

AIC-101 Introduction to Artificial Intelligence
Assessment Activity AIC-101:00010
Applying Knowledge of AI Systems and Concepts

Name: (Win) Thanawin Pattanaphol

Student ID: 01324096

“Artificial Intelligence”, “Machine Learning”, and “Deep Learning”. These three words have created a huge impact on the overall tech industry as well as the global economy since the introduction of “ChatGPT” by OpenAI in November 2022 ¹. This phenomenon can be observed in the increase of product developments by tech companies ² and mentions of “A.I.” by various news outlets, big tech companies and in public conversations. Main examples of this are in mainstream tech companies’ product events and earning calls: Google’s 2023 I/O event in California where, the acronym “A.I.” was mentioned roughly 143 times during the two-hour presentation which approximately equates to the acronym being mentioned 1.153 times every minute ³. Meta and Microsoft also follow this trend where, in their Q1 2023 earnings calls, mentions “A.I.” 56 and 53 times respectively ². As these so-called trends may seem to suggest that “Artificial Intelligence” is the latest and greatest technology that will completely change the world as we know it, it also seems to suggest that there is a re-emerging pattern of the “A.I. Boom” which was last observed during the early 1980s where “expert systems” were actively being promoted, funded, and developed ⁴ but was then followed by at two to three decades long “A.I. Winter” ⁵. This raises a crucial question in the tech and business industry: Is the A.I. “hype” of the 2020s about to disappear, just like in the late 1980s? If so, what will happen, and when will it happen?

To be able to answer the three questions mentioned in the introduction, we must first reflect on the history of the “A.I. Boom” and the “A.I. Winter”. Thomas Haigh’s opinion article titled “How the AI Boom Went Bust” ⁶ argues that the “first A.I. winter” during the 1970s never happened but instead was a time for steady development of A.I. He also states that the A.I. winter instead happened during the late 1980s. Haigh begins by discussing the role of A.I. in education, covering how the subject has evolved parallel to the development of A.I. technology. Haigh highlights the subject’s shift from a marginal subject that was only recommended for master degree students and a small group of undergraduates during the late 1960s to a mainstream subject that is considered “core subjects” in various computer science curricula in the early 1990s. In addition to that, Haigh mentions the increase of A.I. textbooks published by professors from two leading institutions in the field of A.I. i.e., MIT and Stanford and how this reflects the ability of these two institutions to influence what topics are being taught in other institutions. Haigh uses the presence of A.I. in academia to describe the atmosphere of the field during the 1970s where most, if not all, A.I. developments are focused on knowledge-based systems with the goal of understanding human language or also known as “symbolic language models”.

Haigh proceeds to cover the rise of symbolic language models, specifically “expert systems”, and the decline of said technology, mentioning the shift in development from powerful reasoning machines in the late 1960s to a more effective way of representing knowledge by the 1970s. During this period, the general consensus on the goal of symbolic language models is to create a “general purpose reasoning engines driven by collections of individual facts”, essentially a system that involves embedding logical schemas and elements and creating explicit set of rules which enables the program to manipulate these schemas. An example of such program that was brought up is a program called “SHRDLU” developed by Terry Winograd in 1971. Haigh delves into covering the abilities of the program that Winograd claimed, such as it being able to understand natural human language, process commands via teletype and executing those commands in its block simulation and being able to politely respond to the user. However, as impressive as it sounds, Haigh points out the limitations that these types of programs have: “...highly specialized systems, few of which were ever applied

beyond carefully chosen problems.", essentially that such systems need highly specific instructions to be able to create the impressive response that programs like SHRDLU produces.

Following the creation of SHRDLU, comes a more advanced set of programs called "expert systems". The development of expert systems started in the 1960s by Edward A. Feigenbaum where he stated his focus for said systems is to "automate the work of human experts", the introduction of various expert systems were followed by the rapid increase in private investment, startup companies specializing in expert systems and eventually reached its peak when Feigenbaum led a campaign to present the issue of the Japan's potential to overtake the United States in its development of natural language interfaces. In the end, Feigenbaum succeeded in advocating for said campaign and garnered public funding into the project. However, several issues start to become apparent as the development continued. DARPA concluded that several of its autonomous systems failed to reach its expectations, which, combining with budget cuts from the Reagan administration, inevitably ends with DARPA abandoning its work on autonomous systems and companies and startups subsequently lose interest in further developments of expert systems as they are expensive to maintain and is difficult to find new experts to maintain them as well. Mentions of "expert systems" or "artificial intelligence" dropped during the late 1990s until its revival in the 2010s.

I very much agree with the author's overall point on the period of the "A.I. boom" and its consequences. Haigh encapsulates how such technology can be carefully engineered to produce overblown outcomes, leading to a surge in popularity (hype), consequently resulting in huge investments and inevitably ending with a huge disappointment when the technology is unable to reach expectations. However, we must take note on that the hype that is often associated with "A.I." is slightly misleading as they mostly only surround the sub-category of A.I. called "Generative A.I.". These generative models are statistical models that identifies different structures and patterns in their base training data which enables its ability to produce new text, images, videos and other forms of content ⁷; these models are often used in contexts where the user enters a specific prompt and the model produces data from it ⁸, the most frequently talked about in the recent times is the Large Language Models (LLMs) such as OpenAI's "Generative Pre-trained Transformer 4" ⁹ or "GPT-4" ¹⁰ that currently powers its chat bot, ChatGPT and Microsoft's Bing AI.

The hype of these models both have similar characteristics to the "A.I. Boom" of the 1960s. First, they both have the notion of a "new" technology that produces seemingly amazing results e.g. ChatGPT's ability to generate coherent paragraphs and texts ¹¹. Second, they both have wide range of press coverage and publicity whether it is on news networks, articles or magazines. And third, they both receive huge private investments from various computers along with spawning a wide range of new A.I. startups ¹². These three traits show that Generative A.I. is starting to follow the same pattern or timeline as symbolic A.I during its most popular period.

In the next five years, however, I believe that the *hype* of Generative A.I. will continue to follow a similar downward trend as the case in the 1980s as issues surrounding the technology has become more apparent in recent times. Shumailov et al. ¹³ discusses the issue on the increase in unreliability of large language models where as the human-generated data that is being used to train these large language models are gradually being replaced by LLM generated content, the results that are produced by next-generation large language models become less accurate; referred to as a "model collapse". The paper reveals that there are two stages of the model collapse, the early model collapse and the late model collapse. The early model collapse shows the model losing information from the original data and the late model collapse shows the model generating information that has little resemblance to the original data.

Figure 1 shows a histogram of the perplexities of data produced by different large language models which are evaluated by the very first model trained with the original wikitext2 data. The graph shows that over the generations, the models tend to produce the same data as the original model with real data, however, the graph also shows the increase of tails in later generations which describes the models producing new data that the original model would have never produced, which the paper describes as “they start misperceiving reality based on errors introduced by their ancestors”. This demonstrates the risk of future generations of generative A.I. models to produce less accurate information leading to user complaints and distrust.

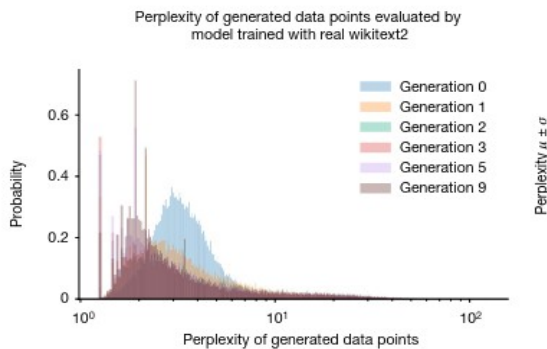



Figure 1: Perplexity of generated data points evaluated by model trained with real wikitext2.¹³

Other researchers also delved into the effects that generative A.I. has on society. Bender et al.¹¹ presents the underlying issue of biases in LLMs in which they discuss the source material that being used to train mainstream LLMs i.e. the internet which is retrieved through the method known as “web crawling”¹⁴. The paper presents how views by younger users and users from developed countries are more likely to be overrepresented by analyzing sample data, in which, the analysis shows that the data source for this training data set comes from user-generated sites such as Reddit where 67% of its users in the United States are men and 64% of its users are between the age of 18 and 29, and Wikipedia where only 8.8 – 15% of its users are women or girls¹⁵. Additionally, the paper also argues that, since the training data set is based on user-generated content sites like Reddit or Twitter, the data that was crawled can contain content that has themes of homophobia, transphobia and harassment towards groups such as people of color queer people, immigrants, and other marginalized groups which leads to LLMs amplifying biases and harms resulting in responses that exhibits racial biases, cultural stereotypes or negative sentiment towards specific groups of people^{11,16}.

Lastly, the training of large language models and other types of A.I. has large contributed negative effects to the planets’ environment. In its 2024 environmental report, Google revealed that its carbon emissions have increased by 48% compared to 2019 and a 13% increase year by year in 2023 which the company credited the figure increase to the “rapid advancements in and demand for artificial intelligence”, also mentioned was its data centers’ electricity consumptions which grew by 17% in 2023¹⁷. Similarly, Microsoft, the company in which OpenAI relies on to operate their GPT models, also revealed in its 2024 sustainability report where its carbon emissions increased by nearly 30% due to the company’s data center expansion¹⁸. In addition these sustainable reports, independent researchers have also revealed more statistics for environmental impacts by LLMs. Li et al. revealed that Microsoft used approximately 700,000 liters of freshwater (from the process of producing electricity) during GPT-3’s training in its data centers¹⁹ which according to the environmental news website, Earth.org, is equivalent to the amount of water used to produce 370 BMW cars or 320 Tesla vehicles²⁰. These figures reflects the effect on the environment that the A.I. hype has caused and as these LLMs continue to become larger, the more carbon will be emitted to the atmosphere, the more water will be consumed to produce electricity for training these A.I. models and more energy from various sources will be needed to run these models.

In conclusion, I believe that, in the next five years, the “A.I. Hype” surrounding large language models such as ChatGPT, Gemini and other mainstream models will eventually grind to a halt due to its under performance, ethical concerns, biases, and environmental effects that became more apparent in recent years. On the other hand, I do believe that the *technology* of language models will continue to be developed over the next five years by open source developers, hobbyists or even tech companies ²¹; a prime example of this is a popular framework called Ollama ²², a community-driven project that enables any individual to run various large languages models such as Meta’s LLaMa with a Docker container. The project currently has approximately 95,000 stars and around 7,500 forks. Large language models are indeed here to stay and I believe that it is best for developers or engineers to continue and combine its development with other machine learning technologies such as text-to-speech, speech-to-text or OCR to maximize the benefits from A.I. while also taking precautions in the ethical and environmental issues surrounding them; thus creating a more open, ethical and environmentally friendly field of A.I. for humanity and the planet.

References

1. Vasani, S. ChatGPT, explained. *The Verge* <https://www.theverge.com/24161231/chatgpt-generative-ai-explained> (2024).
2. Richter, F. Chart: Tech Giants Were All About AI This Earnings Season | Statista. <https://www.statista.com/chart/29979/ai-mentions-during-big-tech-earnings-calls/> (2023).
3. Lumb, D. & Blumenthal, E. Google Said ‘AI’ Over 140 Times in Its 2-Hour Google I/O Keynote. *CNET* <https://www.cnet.com/tech/computing/google-said-ai-over-140-times-in-its-two-hour-google-io-keynote/> (2023).
4. Russell, S. J. & Norvig, P. *Artificial Intelligence: A Modern Approach 2Nd Ed.* (Prentice-Hall Of India Pvt. Limited, 2007).
5. Kautz, H. The Third AI Summer: AAAI Robert S. Engelmore Memorial Lecture. *AI Mag.* **43**, 105–125 (2022).
6. Haigh, T. How the AI Boom Went Bust. *Commun ACM* **67**, 22–26 (2024).
7. Pinaya, W. H. L. *et al.* Generative AI for Medical Imaging: extending the MONAI Framework. Preprint at <https://doi.org/10.48550/arXiv.2307.15208> (2023).
8. A Cheat Sheet to AI Buzzwords and Their Meanings: QuickTake. <https://news.bloomberglaw.com/tech-and-telecom-law/a-cheat-sheet-to-ai-buzzwords-and-their-meanings-quicktake>.
9. What is GPT-4 and Why Does it Matter? <https://www.datacamp.com/blog/what-we-know-gpt4>.
10. GPT-4. <https://openai.com/index/gpt-4/>.
11. Bender, E. M., Gebru, T., McMillan-Major, A. & Shmitchell, S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?  in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* 610–623 (Association for Computing Machinery, New York, NY, USA, 2021). doi:10.1145/3442188.3445922.

12. Investors Pour \$27.1 Billion Into A.I. Start-Ups, Defying a Downturn - The New York Times.
<https://www.nytimes.com/2024/07/03/technology/ai-startups-funding.html>.
13. Shumailov, I. *et al.* AI models collapse when trained on recursively generated data. *Nature* **631**, 755–759 (2024).
14. What is a web crawler? | How web spiders work. <https://www.cloudflare.com/learning/bots/what-is-a-web-crawler/>.
15. Mitchell, M. B., Galen Stocking, Jesse Holcomb and Amy. 1. Reddit news users more likely to be male, young and digital in their news preferences. *Pew Research Center*
<https://www.pewresearch.org/journalism/2016/02/25/reddit-news-users-more-likely-to-be-male-young-and-digital-in-their-news-preferences/> (2016).
16. Fang, X. *et al.* Bias of AI-generated content: an examination of news produced by large language models. *Sci. Rep.* **14**, 5224 (2024).
17. Bartlett, K. Google’s carbon emissions surge nearly 50% due to AI energy demand. *CNBC*
<https://www.cnbc.com/2024/07/02/googles-carbon-emissions-surge-nearly-50percent-due-to-ai-energy-demand.html> (2024).
18. Kimball, S. Microsoft’s carbon emissions have risen 30% since 2020 due to data center expansion. *CNBC*
<https://www.cnbc.com/2024/05/15/microsofts-carbon-emissions-have-risen-30percent-since-2020-due-to-data-center-expansion.html> (2024).
19. Li, P., Yang, J., Islam, M. A. & Ren, S. Making AI Less ‘Thirsty’: Uncovering and Addressing the Secret Water Footprint of AI Models. Preprint at <http://arxiv.org/abs/2304.03271> (2023).
20. McLean, S. The Environmental Impact of ChatGPT. *Earth.Org* <https://earth.org/environmental-impact-chatgpt/> (2023).
21. Team, O. The Growth of Small Language Models. *Open Data Science - Your News Source for AI, Machine Learning & more* <https://opendatascience.com/the-growth-of-small-language-models/> (2024).
22. ollama/ollama. Ollama. <https://github.com/ollama/ollama> (2024).