Generative AI: Developments, Capacities, and Risks

By Ruoyu Liu (L’25)

The field of generative artificial intelligence (“AI”) saw major breakthroughs in recent years. Meanwhile, there has been increasing anxiety about generative AI’s disruptive nature: industry leaders argued that AI poses “risk of extinction” in an open letter;¹ colleges across the world proposed policies to stop students from using generative AI in their assignments;² and screenwriters and actors on strike demanded restrictions on AI in Hollywood productions.³ Al development has created a moment of societal change, and we need to brace for its impact by understanding AI’s potential and mitigating its risks. This essay serves as a primer for how generative AI works, its potential applications, as well as current legal and social concerns surrounding generative AI. A follow-up essay will discuss emerging regulations governing generative AI in different countries and regions.

What is Generative AI?

Generative AI is a catch-all term for AI models that generate new media content based on user prompts. Chatbot AI models like OpenAI’s ChatGPT and Google’s Bard produce text responses in plain language. Other models like Midjourney and Meta’s Make-a-Video produce outputs like images and videos. A generative AI model is essentially a combination of multiple AI algorithms that creates predictive outputs after processing inputs.

How does Generative AI Work?

To best understand how generative AI models work, we will lay out the whole process, from inputting a prompt to generating an output. First, the user inputs a command in plain language (or with an image in certain cases). Today, AI models have the ability of Natural Language Processing (NLP), which analyzes the meaning and context of the plain language input. Additionally, image and video generating programs require AI models to “see” by processing and analyzing images. This is a subject within AI called “computer vision” (CV). Then the pre-trained AI model compares the user input to the training dataset and uses complex architectures such as generative adversarial network (GAN) and variational autoencoder (VAE) to “predict” the output. Finally, the AI model seeks to “recreate” a new output based on the characteristics it identified from the input and patterns it learnt from the training data. In image-generating models, diffusion models generate images by decreasing the noise of the generated images step-by-step through layers.

To process inputs, developers first train these AI algorithms to “learn” patterns within a huge amount of existing data through a process known as “machine learning” (ML). For example, OpenAI’s GPT-3 is trained on “approximately 300 billion ‘tokens’ (or pieces of words) scraped from the web and had over 175 billion parameters, which are variables that influence properties of the training and resulting

Developers then align and fine-tune the pre-trained models with additional datasets to refine their accuracy.

Development of Generative AI

The development of generative AI reflects how artificial intelligence has progressed, through the collaboration of government agencies, private enterprises, and research institutions. The breakthroughs in AI technologies appeared in three waves, accompanied by advancements in computer hardware and data science.⁵

Scientists like Alan Turing founded the field of artificial intelligence, seeking to use computers to imitate human thinking processes and solve real-world problems. The Department of Defense’s Defense Advanced Research Projects Agency (DARPA) became a major funder of AI research in the 1960s, initiating the first wave of AI technology. This wave included earliest generative AI chatbot, ELIZA, that simulated a therapist and responded to user input with pre-defined scripts.⁶ However, AI models in this era could only follow rigid pre-defined rules and provided limited output due to a lack of computing power and learning capabilities.

In the 1990s, the second wave of AI development emerged, focusing on machine learning and data processing.⁷ The accumulation of and access to big data provided ample training data for AI models. More powerful and affordable graphic processing units (GPU) supplied the computing power for more complex AI models. GPUs were initially designed for faster graphic processing with their ability to process massive amounts of data in parallel. However, the enhanced computing power of GPUs made it very suitable for machine learning. Today, industrial-grade GPUs are built specifically for machine learning.

Basic research in computer science and data science enabled more advanced ML techniques, such as deep learning (DL). DL allows AI models to process the data in layers progressively, which has enabled models process complicated data, like speech or images, with greater precision. Neural networks, a type of DL algorithm that automatically adjust the relative weight of data grouped into nodes while processing the data, are advanced enough to mimic human cognitive processes.

Advancement in deep learning paved way for more complex AI models as well. The creation of GAN in 2014 by Ian Goodfellow accelerated the training of AI models by simultaneously generating content and assessing the output with competing algorithms.⁸ AI models can learn not just to identify patterns, but to make predictions from these patterns and create more authentic content. The prediction capacity of GAN, especially the image generation ability, is one of generative AI’s key features. The latest breakthroughs in deep learning like Transformer and Large Language Models (LLMs) have marked significant improvement in processing natural language.

---

⁷ Launchbury, supra, note 5.
With foundational models like GAN acting as pillars, generative AI quickly came to fruition for application as the market recognized its capabilities in processing user input and generating new content. In 2018, OpenAI created General Pre-trained Transformer (GPT), a LLM based on Transformer network that became a “foundational” model. OpenAI subsequently updated GPT into GPT-3.5 and GPT-4. Generative AI models based on these foundational models capable of real-world application like ChatGPT and Dall-E images started to emerge around 2020. Industrial giants like Google and Microsoft also became heavily involved in developing AI applications in the 2010s.

The third and ongoing wave of AI is likely to be classified as focused on reasoning and contextual capability. Current generative AI models are probabilistic models that generate legible texts and images based on data patterns without “understanding” what these outputs mean. Explanatory AI models that possess reasoning and contextual ability would need to understand complex causal relationships. Generative AI can be seen as a primitive model for Artificial General Intelligence (AGI) envisioned by AI pioneers that can perform different types of complex real-life tasks. Future generative AI models will be capable of not only generating more accurate and natural responses, but also tailoring to more nuanced and contextual input from users.

Applications of Generative AI

Recent development in generative AI algorithms have produced more realistic and complex content from plain language prompts, and provided new capabilities like interaction and generation. Based on these new capabilities, generative AI has the potential to significantly boost productivity. This has sparked major investment into the field; in the first half of 2023 alone, generative AI startups have attracted over $14 billion in equity funding. The ability of processing plain language input and generating outputs has provided generative AI models with three key capabilities: customizability, productivity, and interactivity. This section will explore some applications of generative AI from these three directions.

Through providing customized training dataset and refined specific user prompts, generative AI models can provide very specific output tailored to the need of the user, which offers high customizability for complex work products like programming codes and art designs. By finetuning the training datasets and the user prompts, users are able to further refine the output to their satisfaction. Even if the AI-generated outputs may not be immediately ready for use, they can be a basis for further human work. For instance, an advertisement agency may try out different variations of a poster and refine the prototypes generated by AI models, lowering the cost and time to try out new ideas.

In addition to creating a basis for further human efforts, generative AI models can entirely perform some tasks, significantly boosting productivity. At the current stage, generative AI models are best suited to handle two types of tasks independently. Generative AI are highly capable in performing administrative tasks that are repetitive in nature with slight variations, like providing summaries to documents and drafting emails. They are also good at processing large quantities of data faster than humans, like

---

9 Launchbury, supra, note 5.
10 Laurie Harris, Cong. Rsch. Serv., R46795, Artificial Intelligence: Background, Selected Issues, and Policy Considerations (2021)
organizing lab results and analyzing for trends. The types of work that are mostly likely to be impacted by generative AI are associated with programming, translation, and writing.\textsuperscript{12}

In terms of interactivity, the natural language processing capability of generative AI significantly reduced the barrier for the general public to interact with generative AI. For the general public, AI chatbots work as a customizable search tool that provides more direct output than traditional search engines. For businesses, generative AI can be used to provide a customizable experience to customers and to receive feedback.

**Legal and Social Concerns**

Generative AI models today have limitless applications that significantly improve productivity. However, the use of generative AI has also created legal and social concerns that must be addressed in the long run. Concerns surrounding the generative AI emerges in three main aspects, including data collection and model training issues, undesirable output of existing AI models, and the negative social impacts of generative AI use.

**Data Collection and Model Training**

AI models take huge datasets to train, and generative AI developers may run into privacy and copyright issues before launching an AI model if they improperly collecting and using data. Currently, many AI developers use data “crawled” from public websites to train their AI models without the consent of data owners. These publicly-available data can include personal identifying information or even biometric information like face data that may be revealed or misused by the AI model.\textsuperscript{13} Without comprehensive data privacy laws, those affected may lack proper legal redress.

In addition to data privacy issues, the data used to train AI models may also raise copyright issues; especially as generative AI models are increasingly used in artistic creation. When a generative AI model is trained with the artwork of a human artist, it is necessary to first recreate the artwork digitally, which may infringe on the artist’s right to recreate the artwork.\textsuperscript{14} In January, a group of artists initiated a class-action lawsuit against image generating AI tool Midjourney for violating copyrights by scraping artwork online for training without consent.\textsuperscript{15}

The issue of copyright reemerges after the generation of artwork. After a user generates a piece of artwork with an AI model based on the work of a human artist, the user, the human artists, and the AI developer have all contributed to the artwork, and the questions of authorship and ownership of the generated artwork emerge.\textsuperscript{16} A key contention of the SAG-AFTRA strike is the ownership of AI-generated images of actors of actresses.\textsuperscript{17} So far, the US Copyright Office and US Patent and Trademark has declined

---


\textsuperscript{13} CHRISTOPHER T. ZIRPOLI, CONG. R.SCH. SERV., LSB10922, GENERATIVE ARTIFICIAL INTELLIGENCE AND COPYRIGHT LAW (2023).

\textsuperscript{14} KRISTEN BUSCH, CONG. R.SCH. SERV., R47569, GENERATIVE ARTIFICIAL INTELLIGENCE AND DATA PRIVACY: A PRIMER (2023).

\textsuperscript{15} James Vincent, AI art tools Stable Diffusion and Midjourney targeted with copyright lawsuit, VERGE, (Jan. 16, 2023), \url{https://www.theverge.com/2023/1/16/23557098/generative-ai-art-copyright-legal-lawsuit-stable-diffusion-midjourney-deviantart}.

\textsuperscript{16} Tracy Rubin et al., Ten Considerations for Developing an Effective Generative AI Use Policy, COOLEY (Jun. 5, 2023), \url{https://www.cooley.com/top-ten-considerations-for-companies-using-generative-ai-externally}.

\textsuperscript{17} Kevin Collier, Actors vs. AI: Strike brings focus to emerging use of advanced tech, NBC NEWS (Jul. 14, 2023), \url{https://www.nbcnews.com/tech/tech-news/hollywood-actor-sag-aftra-ai-artificial-intelligence-strike-rcna94191}. 

to extend copyright and patent to generated work, but the attitude could change with broader social input on this issue.

Undesirable Outputs of Existing AI Models

Currently, generative AI models largely work as “black boxes” due to their probabilistic nature, i.e., the models do not “understand” the generated output. This means the existing limitations of these models may lead to unintended or undesirable outputs. For instance, AI-generated human images sometimes have hands with six fingers, since the AI model does not have a concept for “a hand.” Current chatbots also suffer the problem of confidently generating inaccurate or even fictional information, known as “hallucination,” that are not in accordance with the training dataset. The six-fingered hand may be easy for human users to notice, but instances of hallucination could be particularly dangerous in some circumstances as human users rely on inaccurate responses that sound increasingly realistic.

AI models trained on problematic data may also perpetuate the bias and inaccuracies reflected in the training dataset. For example, image generative AI tools and chatbots can amplify racial and gender biases due to deficiencies of the algorithms or the training dataset. Currently, mitigating unintended AI outputs requires greater fine-tuning and alignment of AI models from the developer before release. Users should also treat the output of AI models with a grain of salt.

Negative Social Impacts of Generative AI Use

Even if it generates outcomes that are desirable to the user, an AI model may create content that has negative social effects. Generative AI may be used for illegal or unethical purposes. For instance, using past video and audio input, generative AI models could create “deepfake” videos of virtually anyone saying anything. AI developers cannot account for what the output is used, and users may abuse this feature to spread misinformation. A related discussion focuses on the tort responsibility of malicious AI use or undesirable AI outputs, in terms of who should be responsible for the negative outcome of generative AI use.

An even greater concern is the prospect of job displacement, a recurring topic with productivity improvement from automation. A Goldman Sachs study identified that generative AI may impact nearly 2/3 of current occupations in Europe and the United States. OpenAI’s own study also predicted that 19% of US workforce may see 50% of their work impacted by generative AI.

---

Furthermore, the benefits of the increased productivity from generative AI will not be shared universally across society. An econometric study identified that automation has significantly contributed to wage inequality in the United States for the past 40 years,\(^{25}\) and the adoption of generative AI technologies may have similar effects.

**Conclusion**

The boom of generative AI exemplifies the how the academic, the market, and the government interact with each other in the process of technological advancement. Basic research in data science and cognitive science enabled the rapid development of AI, while the market accelerated the application of generative AI. As society embraces the new productivity by utilizing generative AIs, the government also plays the key role of preparing for and mitigating AI-posed risks and social impacts. During a Congressional hearing, OpenAI CEO Sam Altman suggested that mitigating AI-induced unemployment “requires partnership between the industry and government, but mostly action by government.”\(^ {26} \) The follow-up essay will explore the government attempts in the European Union, China, and the United States to regulate Generative AI.

---
