STIFEL | Healthcare

Insecurity and Opportunity:

Where is Our AI-Enabled World Headed Next?

Written by

Tim Opler
Managing Director
Healthcare Investment Banking, Stifel
Email: oplert@stifel.com.

May 2026

Table of Contents

Executive Summary	3
1. Overview	5
2. How We Got Here	11
3. Where are We Now?	15
4. AI Adoption Rates	19
5. Broad Applications are Arriving	21
6. AI Fears	29
7. Form and Structure	43
8. Where Are We Headed?	47
9. Summary	53

Appendices:

About Stifel	56
Accessing Stifel Reports	58

STIFEL | Healthcare

Executive Summary

This essay provides an overview of current developments in AI, their lineage and future implications. Key points include:

Think Boom, Not Doom

While AI will be disruptive to parts of the software, professional services and physical industries, it is far more likely to raise overall productivity and living standards, much as did prior technology revolutions (steam engines, cars, PCs, internet).

From Turing to Transformers and Inexpensive Computation

Today's AI technology traces a lineage from Turing's "Intelligent Machinery" paper of 1948; Shannon's information theory through backpropagation and the 2017 transformer paper to today's foundation models. The ongoing collapse in computation costs (per token processing costs down over 99% since 2020) has enabled mass AI adoption.

Current Frontier Models and Abstraction Layers

Recent software (GPT 5.5, Claude, Gemini, DeepSeek, Kimi, etc.) already meet or exceed human benchmarks in many cognitive tasks; the locus of competition is shifting from raw models to "abstraction layers" and inference time scaling (agentic orchestration, chain of thought, tool use) that turn models into systems capable of multi step reasoning and workflow execution.

Mass Adoption and Valuation Boom

Generative AI has reached majority consumer adoption within three years, organizational AI usage is near ubiquitous (≈88% of organizations, with 70% using gen AI in at least one function). Valuations of the two leading players (Anthropic, OpenAI) are reaching trillion-dollar levels as the lead up to major IPOs.

Rise of Agentic AI and World Models

We are seeing a major shift from "chatbots" to agentic AI that can manage hundreds of autonomous agents across thousands of steps and to world models that enable major improvements in robotics, autonomous vehicles, defense, industrial systems, and more.

Physical AI is a Major Area of Development

A significant new frontier is AI for engineering complex things (think jet engines). There is a major opportunity to improve product development by computationally-aided engineering of new devices, creating the potential to remake industries focused on the development of physical goods.

Job Loss Fears are Overstated

AI doom narratives, envisioning up to 50% employment loss are economically naïve. Technology tools that improve productivity have historically increased demand for services by making them better (Jevons Paradox). Examples of value-added professions that are growing because of new technology include radiology and wealth management. This said, new technologies can be disruptive while building new professions.

Don't Forget What Makes Us Human

AI is very good at packing existing knowledge to address user queries or to improve work product. AI struggles with very high-dimensional problems, dirty or unstructured data, sparse datasets, and fuzzy judgment. AI is not known for intuition, critical thinking, and imagination of the possible. We advocate liberal education and broad thinking over deep specialization as a response to the improving mastery of the routine by machines. If we invest in what makes us great, the brightest days for humanity should lie ahead of us than behind us.

Acknowledgements

I would like to thank the many people who have helped me learn the AI field and who contributed comments and suggestions on this essay. These include Racquel Bracken (Venrock), Ishan Kumar (GA Capital), Lily Li (Eli Lilly), Jeremy Norman (science historian), Elena Viboch (General Catalyst), D.A. Wallach (Time Bioventures), Tal Zaks (OrbiMed) and Masaki Doi, Stephanie Leouzon, Declan Quirke and Vivian Xu of Stifel.

1

Overview

The two core insecurities in today's life sciences industry are the fear that China will overtake the West in molecular innovation and the fear that AI is going to remake the sector in ways that will be bad for incumbents.

We have spoken quite a bit about China before. While big pharma is starting to pick up molecules in China, the real wealth creation from Chinese innovation activity is taking place in the West.

A good example is the recent news about the acquisition of Candid for $2bn upfront by UCB. Candid was built around Chinese molecules but the upfronts paid to Chinese companies were minuscule compared with the gains of investors who funded the strategy.

Things could change; indeed, they will change as China continues to invest more in the life sciences, but it isn't obvious to us at all that Western companies and investors will be worse off.

We think a similar story holds for AI.

In this essay and in a subsequent one to come we will delve into what is going on now in AI across the economy (topic of this essay) and the life sciences (topic of the second essay).

We're going to review the breakneck pace of innovation in AI tools, their broadening applicability and their increasing utility.

We're going to review concerns about AI's impact on the overall economy – with a particular focus on the potential for AI to disrupt industries.

And we are going to take a point of view on what it all means for the future of our society, our species and ourselves - as individuals.

Not being one to hide the conclusion: while ongoing AI innovations may be highly disruptive to certain industries – we see opportunity overall.

AI tools promise to make both knowledge industries and physical industries more productive. AI tools allow greater intelligence to be applied for routine matters (and, someday, not so routine ones) – ultimately, allowing all of us to focus on doing what we want to do.

This might be creating things.

Enjoying ourselves.

Solving big problems. Etc.

Fears of ever better AI somehow taking over the world are facile. Specially, there is concern in some quarters about Artificial General Intelligence (AGI), a hypothetical type of AI that surpasses human intelligence across almost all cognitive tasks, which then takes over humanity on its own accord.

Computers are agents that act on behalf of principals.

That is us.

Presumably, we can always take away the processing chips or pull the proverbial plug.

Overall, while the societal impact of ongoing AI innovation will be profound, we don't see AI as being likely to make society worse off.

We do see AI as being likely to make the large mass of society *better* off.

In a way that is no different than what has taken place with other innovations – whether it be the steam engine, the car or even the humble toaster.

AI tools are likely to make society better off overall and to accelerate wealth creation.

There are some who worry that AI will widen social wealth disparities.

Even to create a new "permanent underclass."

This worry has been expressed by more than a few in Silicon Valley.

We believe that this worry belies a certain lack of sophistication about the role of tools and computation in the trajectory of the economy.

Alan Turing Guided Us Here

A few weeks ago, I had the pleasure of spending a week in the UK. Three days in Cambridge at a wonderful conference on biomedical innovation sponsored by Royalty Pharma and two days in London.

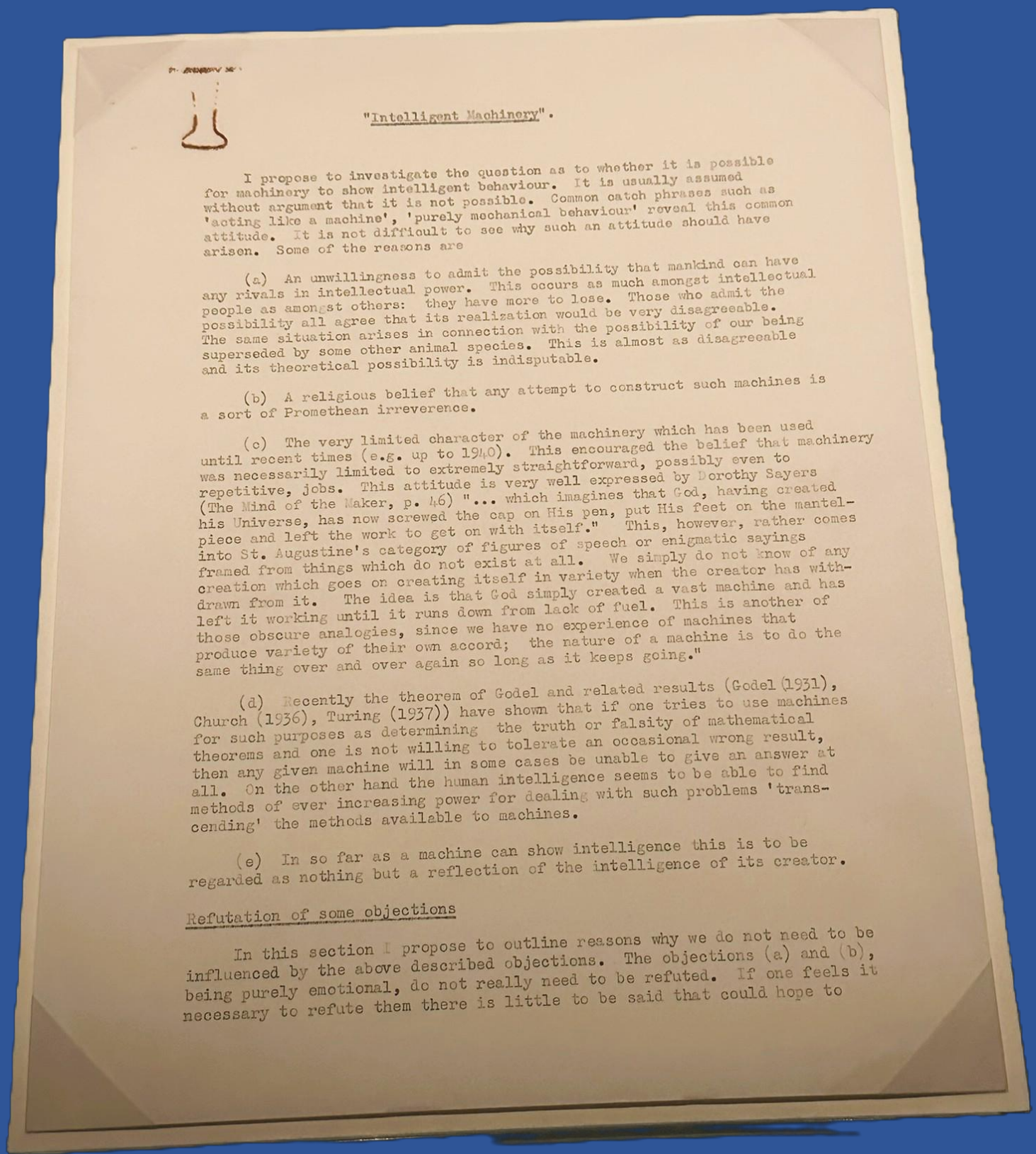
Between client meetings in London, I snuck out for a few hours at the British Library.

Amidst the early editions of Shakespeare, I spent half an hour reading through some of the original documents on the computer age including works by Babbage and Turing.

Alan Turing's 1948 essay *Intelligent Machinery* made for most interesting reading. He wrote:

> "The possible ways in which machinery might be made to show intelligent behaviour are discussed. The analogy with the human brain is used as a guiding principle. It is pointed out that the potentialities of the human intelligence can only be realized if suitable education is provided. The investigation mainly centres round an analogous teaching process applied to machines. The idea of an unorganized machine is defined, and it is suggested that the infant human cortex is of this nature. Simple examples of such machines are given, and their education by means of rewards and punishments is discussed… A great positive reason for believing in the possibility of making thinking machinery is the fact that it is possible to make machinery to imitate any small part of a man."

“Intelligent Machinery” by Alan Turing, 1948



"Intelligent Machinery".

I propose to investigate the question as to whether it is possible for machinery to show intelligent behaviour. It is usually assumed without argument that it is not possible. Common catch phrases such as 'acting like a machine', 'purely mechanical behaviour' reveal this common attitude. It is not difficult to see why such an attitude should have arisen. Some of the reasons are

(a) An unwillingness to admit the possibility that mankind can have any rivals in intellectual power. This occurs as much amongst intellectual people as amongst others: they have more to lose. Those who admit the possibility all agree that its realization would be very disagreeable. The same situation arises in connection with the possibility of our being superseded by some other animal species. This is almost as disagreeable and its theoretical possibility is indisputable.

(b) A religious belief that any attempt to construct such machines is a sort of Promethean irreverence.

(c) The very limited character of the machinery which has been used until recent times (e.g. up to 1940). This encouraged the belief that machinery was necessarily limited to extremely straightforward, possibly even to repetitive, jobs. This attitude is very well expressed by Dorothy Sayers (The Mind of the Maker, p. 46) "... which imagines that God, having created his Universe, has now screwed the cap on His pen, put His feet on the mantelpiece and left the work to get on with itself." This, however, rather comes into St. Augustine's category of figures of speech or enigmatic sayings framed from things which do not exist at all. We simply do not know of any creation which goes on creating itself in variety when the creator has withdrawn from it. The idea is that God simply created a vast machine and has left it working until it runs down from lack of fuel. This is another of those obscure analogies, since we have no experience of machines that produce variety of their own accord; the nature of a machine is to do the same thing over and over again so long as it keeps going."

(d) Recently the theorem of Godel and related results (Godel (1931), Church (1936), Turing (1937)) have shown that if one tries to use machines for such purposes as determining the truth or falsity of mathematical theorems and one is not willing to tolerate an occasional wrong result, then any given machine will in some cases be unable to give an answer at all. On the other hand the human intelligence seems to be able to find methods of ever increasing power for dealing with such problems 'transcending' the methods available to machines.

(e) In so far as a machine can show intelligence this is to be regarded as nothing but a reflection of the intelligence of its creator.

Refutation of some objections

In this section I propose to outline reasons why we do not need to be influenced by the above described objections. The objections (a) and (b), being purely emotional, do not really need to be refuted. If one feels it necessary to refute them there is little to be said that could hope to

Source: Photographed at the British Library, April 2026

Alan Turing

Turing's ideas are remarkable and prescient, anticipating recent developments. It should be possible to design a thinking machine. The way to do this would be to use the human brain as a guiding principle. Thinking is all about inputs and rewards and punishments are required to improve the machine's thinking.

The big breakthrough in AI was the use of neural networks and highly parallel computer chips to find the structure in large amounts of linguistic data, resulting in what is called a "foundation model". This led to OpenAI's ChatGPT-3. These tools have gotten rapidly better, particularly using human feedback (reinforcement learning).

In his 1950 paper *Computing Machinery and Intelligence* Turing described what he called the "Imitation Game" positing that if a human could not tell the difference between a computer and a human's response to questions, challenges etc. then the machine could be said to have "artificial intelligence".

As I write this in May 2026 it is more than clear that Turing's test has been passed, at least, in most important respects by large language models (LLMs) such as ChatGPT.

Not only can machines provide human-like answers but the machines can far surpass humans on some tasks.

A Stanford University report out last month found that today's AI frontier models now meet or exceed human baselines on PhD-level science questions, multimodal reasoning, and competition mathematics.

None of this should be surprising, of course.

Some like to note that LLM's can beat the average doctor at the medical entry exams.

True.

But you too would be able to clobber the average doctor at a medical reasoning question, if you could access a summary of the entirety of the internet in 10 milliseconds or less.

Modern computational power appears to far exceed the capability of any one human brain.

A Stanford 2026 Report Shows that AI Easily Exceeds Humans at Major Intellectual Tasks

Source: https://hai.stanford.edu/ai-index/2026-ai-index-report

2

How We Got Here

The big breakthrough came in 2016 when Jakob Uskoreit simplified natural language processing by using parallel processing to compare each word in a sentence with all the other words nearby, then using the resulting probabilistic structure to predict the next token (basically a word or word fragment) —a mechanism called 'self-attention' (described in a now famous paper entitled "Attention is All you Need")

It so happened that Uskoreit's experiment (and soon that of a famous seven-person team at Google) was facilitated by the massive parallel processing ability of graphics processing units (GPUs) being produced by companies like Nvidia.

The Google team which included Ashish Vaswani, Niki Parmar, Llion Jones, Łukasz Kaiser, and Aidan Gomez built a transformer model, which used only attention mechanisms, shattered translation records while being more efficient to train. This model became the architecture for today's large scale AI models.

It's not a fluke that Anthropic's LLM engine is named Claude – in honor of Claude Shannon.

The idea, of course, is that the problem of token prediction using insights gained from large datasets is a direct application of the ideas of Shannon expressed in his 1948 paper A Mathematical Theory of Communication in the *Bell System Technical Journal*.

Key Paper that Triggered Introduction of LLM's

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com

Noam Shazeer* Google Brain noam@google.com

Niki Parmar* Google Research nikip@google.com

Jakob Uszkoreit* Google Research usz@google.com

Llion Jones* Google Research llion@google.com

Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu

Łukasz Kaiser* Google Brain lukaszkaiser@google.com

Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

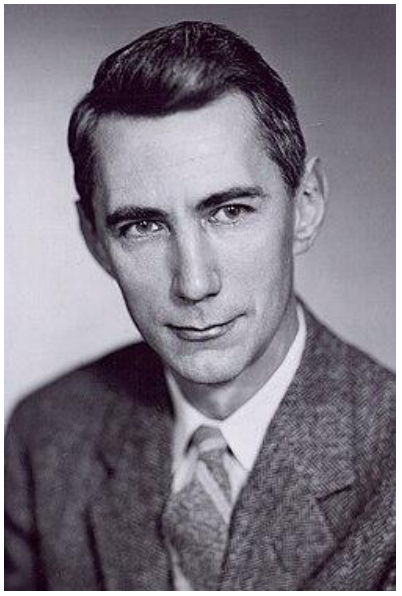
*Equal contribution. Listing order is random. Jakob proposed replacing RNNs with self-attention and started the effort to evaluate this idea. Ashish, with Illia, designed and implemented the first Transformer models and has been crucially involved in every aspect of this work. Noam proposed scaled dot-product attention, multi-head attention and the parameter-free position representation and became the other person involved in nearly every detail. Niki designed, implemented, tuned and evaluated countless model variants in our original codebase and tensor2tensor. Llion also experimented with novel model variants, was responsible for our initial codebase, and efficient inference and visualizations. Lukasz and Aidan spent countless long days designing various parts of and implementing tensor2tensor, replacing our earlier codebase, greatly improving results and massively accelerating our research.

†Work performed while at Google Brain.

‡Work performed while at Google Research.

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

Source: https://arxiv.org/abs/1706.03762



Claude Shannon

Shannon's 1948 paper defined entropy as a measure of uncertainty or information content.

High entropy = high unpredictability = less structure. Low entropy = more predictability = more latent pattern. Shannon showed that any source of symbols (or data) that is *not* maximally entropic contains redundancy — and that redundancy is, in a deep sense, *information about the generative process* that produced the data.

What an unsupervised learning AI model is implicitly doing is building a compressed model of the data distribution — a model that assigns shorter codes to more probable patterns and longer codes to surprises.

Basically, AI models engage in pattern recognition from large datasets.

Extracting signal from noise.

There are numerous details which make today's AI models good.

Geoffrey Hinton

The first is that the underlying neural network technology, developed by many – but most notably Geoffrey Hinton – exploits the idea of back propagation.

Back propagation is the approach that makes training deep neural networks practical.

This works by running two passes through the network: a forward pass, where input data produces a prediction, and a backward pass, where the error in that prediction is propagated back through each layer, allowing the network to compute how much each weight contributed to the mistake. Those gradients are then used to nudge the weights in the direction that reduces error, iteratively over many examples.

In 1986, Hinton co-authored the landmark *Nature* paper with Rumelhart and Williams that popularized back propogation as a practical learning rule for multi-layer networks — demonstrating that hidden layers could learn meaningful internal representations, not just input-output mappings.

Then, starting around 2006, Hinton was central to the "deep learning" revival, showing that very deep networks could be pre-trained layer by layer to initialize weights of elements in layers of a neural net in a sensible region before fine-tuning with back propogation.

The 2012 AlexNet breakthrough (his student Krizhevsky's work) demonstrated definitively that deep networks trained with backprop on GPUs crushed prior approaches on image recognition and essentially launched the modern era.

Remarkable progress in computer processing throughput on GPUs has facilitated today's highly evident progress in AI programs.

In 2007, the best GPU delivered around 497 GFLOPS (one billion floating point operations a second) of single-precision computing power. Today, the best GPUs can push over 100,000 GFLOPS — an astonishing leap driven by advancements in architecture, manufacturing, and the demands of AI and scientific computing.

Nvidia's latest GB300 NVL32 machine for AI data centers delivers astounding performance at a far lower cost per token than was possible just a year ago (shown at left).

Nvidia's GB300 NVL32

That's roughly a 200-fold increase in raw throughput for consumer hardware, and far more for datacenter-class chips.

Computational throughput is now referred to simply as "compute".

The economics of compute have shifted just as dramatically.

Token level processing costs have plummeted because hardware gains compound with algorithmic improvements (better architectures, quantization, speculative decoding, distillation): GPT-4 equivalent performance now costs roughly forty cents per million tokens versus twenty dollars in late 2022 — a 50-fold decline in roughly three years.

A March 2026 paper documented an approximately 600-fold decline in token prices from 2020 to 2026.

As token processing costs have collapsed, use of AI tools has accelerated vastly, and most knowledge workers routinely use AI tools in their everyday endeavors.

3

Where Are We Now?

We find ourselves today locked in a torrent of AI announcements and news.

An almost numbing onslaught of developments promise to change the real economy in profound and unpredictable ways.

In the last ten weeks we have seen more than fifteen major announcements on new tools from leading AI companies, including GPT-5.5, Deepseek V4, Kimi K2.6, Claude Mythos and Gemini 3.1.

It's fair to say that any of the top three LLM providers (Anthropic/Claude, OpenAI/ChatGPT and Google/Gemini) are exceptionally good.

In a recent podcast by Peter Diamandis four commentators remarked on the recent progress, its accelerating pace and what is differentiating the various services.

A key point was that the quality of abstraction layers is becoming more important than the underlying models and raw compute power.

An abstraction layer overallocates compute to a question to figure out what it means and to then break it into subtasks which are then fed to the machine. This is called inference time scaling.

Think of it like the dashboard of a car—drivers don't need to understand the intricacies of the engine to get from A to B, they just need the tools and information provided by the dashboard. Your question needs to be broken into computation tasks and the quality of the system that does this is now paramount.

In many ways, as compute itself has become a commodity, the "creativity" and "mindfulness" of AI is shifting to the engineering of abstraction layers.

A key aspect of abstraction layer engineering involves Chain of Thought (CoT) prompting.

Dave Blundin

CoT prompting enables AI to explain its thought process and walk through multi-step decisions, leading to more accurate and reliable outcomes.

As AI becomes more sophisticated, CoT will be crucial in guiding machines to handle multi-faceted challenges in areas such as scientific research, healthcare, and autonomous systems.

Dave Blundin, Managing Partner at Link Ventures, recently argued that thanks to advanced CoT reasoning and improved raw parameter counts, we are now well beyond the point where AI can match the capabilities of a human.

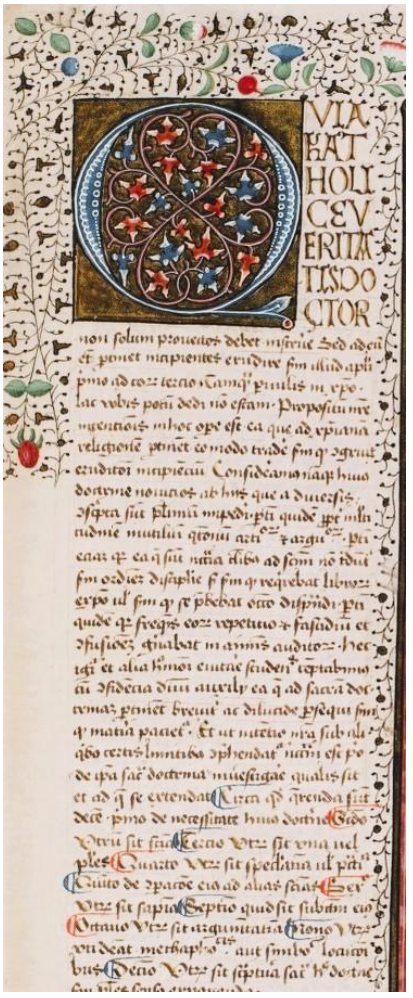
My own sense is that it's not really a contest anymore.

Humans are nowhere close to machines (although see below for more on this).

To illustrate, my somewhat esoteric hobby is studying and collecting antiquarian medical texts. I spend several hours a week plowing through works in medieval Latin related to various aspects of the history of medicine.

This area is sort of the opposite of AI. Not exactly booming. Some would say boring.

This area is so dry for most people that even my supportive wife's eyes glaze over when I describe what I learned that day, say, comparing the 1502 works of Alessandro Benedetti to Gabriele Zerbi. At least Claude doesn't seem to mind when I keep dwelling on the details that differentiate the two.



Today's AI systems can correctly transcribe Latin manuscripts from the Middle Ages and explain their meaning in modern plain English.

What I find most remarkable is the speed with which AI programs can read ancient Latin PDFs, translate them and understand abstruse concepts presented in the language of the time.

This works well even if the Latin is in manuscript form – sometimes in chicken scratch that would match today's worst physician handwriting. Programs like Claude can plow through texts and do philological work in minutes that in the past would have taken up months by classics scholars.

Perhaps most impressive to me has been the CoT reasoning that AI programs use when reading large masses of arcane and ancient texts to understand meaning and context in writings about medicine of more than five hundred years ago. The work is remarkably good.

What's Working Well for Me

For the moment, I have found that Claude seems to do a better job of agentic tasks and answering complex questions while Gemini's Nano Banana excels at creating images, particularly those involving complex scientific concepts in biotech.

ChatGPT, not to be outdone with its recent GPT-5.5 release, now provides much improved answers to hard questions – but is still limiting compute, which can put it behind on hard tasks like image creation.

It's clear that OpenAI is struggling to allocate demand across the compute that they have as I used to get much better results from their GPT models – when they were a little less popular. Anthropic is also struggling with demand for compute.

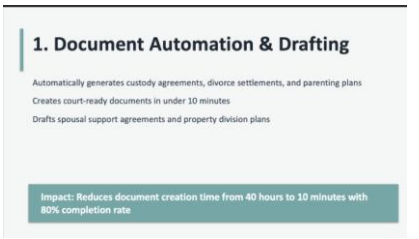
While I write my own material (such as this essay), I do like to use AI tools to check what I have written for factual and grammatical errors. For whatever reason, some of the main AI programs are really bad at this but I have found that Perplexity is an exceptionally good proofreader.

Like many other people I typically will input a query to two or three LLM's at a time and compare results. I must say that Claude just speaks to me in a consistent and high-quality way that is not always matched so well by the other services.

What is Not Working for Me

One of the most common use cases in my world of listening to countless biotech pitches each week is the AI-created PowerPoint presentation.

Half the world seems intoxicated by decks created by Claude or ChatGPT (when you load Canvas). I personally find these decks to be too easily made – often long on form and short on substance.



I was excited the first time I saw a Claude Powerpoint Presentation. Not so excited a few weeks later.

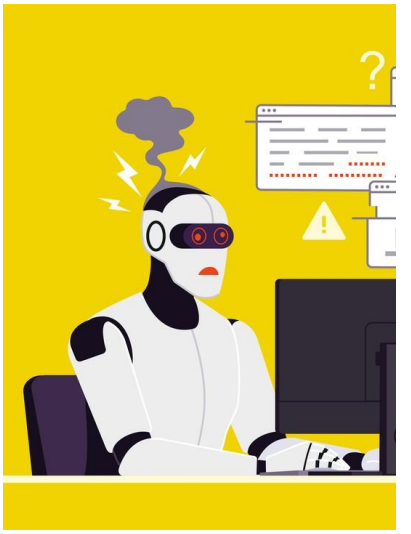
A good pitch involves telling a story in an engaging way that reflects the idiosyncrasies of a situation and LLM's aren't quite there yet on weaving a good narrative. They are a bit too generic still.

Another big issue is hallucinations.

Sometimes AI systems say things that are just wrong or fabricate references that don't exist.

How can a system that seems so smart be so wrong?

Hallucinations happen when the neural network underlying an AI system's answers lacks enough predictive power relative to a specific token and makes a guess.



AI hallucinations have been a problem – but are getting less common.

We'd all prefer that the systems would say "I don't know" but this doesn't seem to be the way things work.

Basically, LLMs are programmed to "fake it" when they don't know an answer.

An important aspect of recent releases from Google, OpenAI and Claude is that hallucination rates are down substantially.

This is very good news.

Recent benchmarking studies show that Google's Gemini has the lowest hallucination rates.

Others worry about sycophancy from AI systems. Sort of like a waiter in a restaurant telling you that you have made a "great choice" – no matter what you pick from the menu.

This hasn't been too big of an issue for me.

Maybe I don't mind a little sycophancy.

4

Adoption Rates

Stanford University's recent *2026 AI Index Report* presented over 400 pages of data on AI trends. Some key items of note included:

AI adoption is spreading at historic speed, and consumers are deriving substantial value from AI tools

Generative AI reached 53% population adoption within three years, faster than the PC or the internet, though the pace varies by country and correlates strongly with GDP per capita. The estimated value of generative AI tools to U.S. consumers reached $172 billion annually by early 2026, with the median value per user tripling between 2025 and 2026.

Organizational AI adoption continued to rise in 2025, up to 88% of organizations, though AI agent use remains early.

Generative AI is now used in at least one business function at 70% of organizations, and China and Europe posted the highest year-over-year increases.

This high adoption has been accompanied by high valuations. There is a bit of a rush to invest in the best AI startups. Pitchbook recently noted that the average Series D valuation of an AI startup is around $4.5 billion (compared to $1bn in 2024). Valuations for Anthropic are now in the trillion-dollar area whereas the last reported value for OpenAI is $880 billion. Both companies are headed for the public markets and we wouldn't be surprised to see public market caps exceed these already heady levels.

Organizational Use of AI is Skyrocketing

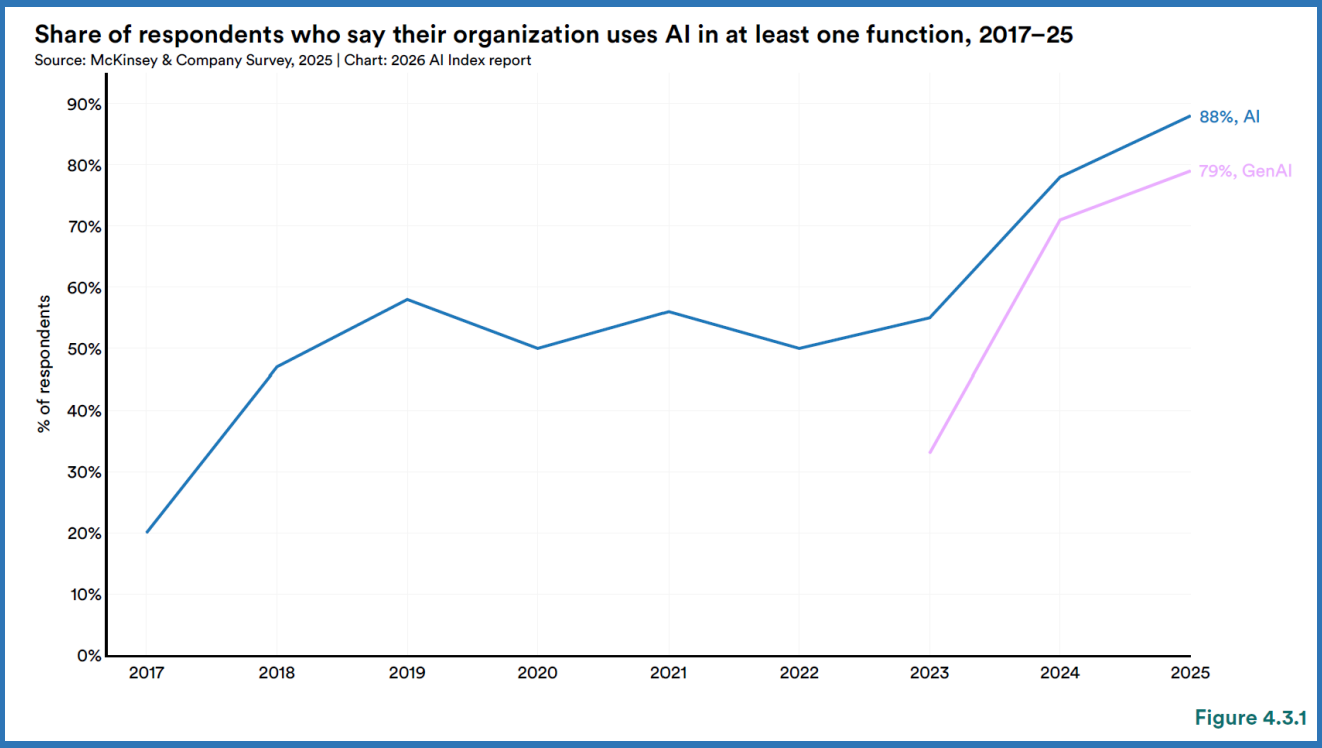
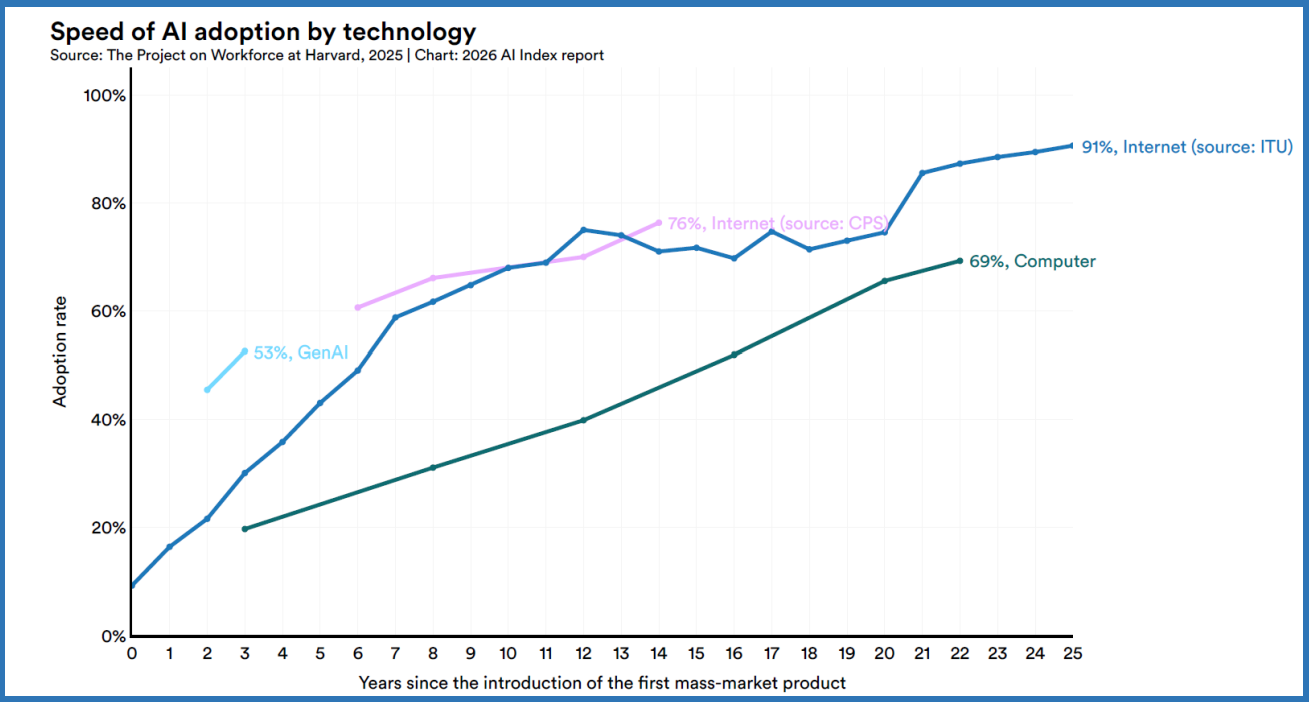


Figure 4.3.1

AI Adoption Pace Outpacing Internet Adoption History



Source: https://hai.stanford.edu/ai-index/2026-ai-index-report

5

Broad Applications are Arriving

The big change in the last three months has been the broader adoption of AI tools in the workplace and in industry – rather than at the level of the everyday consumer like me.

A key idea is that one program can manage multiple agents at once to undertake complex program management.

Rather than being a worker bee with a little help, the new AI models elevate you to be the CEO of countless agents working on your behalf.

We are seeing strange programs that most of us have never heard of before becoming heavily used by enterprise cognoscenti.

These include Fireworks AI, Skippy, and, most recently, Kimi. Kimi2.6, recently released from China in open-source format, was programmed for less than $5 million, can command 300-plus agents, handling up to 4,000 steps per task. Kimi is notable for being able to deliver AI power at vastly lower cost per token.

These approaches appeared just months after the original "Agentic Moment" was triggered this year by the release of OpenClaw.

OpenClaw goes beyond a standard chatbot by executing multi-step workflows—such as managing your inbox, scheduling meetings, writing and deploying code, or booking flights—by directly interacting with your files and connected online services through integration with Large Language Models.

Agentic AI

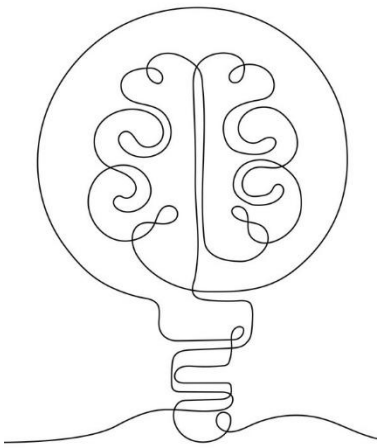
Chain of Thought prompting is advancing rapidly and allows computers to take on far more autonomy than before in the space of actions – rather than the space of symbolic manipulation.

Computers are shifting from being reactive tools to being intelligent agents capable of autonomous decision-making based on complex logic.

Agentic AI refers to systems that operate with autonomy, functioning as intelligent agents capable of understanding context, making decisions, and executing tasks without intermediate instruction.

Unlike traditional AI, which is typically task-specific and reactive, Agentic AI has the potential to act with purpose and self-directed reasoning, analyzing complex environments and acting in ways that mimic human-like decision-making.

Agentic AI is now enabling machines to handle multi-faceted challenges in areas such as scientific research, healthcare, and autonomous systems.



Chain of Thought Prompting Has Been Transformational for AI Applications

This year has seen a definitive shift from "generative" to "agentic" AI, moving beyond models that simply answer questions to systems that autonomously execute multi-step objectives.

This "Action Era" is defined by AI agents that can navigate complex software environments, manage their own memory, and self-correct when they encounter errors.

Instead of receiving a static text response, users are now deploying agents to handle entire workflows—such as conducting a full competitive market analysis or managing end-to-end software deployments—with minimal human intervention.

Agentic AI is soaring in popularity.

Industry-specific adoption has moved from pilot programs to core operational integration.

In software engineering, "coding assistants" have evolved into autonomous software engineers capable of refactoring entire codebases and managing pull requests independently.

The healthcare sector has seen a similar transformation, with agentic systems navigating the "patient journey"— from explaining lab results in plain language to coordinating insurance benefits and post-discharge monitoring.

In professional services like tax and audit, AI is no longer just a search tool; it is a "reasoning colleague" that compiles defensible, client-ready work products by cross-referencing complex global regulations.

We are seeing mind-boggling developments across the global economy on a daily basis.

One of our recent favorites was the announcement from the UAE that it will be shifting half of its government functions in the next few years to be managed by agentic AI.

SeeDance 2.0 from ByteDance Shows the Potential for AI in Creating Hollywood Style Video

AI in Video and Cinema

The future of video production and cinema is rapidly being transformed by advancements in AI, where technology is becoming capable of generating entire commercials, movies, TV shows, and other forms of media. Programs like SeeDance can generate Hollywood quality video in minutes, opening vast new vistas – particularly for “narrowcasting” in areas such as healthcare.

This shift is set to revolutionize the entertainment industry, allowing for unprecedented creativity, efficiency, and customization in content creation.

Even the world of art is changing due to AI. The world’s first AI art museum called Datalab (by Refik Anadol) is soon to open in LA.

World Models

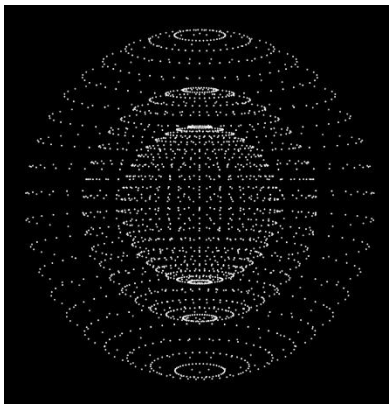
World models are a class of AI systems designed to go beyond the statistical pattern-matching of today's large language models by building internal representations of how the physical world actually works — incorporating laws of physics, the principles of chemistry, appropriate spatial relationships, cause and effect chains, and the arrow of time.

Where current generative AIs learn probabilistic correlations from text, images, and video, world models allow an AI to reason about its environment, plan actions, and predict outcomes, much the way humans and animals do.

Rather than learning to fly by reading a textbook or watching videos, compare to learning through a flight simulator.

The technical architecture involves training neural networks on massive, carefully curated multimodal datasets — petabytes of video and imagery — so that the system can generate realistic simulations of physical environments.

These simulations then serve as training grounds where AI agents learn through reinforcement learning: they act, receive feedback, and iteratively improve, without the cost or danger of real-world trials.



World Models are a Major Force in Today's AI Landscape

Specialized variants include prediction models, style-transfer models, and reasoning models, while large-scale "world foundation models" like Nvidia's Cosmos serve as general-purpose pretrained bases that developers can fine-tune for specific applications.

The applications are already substantial.

Autonomous vehicle companies like Waabi use purpose-built simulated worlds to log millions of virtual driving miles before putting a truck on a real road.

Waabi is Designing Better Trucks Using Simulation

Google DeepMind's Genie 3 can generate photorealistic open-world environments from a text prompt, providing essentially unlimited training terrain for embodied AI.

Robotics, warehouse automation, industrial safety monitoring, and drone systems all benefit from the same underlying capability: the ability to rehearse complex physical tasks in simulation before deploying them in the real world.

The field of robotics is getting heavy investment these days and, in a way, is the perfect embodiment of physical AI. This is an area where China is investing quite heavily and may well pull ahead. With this said, companies like Figure are moving rather quickly to introduce good humanoid robots to the U.S. market.

Havoc's AI solution can be used to control hundreds of autonomous gunboats in a military situation.

Jensen Huang, CEO, Nvidia

Military AI is also an extraordinarily important field and is seeing heavy investment worldwide. Numerous startups such as Havoc AI are doing quite interesting things that could remake the dynamics of modern battlefields.

The longer-term stakes in the physical AI field are quite high.

AI pioneers — including Fei-Fei Li, who has raised $230 million for her world-model startup World Labs, and Nvidia's Jensen Huang — argue that world models are the critical missing ingredient on the path to artificial general intelligence.

Of course, we are getting a rather rapid taste of what complex simulated world models might look like in early efforts from groups like OpenClaw, Google and Meta. Some models are incredibly useful such as a world model that is used to train a self-driving car.

But other efforts remain far from applicable, particularly in contexts where bottom-up modelling of diverse agents is required. The key factor yet to be incorporated in any deep sense are the objective functions and resource bounds of individual agents.

One might imagine that as such efforts progress it will be possible to build better models for stock market investment, for optimizing economic policies and the like.

But we're not there yet in these types of applications although stories of AI agents that exploit arbitrage opportunities in financial and betting markets abound.

Physical Models

One entirely new and exciting area involves the application of AI tools to engineering of physical things.

What's nice is that the physical world is governed by, well, physical laws. Determined by disciplines such as physics, chemistry and the like.

And there are well-worn approaches to the design of complex physical things through modeling and simulation.

We've just never had the type of compute power to apply to these areas in the past as we do now.

Fei-Fei Li of Stanford Has Recently Launched a World Model AI Company Called World Labs

Source: Stanford University

Vik Bajaj, Co-CEO, Prometheus Project

Jeff Bezos, Co-CEO, Prometheus Project

Entire sub professions in engineering are built around the mastery of the art and principles of the applications of physical and chemical rules to real things.

We do not wish to understate the difficulties involved in, say, creating a better building, with a world model rather than through a traditional architect.

But such problems are subject to physical laws and can be thought of as a highly complex optimization problem governed by objective functions and resource constraints.

Examples of startups in this area include World Labs, AMI Labs, Physical Intelligence, and the Prometheus Project. Tech giants like Meta, Nvidia and OpenAI have entered the space, too.

The Chinese are not far behind and are determined to maintain their lead in building really big cool things like giant tunnels, dams and the like.

Perhaps the most visible and interesting effort in this area is the Prometheus Project, of all things, devised by the biotech VC, Foresite through their Foresite Labs team.

We have had the chance to hear the pitch for Prometheus (albeit with a CDA in place) and came away with the impression that this business effort, pioneered by Foresite's Vik Bajaj and Jeff Bezos, is likely to do quite well.

The ideas are big and the intended approach to implementation is sophisticated, thoughtful and optimized to work in world replete with GPUs.

Prometheus has recently completed a $10bn private raise at a $38bn valuation and is focused on designing the most complex physical things using AI tools.

This is a completely different application of AI than its use in computational linguistics to create LLM's – but seems equally if not more useful.

Cristina Criddle and George Hammond of the *Financial Times* described the business on April 20, 2026 as follows:

> "Prometheus, meanwhile, has hired hundreds of staff at its headquarters in San Francisco and in its offices in London and Zurich. It has focused on hiring engineers, AI researchers and people with experience in "building out massive infrastructure projects", one person familiar with its hiring said. The start-up, launched by Bezos last year, is working on AI systems that can operate in the physical world and go beyond the language-based systems behind chatbots such as ChatGPT or coding tools such as Claude Code.
>
> …The company is particularly focused on the industrial sector. It envisions a model that can understand the laws of physics and is trained on data from specific domains, such as jet engine design, one person close to the company said. They added that the company had already "assembled the largest corpus of data on engineering" and how such systems work. Prometheus also plans to amass stakes in companies across sectors such as engineering, aviation, architecture and design. Those deals would include gathering data from these companies, which could be used to improve the start-up's AI model."

We really like what this company is doing insofar as it seems rather obvious that one could improve the design of highly complex physical things like jet engines by applying the principles of physics in AI-governed abstract worlds.

6

AI Fears

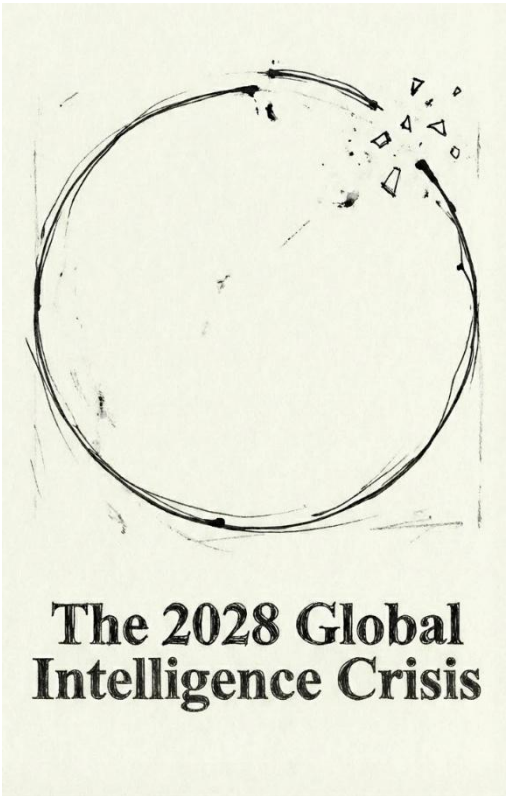
A persistent narrative in recent months has been the notion that AI will end up being destructive to the real economy by displacing jobs and disrupting valuable industries.

A Quinnipiac poll in March found that 70 percent of Americans think that artificial intelligence will lead to fewer job opportunities for human beings, up from 56 percent a year ago. Thirty percent say they're worried for their own jobs. And why not? Warnings of a coming labor market apocalypse feature prominently in the remarks of A.I. leaders.

Perhaps the most influential "scare" essay was published by Anthropic's Dario Amodei on Jan 27, 2026. Entitled *The Adolescence of Technology*, this essay frames AI's current moment as a transitional stage resembling adolescence rather than maturity — systems that display expanding capability alongside volatility in a world not ready for technology change.

Amodei predicted that AI would eliminate 50% of all entry-level white-collar jobs within one to five years, potentially spiking U.S. unemployment to between 10% and 20%, and argued that previous technological shocks had affected only a narrow range of human abilities whereas AI's effects will be both broader and faster — making adaptation far more challenging.


Dario Amodei
The Adolescence of Technology
Confronting and Overcoming the Risks of Powerful AI
January 2026



This Citrini Report Sent the Shares of SAAS Software Companies into a Tailspin

The market reaction was swift and severe.

When Anthropic released 11 open-source plugins for its Claude Cowork tool on January 30, designed to automate tasks in legal, finance, sales, marketing, and data analysis, a $285 billion market selloff followed on February 3–4 — dubbed the "SaaSpocalypse" by market watchers — as investors concluded that AI had evolved from productivity enhancement to existential threat for the software industry.

Not to be outdone, on the evening of Sunday, February 22, 2026, Citrini Research published a speculative retrospective dated from June 2028. The piece, entitled *The 2028 Global Intelligence Crisis*, imagined a "human intelligence displacement spiral" in which agentic AI hollows out software companies, displaces white-collar workers, destabilizes credit and housing, and bankrupts the middle class while AI mega-caps still post record profits.

To quote a hypothetical scenario in the report:

> “Agentic coding tools made a step function jump in capability. Tools like Claude Code or Codex could replicate a mid-market SaaS in weeks, for the price of a few hundred dollars in API credits. CIOs started asking: why are we paying $500k a year for this?”

By Monday's open the memo had gone viral across trading desks. Combined with renewed tariff uncertainty out of Washington, it helped drive the Dow Jones Industrial Average down 822 points — its sharpest drop in months — with software, payments, private credit, and gig economy names leading the slide.

We will not rehash the swift rebuke to basic errors of economic logic embedded in this essay from groups like Citadel and the *Economist*.

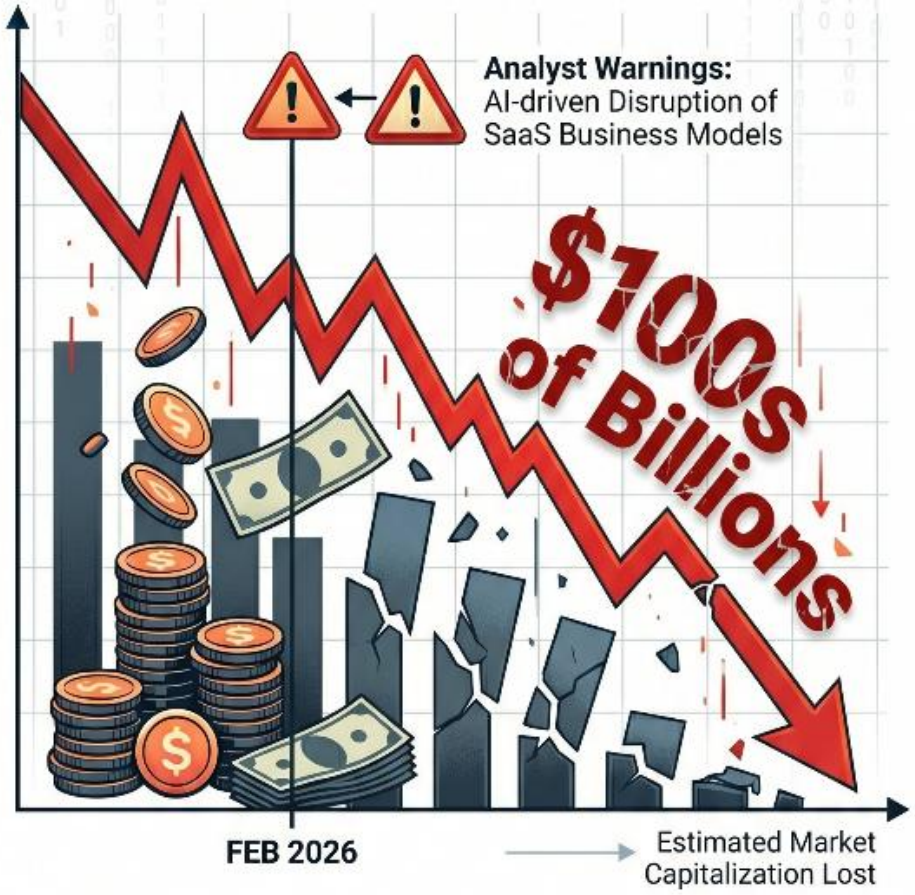
Suffice it to say, it does not follow at all that technological advancement that improves economic productivity would cause the economy to collapse.

Far from it.

Further, the embedded arguments about distribution of economic spoils from AI in both the Anthropic and Citrini reports suggest a basic need to promote better economic education in our society.

SAASpocalypse: The $Billions Melted

Following Feb 2026 Warnings on AI's Impact on the SaaS Industry



Subscription Model Erosion

- AI replacing traditional software tasks
- Reduced seat counts

Revenue Compression

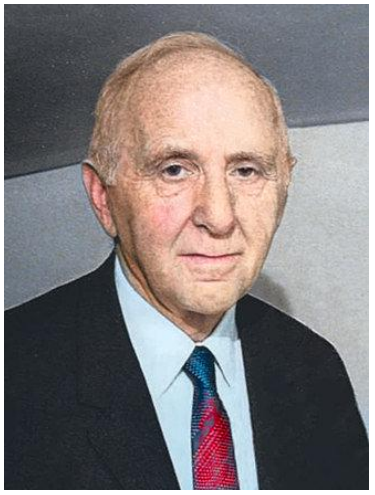
- Pricing pressure from low-cost AI alternatives
- Long-term contract uncertainty

Investor Sentiment Shift

- Re-evaluation of SaaS growth potential
- Capital flight to AI infrastructure

Timothy Bresnahan

Claudia Goldin



Simon Kuznets

The doomsday view is one-dimensional – focused on the lawyer who is displaced by software rather than on the new jobs created from the productivity gained by weeding out an expensive, rent-seeking wage worker.

Indeed, while innovations have always had distributional effects – the notion that they will be ones that widen inequality, is far from a foregone conclusion.

If anything, we see AI tools as equalizers.

Tools that allow economic Davids to slay Goliaths.

Tools that allow the dispossessed – say, a high school student in a third world country – to be more able to compete on the world stage.

The history of innovation tells this story again and again.

It was not so long ago that the world was one of extreme inequality where feudal lords oversaw serfs and the masses led lives that were "nasty, brutish and short".

Our recent report on the history of women's health lays this all out – from a female perspective. Being female and alive in 1200 AD was no picnic.

The last 800 years of technology progress have made both women and men vastly better off despite more than a little economic tumult and Schumpeterian "creative destruction".

While initially creating a "Gilded Age" upper class, the industrial revolution in its many iterations largely elevated the masses to better lives – longer, happier and healthier lives.

Economists Timothy Bresnahan, Claudia Goldin and others have developed the idea that transformative technologies — steam, electricity, computing — follow a characteristic pattern: early adoption is costly and uneven, benefiting only those with capital or skills to deploy the technology, creating a divergence. The diffusion phase then broadens access and compresses inequality.

Similar analyses were carried out by Simon Kuznets in the 1950s, for whom the phrase "Kuznets Curve" is named.

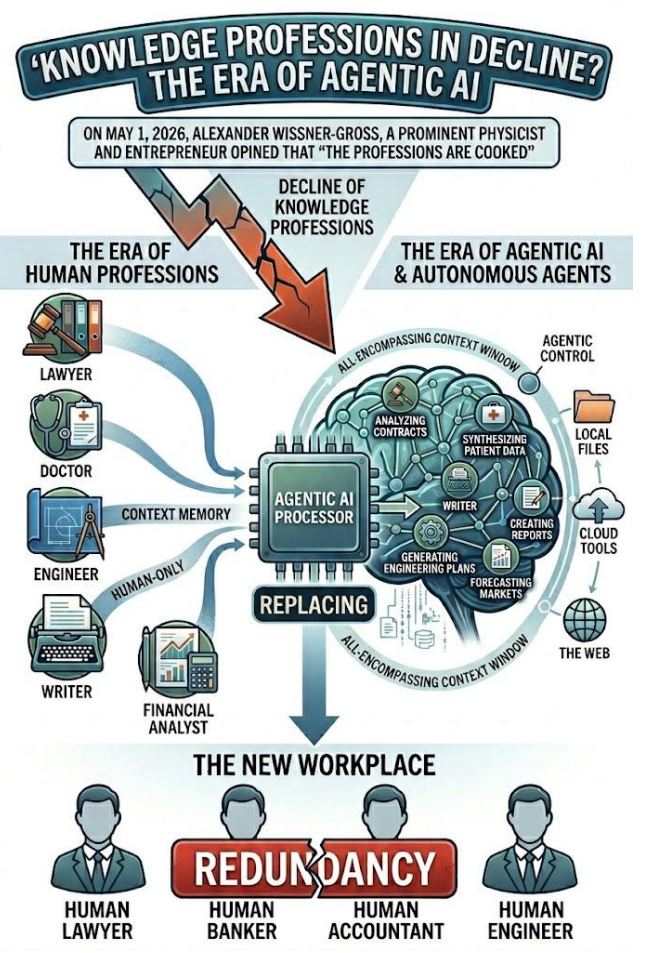
The Knowledge Professions are Cooked. Really?

On May 1, 2026, Alexander Wissner-Gross, a prominent physicist and entrepreneur opined that "the professions are cooked".

He highlighted the fact that OpenAI has mapped out knowledge work verticals and can generate the types of reports that one might normally expect of a management consultant at McKinsey.

We aren't suggesting that there might not be a consultant or two that gets nicked by the AI boom. Or perhaps a finance bro in a Patagonia vest gets decommissioned somewhere.

However, while today's LLMs are very good at aggregating existing data and summarizing it for a user, we have not found that LLMs are particularly good at looking at specific strategic quandaries facing a company and designing solutions based on existing and new data to such problems.



Try it yourself. Ask your favorite LLM what a company like IBM should do differently to grow its business in a volatile environment for IT products.

The answers you get are definite but generic – sort of like what you might get after a week's work by a first-year MBA student. Most likely the board of IBM would be a bit horrified by the lack of depth.

Surface level thinking - not based on deep research about the motivations behind customer buying decisions.

Interestingly, we also have noted that different AI bots give diverse answers to the same question and offer advice, in the case of IBM, that assumes that AI Is going wipe out their largest human-derived segments – which is likely to turn out to be untrue for reasons elucidated below.

If you will, LLM's are great for finding needles in giant informational haystacks, great at writing code, great at computation – but not great at deriving situationally specific insights.

Of course, there are situations where computers should be able to displace knowledge workers with the information that they have in hand (perhaps tax preparers or paralegals, for example).

Indeed, one of the most interesting ironies in professional services today is that the sector is growing rather nicely despite the possibilities of technology substitution.

A battle between robo-advisors and human wealth managers has been ongoing for several years.

Humans seem to be winning despite ever better options to substitute technology-driven approaches.

To illustrate, good options have been available for some time to analyze your financial needs and to make suggestions for shaping your investment portfolio.

So-called robo-advisors have been around for at least five years.

These programs automate investment management by using computer algorithms to build and manage an investment portfolio based on your goals and risk tolerance, charging lower fees than human advisors, which can theoretically translate to higher long-term returns.

The leading robo platforms like Betterment and Wealthfront all operate on broadly similar principles: risk-tolerance questionnaires feed into diversified ETF portfolios, with automatic rebalancing and, on most platforms, tax-loss harvesting.

Despite this theoretical advantage, we are seeing persistent growth in the use of *human* wealth advisors. Private wealth management is not going away anytime soon. McKinsey reports surprisingly that there is a major shortage of good wealth advisors today.

Utilization of robo advisors has been minimal at best.

The score is like a hundred to zero with the machines on the ropes.

There is something intangible and important about delegating your asset management tasks to a human.

Managing someone's retirement savings is a relationship of profound trust, and many investors find it difficult to fully delegate that trust to a software system whose decisions cannot be explained in plain language. In practice, wealthy clients do not feel that AI solutions substitute well for the judgement of a human advisor.

Why Humans Still Beat the Robo-Advisor

Technology can automate portfolio management, but trust, explanation, and human relationships still dominate private wealth management.

THE ROBO-ADVISOR WORLD

LOGIC. AUTOMATION. EFFICIENCY.

Computers can replace some knowledge work in structured tasks.

ETF ALLOCATION

US Stock ETFs	45%
International ETFs	25%
Bond ETFs	20%
Real Estate ETFs	5%
Cash	5%

Robo-advisors build diversified portfolios using algorithms, risk questionnaires, automatic rebalancing, and tax-loss harvesting.

RISK TOLERANCE QUESTIONNAIRE

How would you describe your comfort with investment risk?

1 — 2 — **3** — 4 — 5

Low Risk — High Risk

AUTOMATIC REBALANCING

Keeps your portfolio aligned with your target allocation.

TAX-LOSS HARVESTING

Automatically identifies opportunities to help reduce your tax bill.

Realized Loss
-$4,752
Tax Savings Potential

LOWER FEES Typically 0.10% – 0.25%

24/7 AUTOMATION

SCALABLE AND EFFICIENT

VS.

THE HUMAN ADVISOR WORLD

TRUST. EXPLANATION. RELATIONSHIP.

Managing retirement savings is a relationship of profound trust.
Your goals, values, and life story matter.

Many investors still prefer a human who can explain decisions in plain language.
Context. Clarity. Confidence.

Human advisors provide judgment in complex, uncertain situations.
We plan for life's what-ifs—not just what the model predicts.

We provide accountability and an ongoing relationship.
Someone in your corner through every market cycle.

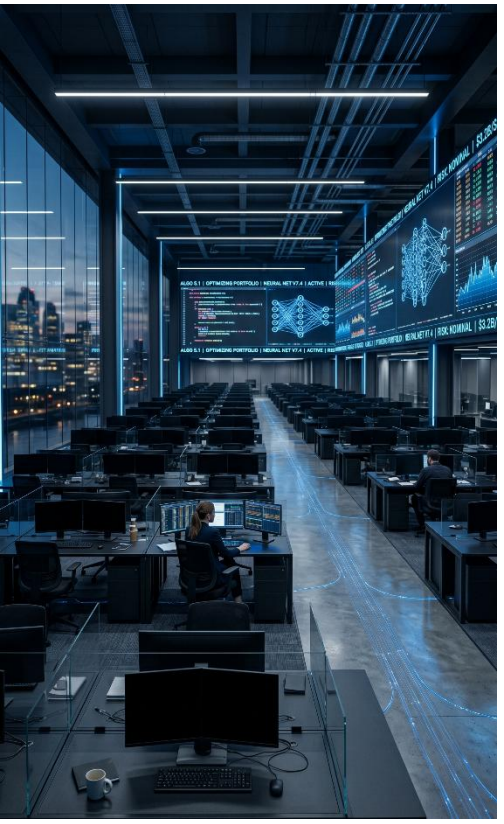
Holistic planning beyond the portfolio.
Taxes. Estate. Philanthropy. Legacy. Aligned with what matters most to you.

Despite lower fees and theoretical advantages, human wealth advisors continue to grow.

At the heart of it all, it's about people, not just portfolios.

When a robo-advisor underperforms, there is no one to call, no relationship to draw on, no accountability structure. When a human advisor underperforms, there is at least a conversation to be had.

Further, most high-net-worth individuals face problems that are considerably more tangled than what a robo-advisor might sort out: estate planning, tax optimization across complex business structures, concentrated stock positions, trusts, charitable giving vehicles, real estate, business sale proceeds, divorce, inheritance. These are not portfolio construction problems — they are life planning problems that benefit from the integration of legal, tax, and financial expertise by humans in ways that no algorithm yet handles well.



Hypothetical investment bank of the future: empty desks everywhere after the AI-driven layoffs.

I happen to work in a classic knowledge industry – investment banking.

Fundamentally, we bankers analyze client needs and situations and attempt to match them up with optimal financing or strategic solutions.

This work requires deep knowledge of one's industry, the client and an intuitive sense of the art of the possible at any given moment.

Relationships are very important in this industry.

Like many investment banks, Stifel's investment banking (IB) department has recently implemented AI tools to undertake financial analyses and obtain assistance with presentation preparation.

These tools are undoubtedly helpful, especially the recent agentic AI capability. Across our industry, it is interesting to see some speculate that with these new AI tools, we will increase the need for junior bankers while others see the opposite.

Time will tell.

We would suggest that the general idea that knowledge workers will be broadly displaced by AI ignores basic economics.

This idea is referred to as the Jevons Paradox.

In 1865 the British economist William Stanley Jevons wrote a book critiquing the then prevalent idea that the best way to solve the country's coal shortage was to make coal production more efficient.

Jevons argued that this solution was not going to work.

He spent an entire chapter in *The Coal Question* arguing that more-energy-efficient technology could not solve England's coal-shortage problem.

In fact, Jevons wrote, greater efficiency of coal-powered technologies would — paradoxically — lead to *more* consumption of coal, not less.

"It is wholly a confusion of ideas to suppose that the economical use of fuel is equivalent to a diminished consumption," Jevons wrote. "The very contrary is the truth."

By analogy, one might imagine that if legal services, financial services, consulting services became more efficient and effective due to the availability of AI, then the utilization of those services could rise.

And productivity enhancing tools might lead to some substitution effect.

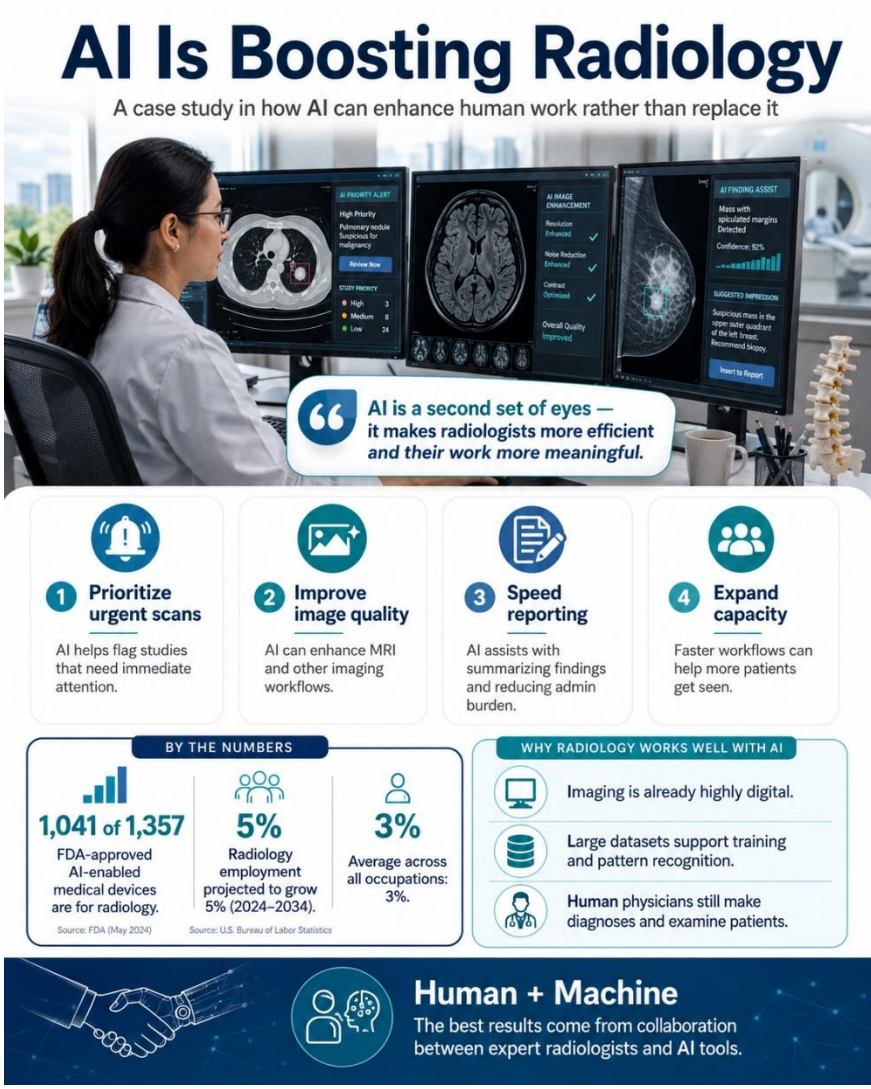
At the margin, some companies might skip using a lawyer, a banker or an accountant because the computer can accomplish the task.

But history tells us that, on average, that Jevons was right.

The Jevons Paradox is evident today in the field of medicine.

AI has made radiology faster, cheaper and better.

So now hospital utilization of *in vivo* scanning has exploded.



William Stanley Jevons, English Economist and Logician

Source: https://en.wikipedia.org/wiki/William_Stanley_Jevons

Rather than putting radiologists out of business, AI has supercharged the profession.

Productivity enhancing innovations are generally associated with greater utilization of the very products where substitution incentives are created.

Interestingly, productivity enhancing innovations can raise wages in other areas of the economy as well.

An important corollary is that if the value of a knowledge-based service can be replaced completely by a computer – which does a better job for less - then the substitution effect would likely outweigh the Jevons augmentation effect. That is, the human worker need have some comparative advantage versus the computational solution.

For example, travel agents are rarely used anymore because the computer can do the same thing more quickly and at a lower cost.

But even usage of travel agencies hasn't fully disappeared. If you are planning a six-week high-end safari to Africa, you might well benefit from the service of a human expert.

Similarly, the availability of productivity-enhancing tools promises to increase the value and insights gained from management consultants, investment bankers and the like.

Try generating your own "McKinsey" report using your favorite LLM. It might look ok but would you be prepared to stake the future of your business on it?

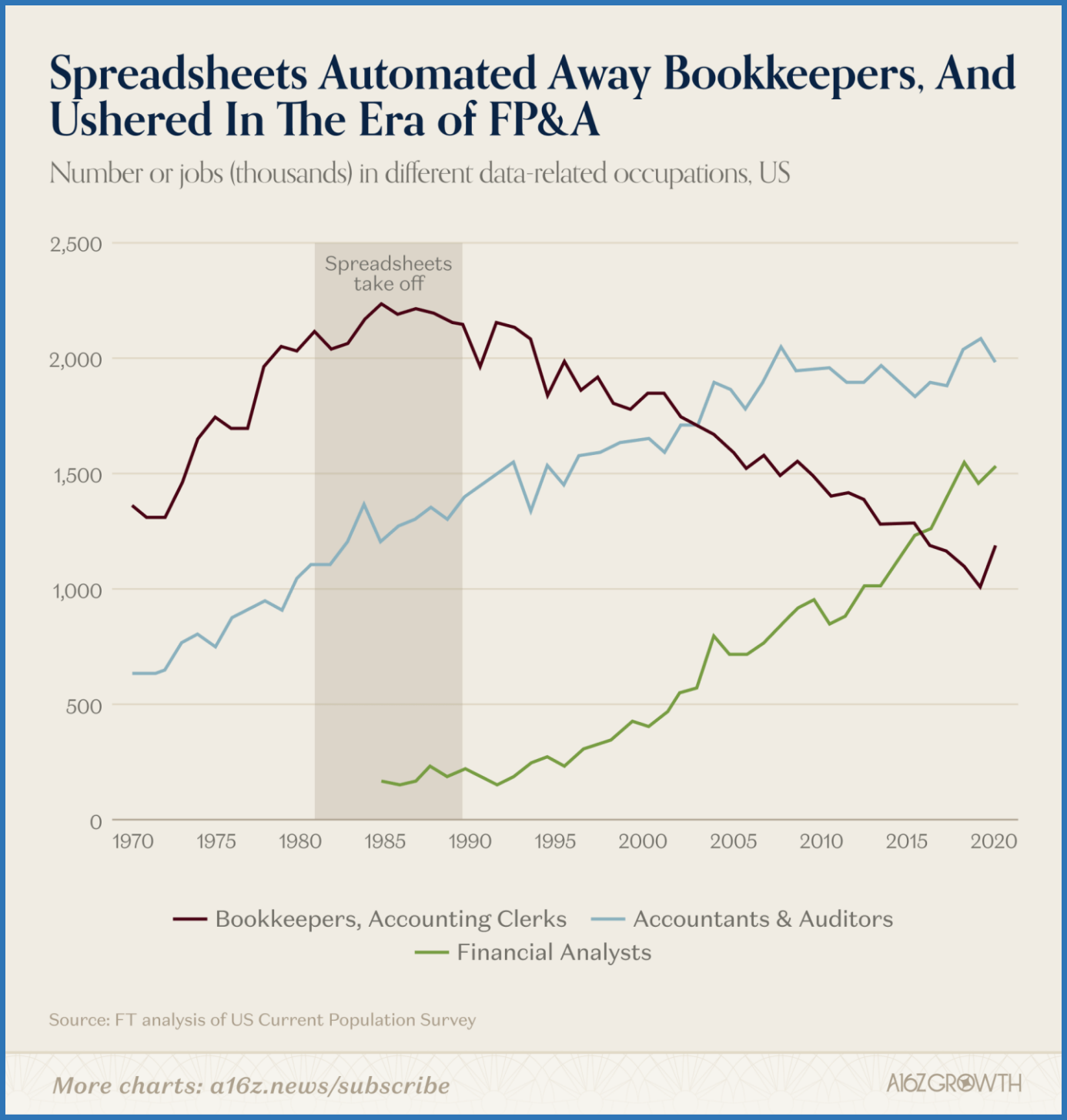
Our general view is that AI tools, if anything, will grow the economy by creating whole new swaths of economic opportunities.

A similar view was voiced by David George of A16Z in his May 6, 2026 essay *The "AI Job Apocalypse" Is a Complete Fantasy*.

He gives a number of examples that show how new technologies have led to *increases* in employment in the involved fields– not the opposite.

We particularly like his example that shows that the advent of spreadsheets was accompanied by rapid growth in the number of accountants and financial analysts but a decline in the number of bookkeepers, whose function was replaced to some degree by spreadsheets (similar to travel agents).

The Employment Effects of Spreadsheets



Source: https://a16z.com/the-ai-job-apocalypse-is-a-complete-fantasy/

Rather than decimating SAAS businesses, we see these businesses getting better because of the availability of AI tools.

We spoke to one CEO of a SAAS business in February who noted that AI was making his service so much better, so quickly that growth was accelerating. Obviously, an online software solution needs to be better than what one can whip up with a program like Claude.

Like the professional services angle, we see AI as facilitating the growth in value-add of customized solutions in the software arena and human-intermediated advice in the professional services field.

The pressure, however, to really add value and to avoid rent-seeking behaviors (e.g., legal overbilling, insight-free consultative reports, generic investment banking advice) will be higher than ever.

The data support our view.

Greg Ip, wrote in the *Wall Street Journal* on May 8, 2026:

> "This [AI] gloom, though, looks like yet another distortion. Some studies see jobs lost to AI, but hard evidence is pretty thin, even in vulnerable occupations such as software development. Private-sector layoff announcements are actually running below levels of a year ago. As for the companies citing AI for job cuts, that probably sounds a lot better than admitting to management failures."

A recent article on compensation and revenue by top law firms in the *Financial Times* showed that usage and revenue from the top firms is going up rather impressively in recent years.

Similarly, despite ample opportunities for substitution we are not seeing anything remotely like a decline in management consulting or financial service industries.

Competition for supposedly endangered IB analysts and associates indeed is fiercer than ever today.

Notably, this earnings season is seeing record results delivered by a number of financial services players.

7

Form and Structure

As we collectively contemplate where AI is taking us next it is worth commenting a little on what it can and cannot do well. The use of neural nets in a high throughput computational environment is a great way to pick out structure in a large dataset.

Human language, for example, is highly structured – which means that the next word you or I might utter is rather predictable based upon context – hence the “miracle” of how well LLM’s seem to work in addressing our various questions and comments.

That underlying structure can be quite complex, of course, and can be stored in some abstract sense in what scientists call foundation models.

This harkens back to Plato’s Theory of Ideas, also known as the Theory of Forms, which allowed us to envision an ideal intellectual reality.

Plato wrote: "The world is a shadow of the true reality that is the world of the Forms".

There are several settings where AI tools based on neural nets break down.

The first is when a problem has a dimensionality that exceeds available computational capacity.

This is more common than one might think. For example, small molecule drug design involves solving incredibly high dimensional problems – well beyond the capability of the even today's most impressive array of GPUs.

The second is when a dataset either lacks structure or there is no obvious way to use an "unlabeled approach" to extract that structure. This is a major issue in healthcare where datasets are not "clean" or will be too costly to clean up.

A third issue takes place in settings where input datasets are too small – which means that brute force construction of foundation models from giant datasets is not possible.

Interestingly, humans seem to be able to learn much better from small datasets than do machines.

Yann LeCun, AI pioneer and Executive Director of Advanced Machine Intelligence states:

> "A 17-year-old can learn to drive a car in about 20 hours of practice, even less, largely without causing any accidents… And we have millions of hours of training data of people driving cars, but we still don't have self-driving cars. So that means we're missing something really, really big."

Yann LeCun

He notes that humans are much better of building mental models of the world from limited observations compared to computers.

We say all of this because it is easy to overestimate what AI can do. Part of our purpose in this essay is to help build a sense of might be possible with AI tools.

This gives one a sense of why "wisdom" built by professionals in service industries (e.g., an expert lawyer) is very hard for computers to replicate.

Sure, the computer can write a nice legal brief but it will struggle with what to say in a specific complex situation.

In an ironic way AI tools commoditize rote thinking. If there are relationships that can be revealed through a dataset that is large enough, then those GPU chips armed with a neural net are going to be hard to beat.

But real insights and wisdom remain scarce – particularly those that might be derived from fuzzy situations or settings where data is sparse.

All innovations place evolutionary pressure on various species, including us as humans.

If there is a takeaway, it strikes us that humans faced with ever better AI, shouldn't respond necessarily by becoming AI experts.

Abraham Maslow

Perhaps we would be better off if we worked to become better at what AI is not so good at.

This includes intuition, imagination, critical thinking skills and wisdom.

Imagination, the ability to envision a yet unseen opportunity, has repeatedly been the critical success factor in knowledge-intensive industries.

This imagination typically emerges from a thorough understanding of current technologies and a sense of what is possible with novel technologies.

This, in turn, often arises from cross-fertilization of ideas from various disciplines, an appreciation of the possible and a deep appreciation of individuality.

In many ways the ultimate meaningful life comes through what Abraham Maslow calls "self-actualization".

Self-actualization is the drive to become *the most that one can be* — not in a competitive sense, but in the sense of fulfilling one's own deepest capacities and nature. Maslow described it simply: *"What a man can be, he must be."* The form it takes is highly individual — for one person it might mean becoming a great parent; for another, a scientist, artist, or leader.

In an age of technology and AI threat, our counsel is to go broad in the pursuit of insight and wisdom and to use emerging tools and technologies to achieve the most for one's inner self and the society around us.

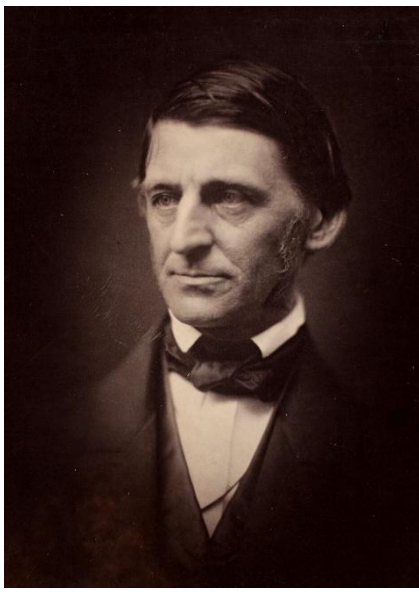
The spirit of Ralph Waldo Emerson in his essay "Self-Reliance" captures the idea of a liberal education as a way of building authentic individuals imbued with uniqueness of perspective.

Emerson wrote. "Trust thyself. Every heart vibrates to that iron string."

Unique insights are the fabric of progress. These often come from the outside. We are reminded of the physicist Erwin Schrodinger's distinctive 1944 book "What is Life?"

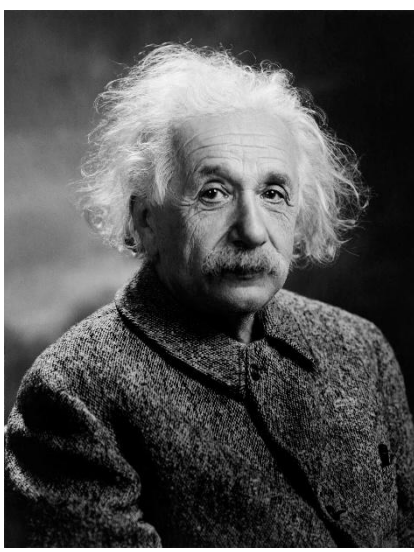
While discussing the order of life in physical concepts his book inspired James Watson in his pursuit of the structure of DNA, indirectly remaking the life sciences in the most profound of ways.

Another example is Einstein's letter to FDR about the potential for an atom bomb. Einstein understood the emerging field of nuclear physics and could see its implications. His letter also had profound implications.



Ralph Waldo Emerson

Einstein himself said: "Imagination is more important than knowledge. For knowledge is limited whereas imagination embraces the entire world, stimulating progress, giving birth to evolution."



Albert Einstein

Former Nebraska Senator Ben Sasse opined on this topic, particularly, the potential of an AI-driven dystopia, in the *Wall Street Journal* on May 8, 2026, writing:

> "Character, whether of an individual or of a nation, is molded by habits and by time. This republic requires men and women to do long-form deliberation, serious thinking, honest humility and daily striving. What good is it to gain the whole world if we forfeit the souls that we're supposed to form? We can't expect to remain free without being virtuous, we can't be bold without being rooted, we can't be great without aiming first to be good. To stave off Huxley's dystopia, we must deliberately shape our children's souls so that they can be creators, doers and thinkers embracing the next frontier."

Our own view is that good ways to build critical thinking skills and imagination is through a liberal education that encompasses the study of history, languages, philosophy, literature and the sciences.

8

Where Are We Headed?

Speaking of imagination, it is worth pausing to speak a bit about where AI might be taking us next.

If you will, today's breathless and often rapid response to the firestorm of AI developments may miss larger opportunities posed by the new technologies.

Count Us in the Ranks of Singularity Skeptics

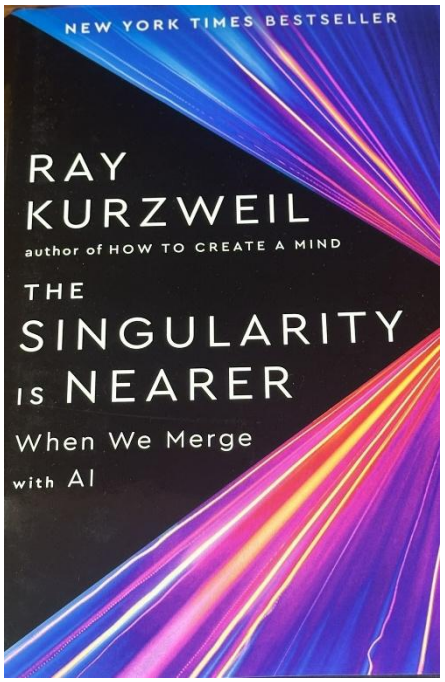
In a recent visit to the Stanford University bookstore, we were impressed to find an entire section on AI. It's obvious that AI has sparked the imagination of many thinkers in our society, Some of the titles struck us as hilarious and there seem to be no shortage of fringe type thinkers looking at the AI field, including several books on the "singularity".

The technological singularity is the hypothesized point at which artificial intelligence becomes capable of recursively improving itself — designing successor systems smarter than itself, which in turn design even smarter ones — triggering an "intelligence explosion" that races beyond human comprehension or control.

The idea rests on a recursive self-improvement loop: once an AI reaches roughly human-level general intelligence (often called AGI), it can be turned on the problem of AI design itself.

AI Section of the Stanford University Bookstore

Source: Photographed in March 2026



Each generation produces a smarter successor faster than the last, compressing centuries of progress into years, months, or — in the most aggressive framings — days.

Beyond that horizon, the future becomes fundamentally unpredictable to current human minds, much as a chimpanzee cannot anticipate human civilization.

Some see human minds merging with that of machine even.

AI leaders such as Sam Altman, Dario Amodei, Elon Musk and Demis Hassabis have all publicly suggested transformative AI may arrive within this decade, and agentic systems that act autonomously across software workflows are widely read as the first tangible rungs of the recursive ladder.

Bear in mind that these are some of the same leaders who argue that AI is going to lead to mass unemployment.

Technical expertise does not necessarily make one an expert on other matters.

This all sort of reminds us of the Y2K scare – the idea that world would end when clocks turned to 2000.

Our own sense is that current architectures hit fundamental ceilings well short of recursive self-improvement, and that the "singularity" is a quasi-religious narrative dressed in technical clothing.

The above points on the lack of real insight and imagination coming from AI tools is incredibly pertinent. Turing's test may have been passed by current systems but there appears to be many miles left to travel before AI takes on truly human thinking abilities.

Agentic AI of Profound Importance

In contrast, it is fair to say that the ability of AI to coordinate actions within environments is of the most profound importance. So-called agentic AI is in its infancy and it strikes us that the ability to delegate tasks to computers is highly valuable and frees up time for higher order thinking.

It strikes us that there is far to travel down this road and that the product of agentic interactions and associated variation in agentic objective functions can be complex, beautiful and deep in implications for the worlds around us.

J. Doyne Farmer

For all the nascent excitement about agents and world models there is a certain naivete to approaches taken – and the notion that these models will lead us to a form of general intelligence.

In early work on complex nonlinear systems (sometimes known as "chaos theory") at places like the Santa Fe Institute in the 1990s, researchers began to experiment with complex agent models.

This work continues today by researchers such as Doyne Farmer, now at Oxford University.

The key point, whether it be in an economy or other type of "world", one needs to think hard about the objective function of individual agents. Farmer's work finds that the results of economic simulations depend very much on the environment, the type of bounded rationality imbued in the various agents and the level of rationality of those agents.

Doyne Farmer's vision of global-scale agent-based economic simulations ("one model could realistically predict the economy of an entire country… the entire world") can be thought of a socio-economic analogue of global foundation models: large, reusable simulators that can be specialized to particular questions.

The big open technical and epistemic questions are similar in both communities: how to validate such models, how to avoid overfitting to past regimes, how to represent rare events, and how to make them interpretable enough for high-consequence decisions.

The possibility to transform industries such as healthcare with agentic AI is obvious. But, there is much that is missing, including efforts to standardize relevant datasets including medical records and API's that bridge cyber interactions to the space of real world actions – like get me a doctor's appointment, pick up a prescription, order labs etc.

Physical / World Models Highly Promising

The work on world models and associated redesign or improvement of designs of physical things is striking. LLM's allow us to gain access to deeper levels of information and to write code, make beautiful images and the like.

If you will, they are quite helpful adjuncts to knowledge work.

But the ability to remake objects in the world around us using advanced computation seems to be of the most fundamental importance as we are defined by and constrained by our environment and the objects within it.

The current talk from groups such as the Prometheus Project is for faster and better jet engines, but one could just as easily contemplate construction of much larger objects such as better cities, better farms, better hospitals and the like.

And, there are problems inherent in our civilization such as the battery life of electric cars that might be transformed with these new AI tools operating at the physical layer of our existence.

The Business World

We have earlier spoken of the implications of LLM's not as a threat to knowledge work but as a valuable adjunct to it. But this only scratches the surface of what might be possible from the application of AI to the improvement of business productivity.

The reality is that some businesspeople are far better than others and it's clear that work ethic, personal skills and intuition about markets drive success.

There is the whole idea of "strategic planning" in business – the notion of thinking through the implications of specific implementations of business ideas to determine which will generate the highest returns in a capitalist system.

In a sense this harkens to Doyne Farmer's ideas about equilibria emerging from complex systems.

The combination of AI reasoning, agentic models and world simulations creates new possibilities for the design, implementation and execution of business plans.

One should be able to grow new constructs around the legacy economy while determining better ways to operate existing businesses, whether they involve airlines, construction or making machines.

Coupled with emerging frameworks for business function (see home.locus for an example), it should be possible to design business models and approaches to markets in ways that have heretofore not been imagined.

In the same sense that one might be able to use AI tools in a constrained optimization setting to build a better jet engine at a far lower cost, it might also be possible to design whole businesses in a similar way, leveraging fresh insights into how coordination can be improved with machine learning and the like.

Rather than chemistry and physics, the frameworks would involve areas such as the principles of microeconomics, consumer choice and functional analysis.

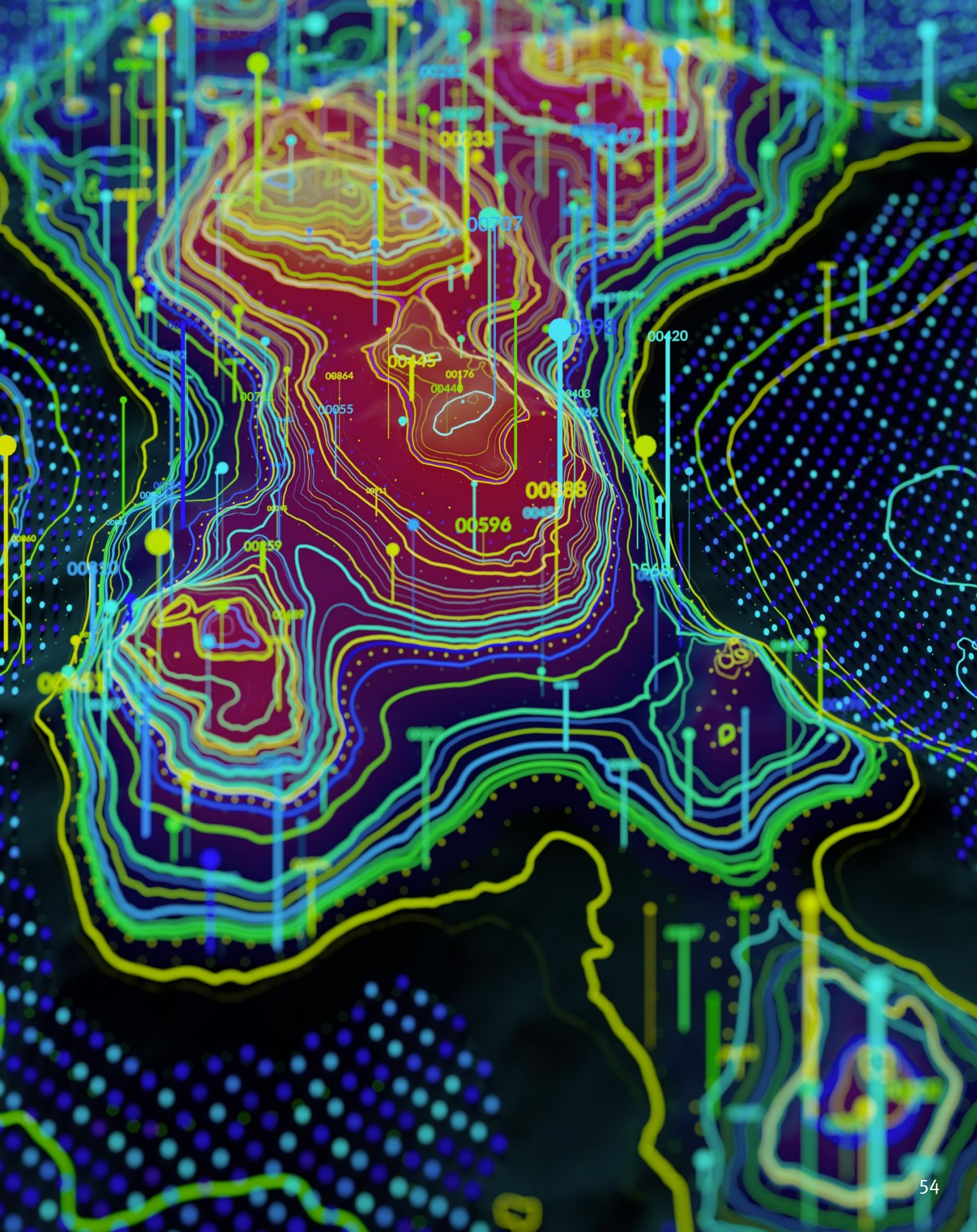
It is so interesting to us that companies like Anthropic have chosen to create applications that compete with SAAS business models (an area where progress is likely to be slow) and have instead not gone after the much more open area of new business design.

9

Summary

This essay has parsed through recent developments in AI tools and has commented on their potential effect on the real economy. Newly arrived AI tools are very good and are changing many aspects of our personal business worlds. Yet, for all of its power, AI has yet to really threaten the ability of humans to generate differential insights, engage in imaginative activities and the like.

In many ways today's financial markets have become more reactive than ever before. Thus, it has been quite unfortunate to see misapprehension and overreaction to recent developments in the agent and tool space, leaving some to believe that areas like the SAAS field is doomed. Similar logic has been applied to knowledge professions suggesting that these areas are "cooked". There is no real data to support any of these contentions. If anything, we see AI as spurring further growth and productivity among the traditional software industry areas and professional services. This said, AI is leading to real and tangible benefits for the total economy, just has happened with past technology innovations such as the internal combustion engine, transistor, the personal computer, the internet and the like. Some of the most interesting developments in recent months in AI technologies, particularly in physical AI, world models and agentic AI promise to bring profound improvements in economic productivity that are likely to make our society better off for generations to come.


00707
00420
00864
00176
00055
00596


STIFEL

STIFEL

Premier Healthcare Investment Banking Group Focused On Growth Companies

Stifel is a full-service investment firm with a distinguished history of providing securities brokerage, investment banking, trading, investment advisory, and related financial services to individual investors, institutions, corporations, and municipalities.

For 135 years, meeting and exceeding the objectives of our clients has positioned our firm for success. We believe that by placing the clients' interests first, they prosper. It is when our clients prosper that the firm succeeds. At Stifel, we can truly say that we predicate our success on the success of our clients. We are nurtured by years of trust and understanding, and by shared goals and shared successes. It is our relationships with, and our commitment to, our clients that enable us to develop innovative financial strategies.

Established in 1890 and headquartered in St. Louis, Missouri, Stifel is a leader in private wealth and investment banking services. Stifel, Nicolaus & Company, Incorporated is a wholly owned subsidiary of Stifel Financial Corp. Stifel Financial Corp.'s publicly traded stock is listed on the New York Stock Exchange under the symbol "SF."

Stifel offers investment banking, wealth management, and an extensive US brokerage presence (with $400bn in client assets).

Stifel's investment banking business is significant. With over 700 investment banking professionals across 13 industry vertical areas, Stifel has 20 offices around the globe.

Stifel offers a full range of investment banking products mainly to high growth companies across the entire corporate lifecycle from venture to public markets.

Stifel's Healthcare Investment Banking business has over 100 professionals and serves clients in the biopharmaceutical, medtech, HCIT and healthcare services sectors around the world. Stifel has prolific underwriting experience with IPO's and follow-on financings and a dedicated M&A franchise. Between 2020 and 2025 Stifel raised over $80 billion for its clients in over 300 transactions. In the same period, Stifel worked on more than 190 advisory assignments for its clients.

Accessing Stifel Banking Reports in Healthcare

There are several ways to access Stifel's Healthcare investment banking thought leadership reports. Many current and past reports are online and can be accessed at the website: https://stifelinstitutional.com/insights/thought-leadership/. From there select reports in healthcare.

To get on the mailing list for this publication feel free to contact Tim Opler (oplert@stifel.com)

Some recent reports include:

Q1 Biopharma Market Update

April 6, 2026

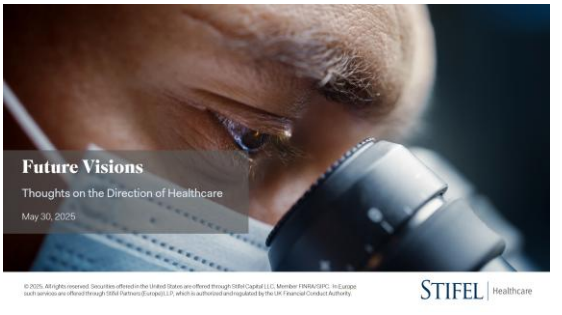
Biopharma Outlook 2026

Jan 8, 2026

Medicine, Progress & AI

Sep 11, 2025

Healthcare Future



May 30, 2025

Further Reports

Some other recent reports include:

Feb 9, 2026 (Biotech Sentiment)
Jan 20, 2026 (Women's Health Pt 1)
Nov 24, 2025 (Biotech Outlook)
Nov 3, 2025 (China Update)
Oct 28, 2025 (Breast Cancer)
Oct 6, 2025 (Biotech Bull Market)
Sep 16, 2025 (Fixing Pharma's Image)
Aug 18, 2025 (Cardiovascular Drugs)
Jul 14, 2025 (Top 40 Pharma)
Jul 9, 2025 (Obesity Drugs)
Jun 23, 2025 (Science and Truth)
May 12, 2025 (MFN Policy)
May 5, 2025 (NIH Cuts, China Tariffs)
Apr 28, 2025 (Eyes on Washington DC)
Apr 21, 2025 (FDA Shifts, Buyside Update)
Apr 14, 2025 (Wild Week in Market)
Apr 7, 2025 (Biotech Market Break)
Mar 31, 2025 (China Biotech Update)
Mar 24, 2025 (Healthcare Reform)
Feb 24, 2025 (Retail Pharma Trends)
Feb 10, 2025 (Pharma Earnings)
Jan 27, 2025 (Women's Health, Obesity)
Dec 17, 2024 (Biotech Blues)
Nov 25, 2024 (Biotech Balance Sheets)
Nov 18, 2024 (New Administration)
Nov 4, 2024 (Election, Obesity)
Oct 21, 2024 (China, Pfizer)
Oct 7, 2024 (VC update)
Sep 23, 2024 (The Fed Rate Cut)
Sep 9, 2024 (Sector Outlook)
Aug 12, 2024 (Biotech Market)
July 8, 2024 (Obesity Market Update)
June 17, 2024 (Lab Market)
June 8, 2024 (Oncology Review)
May 27, 2024 (GLP-1's)
May 20, 2024 (Returning Capital)
May 13, 2024 (Brain, AlphaFold 3)
May 6, 2024 (Earnings, Obesity)
April 29, 2024 (M&A, Japan)
April 22, 2024 (Pharma Pricing)
April 15, 2024 (AI in Pharma)
April 8, 2024 (The Buyside)
April 1, 2024 (Biotech Balance Sheets)
March 25, 2024 (Women's Health)
March 18, 2024 (Inflammasome)
March 11, 2024 (IRA, Immunology)
March 4, 2024 (Biotech Employment)
Feb 26, 2024 (Biotech Strategy)
Feb 19, 2024 (Big Drugs, Autoantibodies)
Feb 12, 2024 (Fibrosis, Endometriosis)
Feb 5, 2024 (Severe Disease in Women)
Jan 29, 2024 (Pharma R&D Productivity)
Dec 18, 2023 (Expectations for Future)
Dec 11, 2023 (ASH, R&D Days)
Dec 4, 2023 (Big Pharma, CEA)
November 20, 2023 (M&A)
November 13, 2023 (AHA, Bear Market)
November 7, 2023 (Unmet Needs)
October 30, 2023 (ADCs)
October 23, 2023 (ESMO Review)
October 16, 2023 (Cancer Screening)
October 9, 2023 (Biosimilars, M&A)
October 2, 2023 (FcRn, Antibiotics)

Disclosure

Stifel collectively refers to Stifel, Nicolaus & Company, Incorporated and other affiliated broker-dealer subsidiaries of Stifel Financial Corp. The information and statistical data contained herein have been obtained from sources that Stifel believes are reliable, but Stifel makes no representation or warranty as to the accuracy or completeness of any such information or data and expressly disclaims any and all liability relating to or resulting from your use of these materials. The information and data contained herein are current only as of the date(s) indicated and Stifel has no intention, obligation, or duty to update these materials after such date(s). These materials do not constitute an offer to sell or the solicitation of an offer to buy any securities and Stifel is not soliciting any action based on this material. Stifel may be a market-maker in certain of these securities and Stifel may have provided investment banking services to certain of the companies listed herein. Stifel and/or its respective officers, directors, employees and affiliates may at any time hold a long or short position in any of these securities and may from time-to-time purchase or sell such securities. This material was prepared by Stifel Investment Banking and is not the product of the Stifel Research Department. It is not a research report and should not be construed as such.

Stifel, Nicolaus & Company, Incorporated | Member SIPC & NYSE | www.stifel.com

STIFEL | Healthcare

STIFEL | Healthcare