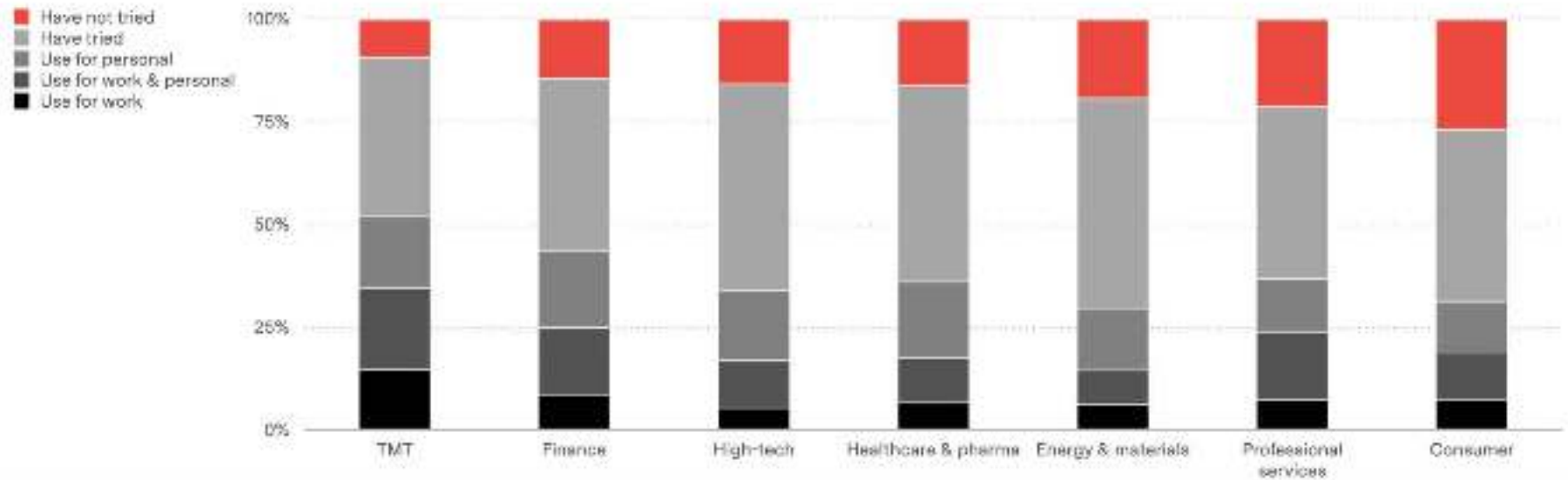


Source: Companies, Goldman Sachs
*including capital leases

“What’s our AI strategy?”

Huge interest in old economy companies after just a few months

Corporate management exposure to generative AI, April 2023



Source: McKinsey

Benedict Evans — December 2023 10

100m ChatGPT weekly active users!

OpenAI has \$1.3bn revenue!!

And a rush to regulate?



Yes, but what *is* this?

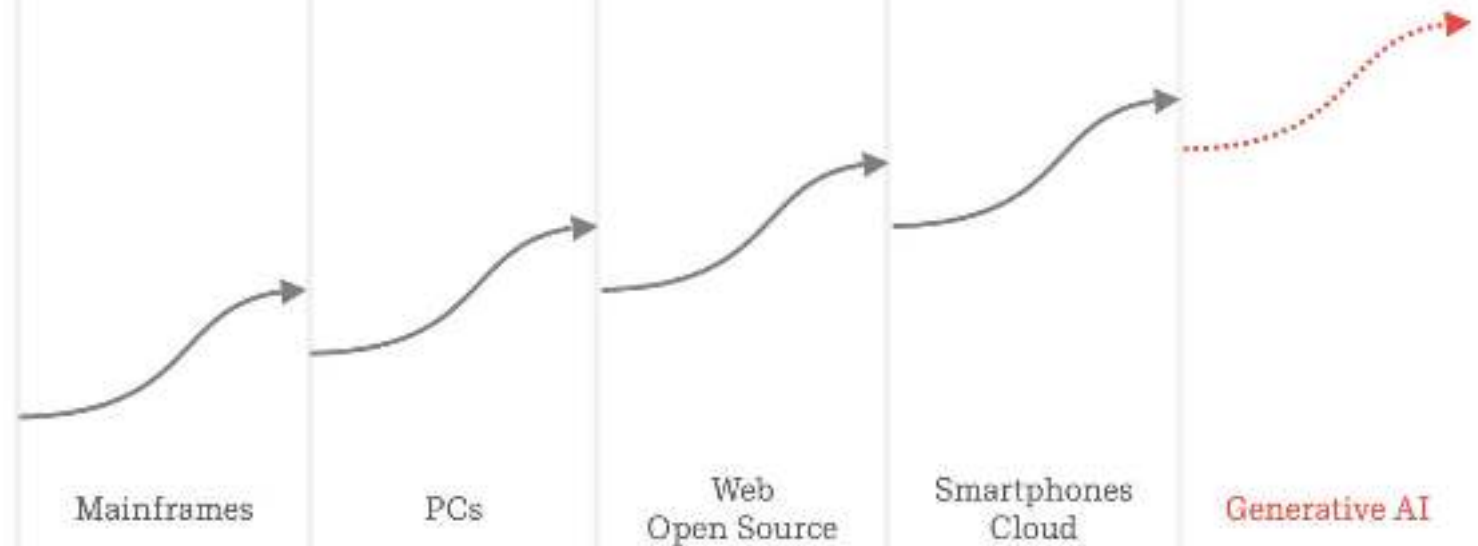
Almost everyone in tech agrees this is The Thing, but it's much less clear what that means



A platform shift?

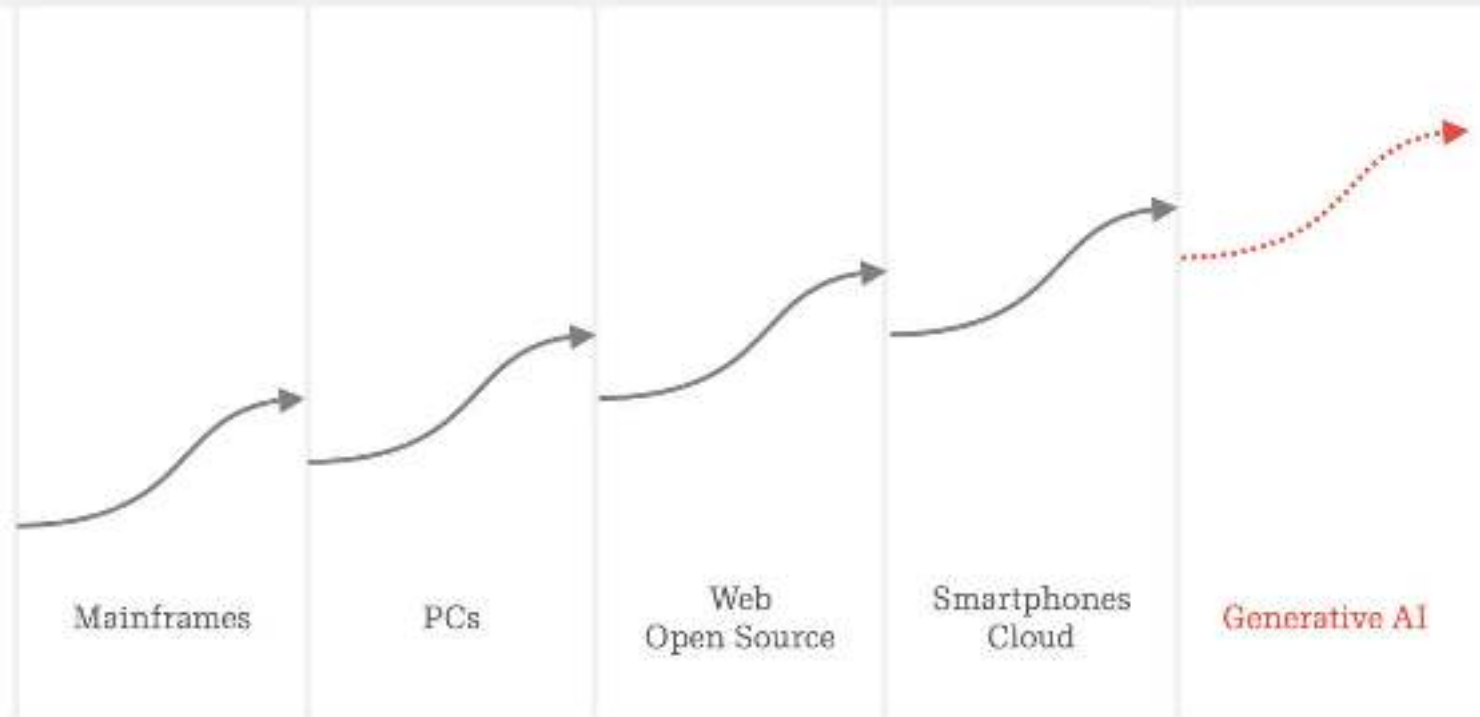
The tech industry moves in platform shifts

Every 10-15 years, a new platform resets the tech industry



The tech industry moves in platform shifts

Every 10-15 years, a new platform resets the tech industry



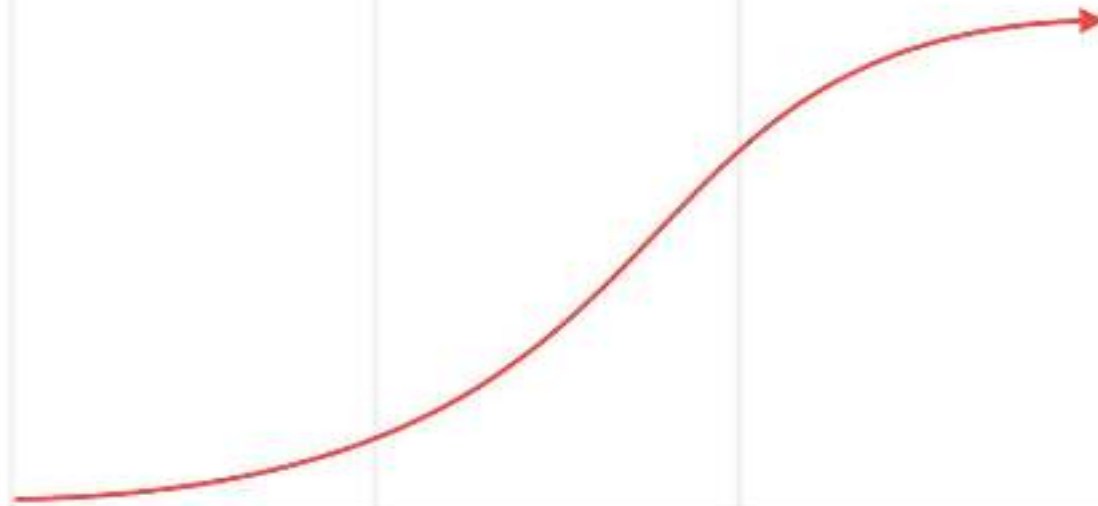
New technologies come in S Curves

New tech generally goes from stupid to exciting to boring

Why is this useful?

Exciting!

Boring



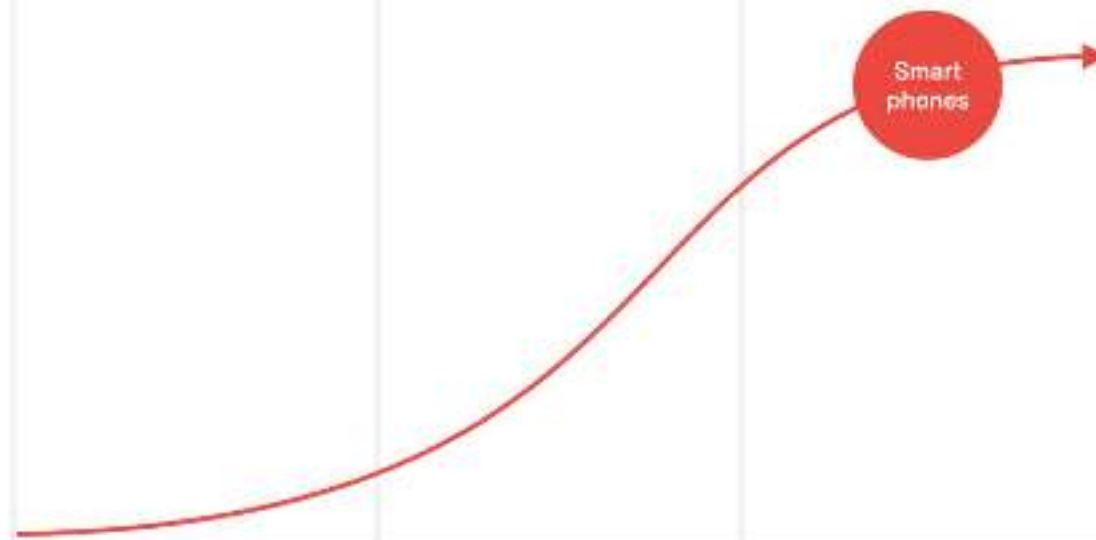
New technologies come in S Curves

Smartphones are mature, which means they're amazing and boring

Why is this useful?

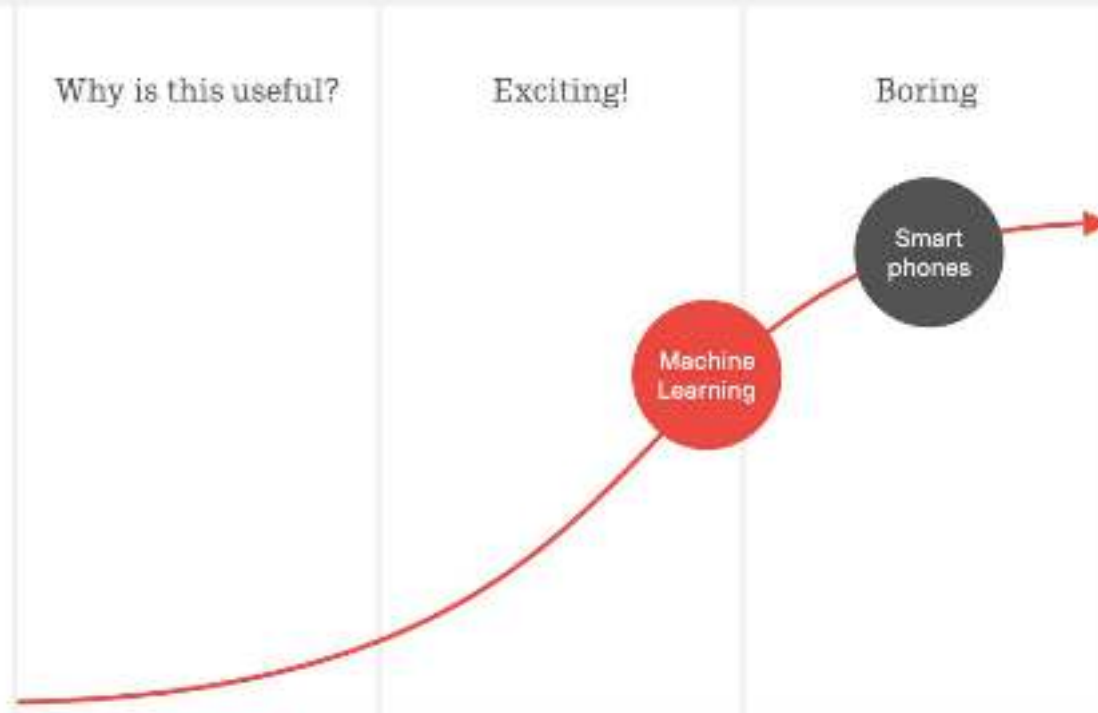
Exciting!

Boring



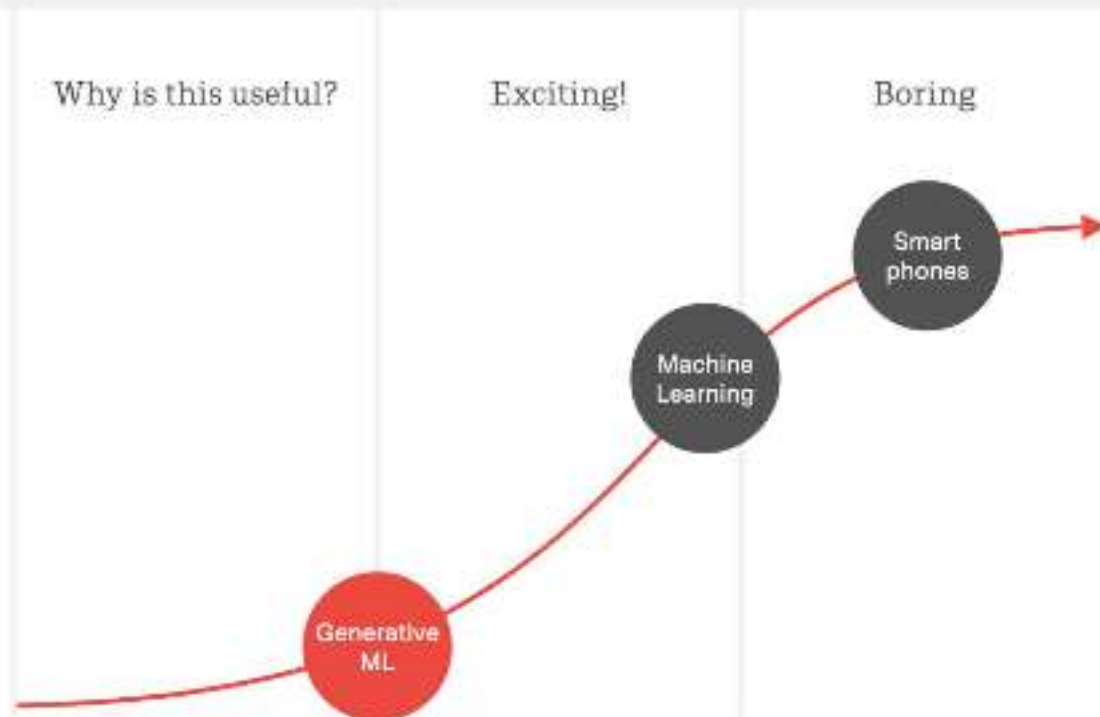
New technologies come in S Curves

And after ten years, machine learning is well on the way to 'boring'



New technologies come in S Curves

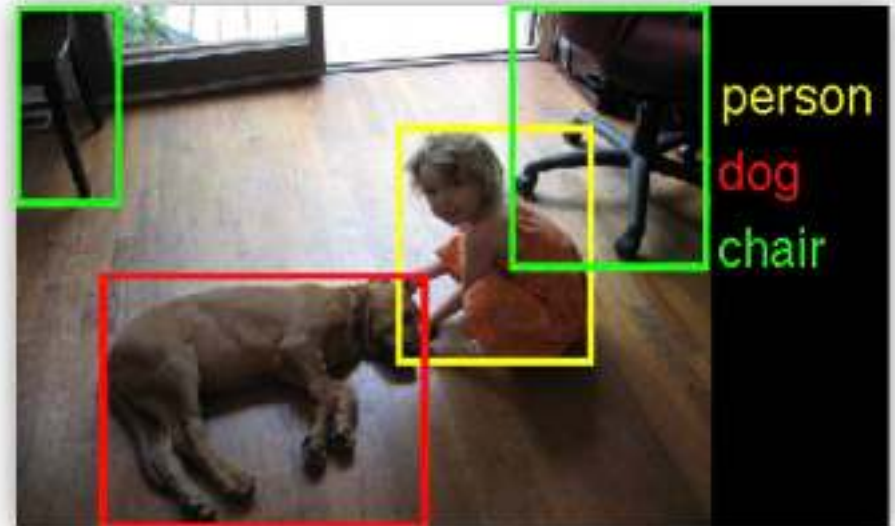
And now we have a new cycle?



2013: why is Machine Learning useful?

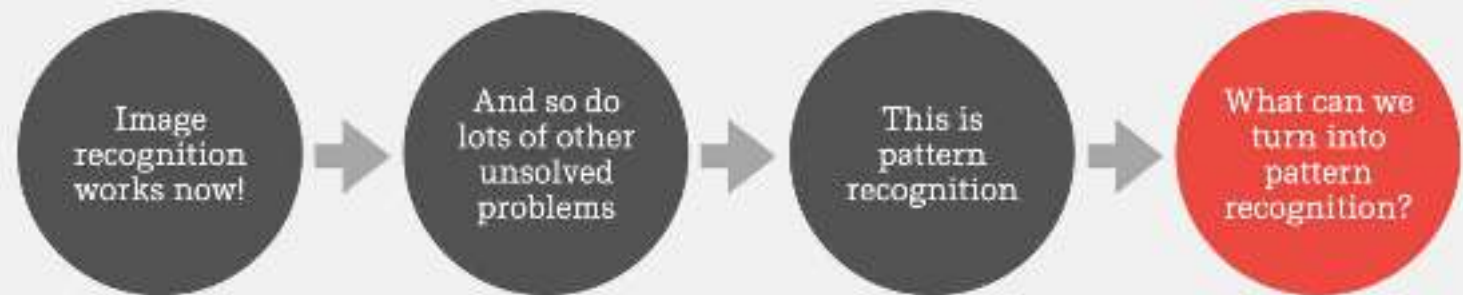
"That's clever, but so what?"

Source: Imagenet



2013: why is Machine Learning useful?

What's the right level of abstraction to understand this?



“Artificial intelligence is whatever hasn’t
been done yet”

Larry Tesler, early 1970s

2023: why is Generative Machine Learning useful?

"That's clever, but so what?"

Source: Benedict Evans / Midjourney



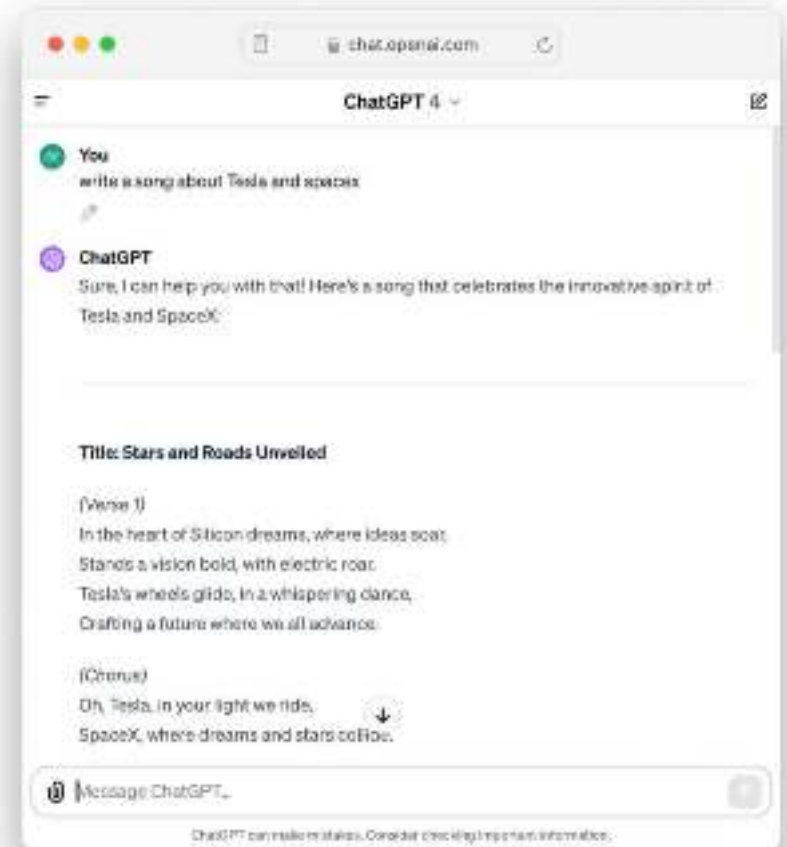
Benedict Evans — December 2023 25

The cool demo

"write a song about Tesla and spacex"

Cool, but why is this useful?

Source: Benedict Evans / OpenAI



Benedict Evans — December 2023 24

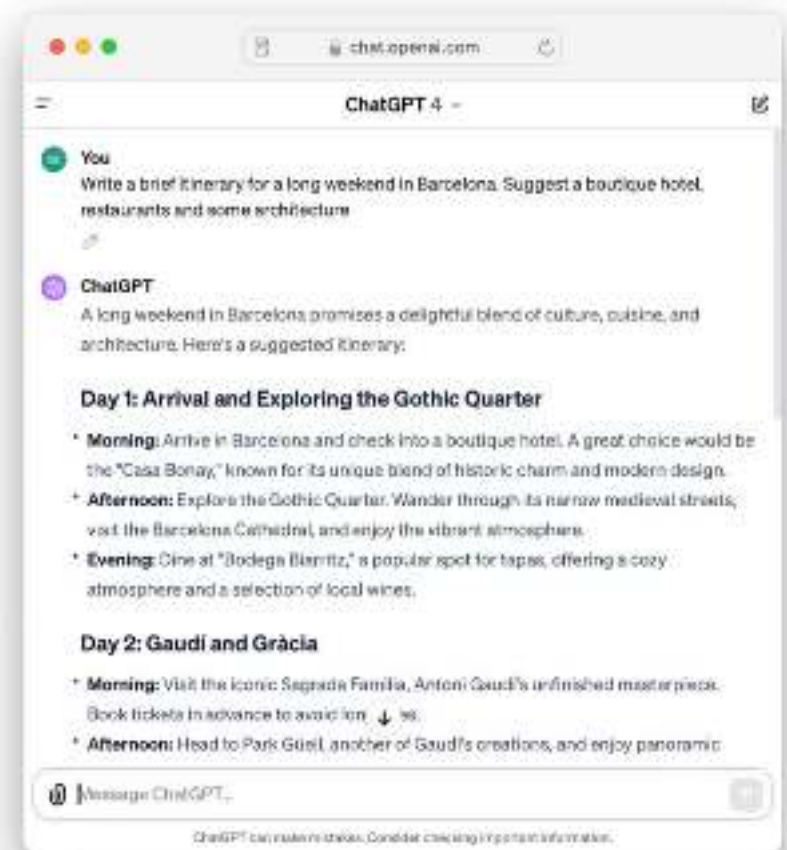
Generative search?

What does this do to search?

The web?

Everything else?

Source: Benedict Evans / OpenAI



Benedict Evans — December 2023 25

However...



The screenshot shows a web browser window displaying a New York Times article. The browser's address bar shows 'nytimes.com'. The article's title is 'The ChatGPT Lawyer Explains Himself'. The sub-headline reads: 'In a cringe-inducing court hearing, a lawyer who relied on A.I. to craft a motion full of made-up case law said he "did not comprehend" that the chat bot could lead him astray.' Below the text are icons for sharing, bookmarking, and commenting (207). A photograph of a man in a suit is visible below the text. At the bottom right, the author is identified as 'Benedict Evans' and the date as 'December 2023'.

nytimes.com

The New York Times

A.I. and Chatbots > How the A.I. Race Began One Year of ChatGPT Key Figures in A.I. How A.I.

The ChatGPT Lawyer Explains Himself

In a cringe-inducing court hearing, a lawyer who relied on A.I. to craft a motion full of made-up case law said he "did not comprehend" that the chat bot could lead him astray.

Share this article

207

Benedict Evans — December 2023 26

This paper does not exist

What is an answer likely to look like? Well...

Economics history papers have the most citations

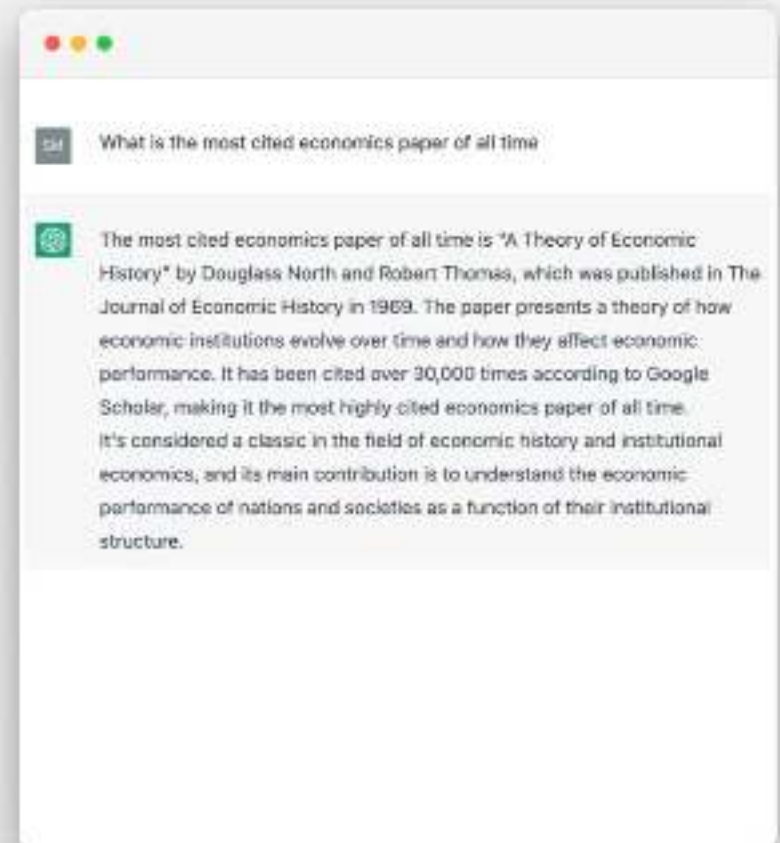
Douglass North is one of the most highly cited authors

Douglass North's most-cited co-authored paper was published in **The Journal of Economic History**

Economics papers most commonly have **two co-authors**...

So what's the pattern?

Source: David Smerdon



Answer this question

~~Answer this question~~

What sort of text in the training data seems likely to be related to questions that look like this?

~~Answer this question~~

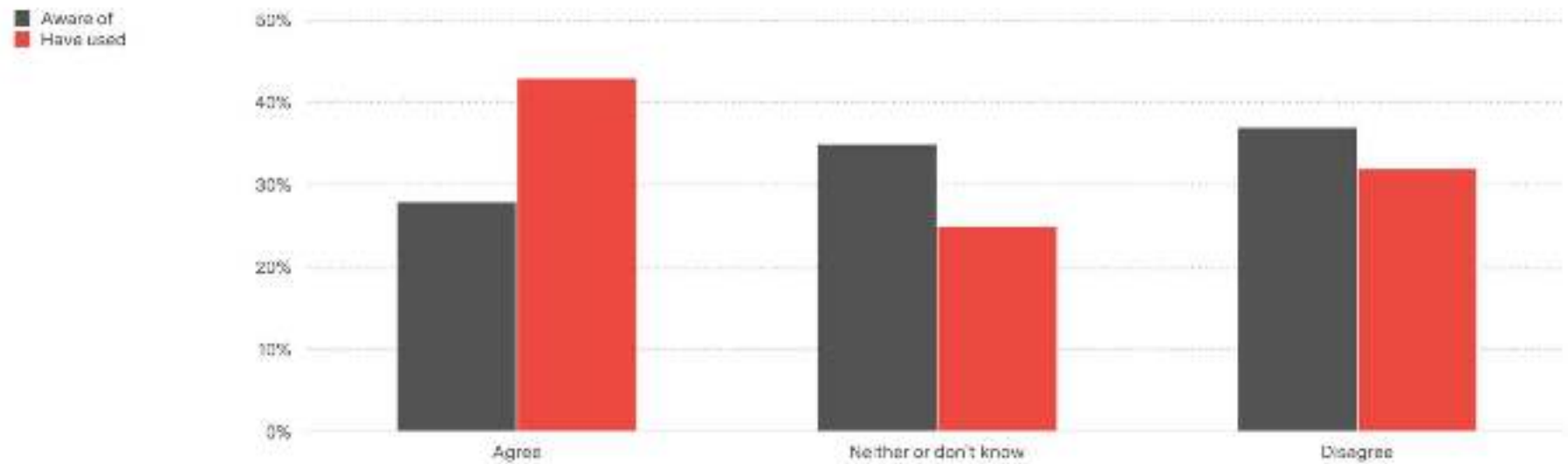
~~What sort of text in the training data
seems likely to be related to questions that
look like this?~~

What would an answer to this *look like*?

A little exposure can be misleading

Generative ML results are very likely to *look* correct, and that can be persuasive if you're not careful

"Generative AI always produces factually accurate answers" (UK, June 2023)



What is this? We're still working it out

How do we understand what these systems are doing?

Not a database!
Not Napster
Not predictable
Does not 'understand'
Bad at things computers
are good at

Probabilistic
Trained on everything (but
half-remembers it)
Looks like it 'understands'
Good at things computers
are bad at

Machine learning, automated learning

Instead of telling a computer how to think, give it the data and get it to work it out for itself

2013: ML 1.0

How do we tell a computer how to recognise a cat?



Give the computer a million pictures of cats

2023: ML 2.0

How do we tell a computer how to reason and understand?



Give the computer the entire written output of all human intelligence (!)

2023: how should we understand generative ML?

What's the right level of abstraction to understand this?



2023: how should we understand generative ML?

What's the right level of abstraction to understand this?



2023: how should we understand generative ML?

What's the right level of abstraction to understand this?



“AI gives you infinite interns”

2023 use cases

Step one: duct-tape the new tool to the problems you already understand. Step two:...

Early wins

Code assistants
Brainstorming
Autosuggest

'Thin wrappers'

Add some UI and
data linkage to
OpenAI APIs

Synthesis

Point an LLM at
your corpus and
get answers

2024?

What new
problems can we
discover?

2023 use cases

Step one: duct-tape the new tool to the problems you already understand. Step two:...

Early wins
Code assistants
Brainstorming
Autosuggest

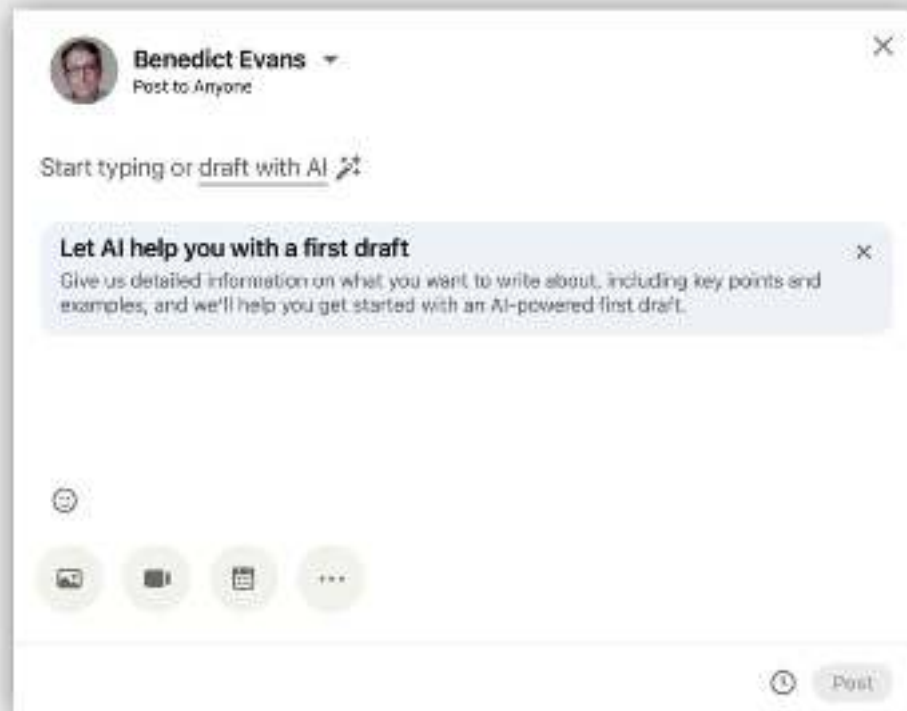
'Thin wrappers'
Add some UI and
data linkage to
OpenAI APIs

Synthesis
Point an LLM at
your corpus and
get answers

2024?
What new
problems can we
discover?

“Every text box on the internet will get an LLM”

Solving LinkedIn's shortage of generic self-promotion



Source: LinkedIn

Benedict Evans — December 2023 37

Infinite interns (to make your product shots)

"Every single team at Amazon is exploring generative AI" - Amazon CEO



Source: Amazon

How is value captured in a platform shift?

If generative AI is the new platform - this is the pattern for how new platforms work



“There are two ways to make money.
You can bundle, or you can unbundle”

Jim Barksdale

And Adobe is trying to do both

Adding Generative AI features to Photoshop while launching Firefly (and buying Figma)

Make it a feature
Generative Fill etc
in Photoshop



Unbundle
New dedicated app
in 'Firefly'

But then, new tools change the nature of the work

First we make the new tool fit the work - then we change the work to fit the tool



Source: The Apartment, 1960

Benedict Evans — December 2023 45

Platform shifts tend to change everything else

New kinds of building blocks can go a lot further than people expect

"SQL makes arbitrary data queries easier"



Just-in-time supply chains...

"We put multitouch and macOS on a phone"



Uber, Instagram, Tiktok... and 5bn people online

"We made a thing that can predict your next word"



???

And create new gatekeepers

New channels and new platforms mean new points of control and leverage

Source: Apple



Benedict Evans — December 2023 44

The end of links?

From discovery to aggregation to synthesis - what's the SEO for ChatGPT?



But beware of extrapolating

Generative AI comes to us first as chat bots, but don't presume that's the right or only model

Source: RIM



Benedict Evans — December 2023 48

Posit: all AI questions have one of two answers

Most AI questions match to a pattern - unless you want to know if the pattern will be broken



More than a platform shift?

“In my lifetime, I’ve seen two demonstrations of technology that struck me as revolutionary... the GUI and ChatGPT”

Bill Gates, March 2023

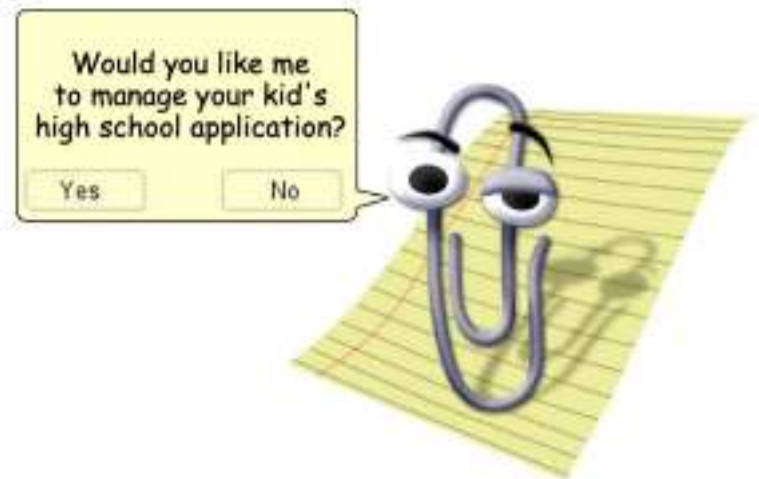
A step change in generalisation?

Is generative AI a change in how broad software can be?



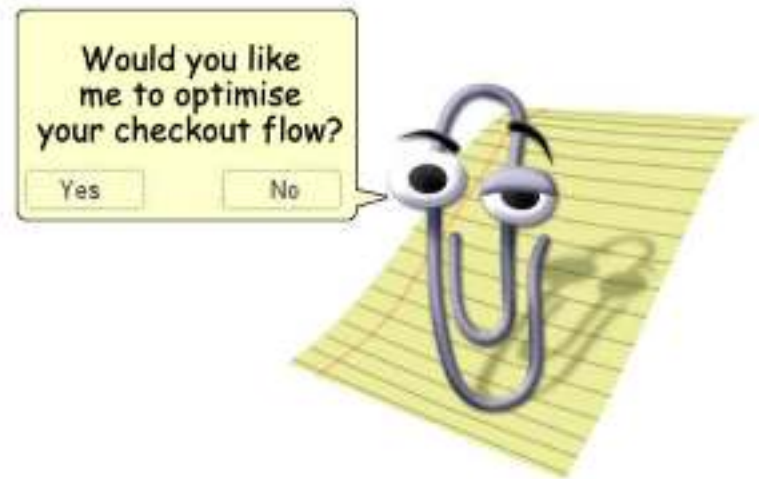
Return of the 'agent' (or the mainframe)?

The 30-year dream of the general purpose 'AI Agent' that can do anything for you



Return of the 'agent' (or the mainframe)?

The 30-year dream of the general purpose 'AI Agent' that can do anything for you

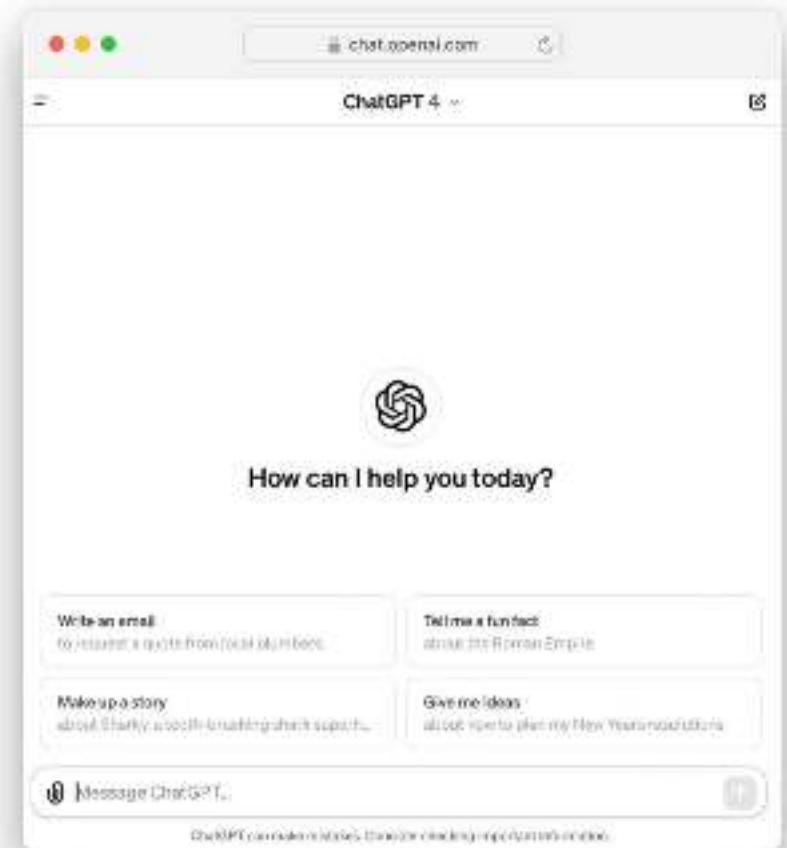


“There are two ways to make money.
You can bundle, or you can unbundle”

Jim Barksdale

Is this a product?

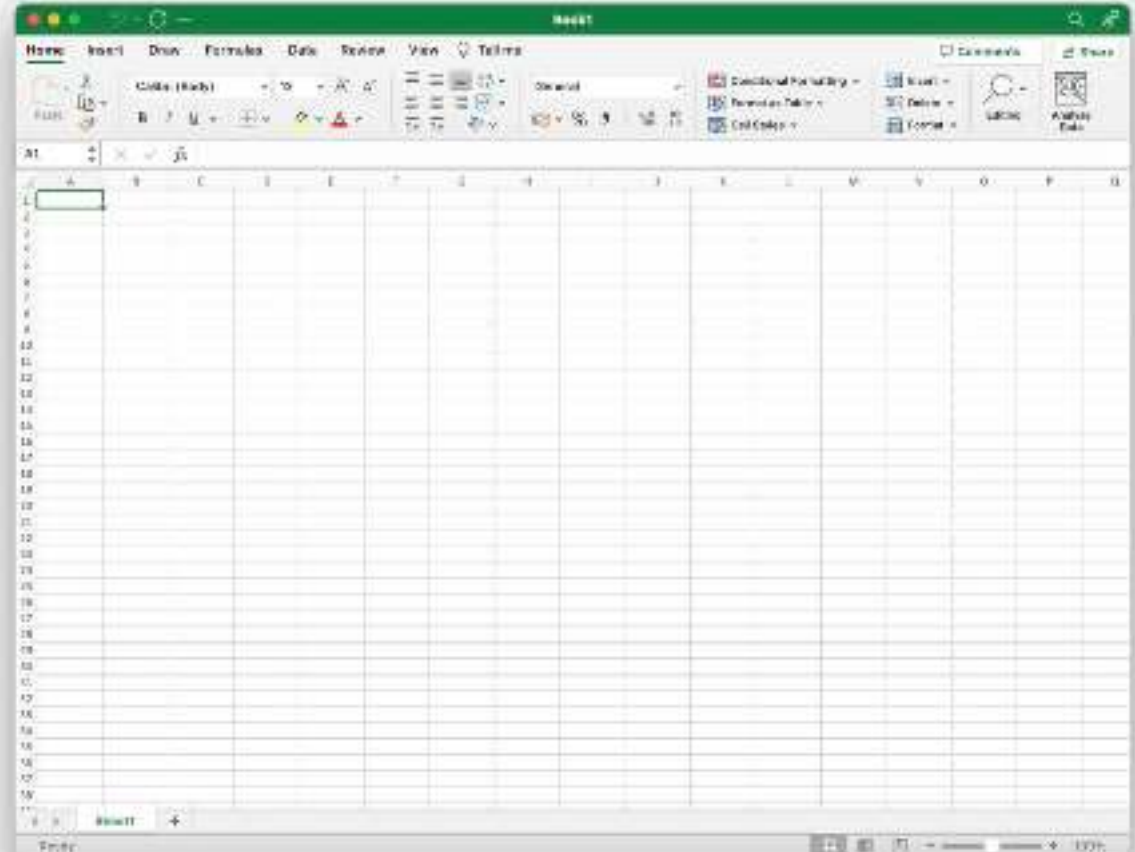
What does it mean when a 'universal interface' has to tell people what they could do?



The blank canvas problem

If you can ask 'anything'...

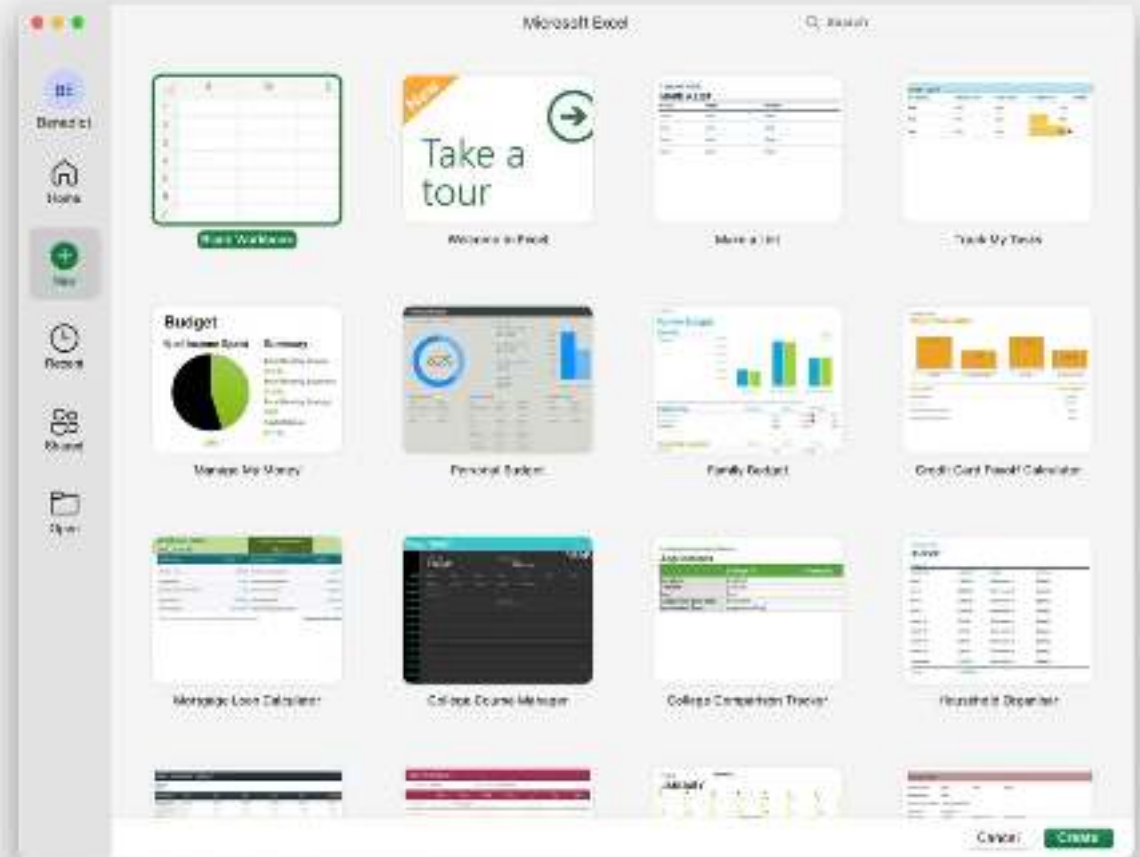
What do you ask?



Unbundling

How do you turn a blank canvas into solutions?

But then, how do those solutions get unbundled into new companies?



Every Unix function became a company

Every Excel template became a company

Every Notion template...

Every ChatGPT suggestion...

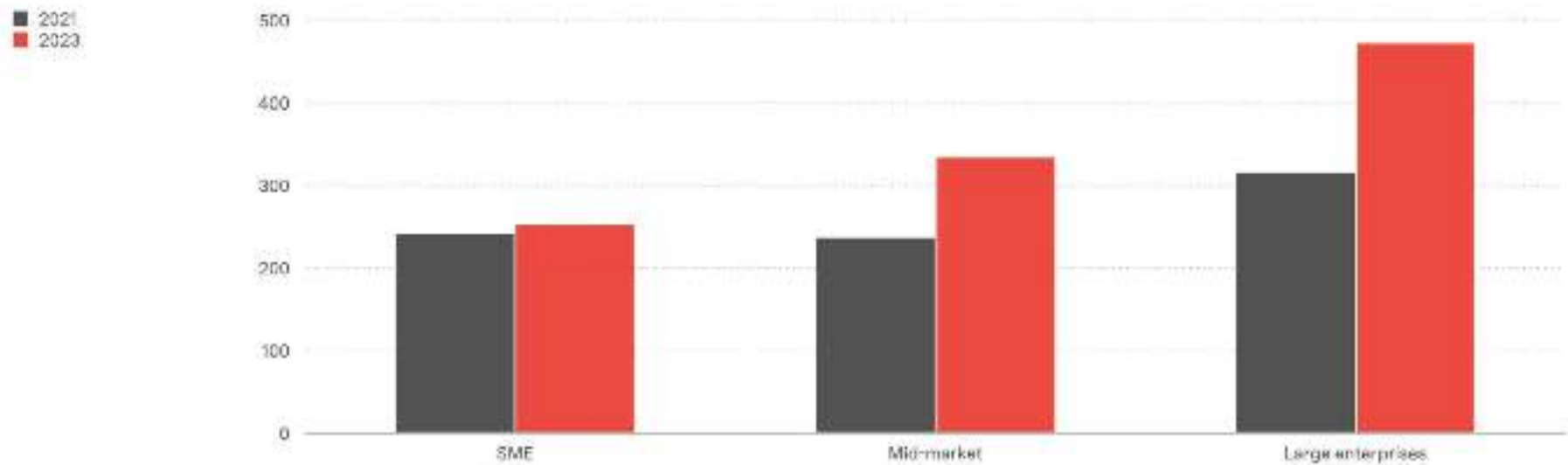
“Half my jobs are telling people who use
Excel to use a database...

and the other half are telling people who
use a database to use Excel”

Bundling or unbundling?

Do LLMs mean all these apps are swallowed by a few big models - or do they enable an explosion of new apps?

Average number of SaaS apps used per company



A product paradox, perhaps?

Does 'general-purpose' AI mean general-purpose tools, or must it be wrapped in single-purpose product?



Unless, of course...



Source: Benedict Evans / Midjourney

Benedict Evans -- December 2023 61

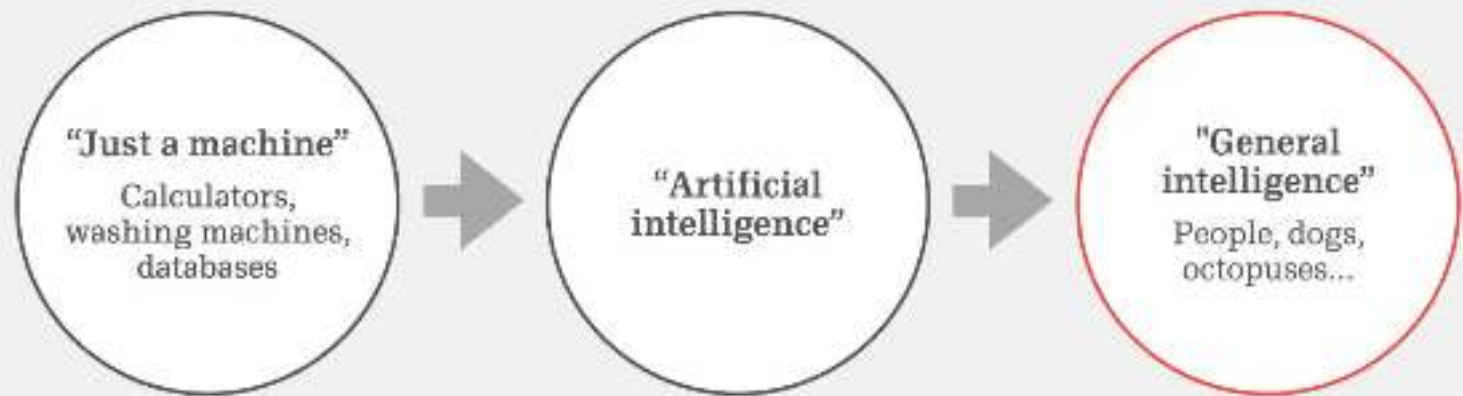
A path to AGI?

“Artificial intelligence is whatever hasn’t
been done yet”

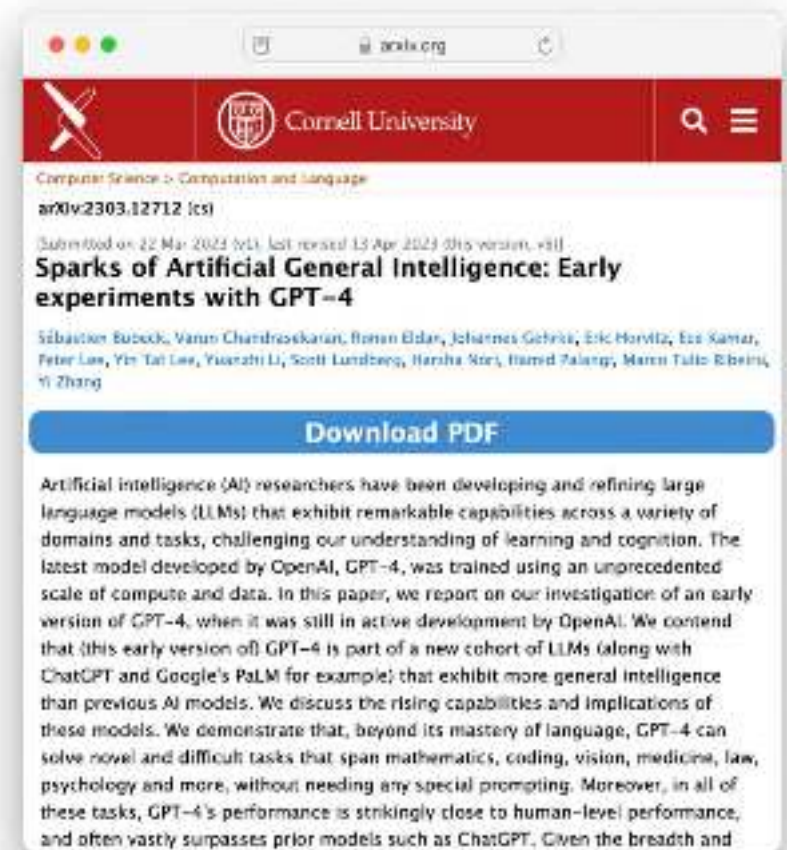
Larry Tesler, early 1970s

What are we trying to build when we say 'AGI'?

"Artificial General Intelligence" is a thought experiment as much as a technology

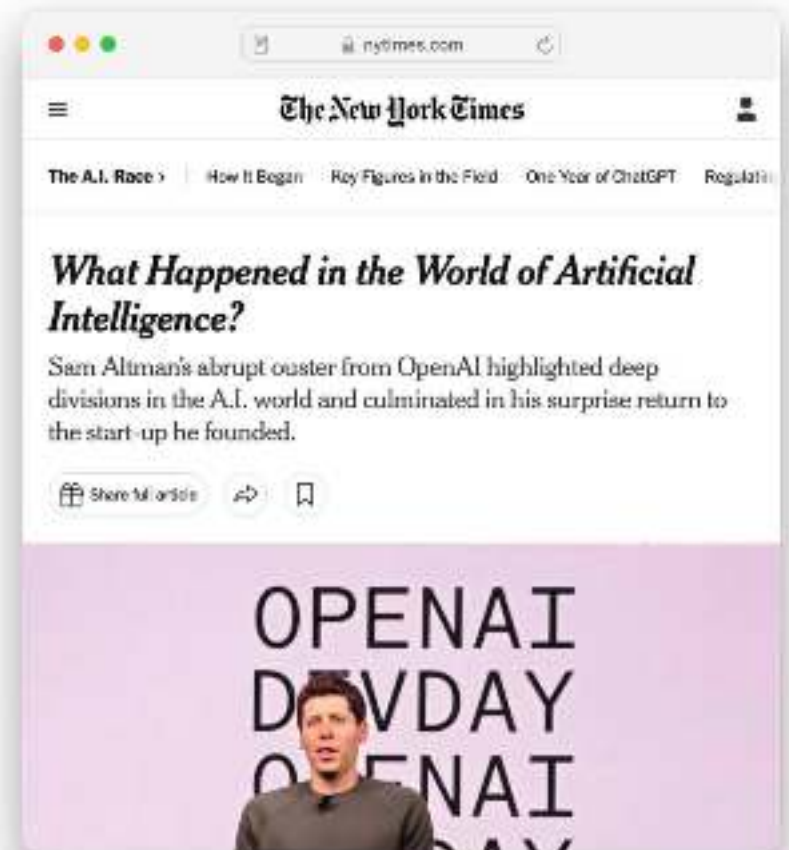


And generative AI *might* be on a path to that...



Hence the OpenAI mess

"Slow down and worry" versus "speed up and worry"



The problem: again,
Generative ML seems
brilliant...

Source: Benedict Evans / Midjourney



Benedict Evans — December 2023 67

But can it “understand”?

Where’s the door?

And why are there two steering wheels?

Source: Benedict Evans / Midjourney



Benedict Evans — December 2023 68

Three views of understanding in LLMs

Can LLMs get to the kind of 'structural understanding' that we have? Maybe. Opinions vary

This will
emerge with
scale?

With enough
scale, it won't
matter?

More,
unknown
breakthroughs
are needed?



(Maybe people don't have
'general intelligence'
either?)



What are we measuring?

Is this intelligence? Reasoning?

Or information retrieval and pattern matching?

Or - the ability to *look like* people?

Should we be benchmarking against things that people are bad at? Or against things that people are good at?

Source: OpenAI

Exam results (ordered by GPT-3.5 performance)

Fabricated percentile score bins (score) and

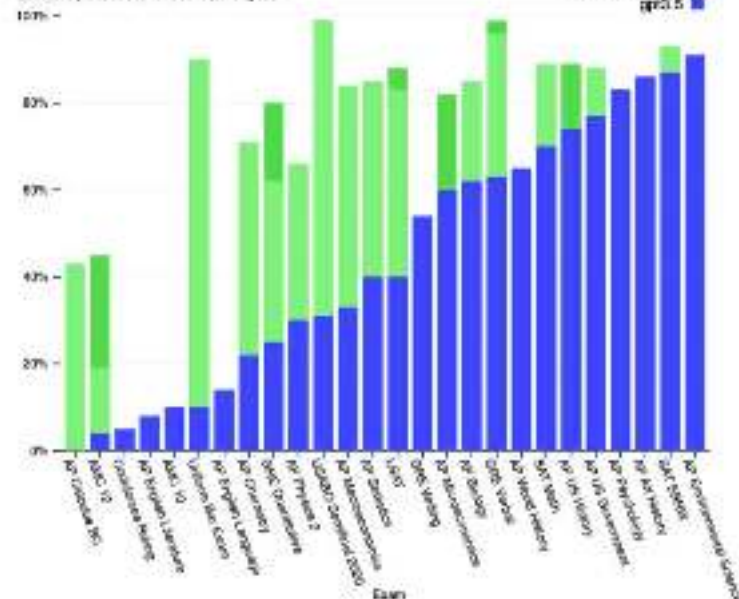


Figure 4. GPT performance on academic and professional exams. In each case, we simulate the conditions and scoring of the real exam. Exams are ordered from low to high based on GPT-3.5 performance. GPT-4 outperforms GPT-3.5 on most exams tested. To be conservative we report the lower end of the range of percentiles, but this creates some artifacts on the AP exams which have very wide scoring bins. For example although GPT-4 attains the highest possible score on AP Biology (55), this is only shown in the plot as 85th percentile because 15 percent of test takers achieve that score.

We don't know

The engineers on the Apollo Program knew the distance to the moon and the rocket's thrust...

But we don't know what AGI is, nor how far away it is, nor how to get there, nor how close large language models are, nor how much better they can get

We don't have a chart for AGI

So how do you, personally, prefer to think about risk?

Source: NASA



And everything else

From vision to deployment

What's already big, what's being built, and what comes next?



(Meta is still metaversing)

Meta still believes in VR - it invested \$16bn in the last 12 months alone

Meta 'Reality Labs' quarterly financials (\$bn)



Source: Meta

And all the externalities

Adding 'AI' to the spread of complex and conflicting regulatory questions...

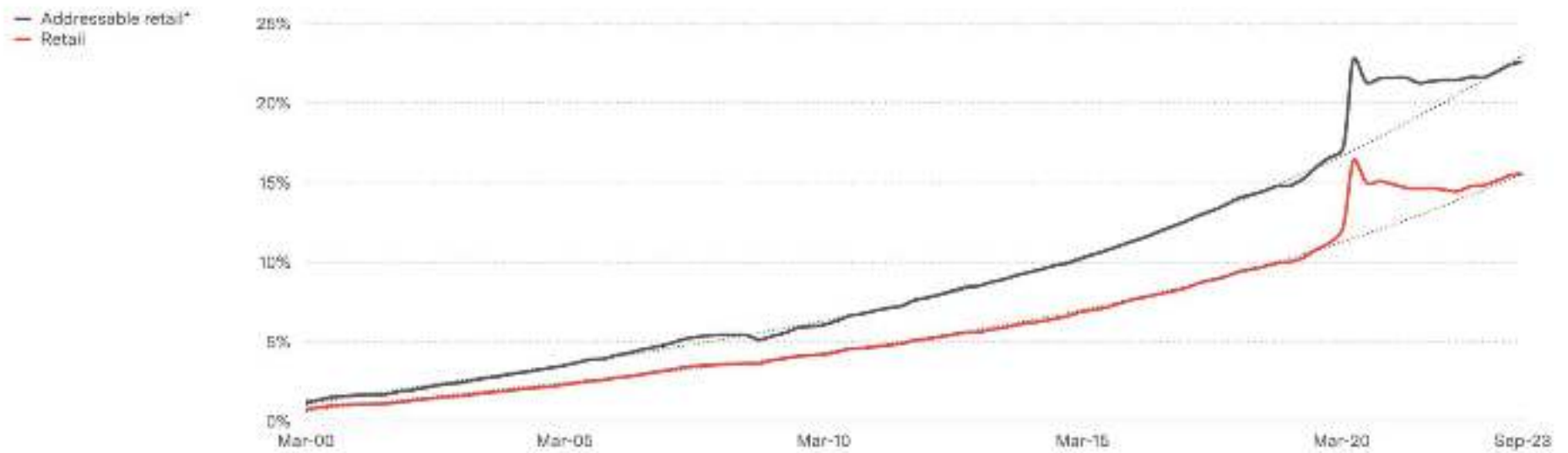


Privacy	Competition	Harmful content
Encryption	AI Bias	Intellectual property, liability, patents...
Geopolitics	Semi-conductors	And more...

Meanwhile, “back to the trend line”

But a trend line at 15-20% and growing!

Ecommerce as % US retail

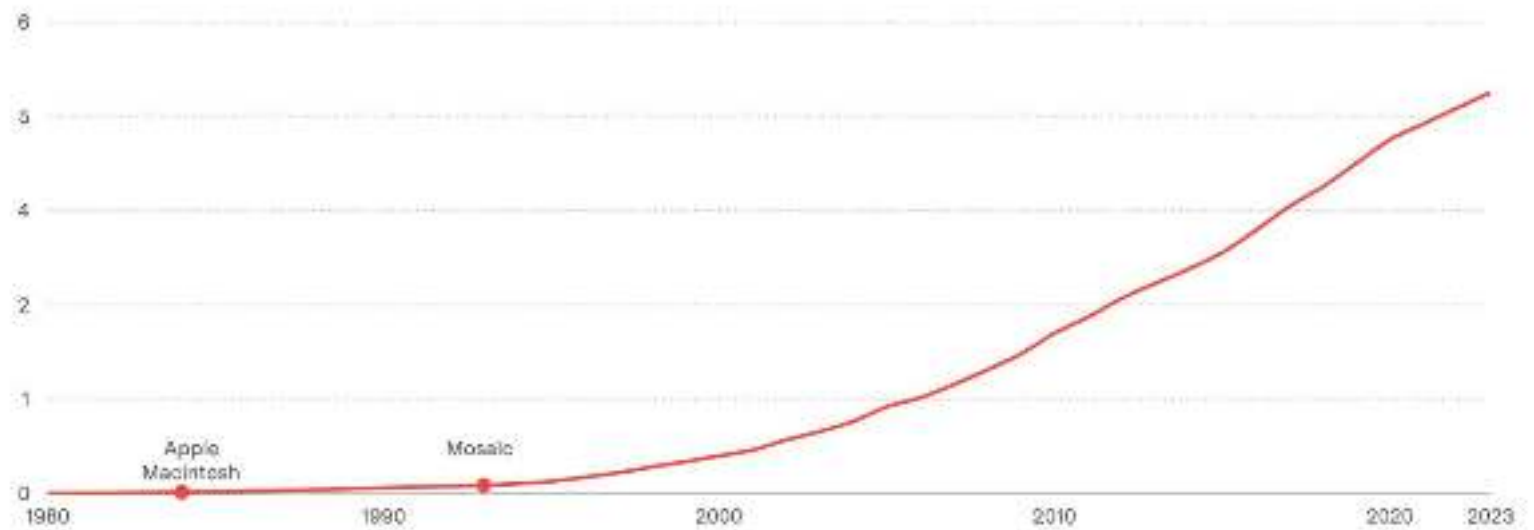


Source: US Census. Seasonally adjusted
* Excluding cars, car parts & fuel

From exciting to universal

There were less than 100m PCs on Earth in 1993 - now 5bn people are online

Global computer users (bn)



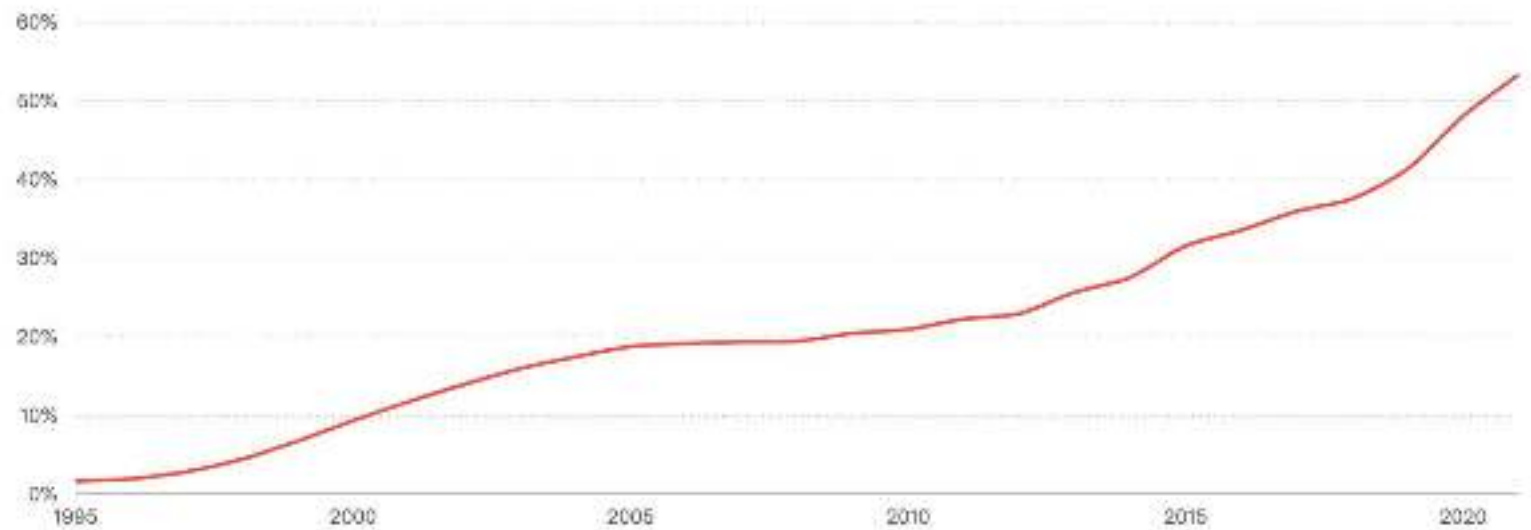
Source: Morgan Stanley, Gartner, ITU, Apple, Google, CNNIC, Benedict Evans

Benedict Evans — December 2023 78

The new normal

Over half of new relationships in the USA begin online

US heterosexual couples who met online, by year of meeting

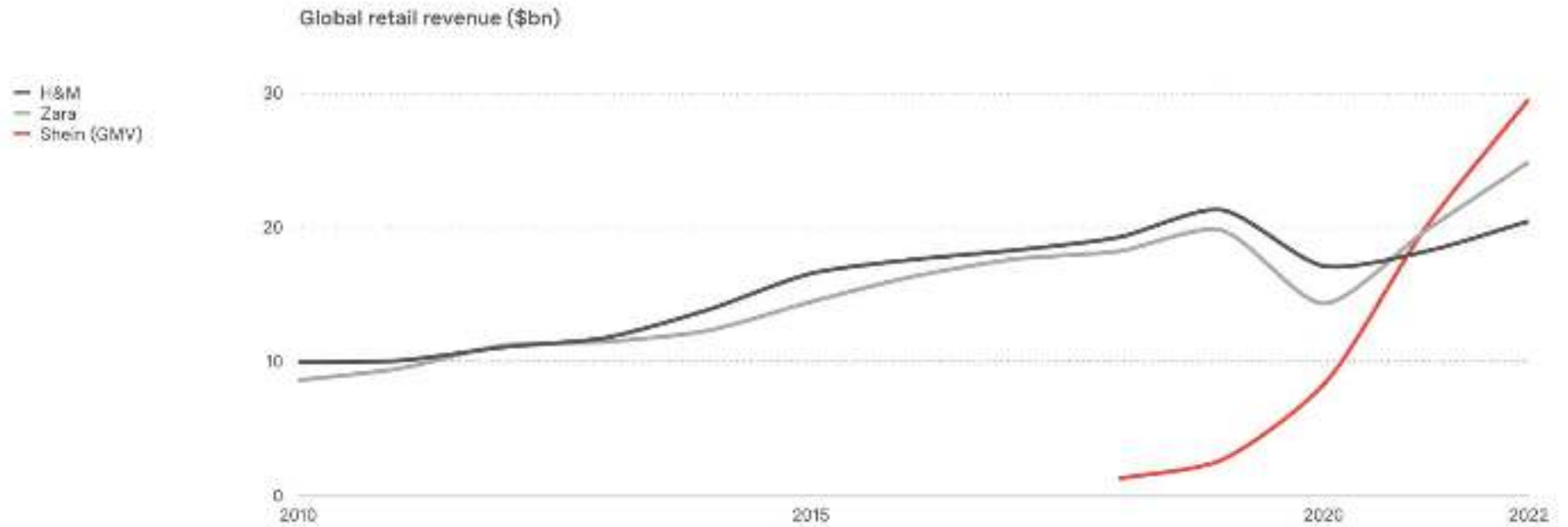


Source: Stanford/GfK

Benedict Evans — December 2023 72

New channels

A new bundle and a new channel build a new global fast fashion business



Source: WSJ, Companies

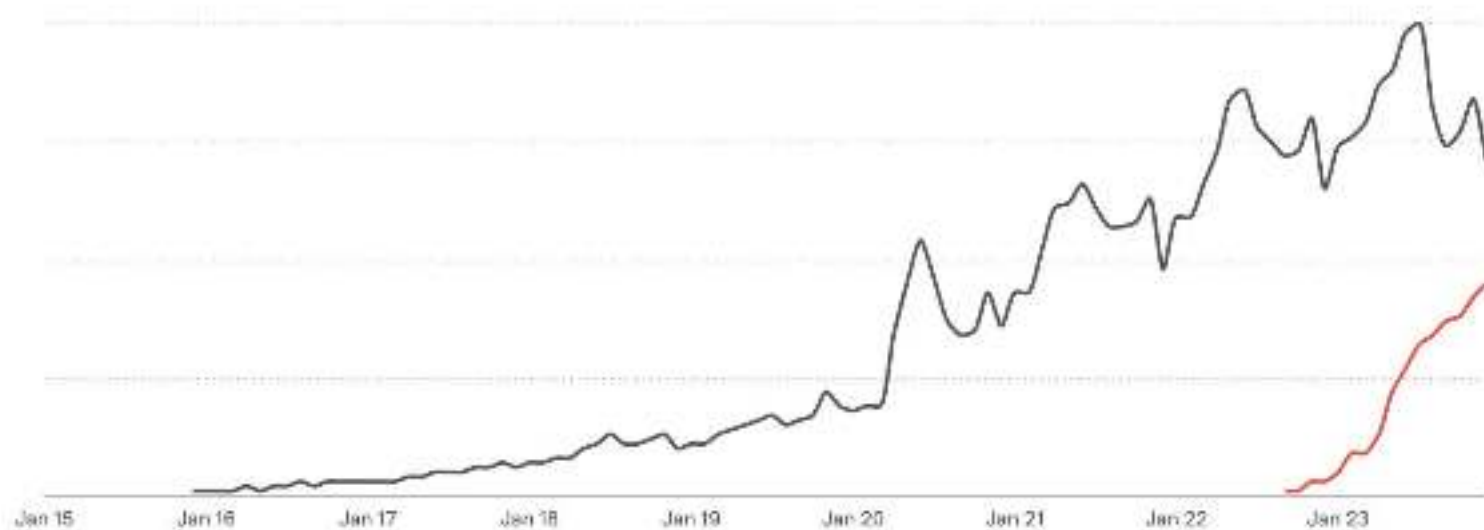
Benedict Evans — December 2023 . 80

And whatever comes next

New models will keep coming, especially from China (and buying customers on Meta)

Global search interest (relative)

— Shein
— Temu



Source: Google Trends

Benedict Evans — December 2023 81

New questions

Ikea builds a flagship in Central London...

(Taking the old 'Top Shop' fashion retailer's space)

Is this logistics, retailing, marketing or advertising?

Source: Benedict Evans



But what kinds of questions?

As software eats the world, most of the questions are not software questions

Will TV be re-aggregated? Does streaming mean a structurally lower profit pool?



Ask a TV analyst

Shein has disrupted fast fashion and may sell \$40bn this year. Is that sustainable?



Ask an apparel analyst

Will EVs let the Chinese car industry do in the 2020s what the Japanese did in the 1980s?



Ask a car analyst!

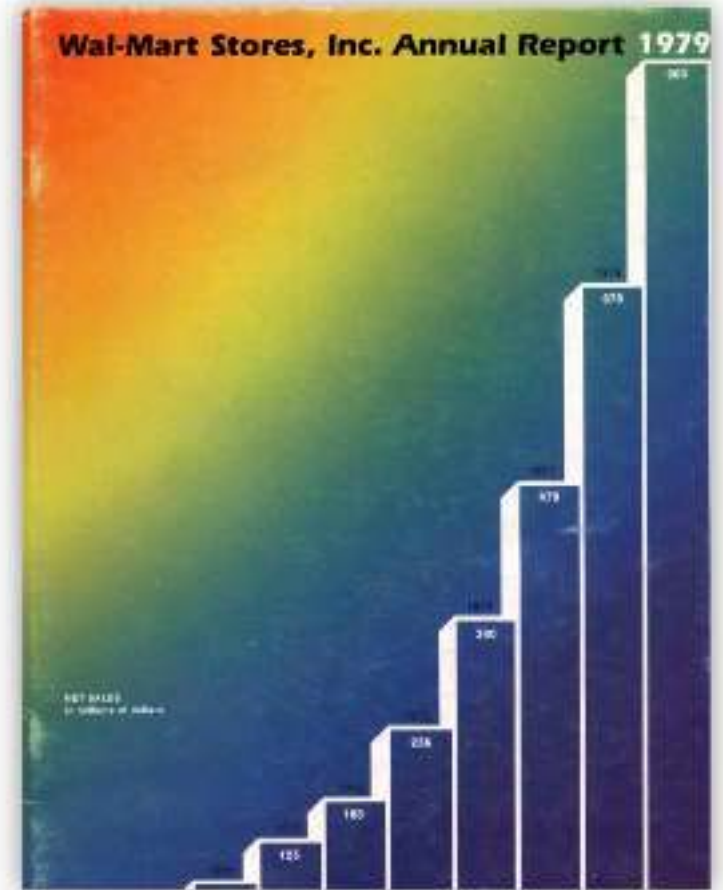
The last time this happened

First 50 years: what's a car? What's a car company?

Second 50 years: how do things change when everyone has a car?

What new kinds of company and business are created?

Source: Walmart



Benedict Evans — December 2023 64

About

What matters in tech? What's going on, what might it mean, and what will happen next?

I've spent 20 years analysing mobile, media and technology, and worked in equity research, strategy, consulting and venture capital. I'm now an independent analyst. Mostly, that means trying to work out what questions to ask.

I write essays about things I'm trying to understand, and a weekly newsletter with 175,000 subscribers, and give presentations pulling together these ideas.

For more, see www.ben-evans.com

Thank you

Benedict Evans

December 2023

www.ben-evans.com
