DEC 16, 2022

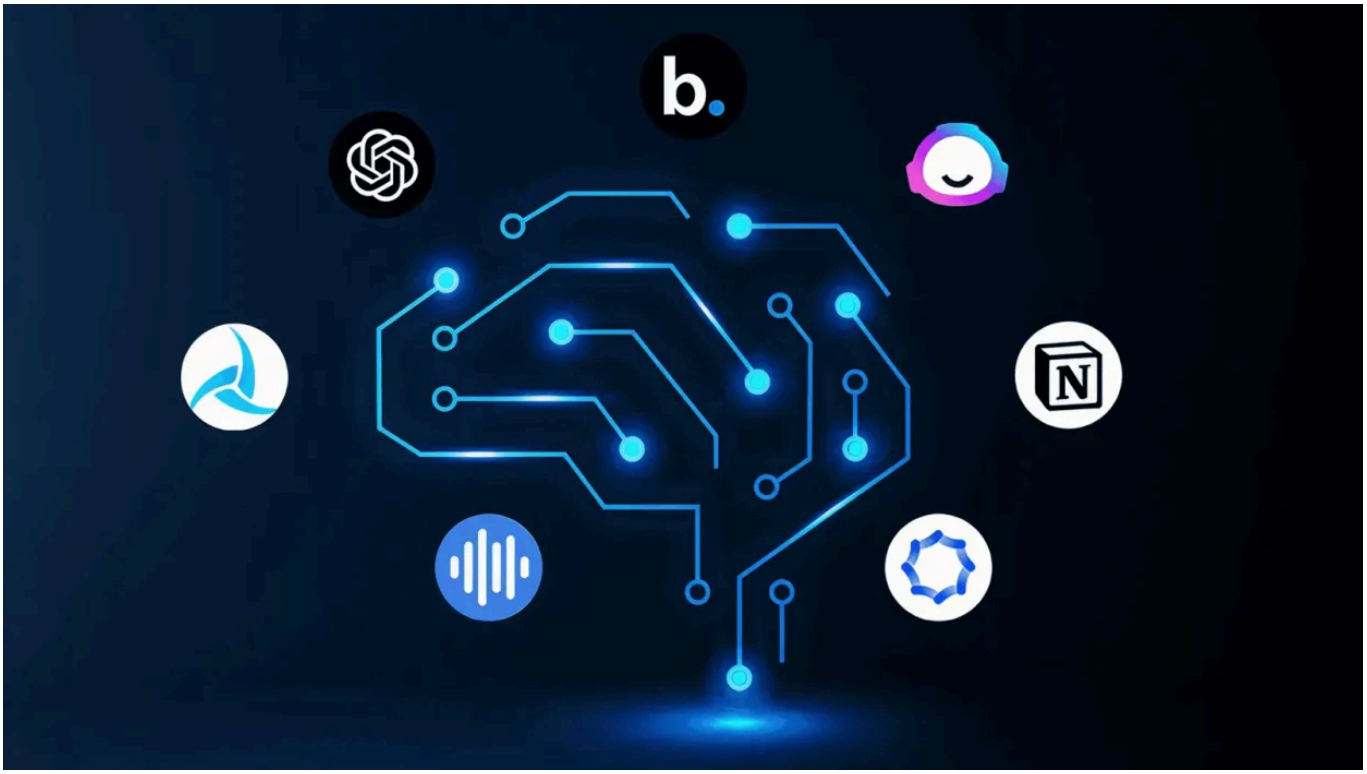
Generative AI: State of the Market, Trends and Startup Opportunities

Our take on the latest industry trends and advice for startups.



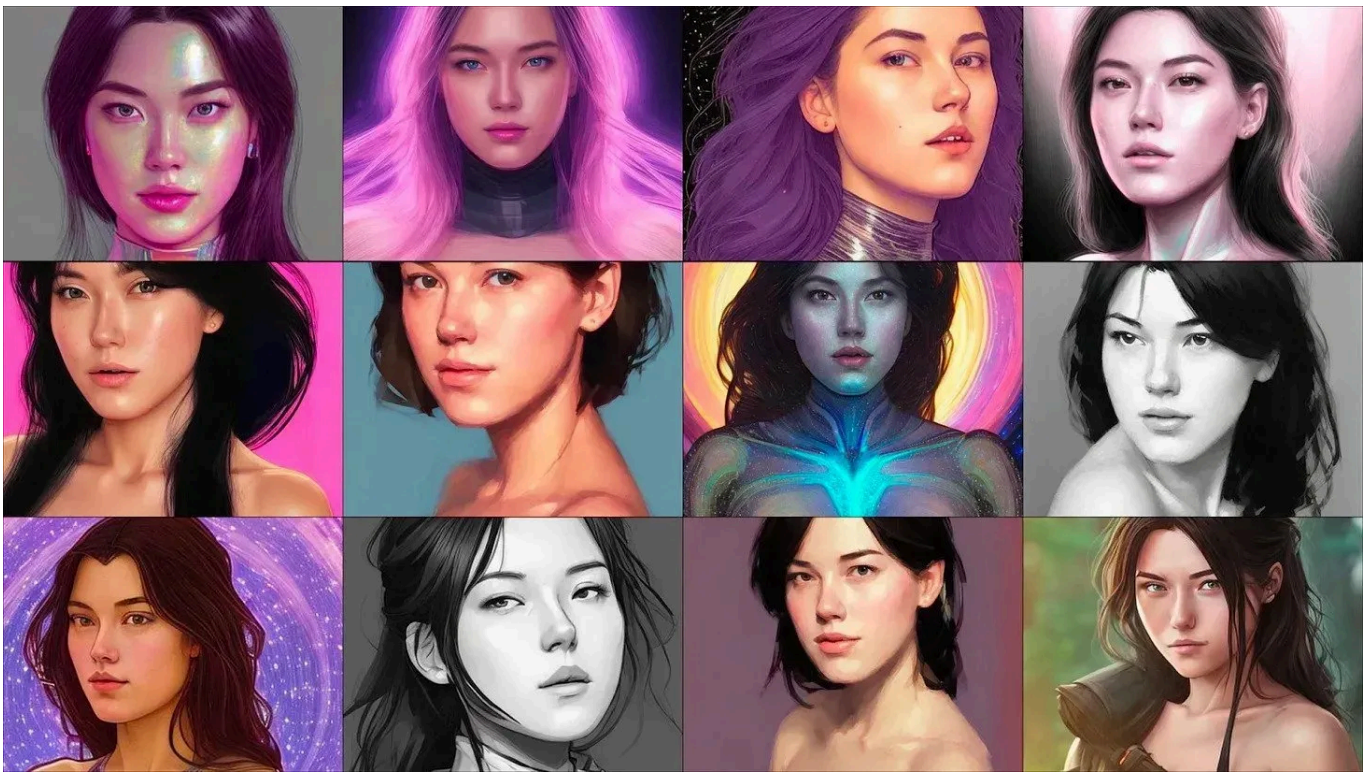
By Jenny Xiao and Jay Zhao

Leonis Capital · leoniscap.com/research/generative-ai-state-of-the-market-trends-and-startup-opportunities/



From Stable Diffusion to ChatGPT, generative AI models have become the spotlight of Silicon Valley. VCs are rushing to invest in AI start-ups as the growing hype around generative AI fills the void left by their previous web3 and crypto ventures. The recent leaps in AI technology have empowered apps that can write scripts and generate art in a matter of seconds. This has created a rare exception in a tech landscape overshadowed by plummeting valuations, job cuts, and web3 pessimism. [1]

But what exactly is generative AI? And what sort of opportunities does it present for entrepreneurs and investors?



If your social media feed is filled with these AI-generated selfies, you are not alone. Image source: MIT Technology Review

Generative applications are driven by technological breakthroughs in large-scale pre-trained models, or "foundation models." These models are distinct from previous-generation AI models in that they have many more parameters, perform better on a wide range of tasks (e.g., text and image generation), and exhibit new capabilities – even going as far as video production! **Understanding the underlying technology is critical to understanding the success and failure of startups in this space.**

In this post, we want to walk you through the current generative AI landscape -- the applications, use cases, and players in the field. We also want to give you a little bit of background about foundation models, so you can see the potential of this technology. Finally, we offer an assessment of what we think are overhyped and underrated opportunities in this fast-moving area.

The Current Generative AI Landscape

Generative AI Market Map

LEONIS CAPITAL

RESEARCH

Model Infrastructure

OpenAI · stability.ai · cohere · ANTHROP\C · AI21labs · Hugging Face · generally intelligent · RAIN · craiyon · Midjourney · DeepMind · Inflection · Adept · CONJECTURE · EleutherAI · LightOn · CROMA

TEXT

Copyediting/Marketing

Jasper · copy.ai · Writesonic · Rytr · anyword · INK · Bertha.ai · sudowrite · wordtune · Hypotenuse AI · regie.ai · CONTENDA · OTHERSIDE AI · WRITER · HyperWrite · COMPOSE AI · LEX · LAIKA · Moonbeam · frase · copysmith · Clickable · letterdrop · Peppertype.ai · Simplified · Omneky

Chatbots

Character.AI · inworld · heyday · chai · Convai · alethea · gamescale · Replika · circle labs · Quickchat

Sales/Customer Support

Outplay · Smartwriter.ai · Cohere · Twain · Creatext · Typewise · XOXind · CRESTA · LAVENDER · KAIZAN

Knowledge Organization

mem · glean · Rewind · ought · Omniscience · BearlyAI · metaphor

Search Engine

YOU · Andi

Education

Speak

Legal

Harvey · PatentPal

Narration

NovelAI · AI DUNGEON · Subtext

Productivity

motion* · reclaimai

Models: ChatGPT (OpenAI Dec 2022), ACT-1 (Adept Sep 2022), Bloom (Hugging Face Aug 2022), Jurassic-X (AI21 Labs May 2022), PaLM (Google Mar 2022), Chinchilla (DeepMind Mar 2022), GPT-J (EleutherAI May 2021), GPT-3 (OpenAI June 2020)

IMAGE

Image Generation

OpenAI · stability.ai · Midjourney · ROSEBUD.AI · OpenArt · Bertha.ai · Jasper · mage.space · KREA · Picsart · playground · artbreeder · SPELLBRUSH · PhotoRoom · Aragon

Image Database

lexica · Nyx · gallery · unsplash · ROSEBUD.AI

Selfies/Avatars

AVATAR AI · PRISMA LABS · ProfilePicture.AI · OASIS

3D Images

KAEDIM · mirage · blackshark.ai · HYPOTHETIC · CSM · Spline*

Design

Diagram · alpaca · Poly · VIZCOM · uizard

Fashion

LALALAND · CALA

Consumer

INTERIOR AI · renoval

Material

BariumA · ARMORLAB

Models: Midjourney (July 2022), Stable Diffusion (StabilityAI Apr 2022), DALLE-2 (OpenAI Apr 2022)

VIDEO

Video Generation/Editing

runway · WOMBO · descript · OPUS · ROSEBUD.AI

Digital Humans

synthesia · Rephrase.ai · Movio · Colossyan · Hour One. · Synths Video · elai.

Avatars/Motion Capture

Kinetix · DEEPMOTION · RADICAL STUDIO · plask

Personalized Videos

tavus · Rephrase.ai

Media

WONDER DYNAMICS

Models: Make-A-Video (Meta Sep 2022), Imagen Video (Google Oct 2022)

AUDIO

Synthetic Voice

SONANTIC · VOICEMOD · PolyAI · REPLICA · coqui · WELLSAID · readspeaker.ai · Fliki · podcast.ai · Listnr · RESEMBLE.AI · Dubverse

Music

soundful · USICO · SONIFY · Mubert · Aiva · SPLASH · Harmonai · Endel · INFINITE ALBUM · boomy

Models: Whisper (OpenAI Nov 2022), data2vec (MetaAI Jan 2022), Jukebox (OpenAI Apr 2020)

CODE

Coding Assitant

replit · tabnine · cogram · GitHub Copilot · MUTABLE AI

Platforms

sematic* · Hugging Face

Documentation

Mintlify · Stenography

Web App

Debuild · durable

Data Analysis

seek · Maya AI · kubit*

Models: CodeGeeX (Tsinghua University Sep 2022), Codex (OpenAI July 2021)

GAMING

NPCs/Avatars

REPLICA · rct.ai* · ROSEBUD.AI · charisma.ai · Kinetix · circle labs · alethea · DEEPMOTION · RADICAL STUDIO · plask · Replika · VOICEMOD · coqui

Game Platforms

latitude · The Culture DAO · REGRESSION

Game Assets

mirage · KAEDIM

Narration

hidden door · AI DUNGEON

Worldbuilding

PROMETHEAN AI · mixAR · blackshark.ai

Background Music

soundful · Aiva · Mubert · SPLASH · INFINITE ALBUM

Models: ChatGPT (OpenAI Dec 2022), Imagen Video (Google Oct 2022), Midjourney (July 2022), Stable Diffusion (StabilityAI Apr 2022), Whisper (OpenAI Nov 2022)

BIOTECH

Biomedicine

Generate:Biomedicines · Antiverse · Ainnocence · JURA Bio

Protein Design

evozyne · menten AI · Cradle.bio · ProteinQure · HELIOPOLIS BIOTECH

Models: RoseTTAFold Diffusion (UWashington Nov 2022), AlphaFold (DeepMind July 2021)

Companies marked with asterisks (*) are ones that Leonis Capital has invested in.

Since the GPT-3 API was made public in September 2020, there has been an explosion of generative AI companies, covering areas such as copyediting, marketing, sales, and knowledge organization. The release of image generation models like Stable Diffusion (April 2022), DALLE-2 (April 2022), and Midjourney (July 2022) further fueled the growth of startups and applications in the visual design and image creation space. (We included links to their demos so that you can try these models out yourself!)

DeepMind's decision to open-source its AlphaFold model also democratized access to protein-folding technology and enabled scientists to build biomedical startups using targeted protein design. The biotech startup Generate Biomedicines released an AI program called Chroma, which the company calls "the DALLE-2 of biology."

The current generative AI startup landscape is driven by the democratization of foundation models – either through APIs or open-source models. This means that a key characteristic of generative AI startups is that they **face fierce competition**, as other developers have access to the same underlying models. Even relatively established companies in this arena don't enjoy a significant technological, product, or data moat but need to constantly innovate to keep up with the release of newer models.

But **the centralization of power at the foundation layer also creates vulnerabilities for application-layer startups**. This is why many application-layer companies are eager to develop their own models. Jasper, for instance, is already training its own models on Cerebras supercomputers. In this way, it can reduce reliance on OpenAI's models, better fine-tune the models for their specific use cases, and keep the data generated by their models.

At the same time, **more generally-capable AI models will likely undermine previous vertical applications**. It's not hard imaging ChatGPT (also known as GPT 3.5) outperforming specialized marketing AI models, such as Lavendar.ai or Smartwriter.ai, many of which are built on a finetuned version of GPT-3. A key trend in the foundation model revolution is that newer models typically perform even better than specialized models. Startups on the application layer will likely iterate between using more powerful general models and building their own vertical models.

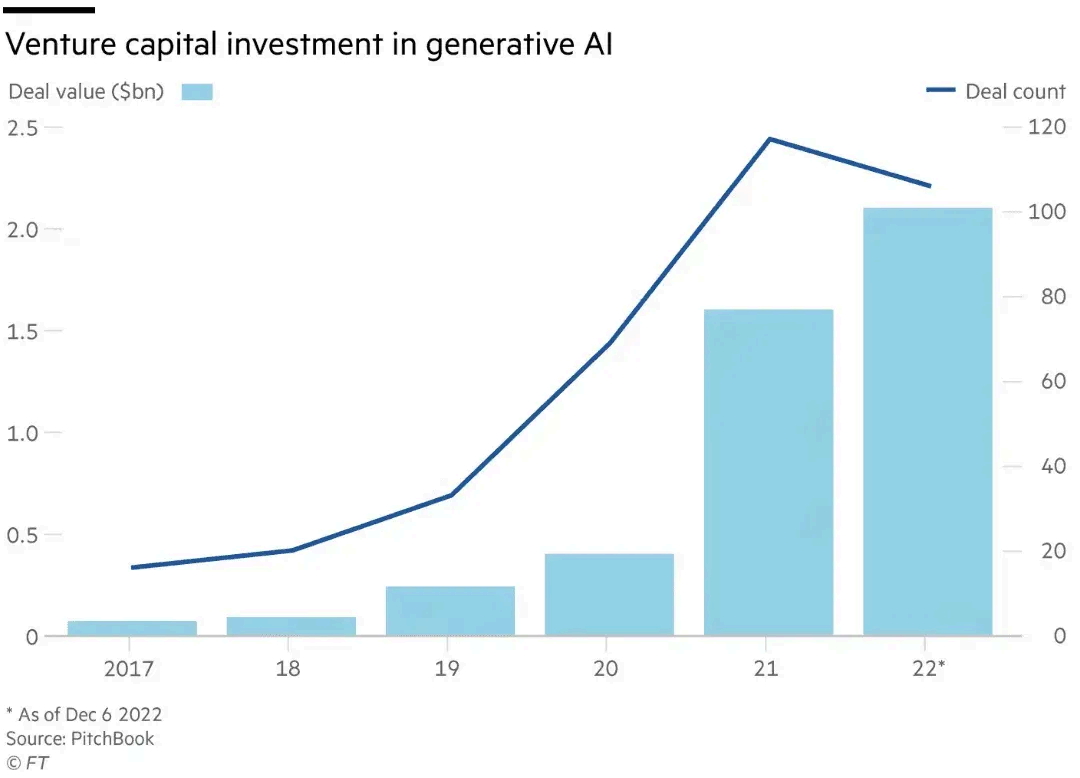
We expect to see constant yet healthy tension between Infra providers' general models and application startups' finer-tuned vertical ones. Given text AI is the area that has been researched and invested in the most, the dynamics seem to be the most fluid and fast-changing.

In other areas, such as video, audio, and code generation, although there aren't open-source or API-based models readily available, startups have managed to build their own AI models using the same foundation model architecture as GPT-3 and Stable Diffusion. The video-generation

startup Rephrase.ai built a proprietary AI model that maps text to voices and videos, enabling marketing teams to easily create hyper-personalized ad videos.

At Leonis Capital, we are optimistic about seeing how AI is boosting productivity in other media formats beyond text, although it is remained to be seen how feasible new entrants can come to the space with a long-lasting moat.

Another factor that makes the generative AI landscape ultra-competitive is that the technology **has become a consensus almost as soon as it took off**. Typically, major tech revolutions evolve slowly because most people are skeptical at first, great examples include the PC and smartphone revolutions. But just as GPT-3 was released two years ago, VC investment in generative AI has increased more than 400% since 2020, amounting to a staggering $2.1 billion this year.



This is probably because generative AI is **creating (pseudo) new market categories by providing aggressive ROI value propositions.** [2] Companies like copy.ai or Jasper are augmenting copying writing, marketing, and sales - with high software scalability, repeatability, and AI-powered low cost. Similarly, Midjourney and Stable Diffusion are amplifying the speed of media / art / entertainment creation while Mutable.ai and Github Co-pilot are helping programmers increase productivity by an order of magnitude. These types of adoption (and frankly buzz) were unseen before, which triggered the VC community to rush for the “next hot thing”. [3]

Because of that, **investors are also paying for the hype**. Recently, Coatue and Lightspeed Ventures led a 101 million seed (!) round for StabilityAI, the company behind the popular Stable Diffusion model, essentially valuating the company at over $1 billion post-money. Whether this valuation is justified remains to be seen. But the fact is that StabilityAI was just an open-source program at the time VC capital was injected.

Importantly, **not all "generative AI" companies use state-of-the-art generative AI models**. As a result, these applications tend to be a lot less "impressive" than those that do. Motion-capture startups, for example, are not technically using "generative AI," and many video generation companies do not use the text-to-video generation empowered by DALLE-like diffusion models. These non-generative companies are included in the map because they are ripe for disruption by newer models.

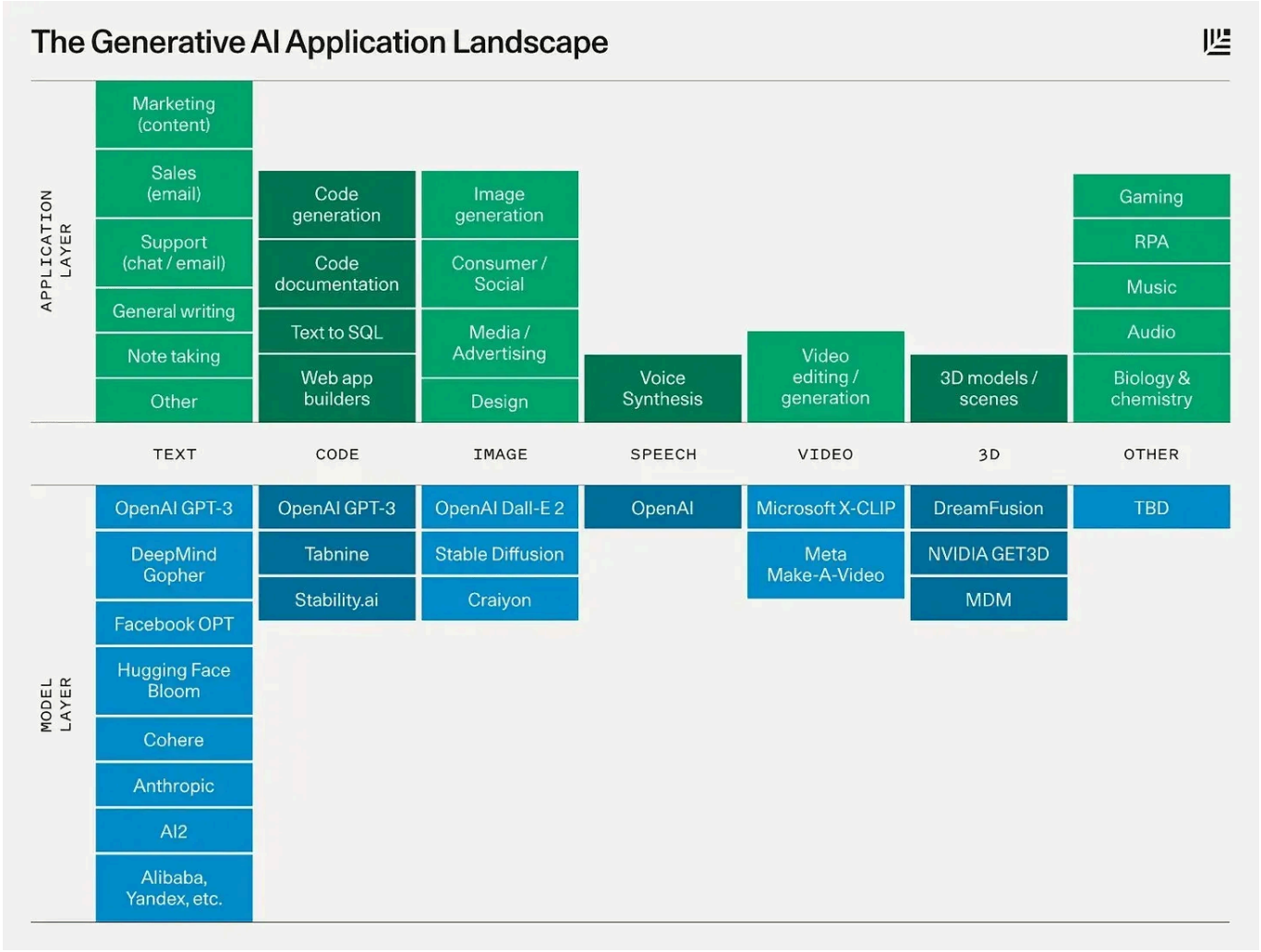
Interestingly, however, these companies might also benefit from the generative AI hype because investors tend to **lump them into the "generative AI" bucket**. However, they are unlikely to capture the true value created by the foundation model revolution unless they innovate their underlying technology.

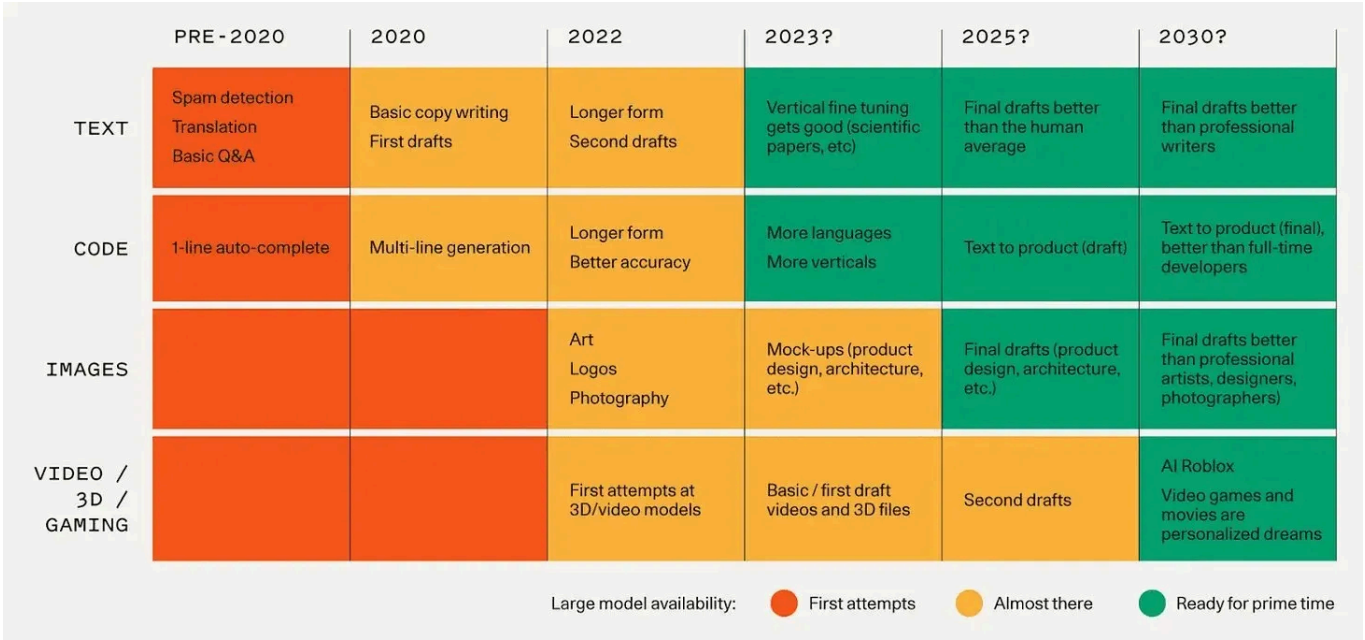
Technology Trends: Limitations, Misconceptions, and Just How Good Will Generative AI Be?

In the nascent field of generative AI, **the underlying technology determines what products can be built**.

The release of GPT-3 two years ago fostered the creation of text-generation startups, such as Jasper.ai and copy.ai. Now, the release of image- and code-generation models provides the foundation for new marketing, sales, design, and coding assistant apps. Following the success of the AI avatar app, Lensa.ai, a new wave of startups is building AI image-generation apps.

But just how good is the underlying technology for each type of application? In the next few paragraphs, we want to give you an overview of the model layer without overwhelming you with the technical specifics.

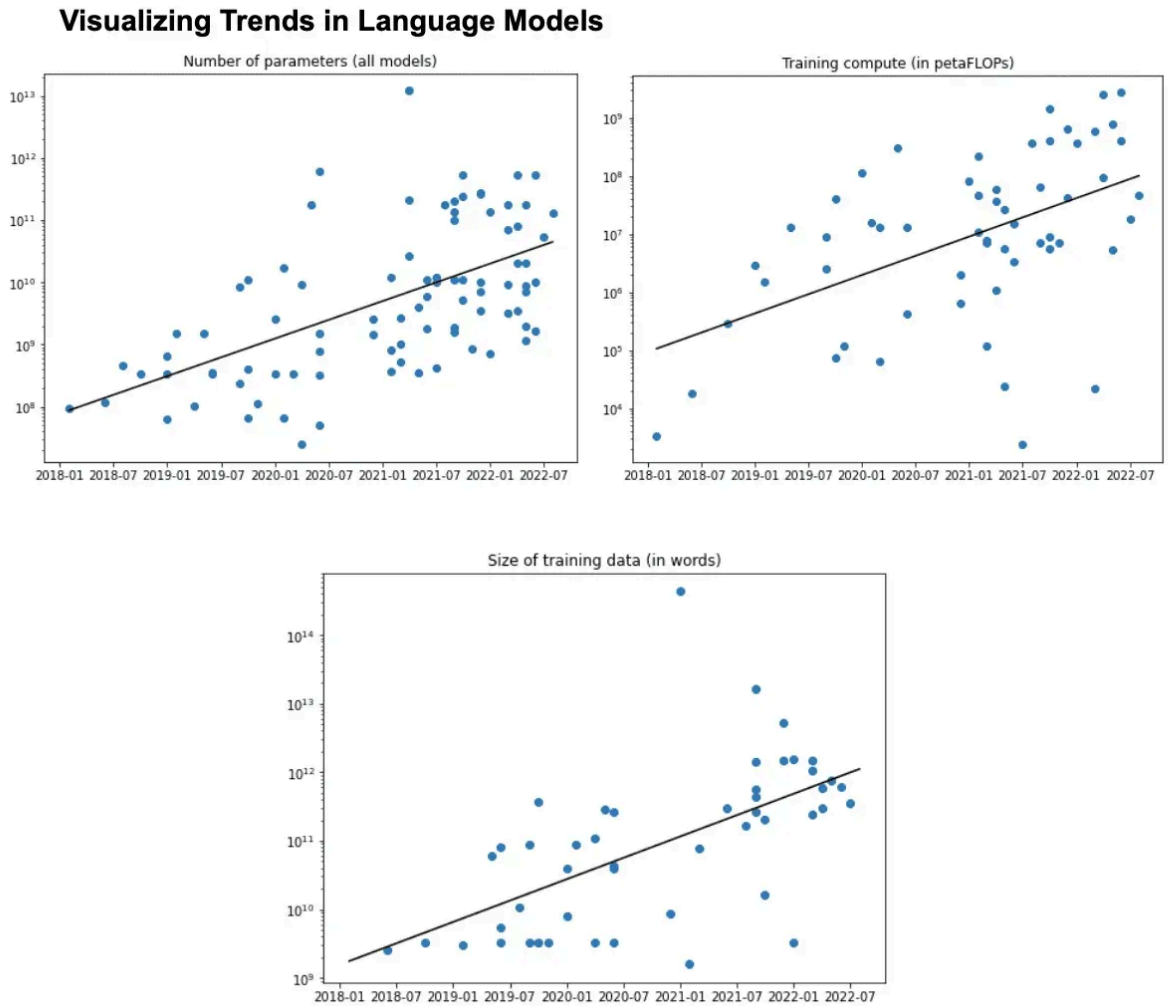




Source: Sequoia Capital Generative AI: A Creative New World (2022)

Text models are the most mature type of generative AI model, and they are also the earliest models to be developed. There are also many more text models available than any other type of generative AI model and there are also more available APIs and open-source models. Aside from well-known labs like OpenAI and DeepMind, newer entrants are also contributing to the AI language model infrastructure layer, including the Israeli AI lab AI21 and the Canadian startup Cohere.

We created this visualization of all large-scale language models (LLMs) released since 2018 (there are many of them!) The **quick conclusion** is that these models are becoming bigger and more compute- and data-intensive at an exponential rate. Foundation model "scaling laws" predict that **model capability will increase with model size**.



The release of popular models like DALLE-2, Stable Diffusion, and Midjourney has brought image-generation models to the public's attention. We are used to seeing these impressive

artworks generated by AI, such as the now-iconic horse-riding astronaut image generated by DALLE-2 or the impressively detailed paintings created by Midjourney.



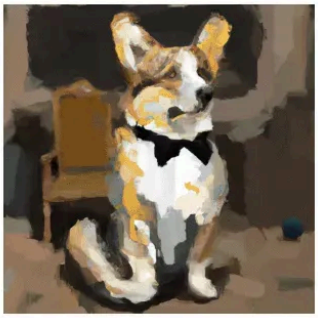
"An astronaut riding a horse in a photorealistic style," generated by DALLE-2

However, the fancy images that we see online are not representative of all AI-generated visuals. **Image-generation AI models still suffer from controllability issues and struggle to respond to human commands.** They will often miss critical information in human instructions, creating a technical barrier to broader commercial applications.

Here is a small (but fun) experiment with three state-of-the-art AI image-generation models.

The first image generated from each prompt is selected for the sample. We can see that DALLE-2 and Stable Diffusion 2.0 exhibit similar levels or responsiveness to human commands (e.g., generating a *realistic* image of a cat or a corgi in Dalí's style). As a smaller model trained on a more specialized dataset, Midjourney does a good job at creating artistic images but often entirely ignores human commands, generating a cat that is not realistic and generating a corgi in

a style quite different than that of the Spanish artist Salvador Dalí. None of the three models responded well to the third prompt "paying for a quarter-size pizza using a pizza size quarter," which aimed to test the model's language comprehension. The two models that generated a human hand trying to pay created weird-looking fingers.

	"A realistic image of a cat playing with a ball of yarn"	"A corgi painted in the style of Salvador Dalí"	"Paying for a quarter size pizza using a pizza size quarter"
DALLE-2			
Stable Diffusion 2.0			
Midjourney			

Foundation models also struggle with video generation. Videos generated by foundation models suffer from low realism and low resolution. Below are some images generated by Google's Imagen Video model (October 2022), a model that was widely lauded by technical experts as setting a record for "high fidelity, controllability, and real-world knowledge." It shouldn't take too long for observers to tell that the technology is not yet ready to replace Hollywood films.

It might take at least 2-3 years for text-to-video models to fool their human viewers and probably 3-5 years for such models to be commercially useful. In the meantime, these models might be applied to scenarios that require less fidelity and controllability.



Figure 1: Imagen Video sample for the prompt: "*A bunch of autumn leaves falling on a calm lake to form the text 'Imagen Video'. Smooth.*" The generated video is at 1280×768 resolution, 5.3 second duration and 24 frames per second.



Campfire at night in a snowy forest with starry sky in the background.



An astronaut riding a horse.

Figure 3: Videos generated from various text prompts. Imagen Video produces diverse and temporally-coherent videos that are well-aligned with the given prompt.

Although the AI research community is bullish about foundation models, the AI startup community is still debating **whether bigger models or more specialized models are the way to go**.

An example is the different approaches that OpenAI and Tabine took to building code-generation models. While OpenAI, in its usual fashion, opted for building a large, diffusion model to power Github's code-generation app Copilot, Copilot's main competitor Tabine took a different approach, building a series of specialized models for each programming language (over 30!). Which of these approaches will succeed depends on the progress of large-scale foundation models.

As a result of the imperfections in the current AI models, one promising application area is gaming, a low-stakes, fast-paced environment where speed and innovation matter more than model accuracy and reliability. For instance, synthetic speech still sounds quite robotic and human listeners can easily distinguish between human and AI-generated speech. This means that AI-generated speech might not yet be suitable for sales or marketing but gamers would tolerate and even love a slightly robotic NPC. In addition, AI-generated music remains far behind the top hits on Spotify but could be great background music for games. Creating visual effects and music for games is also very expensive, so the advent of generative AI offers a solution to dramatically reduce the cost of game production.

Beyond gaming, current-generation AI models are well suited for tasks that are **highly repetitive but highly paid** and those where **imperfection is allowed with humans in the loop**. Coding, marketing, and video editing are perfect tasks where AI could assist human experts, allowing them to do things faster and better.

Opportunities: The Overhyped and the Underrated

Doing research on the previous GPT-3-driven AI hype and the current AI landscape, we think that some areas in generative AI are definitely overhyped. But there are also underrated and overlooked trends. Here, we offer our humble opinion about the market opportunities in this emergent field:

Overhyped

Overcrowded fields - don't follow the crowd.

By this point, there are already over 20 companies that do AI copyediting and marketing, making the field extremely competitive. The concentration of startups in this area is a result of the specific capabilities of GPT-3 and similar language models. When these models were released around two years ago, they were already pretty good at editing human drafts but not sophisticated enough to create original content or engage in meaningful conversations. So, unsurprising, startups founded in that era crowded the field of marketing and sales.

Some crowded fields might also exhibit a strong first-mover advantage, where early entrants with high user engagement accumulate more data to fine-tune their models and further improve their models and user experience. [4] At the same time, newer players have a hard time breaking in, making it harder to make successful later investments in the field. However, it is possible that the release of newer models might create opportunities for new entrants to break in.

AI products that overpromise.

Whenever an impressive generative AI model is released it always generates a lot of buzz and excitement in Silicon Valley. The release of DALLE-2 and Stable Diffusion has led to talks about AI-generated films replacing directors and actors, and the ChatGPT debut created rumors about the advent of AGI or the displacement of Google. Such sentiment lures investors to back extremely ambitious projects that often overpromise – the technology is just not good enough yet to be useful in an intended way.

This is particularly true in high-stakes and highly regulated areas like automobiles, law, and medicine. Self-driving car technology is already highly mature but remains underutilized due to regulatory restrictions. The last 5%, or even 0.001%, of performance improvement is always the hardest to achieve for AI models. This is why FedEx abandoned its last-mile delivery robot – the cost savings from 99% of the deliveries simply aren't enough to make up for 1% of mistakes.

Hammers looking for nails.

AI products that focus too much on AI and not enough on customers and market size are essentially hammers looking for nails. They might be cool at first but soon lose traction as similar products emerge or when consumers get used to the AI models.

An example is AI Dungeon, one of the first apps built on GPT-3. The app generated a lot of hype during its release of the GPT-3 version in July 2020, largely because this was the only way for ordinary users to get access to GPT-3. But since mid-2021, the app's Google Store rating has been tanking from its previous high of 4.8 to under 2.6, with users upset about its overwhelming content moderation. Users have since migrated to similar but uncensored platforms such as NovelAI, a story-telling app powered by GPT-Neo.

Open-source projects with no product.

Investors often underestimate how easy it is to replicate even seemingly impressive AI models. Although Stable Diffusion is one of the most widely used image-generation models, its training cost only around $600,000, a price range that is acceptable to many companies. With ever more academic researchers entering the field, generative AI also has an abundant talent base. This means that hyped-up AI model startups such as Stability.ai might have more buzz than moat.

In fact, about a year ago, an independent research collective called Eleuther.ai trained and open-sourced language models GPT-J and GPT-Neo that are similar in performance to the smaller versions of GPT-3, Ada, Babbage, and Curie. However, unlike Stability.ai, they did not create as much hype and thus did not get much investor attention.

Underrated

AI tools integrated into existing products.

Investors love new startups and apps. This might be why AI tools developed by established companies are receiving a lot less attention. As Silicon Valley marvels at the capabilities of ChatGPT, Notion's new AI writing assistant gets much less buzz. However, it can be argued that this product could be a much bigger competitor to AI writing tools than people realize. The built-in AI text editor could make it much more convenient to use than standalone web apps. At the same time, AI apps that are plugins for existing software find that such a strategy is a great wedge into the user market.

Building the business before the technology.

A new business strategy might be to build the company first and then wait for the release of more powerful AI models. In fact, many generative AI companies were built prior to the release of their underlying models. Lensa.ai started as a photo editing tool in 2018 but quickly integrated Stable Diffusion when the model was released in April 2022. AI Dungeon rolled out in 2019 and initially used GPT-2 before the more powerful GPT-3 was released. Founders can build in adjacent fields before pivoting to generative AI, and Investors can anticipate where AI technology is heading and be the first mover, much like the Uber investors who predicted the rise of ride-sharing.

Niche verticals.

While doing research for the market map, it became clear very quickly that some fields are overcrowded, while certain niche verticals are overlooked by founders. Education, for example, is an area that has clear use cases for generative AI models. Children's education and foreign language education do not require very sophisticated AI models – the writing skills of GPT-3 and the math skills of quantitative reasoning models like Minvera (Google June 2022) far exceeds those of average children (in both cases) and foreign language learnings (in the case of GPT-3).

Having said that, it's also important for founders to realize the market potential whether it's venture-backable or not - because, with the leverage and value proposition of AI, an entrepreneur can build business outcomes that are life-changing, potentially with or without outside capital.

3 Tips for Startup Founders

First of all, congrats on reading this far. At this point, hopefully, we have walked you through the state of Generative AI, major players, the underlying technology/ AI models, the upcoming trends, and AIGC's current limitations and misconceptions.

What do all these mean for a startup founder?

1. MVP, PMF, GTM... Startup fundamental has not changed.

Despite the impressive performance of AI, entrepreneurs should resist the urge on both ends of the spectrum - over-promising before delivery and over-building before knowing there is a real market.

AI is a new way of programming. It's a powerful tool for entrepreneurs to build new products to solve society's problems - in a thoughtful, disciplined manner. Founders should have the courage to ignore the general noise from the media and even the capital market; instead, focus on the specific use case from a set of customers.

The MVPs for AI-first companies might be different than the previous generation of software companies but as the infrastructure layer and the AI models keeps getting more mature and cost-efficient, we would expect many more innovations to be tried and launched in aspects of our day-to-day life. By then, AI would become more "invisible" - just like the Internet, Cloud, and Mobile.

And just like those underlying tech enablers, AI is a tool, with which we can build the next generation of software products.

2. Distribution, distribution, distribution.

Given the "openness" of the infrastructure level, it's important for AI companies to be thoughtful on the go-to-market side and even get creative on partnerships with existing players if possible. One can argue in that in crowded areas such as text-based AIGC where the output qualify of the AI is more subjective, distribution / GTM / brand is all that matters. We think there is truth in that argument - however, distribution is a place for AI businesses to start but not sustain. This leads us to the last point.

3. Data is still the king.

In a typical software business, founders should have unique insights in either the technical breakthrough or the business pain point, or sometimes both; In an AI-powered product/ business, in addition to the insights, AI startup founders would also need to have unique approaches in acquiring data in a sustainable fashion.

Over the long term, companies with better (perhaps more niche) data for fine-tuning their AI models will show superior outputs for the end customers. That creates a sustainable moat.

As the cost of leveraging infrastructure AI models comes down, we expect more startup resources and investors' capital will go towards accessing more and better fine-tuning data and acquiring users who can continuously contribute such data. [5]

Maybe it would be a flywheel.

Calling for Builders

Generative AI technology is still in its early stages. As the underlying technology matures, there will be ever more opportunities for founders and investors. In the meantime, we don't need superhuman AI to create useful products -- there will be immense value to be created in the medium-to-short term. Decades into the future, we can easily imagine a world where generative AI is deeply embedded in human creativity, with the internet instantly creating content that human users dream up, factories producing 3D-printed clothes and furniture designed by AI, and a few clicks on a computer generating entire AI-powered apps.

At Leonis Capital, we are long-term bullish on the next generation of "supercycle companies" that are powered by AI and decentralized protocols. We have already made several investments in generative AI (even before it's hot!). Our portfolio companies include Mutable.ai, Spline, Kubit.ai, Sematic, RCT.ai, Cast, Motion, Semiotic.ai, etc.

If you are a founder and would like to meet or would like us to add your startup to our market map, please reach out to us at Jenny AT leoniscap.com or Jay AT leoniscap.com.

Partner with us now to build this future. It's just getting started. 

FOOTNOTES

1. At Leonis Capital, we were bullish on AI in 2020 when it was not as hot. The truth is, nothing fundamental has changed. We see both AI and decentralized protocols as “supercycle technologies” from a long-term point of view. We will elaborate more on this thesis in the upcoming essay and explain how these underlying technologies are creating new ways for human society to generate, manipulate; store and verify data, therefore impacting our society in the coming decades.
2. The reason these are “pseudo new markets” is that AI applications/companies are solving copying writing, sales, content creating, and coding - the same problems, with a much more efficient approach. Those are not completely new markets but those are the markets that can be expanded with the new type of tooling that’s provided by AIGC.
3. Interestingly, as the output of those AIGC companies gets into the mainstream media market (AI selfies, videos, and chatGPT), it creates more buzz and hype for VCs to pour even more capital into the space.
4. This is why leading investors, such as Sequoia Capital, believe that there will be sustained category leadership.
5. There is an interesting intersection between data ownership/identity in AI and web3 decentralized protocol, which we plan to discuss in our upcoming post.