

Can Artificial Intelligence and humans be friends at work?

Authors: Matthew Morrison & Pat Gudhka

Australian River Restoration Centre

Canberra ACT 2602

matt.morrison@arrc.com.au, pat.gudhka@arrc.com.au

April 23, 2024

Abstract

This paper explores the role of generative AI, specifically natural language processing, in the context of human work and collaboration. The paper provides a brief history of the development of large language models, their capabilities and limitations, and their ethical and social implications. The authors propose a shift in perspective toward viewing AI with curiosity rather than fear or denial, highlighting the potential for generative AI to act as a supportive tool, enhancing human productivity and creativity. A case study is provided to demonstrate generative AI's assistance in the communication of a complex environmental water project, highlighting the benefits and challenges of working with generative AI. The paper concludes by proposing a strengths-based approach to leverage the complementary skills of humans and AI, offering a mental model on how to use generative AI effectively and responsibly.

Introduction

Humans have an innate tendency to seek familiarity, and this inclination is deeply rooted in our neurobiology, stemming from our primal instinct for safety (Galdo et al., 2022). Thanks to our evolutionary need for survival, our brains are inherently inclined toward routines and habits that provide stability, predictability and reliability, and we are predisposed to reducing cognitive effort wherever possible.

Ironically, it is our remarkable ability to adapt that has allowed us to continually evolve in a rapidly changing world (Massey, 2013). Our ability to think creatively, create and pass on knowledge socially, and imagine novel solutions to survival threats has all demonstrated our capacity to innovate, learn and evolve. Therefore, imperative to our progress is our ability to step out of our comfort zones and challenge ourselves to explore new possibilities – this is particularly relevant in the realm of technology, where developments in AI are creating a resounding “uh-oh” among societies across the world (Merisotis, 2023).

In this paper, we look at a particular type of AI – generative AI – and examine its history, applications and implications for our working lives.

Unpacking and understanding generative AI

Among our accomplishments, technology stands out as one of the most transformative and pervasive achievements. Technology has revolutionised every aspect of human life, from democratising information and empowering individuals, to reshaping industries and fostering global collaboration (Brooks, 2023; Merisotis, 2023). More recently, the introduction of Artificial Intelligence (AI) has ushered in a new era of capabilities and possibilities.

One such capability is AI's ability to understand, interact with, and generate human language, also known as 'natural language processing' (Elastic, 2024; Google for Developers, 2023). The history of natural language processing dates back to the 1950s, with the development of language models that began with simple tasks such as predicting the next word across sentences. However, machine learning back then had two basic limitations – small amounts of data storage and painfully slow processing speeds. By the late 80s, computational power had increased significantly and the internet, known at the time as the 'World Wide Web' (gosh, that makes us feel ancient), offered a massive source of data. This spurred the development and progress of large language models (LLMs) – machine learning algorithms that mimic the human brain's neural networks – which became better and better at understanding and responding to human language (Kaplan et al., 2020), noting that the question of whether LLMs actually understand anything is in dispute (Blackman et al., 2024).

With more data being used to train large language models, new capabilities were becoming possible which had not been seen in smaller models. Referred to as 'emergent behaviour', which includes generating plausible, or at least convincing, answers to our questions, creating summaries and classifying text, this raised both possibilities and risks for continued model growth and development (Google for Developers, 2023; Wei et al., 2022). By 2018, large language models were being used in every industry, and more notably, the introduction of ChatGPT in late 2022 brought the world of 'generative AI' to our fingertips (Cardon et al., 2023).

These significant developments did not come without controversy. As large language models became more sophisticated, they led to global concerns due to their potential to automate tasks traditionally undertaken by people, threatening job displacement in content creation, data analysis and customer service roles to name a few (Merisotis, 2023). Some of these fears have been realised as organisations opted for cost-effective AI solutions over human labour. Furthermore, deep learning AI has posed ethical dilemmas around algorithmic bias, accountability and control (Shah & Sureja, 2024). Additionally, large companies developing LLMs, most notably OpenAI, Microsoft, Google, and Meta, have been accused of cutting corners regarding legal and corporate governance in their race to train models on massive quantities of data (Metz et al., 2024). This is because these companies have assumed that using publicly available information to train LLMs is protected under fair use, a position which may or may not be upheld in court. Aside from the abundant copyright concerns, privacy concerns abound regarding the nature, volume, and frequency with which private or personal data is fed into AI systems in the pursuit of more useful features (Patel, 2024).

Can generative AI and humans work together?

This begs the question – is deep learning AI to our benefit or our detriment? We believe the answer isn't black and white, and investigating the nuance enables us to see the grey, especially considering generative AI is here to stay – no rhyme intended (Cremer & Kasparov, 2021; Lakhani, 2023). By taking a strengths-based approach, we have found some ways to work effectively with AI without feeling overwhelmed.

With the emergence of highly capable large language models, it's easy to be fearful of the features and capabilities that this AI brings such as fast and efficient data processing and synthesis. Most people tend to fall into two extremes; the 'We don't have time to look into AI' denial camp or the 'AI is going to steal my job' camp, but there is a third option in the middle – the “AI will help me get things done” support camp (Kniberg, 2024).

This camp is founded on what we as humans are good at — nuanced contextual understanding of content, emotional intelligence, critical thinking that helps address errors and biases, creativity and our ability to think and act ethically (Cremer & Kasparov, 2021; Kniberg, 2024; Lakhani, 2023; Merisotis, 2023). By focusing on our strengths, we in turn direct AI in a supportive role that can help us deliver better results, more quickly. We can also use AI to help us identify possibilities or areas of risk that we may not have considered, which helps us to be more thorough. It's as if we have a personal assistant who has effectively a superior IQ and knowledge of almost every topic imaginable, but who requires clear direction and is known to hallucinate at times.

EEWD: A Case Study

The Enhanced Environmental Water Delivery project combines 20 years of practice in river management that crosses over three states and four river systems in the Southern Connected Basin, is interdependent with another major project and ultimately, is a step change for environmental water delivery. The project operates with four partners across three jurisdictions, with the inherent complexity of coordinating messages across states, mixed community and political support for environmental watering activities and the emerging challenge of communicating the program simply, yet clearly, to key stakeholders. In addition, it's important to note that the project name comprises of four big words that in turn create a sci-fi-esque acronym – EEWD.

We were brought on to help design and develop a package of communication products that would effectively communicate the EEWD project to two core audiences:

- senior executives and management in the water management sector, briefing ministers and decision makers
- the multiple end users that would use the EEWD project outputs, such as river operators

We started by poring through a mountain of communications documentation for various audiences produced by the EEWD team over the past several years. Some of these documents were fed into generative AI to gain a faster synthesis of the key benefits and outcomes mentioned multiple times across documents. All AI outputs were critically reviewed against our own findings from the initial documentation review.

This was followed by drafting a 2-page summary of the project which provided a story arc that explained the project's need, activities and expected outcomes, in the context of the environmental water sector and the challenges being addressed. The summary information was then fed into a generative AI to assist in the synthesis of messaging into shorter, simpler lengths for different audiences, that would underpin the various communications products developed. In addition, the AI was used to road test the development of communication devices such as metaphors that could emotionally engage the EEWD audiences.

What happened? AI was significantly helpful in assisting the team to efficiently consume and synthesise a large volume of communication documentation, and develop shorter, synthesised versions of the developed EEWD story arc, ultimately helping navigate the delivery of simplified messaging for multiple audiences. At the same time, AI's metaphorical suggestions, upon team review and discussion, left much to be desired and were unsuitable for the communication products developed. For example, our team had introduced the metaphor of a 'heartbeat' to convey the Southern Connected Basin as a living entity, and rivers as the 'veins' that brought life to our ecosystems, with the EEWD project designed to strengthen this heartbeat which had weakened over the past several decades due to river regulation. Unfortunately, AI considered the project a 'pacemaker' for the river system – certainly an apt metaphor, but also missing the point!

Conclusion

In the discourse surrounding generative AI and its implications on our jobs and work, it's important to develop a robust mental model that views the role of large language models as complementary to human expertise, not a replacement of it. As evidenced by the practical application of generative AI in the EEWD project, AI supported the synthesis of large volumes of complex documentation and enhanced our messaging, although it fell short in applying contextually rich and emotionally engaging metaphors and language.

The true strength of generative AI lies in its unparalleled ability to process vast amounts of information swiftly, offering insights and serving as a catalyst for decision-making. It lacks the nuanced understanding, emotional intelligence and ethical reasoning that humans bring to complex situations. Ultimately, it is our imagination and ability to communicate with AI – from deciding on what questions to ask, how to formulate the prompts, what context to provide and how to evaluate the results – that unlocks its true potential (Kniberg, 2024). Generative AI is bound to make errors and have algorithmic biases that require scrutiny and governance that only humans are capable of. And for this to occur, we need to truly understand the strength and pitfalls of these large language models.

By embracing generative AI as a supportive assistant, or 'co-pilot', if you will, we can harness its capabilities to enhance our productivity and creativity. The more familiar we become with this technology – beyond the hype – as individuals, teams or organisations, the better equipped we will be to survive and thrive in the age of AI. And as we navigate the AI landscape in the years to come, healthy scepticism is encouraged, however we emphasise the importance of curiosity to explore the unfamiliar so that we can continue to adapt and thrive.

References

- Blackman R., & Grzankowski A. (2024). *AI Understands. A Little*. Retrieved July 8, 2024, from <https://www.reidblackman.com/podcast/ai-understands-a-little/>
- Brooks D. (2023). In the Age of A.I., Major in Being Human. *New York Times*. Retrieved April 13, 2024, from <https://www.nytimes.com/2023/02/02/opinion/ai-human-education.html>
- Cardon, P. W., Getchell, K., Carradini, S., Fleischmann, C., & Stapp, J. (2023). *Generative AI in the workplace: Employee perspectives of ChatGPT benefits and organizational policies*. <https://doi.org/10.31235/osf.io/b3ezy>
- Cremer, D. D., & Kasparov, G. (2021). AI should augment human intelligence, not replace it. *Harvard Business Review*. Retrieved April 13, 2024, from <https://hbr.org/2021/03/ai-should-augment-human-intelligence-not-replace-it>
- Elastic. (2024). *What is a large language model (LLM)?* Retrieved April 15, 2024, from <https://www.elastic.co/what-is/large-language-models>
- Galdo, M., Weichart, E. R., Sloutsky, V. M., & Turner, B. M. (2022). The quest for simplicity in human learning: Identifying the constraints on attention. *Cognitive psychology*, 138. <https://doi.org/10.1016/j.cogpsych.2022.101508>
- Google for Developers. (2023). *Introduction to Large Language Models*. Retrieved April 16, 2024, from <https://developers.google.com/machine-learning/resources/intro-llms>
- Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., & Amodei, D. (2020). Scaling Laws for Neural Language Models. *ArXiv.org*. <https://arxiv.org/pdf/2001.08361.pdf>
- Kniberg, H. (2024). *Generative AI in a Nutshell - how to survive and thrive in the age of AI*. Viewed April 13, 2024. <https://www.youtube.com/watch?v=2IK3DFHRFfw>
- Knight M. (2023). *How Generative AI will transform business*. Dataversity. Retrieved April 16, 2024, from <https://www.dataversity.net/how-generative-ai-will-transform-business/>
- Lakhani, K. R. (2023). AI won't replace humans – but humans with AI will replace humans without AI. *Harvard Business Review*. Retrieved April 13, 2024, from <https://hbr.org/2023/08/ai-wont-replace-humans-but-humans-with-ai-will-replace-humans-without-ai>
- Massey, N. (2013). Humans may be the most adaptive species. *Scientific American*. Retrieved April 15, 2024, from <https://www.scientificamerican.com/article/humans-may-be-most-adaptive-species/>

- Merisotis, J. (2023). In the age of artificial intelligence, we need our human skills to keep it real. *Forbes*. Retrieved April 13, 2024, from <https://www.forbes.com/sites/jamiemerisotis/2023/02/27/in-the-age-of-artificial-intelligence-we-need-our-human-skills-to-keep-it-real/>
- Metz, C., Kang, C., Frenkel, S., Grant, N., & Thompson, S. A. (2024). How Tech Giants Cut Corners to Harvest Data for A.I. *New York Times*. Retrieved April 15, 2024, from <https://www.nytimes.com/2024/04/06/technology/tech-giants-harvest-data-artificial-intelligence.html>
- Patel, N. (2024, February 16). *How AI copyright lawsuits could make the whole industry go extinct*. The Verge. <https://www.theverge.com/24062159/ai-copyright-fair-use-lawsuits-new-york-times-openai-chatgpt-decoder-podcast>
- Shah, M., Sureja, N. (2024). A Comprehensive Review of Bias in Deep Learning Models: Methods, Impacts, and Future Directions. *Archives of Computational Methods in Engineering*. <https://doi.org/10.1007/s11831-024-10134-2>
- Wei, J., Tay, Y., Bommasani, R., Raffel, C., Barret Zoph, Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E. H., Hashimoto, T., Vinyals, O., Liang, P., Dean, J., & Fedus, W. (2022). Emergent Abilities of Large Language Models. *arXiv.org*. <https://arxiv.org/abs/2206.07682>