

### **Annotated Bibliography Assignment**

As a computer science major, I wanted to further my knowledge about the progression of Artificial Intelligence, also known as AI. To do this, I chose a research paper in 2022 about the early(ish) development of AI, a paper released two weeks ago about current developments in AI, and a paper released a year ago that shows the development of AI through the lens of a use-case.

#### **Source 1 -**

Fei, N., Lu, Z., Gao, Y., Yang, G., Huo, Y., Wen, J., Lu, H., Song, R., Gao, X., Xiang, T., Sun, H., & Wen, J.-R. (2022, June 2). *Towards Artificial General Intelligence via a Multimodal Foundation model*. Nature News. <https://www.nature.com/articles/s41467-022-30761-2>

This article, *Towards Artificial General Intelligence via a Multimodal Foundation model*, covers a theoretical framework for obtaining AGI (Artificial General Intelligence, a “goal” within Artificial Intelligence development) back in 2022. To reach AGI, there are common goals that the industry has established for Artificial Intelligence: It matches or exceeds human performance in broad tasks, it has the ability to approach problems differently from its creators, and it can transfer knowledge from one context to another. The researchers of this study focused on the first aspect; outlining an experimental approach to combining visual and textual context that is used in AI models now in 2025. While it may seem obvious when you don’t think too deeply about it, we are often presented captions that relate to an image, but don’t *describe* it. For example, as provided in the study, if someone said, “Happy Birthday! Make a wish!”, you would imagine that there is a potentially cake in front of you, with candles. However, AI at this point

was not able to “imagine” it itself. To solve this problem, the team working on this early model called BriVL, and trained it on images with captions that have a weak semantic connection (Not directly describing what is in the image). Through developing an early model like this, the team was successfully able to make text that depicted images more accurately than before. As a side effect of this, through tuning of the model, BriVL was also able to answer multiple choice questions about certain images to a very accurate degree at this time.

Overall, this paper effectively informed me on some of the early development within the AI industry that has laid the foundation for AI to progress as far as it has recently. If you present ChatGPT with images today, it can describe what is contained in the photo and also answer questions about the photo much more effectively than shown in this early model. This article being released in 2022 shows that despite it being released so recently, AI has developed at a rate that makes the articles from only years ago show age. Using these articles from the earlier developments of AI can allow interested individuals such as myself to learn more about Artificial Intelligence in a more approachable manner.

## Source 2 -

Guo, D., Yang, D., Zhang, H., Song, J., Zhang, R., Xu, R., Zhu, Q., Ma, S., Wang, P., Bi, X., Zhang, X., Yu, X., Wu, Y., Wu, Z. F., Gou, Z., Shao, Z., Li, Z., & Gao, Z. (2025, January 23). *DeepSeek-R1: Incentivizing reasoning capability in LLMs via reinforcement learning*. DeepSeek-AI.  
[https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek\\_R1.pdf](https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek_R1.pdf)

This article, *Incentivizing reasoning capability in LLMs via reinforcement learning*, is an overview of DeepSeek’s new R1 model by the researchers behind DeepSeek. For context,

DeepSeek is essentially China's response to ChatGPT; outperforming it on many metrics and also reportedly also being far more efficient than ChatGPT's o1 model. Most importantly, it is also *open source*, which means DeepSeek shares some information regarding how they developed this AI for the public, which resulted in this paper being published (or pushed!). Basically, DeepSeek took a whole different approach from the standard AI model. While most AI models use supervised data, which is data that has been fine tuned for AI to interpret it, DeepSeek-R1-Zero, the first version developed, used pure Reinforcement Learning (RL). This yielded very strong results for very little cost, comparable to early ChatGPT o1 models. However, problems arose where it would randomly mix languages for no apparent reason. To solve this problem, the researchers tried to use a small amount of data from earlier versions as a "Cold Start". This data consisted of the Chain of Thought (the "thinking" process of the AI) of earlier DeepSeek AI models, and by using that combined with RL, DeepSeek-R1 was able to surpass ChatGPT's best model recently in multiple aspects, and also remain cost efficient..

This paper helped me understand the methodologies of current AI development, and also incentivised me to do some background research on terms heavily involved with them (Reinforcement Learning and Supervised Fine Tuning, for example). I find it fascinating how DeepSeek took a whole new approach to LLMs that focused primarily on efficiency rather than blind power. On a much grander scale, however, it also clued me in on how large the AI industry has gotten in the past few years, and I learned a lot about how other countries are responding to AI's rapid growth in the United States. American companies like Nvidia and Broadcom directly profit off of AI's need for performance by developing more powerful chips. By "showing" the US that AI can be done more efficiently, DeepSeek undermined these American companies and casted doubt upon the need of the newest chips that these companies offer.. Not only that, but a

free open-source model that is equal (and better in some cases) to performance in OpenAI's recent models, which cost as much as 200 a month, casts a lot of doubt on the U.S. company itself. By reading this article, I was able to better understand the reason behind these new developments.

### Source 3 -

Tsai C-Y, Cheng P-Y, Deng J-H, Jaw F-S, Yii S-C. *ChatGPT v4 outperforming v3.5 on cancer treatment recommendations in quality, clinical guideline, and expert opinion concordance*. DIGITAL HEALTH. 2024.  
<https://journals.sagepub.com/doi/full/10.1177/20552076241269538>

This article is a comparison between ChatGPT-3.5 and ChatGPT-4 in terms of giving medical advice for cancer treatment. The researchers conducted an experiment where both of the AI models were prompted with generating a treatment plan for different types of cancer. The response was then evaluated by three urologists, who gave a 1-5 score on aspects such as correctness, comprehensiveness, specificity, and appropriateness, and by comparing the results to the National Cancer Comprehensive Network's (NCCN) guidelines. The results of this experiment were that ChatGPT-4 produced recommendations that were better in quantity (# of recommendations, and amount of elaboration) and quality (rated score and NCCN alignment) than ChatGPT-3.5. However, given that the alignment with NCCN is still only 76.8%, it is *still* much more recommended that patients seek a medical professional rather than using ChatGPT as a replacement.

This article provided me with a great insight into the improvement of AI within use-cases such as the medical field. Although the results show that it is not quite there in terms of replacing medical professionals' expertise, the rapid improvement in recent years could indicate that there is a potential use for AI to be a tool for individuals such as cancer patients. There are often large wait times for these cancer patients to see a medical professional such as a urologist, and cancer patients using AI for general advice in the meantime could provide them with useful information in the meantime. With how much the improvements show, it could also mean that medical professionals work in conjunction with AI in the near future, using it as a tool to help patients as well. By reading this article, I became more aware of the fact that development in Artificial Intelligence can lead to a massive change in the way that other industries operate.