Artificial Intelligence and education: The role of librarians in promoting ethical and responsible AI

By Leon Furze

Snapshot

Leon Furze, an educator currently exploring the world of AI and education for his PhD, considers the evolving role school librarians can play in supporting teachers and students as they navigate ongoing developments in the area of artificial intelligence.

A tsunami of Artificial Intelligence applications is about to wash over our schools: Lucky for us we have librarians.

Before starting my PhD in AI and education, I was a secondary school leader and English teacher. I've taught for over fifteen years, and in that time, I have been a Head of English and Digital, and a council member for the Victorian Association for the Teaching of English. Throughout my time in the classroom there was one constant: the support of school library staff and teacher librarians in helping navigate technological changes.

Throughout my time in the classroom there was one constant: the support of school library staff and teacher librarians in helping navigate technological changes.

The role of the librarian has changed throughout history, leaning much more towards a technology support role in recent years (Wine, 2016). As schools have adopted wave after wave of new technology – from audio/visuals like VHS and DVDs through to Remote Learning technologies and now AI – librarians have worked to integrate the technologies, provide support to teachers and students, and to teach responsible, ethical research skills.

With the release of OpenAI's ChatGPT in November 2022 it seems as though everything has changed and yet, from my perspective now working with teachers, librarians, and academics, some important things have stayed the same. In a recent podcast with UK librarian and educator Elizabeth Hutchinson, we spoke about how the inquiry skills and ethical understandings of librarians make them well placed to navigate this rapidly developing technology.

Understanding AI

Artificial Intelligence comes in many forms. From the predictive algorithm which recommends your Netflix movies to the systems that guide self-driving cars and remote drones, AI has become a large part of our lives. ChatGPT, however, represented a watershed moment in the public perception of AI. Suddenly, every newspaper from the Herald Sun to the Guardian, local news to international was reporting on the AI chatbot. Many reports branded the application a 'cheating machine', and in January and February – before many students had even returned to classes

and lecture halls – there was rampant speculation about students using ChatGPT to write their essays and assignments.

Secondary school systems and universities have reacted differently, ranging along a spectrum from banning and blocking the application to writing new policies to allow and even encourage the technology's use. Much of the din around cheating has started to abate, but the pace of development in these AI apps has only accelerated. Towards the end of February, Microsoft started to release its new Bing chat and Google has teased its own chatbot, named Bard.

At the start of March, OpenAI released its ChatGPT model to developers, meaning that anyone can integrate the wildly successful chatbot into their own apps. For a much cheaper price than its previous models, OpenAI has basically made the technology underpinning its most successful project available to anyone with even a modicum of programming knowledge. Schools need to be prepared for the oncoming tsunami of applications – some of which may be legitimate, many of which will be simply capitalising on the AI hype.

There are also serious ethical concerns with the AI industry as a whole. I wrote a blog post about AI ethics at the start of the year and have since started to break down each of the nine areas I highlighted into more detail. Here is a summary of the areas I have identified in my research, and from books like Kate Crawford's incredible Atlas of AI:

Bias: AI systems are inherently biased due to the data they are trained on, which can contain discriminatory language from sources like Reddit and Twitter. Biases can also come from the methods of training and reinforcement used when developing the AI systems. Predictive policing algorithms, for example, have been found to disproportionately target certain communities, reinforcing systemic biases. Because the large dataset of AI is from a predominately white, male, heterosexual and Anglocentric perspective, the output of AI can further reinforce societal biases.

Environmental impact: The technology industry has a significant impact on the environment, with the mining and refining of rare earth minerals used in devices like smartphones and laptops leading to soil erosion, water pollution, and greenhouse gas emissions. Cloud computing, which is increasingly used for AI computing, relies on data centres and infrastructure that consume energy and produce waste. Companies have pledged to make their data centres carbon neutral, but this often involves carbon-trading or offsetting schemes rather than actual reductions in emissions, leading to criticisms of 'greenwashing.'

Academic integrity and 'truth': There are concerns that AI language models like ChatGPT will be used to cheat in academic settings, raising questions about academic integrity and honesty. However, it is still unclear to what extent using AI constitutes cheating, as the output of the model is an original creation generated 'probabilistically.' There are also concerns that AI will be used to produce fake news or deliberately harmful

...it is still unclear to what extent using AI constitutes cheating, as the output of the model is an original creation generated 'probabilistically.'

media, either unintentionally due to current flaws in most language models or intentionally for malicious purposes.

Copyright and intellectual property: AI image generators like Midjourney, Stable Diffusion and DALL-E 2 have been criticised for infringing on artists' intellectual property rights by using their 'styles' in AI image generation without permission. Similarly, language models like ChatGPT incorporate large amounts of other writers' work, which raises questions about who owns the copyright to materials produced by AI systems.

ChatGPT incorporate large amounts of other writers' work, which raises questions about who owns the copyright to materials produced by AI systems.

Privacy and security: Privacy is a major concern in the development and use of AI systems, with concerns about the collection and use of personal data, data breaches, and lack of transparency in AI decision-making. Facial recognition technology has been criticized for its potential to violate individuals' privacy and civil rights, with higher error rates for people with darker skin tones and use in targeting and monitoring marginalized communities. Targeted advertising based on personal data raises concerns about data privacy, data breaches, and the use of personal data for commercial gain.

Data collection and 'datafication': The collection of data for AI algorithms raises concerns about data privacy, exploitation, and bias. 'Datafication' turns all aspects of our lives into data points, which can lead to exploitation as users become commodities generating capital for platform owners. 'Big Data' contributes to issues such as bias and discrimination, as not everyone has access to the technologies producing the data.

Affect recognition: Affect recognition, or interpreting emotions through facial expressions, body language, speech patterns, and actions, is a controversial practice with accuracy concerns and questions about whether emotions should be 'datafied' at all. Built into surveillance technology, it raises privacy concerns, and like bias, it can perpetuate discrimination, as an algorithm trained to identify possible 'terrorist behaviour' resulted in racial profiling.

Human labour: There are concerns about AI replacing human jobs in white-collar industries, but current AI systems rely on low-paid human labour, as seen in the harsh conditions of Kenyan workers employed by OpenAI to label inappropriate data for language models. These workers were paid dollars a day to label and classify harmful text including violent and abusive language, in order to 'clean up' the models output for users. In a similar fashion, companies like Amazon use low paid human workers to label data for AI systems like its Alexa voice assistant.

There are concerns about AI replacing human jobs in white-collar industries, but current AI systems rely on low-paid human labour...

Power and hegemony: The data AI models are built on represents a static worldview that encodes existing power and hierarchies in society, which can further oppress and marginalises disadvantaged people. AI reinforces global hegemonies, as powerful AI is concentrated in the hands of those who already have the most, entrenching the divide between wealthy and poor countries. Even efforts to make 'fair' synthetic data have still been found to reproduce biases.

AI reinforces global hegemonies, as powerful AI is concentrated in the hands of those who already have the most...

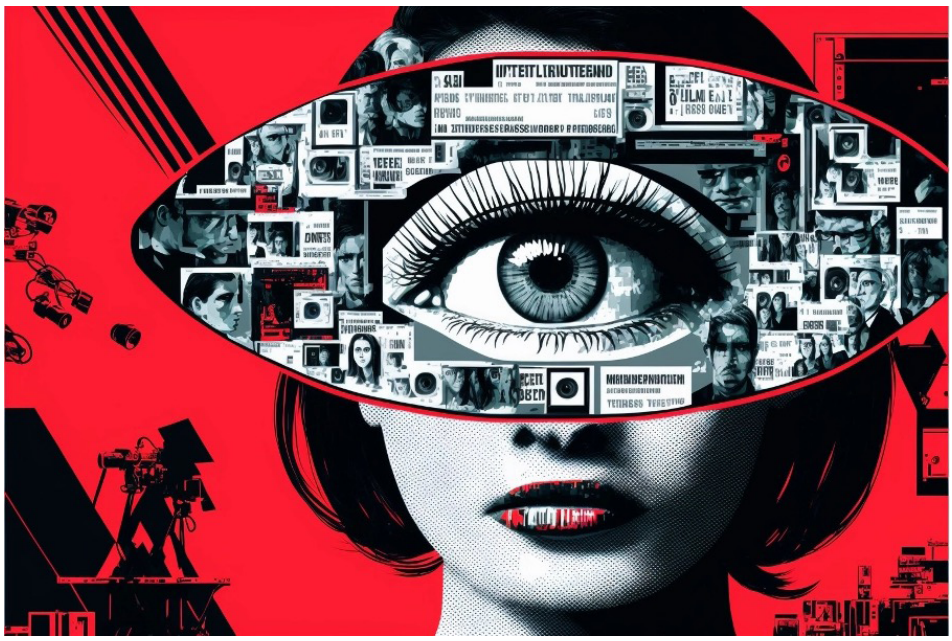

Image by the author in Midjourney AI image generator. Image generation prompt: dramatic feature article head image collage of surveillance technologies. red, white, and black. in the style of an editorial header image. Techno. CCTV. Privacy and data breaches. Digital collage. --ar 3:2 --q 2 --v 4

How librarians support the integration of AI technologies

Looking at the ethical concerns and based on my own experience of working closely with librarians and teacher librarians, I can see several ways in which library staff can support schools and help navigate the AI storm. Many of these are, in fact, things which school librarians are already doing:

Building ethical understanding

By fostering ethical reflection, librarians can help to ensure that AI is deployed in ways that align with the values of the school community.

Librarians and teacher librarians can play a crucial role in building ethical understanding by leading discussions and creating resources that help students and teachers engage with ethical questions around AI. This could involve everything from leading workshops on bias in datasets to curating collections of literature that explore ethical themes related to AI. By fostering ethical reflection, librarians can help to ensure that AI is deployed in ways that align with the values of the school community. I have found school libraries to be

some of the most inclusive, diverse, and welcoming places within schools. Library staff are often champions of the kinds of marginalised communities who are disadvantaged by AI bias. Promoting an ethical approach to AI is an extension of this great work.

Teaching responsible research practices

As experts in research, librarians can also help to teach responsible research practices that take account of the limitations of AI. This could involve teaching students how to critically evaluate the sources of information they encounter, including AI-generated content. AI has a reputation for producing false and misleading content. ChatGPT in particular has no 'ground truth': no access to a source for checking the voracity of its output. In fact, using ChatGPT as a research tool is highly problematic. There are, however, other applications which use the technology and might be used for research – they still require a critical approach. Supporting research could also involve developing strategies for identifying and mitigating bias in research projects that use AI tools, for both staff and students. By embedding responsible research practices, librarians can help to ensure that AI is used in ways that are rigorous and reliable.

By embedding responsible research practices, librarians can help to ensure that AI is used in ways that are rigorous and reliable.

Guiding and curating apps and services

With so many AI-powered apps and services available, it can be difficult for schools to navigate the landscape and make informed decisions about which tools to use. Here, librarians can help by offering guidance on the selection and use of AI-powered apps and services. This could involve vetting apps for reliability and safety or providing recommendations for the most effective and appropriate tools. Librarians can play an important role in curating technologies and providing recommendations for the most reliable and effective AI tools for a given purpose. This could involve keeping abreast of new developments in AI, assessing their potential impact on schools, and making recommendations to teachers and school leadership.

Librarians can play an important role in curating technologies and providing recommendations for the most reliable and effective AI tools for a given purpose.

As the oncoming tsunami of AI apps and services washes over our schools, it's important to acknowledge the ethical concerns and to implement responsible research practices that align with the values of the school community. School librarians are uniquely positioned to support schools in navigating the ethical minefield of AI technologies. They can help to build ethical understanding, teach responsible research practices, and guide the use of AI-powered apps and services.

School librarians are uniquely positioned to support schools in navigating the ethical minefield of AI technologies.

Librarians and teacher librarians can also champion the needs of marginalised members of the community who are most at risk of being further disadvantaged by AI bias. The role of the school librarian has always been about more than simply providing access to books and information. As librarians continue to adapt and grow with the technological changes, they remain an invaluable resource for school staff and students in navigating the ever-changing landscape of AI. Together, with their knowledge and expertise, staff and students can ensure that AI is used responsibly and in ways that benefit all members of the school community.

References

Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.

Furze, L. (2023). *Teaching AI Ethics*. [online] Leon Furze. Available at: **https://leonfurze.com/2023/01/26/teaching-ai-ethics/**.

Wine, L. D. (2016). School librarians as technology leaders: An evolution in practice. *Journal of Education for Library and Information Science, 57*(2), 207-220.

Leon Furze *is a PhD student, experienced educator, consultant, and educational writer. He is author of 'Practical Reading Strategies' and Jacaranda's new 'English' series of textbooks, and a VCE assessor. Leon provides professional learning and strategic planning for curriculum, literacy, and digital technologies.*

Leon has taught English, Literature and Digital Technologies in Australia and the UK for over fifteen years and was formerly Director of Learning and Teaching at Monivae College Hamilton. In 2016 he completed his Master of Education at the University of Melbourne, focusing on how Professional Learning can mitigate the risk of burnout in teachers.

His PhD is focused on Artificial Intelligence in education, and particularly the impact large language models like GPT will have on writing.