



Image by rawpixel.com on Freepik, CC BY 4.0

The Rise of Artificial Intelligence: Balancing Innovation and Responsibility

Imagine a world where Artificial Intelligence (AI) chatbots, like ChatGPT, transcend their current role as mere tools and evolve into collaborators, trusted advisors, and even friends.

It's a vision that's both exciting and daunting, as the potential consequences of unethical AI could be catastrophic in a world where humans and machines operate as one. To prevent these negative outcomes, there are three key strategies that we must implement: transparency, awareness, and legislation.

But how can AI be unethical? Timothy McDermott, a Lead AI Engineer at Spark clarifies that AI algorithms aren't inherently biased, but rather, the data from which they are created can introduce biases that the AI may unknowingly amplify.

"Too often AI's reasoning and decision-making processes are not accounted for," states McDermott when asked about the leading cause of unethical AI use. McDermott explains that when developing AI for systems like credit checks, using protected characteristics like race or gender to make assumptions about someone's income can lead to biased and unfair conclusions that are often overlooked.

This means that without careful consideration of ethics during the design and implementation of AI or legislation guiding its use, we may perpetuate discrimination and inequality, amplifying existing biases and power imbalances. As we continue to integrate AI into our daily lives, these negative outcomes will only become more prevalent and impactful.

To ensure that AI is developed and used in a responsible and ethical manner, McDermott, as well as studies conducted by Riccardo Guidotti for [Association for Computing Machinery](#) and

This is a good opening sentence - id recommend starting a new paragraph so you don't bury this opening in a dense paragraph.

Reuben Binns for [Oxford](#), suggest there are three key strategies to consider: transparency, awareness and legislation.

Transparency

Transparency is vital as it encourages a healthy dose of scepticism towards AI outcomes, empowering users to proactively identify and challenge any unfair or misleading results. According to a survey conducted by Ipsos for the [World Economic Forum](#), 49% of people worldwide trust AI to make decisions. This statistic highlights the pressing need for transparency in AI decision-making processes. As without a clear understanding of how these algorithms arrive at their conclusions, users are left to blindly accept their output.

Awareness

Preventing biased data from being used to create AI models or mitigating biased outcomes in its output can be achieved by educating developers and businesses about the impact of design decisions on AI. For example, in an AI model used to select applicants for a job, we can require an equal number of female and male candidates to be selected, mitigating the amplification of gender bias. By being aware of and implementing these preventive measures, companies and developers can create AI systems that are more equitable and prevent the amplification of discrimination and inequality in existing and future AI systems.

Legislation

Enacting legislation is a crucial step to prevent the direct misuse of AI. For instance, the lack of legislation around mass surveillance and autonomous weapons creates a lack of accountability for any accidental or intentional harm and infringement of privacy. Several countries, including New Zealand, have already taken steps to regulate the appropriate use and privacy requirements of AI. Other countries must follow suit to ensure that AI is used in a safe and ethical manner.

As we continue to embrace the immense potential of AI to revolutionise our lives, it's important to remember that with great power comes great responsibility.

The rise of AI poses significant ethical challenges, from perpetuating biases to potential misuse. However, by promoting transparency, raising awareness, and enacting legislation, we can ensure that AI transforms our lives for the better.

Article Proposal

Topic

Presently, artificial intelligence (AI) is an increasingly popular technology with a broad range of applications in various industries. For my first project, I will investigate "Ethical considerations in the creation and deployment of artificial intelligence".

Angle

This article will concentrate on the ethical repercussions of AI and how programmers can help ensure that their products align with ethical principles. I will discuss the possible effects of using AI and the need to establish guidelines to prevent undesirable outcomes.

Furthermore, I will also emphasise the significance of diversity and inclusivity in AI development to avoid perpetuating biases and discrimination.

Supporting Evidence

I intend to conduct an interview with lead AI engineer Timothy Mcdermott to gain their professional viewpoints and gain a better understanding of the ethical issues facing the AI domain. I'll also use academic publications and studies to reinforce my arguments and findings.

Outlet

The outlet I propose for my article is the Verge, an established technology news website. Verge's audience is tech-savvy and knowledgeable, and they boast an excellent reputation for covering cutting-edge technological issues. The tone and style of my article will be tailored to The Verge's audience and will explain complex ideas in non-technical terms to be consistent with present offerings.

Target Audience

The target audience for my article is those who may be interested in the relationship between ethics and technology. The article will appeal to readers who want to improve their understanding of the societal implications of AI and expand their knowledge of how technology can be developed and deployed ethically.

Critical Reflection

In this assignment, I wrote an article aimed at technology publication 'Verge' and intended for a general audience who likely have a science/technology background and an interest in ethics. This reflection will address the highlights, difficulties, and outcome of the article and the process of its development.

Throughout the development of my article, there were several aspects that I found enjoyable and rewarding.

One of the highlights was conducting an interview with Timothy Mcdermott, which proved to be a valuable addition to the article as it reinforced the urgent need to address ethical concerns in Artificial Intelligence (AI) development. Most importantly the process taught me the key of a successful interview is preparing a diverse range of questions to avoid constraining the interviewee's viewpoint and to allow for a wider breadth of content to be discussed.

For instance, while I initially thought that biased data was the only ethical concern in AI, Mcdermott's insights raised new ideas.

By keeping an open mind, I was able to expand my understanding of the topic beyond my initial assumptions and create a more comprehensive article.

Participating in the workshop was another valuable aspect of the assignment. The feedback I received allowed me to reimagine how a general reader might consume the article and make modifications to make it more appealing to my audience. Through the workshop, I identified the strengths and weaknesses of my article by reading other people's articles and improved it based on the feedback received by my peers.

Some common critiques included the choice of an intimidating image for the topic, lack of common examples, and absence of subheadings.

In response, I simplified the image to be less confronting, added more relatable examples such as ChatGPT and gender bias, and included subheadings to highlight the three key strategies presented in the article. The workshop experience was instrumental in helping me refine my article and make it more reader-friendly.

Another aspect I enjoyed was transforming academic content into an engaging article. This task can be challenging because academic content is typically dense and technical, and it can be difficult to make it accessible to a broader audience.

However, by reworking the content and using creative approaches, I was able to create an article that was not only informative but also engaging and enjoyable for readers.

My success in this endeavour was evident in the positive feedback I received during the workshop. I was praised for my use of simple language and the effectiveness of how I communicated the ideas.

While there were many aspects of the project that I enjoyed, I found it challenging to navigate the word restriction.

The broad range of topics covered in the article made it challenging to decide which ideas to prioritise and discuss in detail. I had to consider which ideas were most essential to the overall message of the article and which ones could be explored within the limited space available.

This meant making difficult decisions about which ideas to omit or only briefly touch upon. Ultimately, I prioritised the most important concepts to present a comprehensive and engaging overview of the topic.

The strict word count for the article also made it difficult for me to include both examples and explanations of complex ideas while keeping the article concise.

This required me to strike a balance between providing enough detail to make the ideas understandable and staying within the word limit, often sacrificing certain details in favour of other essential content.

Hence in some places in the article such as the introduction I had to omit more common examples of AI uses and a more detailed discussion of unethical AI consequences which I felt may have weakened the communication of the article.

Overall, I am confident that my article is well-suited for the target audience and publication.

The focus on the topic of AI is timely and relevant to current events, and I made sure to present the information in a clear and accessible manner, using plenty of explanations and examples to aid understanding of this new technology.

The inviting image I chose helps to make the article more approachable, especially for readers who may feel uncomfortable with the development and popularisation of AI.

Additionally, despite the complex and broad nature of the subject matter, I was able to convey the key ideas in a concise and engaging manner, which I believe makes the article both informative and entertaining.

As a result, I am pleased with the final product and hope that it will be well-received by Verge's readers.

If I were to revisit this assignment, I would allocate more time to conduct extensive research on AI ethics to gain a more comprehensive understanding of the different perspectives/viewpoints.

By doing so, I would be better equipped to delve deeper into areas that I may have lacked knowledge during the interview and generate more valuable insights for the article.

Furthermore, if the word count allowed, I would elaborate on the potential implications of unethical AI in the introduction, providing specific examples of its impact on society.

I would also provide a more detailed explanation of the legislation that needs to be enforced and suggest practical ways to implement them in the key strategies section.

This would make the article more informative and actionable, providing readers with a clear and concise understanding of how to promote ethical AI.

By implementing these improvements, the article's overall quality would improve, making it more accessible and engaging to readers.

Source

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys*, 51(5), 93. <https://doi.org/10.1145/3236009>

Binns, R. (2018). *Fairness in Machine Learning: Lessons from Political Philosophy*. Department of Computer Science, University of Oxford, United Kingdom, 1-11, 81. <https://arxiv.org/pdf/1712.03586.pdf>

Myers, J. (2022). 5 charts that show what people around the world think about AI. World Economic Forum. <https://www.weforum.org/agenda/2022/01/artificial-intelligence-ai-technology-trust-survey/>