

# Navigate Generative AI



**Author: Jen Zhu Scott, Founding Partner, IN. Capital**

**Researcher: Yong Jin Lee, Analyst, IN. Capital**

**March 2023**

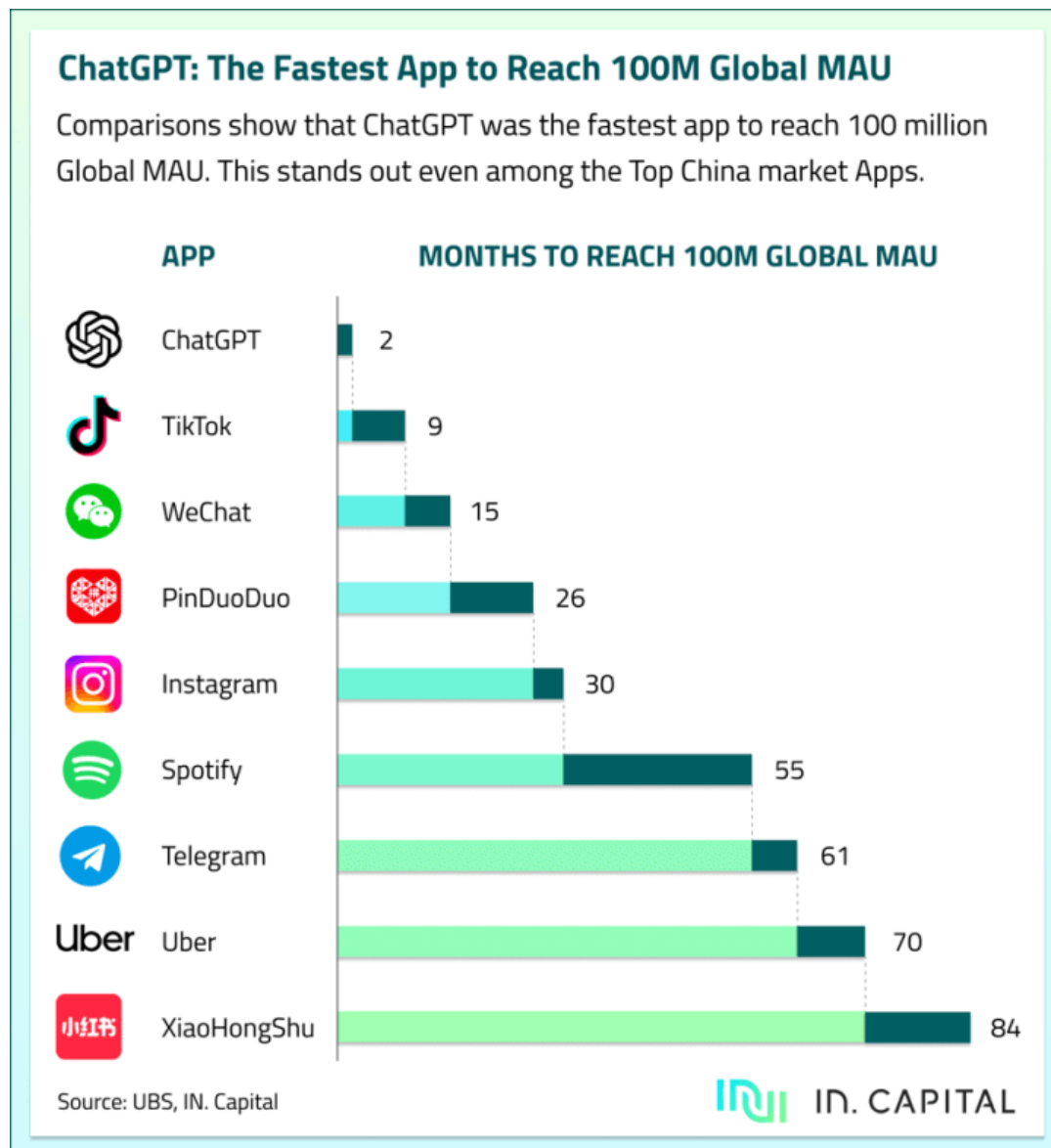
We are living in a moment of history when productivity upgrade is as profound as the invention of the computer and the Internet. Except for this time, the speed of adoption and scaling is exponential and more breathtaking than all the previous technological turning points combined.

Driven by the recent breakthroughs of Generative Artificial Intelligence, we are at the beginning of a cognitive revolution that is poised to transform many aspects of our work and life.

AI's multi-decade journey hasn't always been this exciting. From the 1960s to 2010, the progress of AI was linear. From 2010 to 2020, we witnessed transformative milestones highlighted by AlphaGo beating the human Go Game champion (then itself soon after) for the first time in 2016 (*Reinforcement Learning*) and the autonomous vehicles in 2018. In the past few months, Generative AIs (Gen AI) led by image generators such as MidJourney and chatbots such as OpenAI's ChatGPT have created applications that are attracting lightspeed adoptions. These user-friendly and affordable productivity superchargers are empowering millions of early adaptors to do more, better, faster, but cheaper.

The catalyst of the Gen AI wave is OpenAI's ChatGPT. Launched on 30 November 2022, ChatGPT (stands for Generative Pretrained Transformer) quickly established itself as the world's most dominant chatbot. It's not only caught users' imaginations as a tool, its powerful, human-like natural language capability has also inspired much fascination and sometimes, fear, as it is occasionally regarded as a living entity. ChatGPT gained 1 million users in 5 days and 100 million in just two months. This broke TikTok's record of 9 months and Instagram's previous record of 2.5 years.

So what exactly is Gen AI?



## Parrots and Crows

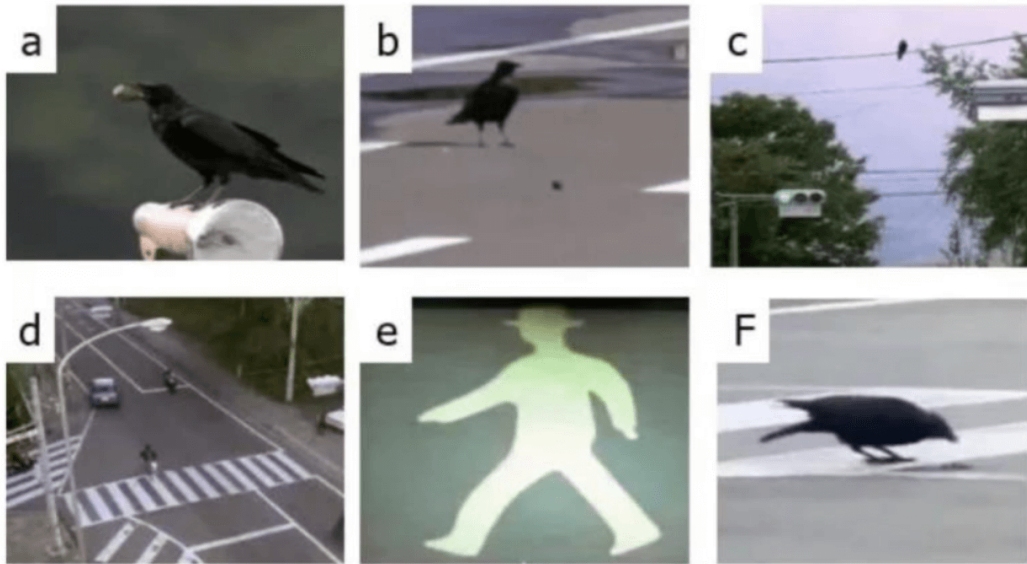
I asked the world's first native Gen AI search engine [You.com's](https://you.com/) search bot, YouChat, this question. Here is the answer to the definition of Gen AI by a Gen AI:

*"Generative AI is a form of Artificial Intelligence that uses deep learning and Generative Adversarial Networks (GANs) to generate new content from existing data. It is used for a wide range of applications, from creating new images and music to generating text and providing more accurate search results. Generative AI models are trained on a large dataset and can learn to recognize patterns and create new content based on the data they have been given."*

To put into simple words, Gen AI uses large datasets to recognize the contextual patterns and mimic the 'creation' process to generate new text, images, music, coding, audio, or videos. Such capability is also extended into translating one form of content to another (*speech-to-text, text-to-image, etc*). But AI has always been about processing and analysing large datasets for pattern recognition. In fact, I helped to lead a speech-to-text, deep indexing project way back in 2009. Why is Gen AI so different?

To put into plain words, previously AI are mainly parrot-like 'single domain optimizers' or at times, fancier linear regression models, missing the 'creative' layer. Machine Learning and Deep Learning models are ultimately based on  $Y = f(X)$  with some predetermined optimizer targets. Even if it didn't know  $X$ , based on the patterns the models can figure out what  $Y$  is or what  $Y$  could be. But Computer Vision with such an approach, a model that is trained with pictures of cats and dogs would not be able to recognise an airplane. A parrot can repeat what you say, but cannot 'make sense' of what you say and come up with its own content.

Renowned AI researcher [Professor Zhu Songchun](#) of UCLA uses crows' behaviour to explain the next generation of AI. A crow, as a wild animal, found a nut in a city. It does not know how to open the nut and nobody is there to teach it. The crow observed that the cars on the road can crush the nut, but it is too dangerous to eat the nut with cars driving past constantly. It also learned that the stripy part of the road with traffic lights has cars driving pass but stop when the red light comes on. So with the objective of 'crush and eat the nut', the crow learns two things that seem to be unrelated but can help it to achieve the goal: a. cars can crush the nut; b. traffic lights in front of the stripy part of the road can control and stop cars.



Source: [https://mp.weixin.qq.com/s/-wSYLu-XvOrsST8\\_KEUa-Q](https://mp.weixin.qq.com/s/-wSYLu-XvOrsST8_KEUa-Q)

Gen AI uses massive natural language, image, and other media datasets, analyses them with pre-determined parameters effectively, recognises the patterns, 'makes sense' of different patterns and the relations in between, generates fresh content in same media, or even fresh content in different media, and produces often highly *useable* results. An average user can ask ChatGPT to create a diet plan that makes sense, YouWrite to write an essay about ESG investing strategy that is plausible, and MidJourney to create a Banksy-style art with an astronaut landing in the war zone in Ukraine by only typing in these words. The human-machine co-creations have never been this intuitive, dynamic, and integrated.

To summarise, the previous Machine Learning models required 'feeding' in order to 'imitate' (parrots), Gen AI models required 'teaching' in order to 'make sense' and 'understand' as the internal logic (the crow).

Such models built using massive datasets are Large Models. LMs have a growing variation and each is specialized in a certain modality - a shorthand phrase for "a particular method or procedure" often categorized by the content the AI creates. For example, AI for image or AI for text. Dall-E and Stable Diffusion specialize in generating images (*modality: image*). GPT-3 specializes in generating texts (*modality: text*), and Whisper AI excels in voice recognition (*modality: voice, text*), converting voice to text.

Building on top of these Large Models, applications further fine-tune the underlying models to target a set of end-users with specific scenarios. Zeemo, for example, utilizes the underlying voice recognition and translation models to generate multilingual subtitles in real time. Ultimately, the interoperability between different modalities opens up possibilities for multi-modal applications where boundaries do not exist for converting text, image, or voice. This is changing how humans interact with machines, and in return, how machines learn about humans.

Excitements aside, Gen AI isn't without limitations. One scientific understanding is fundamental to appreciate the limitations of the Large Models approach: when we study infinitely complex phenomena such as human intelligence, it is impossible to capture every last parameter and bit of context. So we create simplified models (with restricted perimeters) to capture the key elements while maintaining full control of the system. Therefore, the final results are at the mercy of what has been kept and what has been given up in the predetermined parameters. This is why Gen AI produces such familiar, plausible results that are also often so ridiculously wrong. When a non-human intelligence produces human or super-human level results while being wrong on some most basic things to the human eyes, it is understandable why such strange beings could provoke controversies and human fear.

As this article goes live, OpenAI just released ChatGPT API at 10% of the cost but at 5 times the speed compared to the previous version. The underlying capabilities combined with such efficiency improvement at orders of magnitude will create a tsunami of applications that will change our work and life. We anticipate the efficiency will continue to grow while the cost will continue to reduce in the coming years. Gen AI led by OpenAI's latest release is an undeniable productivity revolution.

## The Implications

In the coming 12 months, Gen AI tools and capabilities will be embedded everywhere, especially in aspects of work and life that require content generation, including image, text, audio, video, code, simulations, or a translation from one format to another. If you are an entrepreneur, Gen AI can help you test a business idea,

create a logo, generate content for your website and social media, create your website, pitch deck, business plans, and legal documents, and summarize interviews or any digital content, including podcasts. If you are an artist, Gen AI can help you write lyrics, compose music, and create still or dynamic images. If you are a trader, Gen AI can help you analyse stocks, create portfolios, dynamically compare your portfolio to other funds, and manage your portfolio.

Our own imagination will be the biggest barrier to adopt and utilise Gen AI tools. The sudden unleashing of productivity gains at an affordable cost means Gen AI will make us rethink how we organize our work and life in the coming years. Our calculation of opportunity costs changes when so much time and costs can be saved by having AI to conduct complex, repetitive, and time-consuming tasks.

However, many of such tasks still require human-in-the-loop (HITL) to achieve the optimized result. Think machine plus or multiplied by human or vice versa, instead of machine vs. human. AI will not be fully replacing humans yet, but individuals who understand how to utilise AI will gain a significant edge against other peers. Employable skills in the coming years will be drastically different, and new professions will emerge from this era.

End-users need to adopt new mindsets and tools to remain competitive. Parents and educators need to rethink how best to prepare the next generation. Shifts of this magnitude also suggest that there will be significant value generated by companies that master the art of Gen AI, while rendering slow adapters obsolete.

## Navigate Gen AI

There are hundreds of start-ups that have already mushroomed in the past few months, offering intuitive tools for everyday tasks. We anticipate many more will show up at a dizzying pace. I recommend a simplified framework to help investors to navigate this highly dynamic space.

### **Application Stack**

- Applications with proprietary models

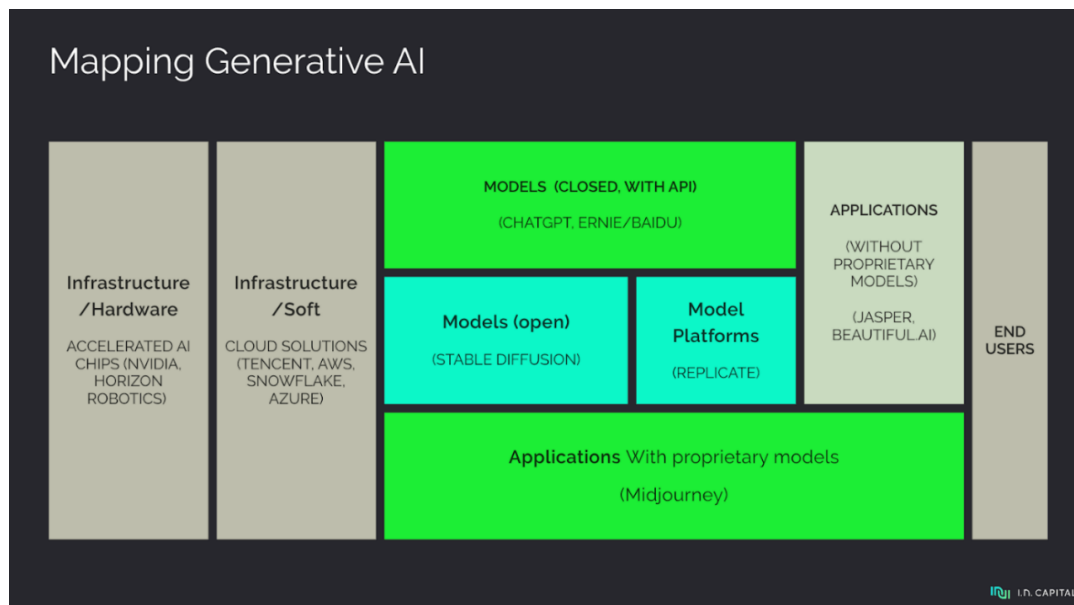
- Midjourney
- Applications without proprietary models
  - Jasper, Beautiful.ai

## Models Stack

- Closed-source large models with API
  - ChatGPT, Ernie from Baidu
- Open-source large models
  - Stable Diffusion, Stability
- Model Platforms and Hubs
  - Replicate

## AI Compute Infrastructure

- Soft infrastructure (Cloud)
- Hard infrastructure (Processor Chips)



Source: IN. Capital

**Applications without proprietary models** rely upon 3rd party AI models despite limited configurability either out of necessity or by choice. Training a model is costly, and if the application's end product can be fully enabled by the existing AI solutions, the application can benefit from the ease of development offered by the 3rd party AI models.

Without proprietary models or underlying tech, applications can only stay ahead if it can outperform competitors on market adoption. This requires relentless differentiation and retention.



Given the relatively lower barrier to entry, the constant competition will make the companies in this layer rather risky. Even for companies that have shown impressive initial momentum, their defensibility isn't guaranteed as competition heats up.

On the other hand, such competition is highly beneficial to the end users. We should leverage all the cheap or free AI applications to become fluent AI tool users early. The IN. Capital team has compiled a list of current AI tools [here](#) and we aim to update it continuously to help our followers to navigate the space, including the applications in the English speaking universe as well as those from China, which might be less visible for many international investors and observers.

**The soft and hard infrastructure segments** are dominated by the large established players where we see little venture opportunities currently.

In our view, the most defensible opportunities are within the mid-layers, where **applications with proprietary models, close-sourced models**, and **model 'marketplaces'** occupy. Apart from the projects that are now well known by the global audience, Chinese players such as Baidu have also entered the mid-layer space.

Baidu began its development of a Gen AI chatbot, Ernie, in 2019 and plans to integrate in March this year. Interestingly, Ernie is built based on Baidu's own Deep Learning framework, PaddlePaddle, heavily contributing to the early stages of AI developments. We carefully predict that Ernie will likely lead China's Gen AI space.

It is worth highlighting that **applications with proprietary models**, such as MidJourney, naturally have stronger defensibility but face a unique challenge compared to other players in the mid-layer. They will have to reconcile between low or no cost products to collect users and data to feed to their models and having an overall commercially viable business model to become investable. Companies like Baidu, though, could use its legacy advertising/search engine model to subsidize the data collection.

Overall, it is important to caution that we are still in the early days of the Gen AI hype cycle. A clear, systematic pattern has not fully



emerged yet. Given its versatility, multitudes of productivity gains, and astonishing cost reduction, Gen AI has the potential to impact an extremely wide range of human activities. It would therefore be every investor's responsibility to understand the space and watch it closely.

## Lost in Translations

As mentioned above, Gen AI's foundation is Large Models, including Large Language Models (LLMs). LLMs are good examples to examine Gen AI's limitations. Human language patterns have infinite factors that contribute to the final result of the content and the implications of such content. Cultural context and nuanced human emotions are impossible to be exhaustively included in LLMs predetermined parameters. LLMs absorb the text content available online or other forms within the models and generate new content, but only based on the parameter pre-specified by LLMs. We humans with similar parameters often miscommunicate and misunderstand each other. Imagine someone using a drastically simplified framework to understand what you say. If 'lost in translation' is common between human brains, mistakes and limitations with Gen AI are a feature, not a bug at this point, and for the foreseeable future.

Take a few top applications (ChatGPT, Stable Diffusion) as examples, the errors and limitations are as 'impressive' as their capabilities:

- **Math:** ChatGPT is notoriously bad at math. Computers Arithmetic Logic Unit (ALU) is extremely good at arithmetic operations but LLMs are not ALU. LLMs process natural languages and recognise language patterns, which does not include arithmetic logic. ChatGPT's math answers often sound like a polished speaker who is convinced he/she is good at

math but without the most basic math skills.



If an orchestra of 30 takes 120 minutes to play a concerto, how many minutes would it take for an orchestra of 60 to play the same concerto?



If an orchestra of 30 musicians takes 120 minutes to play a concerto, then each musician plays for  $120/30 = 4$  minutes. If an orchestra of 60 musicians were to play the same concerto, then each musician would play for the same amount of time, or 4 minutes. Therefore, it would take 60 musicians a total of  $60 * 4 = 240$  minutes to play the concerto.

- Using Stable Diffusion to generate an image with my prompt “6 hands holding a globe”, these are the results – none of them produced results with correct number of fingers:



- **Plausible incorrect content:** ChatGPT or Gen AI based bots are proficient at producing plausible answers that sound convincing but factually incorrect. This is also because the datasets for the LLMs are mixed with both true and fake information. There is no fact check functionality within AI generated content yet. If we think the last ten years the internet was full of fake news, we certainly have not seen anything yet.
- **Privacy:** many LLMs use online contents. Depending on your privacy agreement with each platform you use - social media, email, messaging apps - your private content could be used as the raw material for Gen AI to analyze and create new content. For ChatGPT, long term content saving is at least 30%. Many users who are acutely aware of privacy issues have started to advocate limiting interactions with ChatGPT with sensitive personal information. Gen AI will become a privacy nightmare and the regulators will have a hard time to play a catch-up game.
- **ESG concerns:** the carbon emission of Gen AI is unclear at this point but worth examining. The implicit biases of Gen AI could be a risk for corporates that want to do a better job on the social and gender equality fronts.

I am certain there will be new concerns emerging in the coming years that we are ill prepared for right now. At IN. Capital, we will continue to learn and share knowledge in this space. In this way, we can empower our investors to capture the monumental investment opportunities that are coming in this space while being responsible for what we are enabling.

## Useful Readings

Please feel free to explore our compiled list of useful articles in the following link.

<https://i-n-cap.notion.site/Useful-Readings-07cf90152a564a73b7ce09fe4db8c21b>

## List of AI Tools & Applications

<https://i-n-cap.notion.site/IN-sights-Series-1-ChatGPT-AI-4a74d7f75ec74f6fa5146f2714b085e5>

## Glossary of Key Jargons

<https://i-n-cap.notion.site/AI-Glossary-60de821b92ab45d284226a58d9ea5a72>

*Are you building the next great AI company? Would you like to learn more about investing in AI with us? Contact us [here](#).*



Disclaimer: The content and functionality on the Website are provided with the understanding that I.N. Capital is not herein engaged in rendering professional advice or services to you. No content on the Website is intended to serve as or shall be deemed investment, legal, tax, accounting, or other regulated advice, and that you shall remain solely responsible for your use of ALL Website content and acknowledge that any reliance upon the content shall be entirely at your sole will and risk. All content and functionality on the Website is provided "As Is," without warranty of any kind, including, without limitation, implied warranties of merchantability and fitness for a particular purpose. I.N. Capital and its third-party content providers make no warranties, as to the ownership, accuracy, or adequacy of the content on the Website. I.N. Capital shall have no liability or responsibility for any information published on the Website or any linked websites. Neither I.N. Capital nor its third-party content providers shall be liable for any indirect, incidental, consequential, or punitive damages or losses.