AI Has Finally Become Transformative

After a decade's worth of innovation, new models can change the world the way the internet did.

By Martin Casado

Aug. 3, 2023 12:47

PHOTO: ALAMY STOCK PHOTO

Artificial intelligence has generated tremendous value across many applications over the last decade, including search, ad targeting and recommendations. But nearly all these gains have gone to tech giants such as Google and Facebook. Despite the hoopla—and a lot of related startup activity—AI hasn't brought a market transformation similar to the internet or mobile, in which an entire new class of companies emerge and become household names. That may soc change.

Despite AI's enormous capabilities, the economic realities for using it haven't been great for startups. Often the amount of value a company gets from AI diminishes quickly over time, and therefore requires significant continuing investment. And while the benefits are tangible, it is hard for startups to maintain growth and distance themselves from more-standard approaches As a result, AI's primary value has been to improve existing operations for incumbents who have the resources to invest at the required levels.

A common failure scenario in earlier-generation AI startups (which I call the AI mediocrity spira highlights a few major factors at play.

In order for a startup's AI-based application to have sufficient accuracy early on, the company hires humans to perform the function it hopes the AI will automate over time. Often, this is part

of an escalation path where a first cut of the AI will handle the most common use cases, and humans manage the long tail of less-common ones.

Early investors tend to be more focused on growth than on margins. To raise capital and keep the board happy, the company continues to hire people rather than invest in the automation—which proves tricky anyway because of the aforementioned complications. By the time the company is ready for growth-level investment, it has already built an entire organization aroun hiring humans in the loop, and it's too difficult to unwind. The result is a business that shows relatively high initial growth, but maintains a low margin and, over time, becomes difficult to scale.

Fortunately, this doesn't seem to be the case with the current wave of generative AI applicatio such as ChatGPT and the foundation models such as GPT-4 that power them. While still very early, we're already seeing use cases in large existing markets with orders-of-magnitude improvement in time, cost and performance. This has led to some of the fastest-growing technology and product adoption in the history of the software industry. We may be experiencing what is likely the start of a new supercycle on par with the advent of the microchi or the internet.

For one reason, accuracy isn't that important for many applications. When a model is generating novel images or engaging in entertaining banter, being correct simply means appealing to or engaging the user. In other popular uses, like helping developers write code, th user is the human in the loop—iterating and providing the feedback to improve the generated answers.

Another big reason things are different now is that generative AI is facilitating uses, from companionship to therapeutic art communities, previously impossible for computers in any meaningful way. We don't really have a good understanding of what the behaviors will lead to, nor what the best products are to fulfill them. Amazingly, while the use cases for these new behaviors are still emerging, millions of users have already shown a willingness to pay.

This all means opportunity for the new class of generative AI startups to evolve along with users, while incumbents focus on applying the technology to their existing cash-cow business lines.

Generative AI can bring real economic benefits to large industries with established and expensive workloads. Large language models could save costs by performing tasks such as summarizing discovery documents without replacing attorneys, to take one example. And there

are plenty of similar jobs spread across fields like medicine, computer programming, design ar entertainment.

Consider the task of creating an image to use for marketing content or for a movie poster. For companies running their own version of an open-source model like Stable Diffusion, it costs roughly 0.1 cent and takes around one second to generate an image. Hiring a graphic designe or a photographer would cost hundreds of dollars and take hours or days. Even if, for simplicity's sake, we underestimate the cost at $100 and the time at one hour, generative AI is 1/100,000th the cost and 3,600 times the speed of the human alternative.

For generative AI to remake our economies and lives to the same degree does assume a continued pace of innovation, but many experts believe we're very likely to see continued progress. There might be growing pains like the Hollywood strikes along the way, but the end result is more jobs, more economic expansion and better goods for consumers. This was the case with the microchip and the internet, and it will be with generative AI, too.

Mr. Casado is a general partner at Andreessen Horowitz.