Introduction

Author Arthur C. Clarke famously argued that in science fiction literature “any sufficiently advanced technology is indistinguishable from magic” (Clarke). On 30 November 2022, technology company OpenAI publicly released their Large Language Model (LLM)-based chatbot ChatGPT (Chat Generative Pre-Trained Transformer), and instantly it was hailed as world-changing. Initial media stories about ChatGPT highlighted the speed with which it generated new material as evidence that this tool might be both genuinely creative and actually intelligent, in both exciting and disturbing ways. Indeed, ChatGPT is part of a larger pool of Generative Artificial Intelligence (AI) tools that can very quickly generate seemingly novel outputs in a variety of media formats based on text prompts written by users. Yet, claims that AI has become sentient, or has even reached a recognisable level of general intelligence, remain in the realm of science fiction, for now at least (Leaver). That has not stopped technology companies, scientists, and others from suggesting that super-smart AI is just around the corner. Exemplifying this, the same people creating generative AI are also vocal signatories of public letters that ostensibly call for a temporary halt in AI development, but these letters are simultaneously feeding the myth that these tools are so powerful that they are the early form of imminent super-intelligent machines.

For many people, the combination of AI technologies and media hype means generative AIs are basically magical insofar as their workings seem impenetrable, and their existence could ostensibly change the world. This article explores how the hype around ChatGPT and generative AI was deployed across the first six months of 2023, and how these technologies were positioned as either utopian or dystopian, always seemingly magical, but never banal. We look at some initial responses to generative AI, ranging from schools in Australia to picket lines in Hollywood. We offer a critique of the utopian/dystopian binary positioning of generative AI, aligning with critics who rightly argue that focussing on these extremes displaces the more grounded and immediate challenges generative AI bring that need urgent answers. Finally, we loop back to the role of schools and educators in repositioning generative AI as something to be tested, examined, scrutinised, and played with both to ground understandings of generative AI, while also preparing today’s students for a future where these tools will be part of their work and cultural landscapes.

Hype, Schools, and Hollywood

In December 2022, one month after OpenAI launched ChatGPT, Elon Musk tweeted: "ChatGPT is scary good. We are not far from dangerously strong AI". Musk’s post was retweeted 9400 times, liked 73 thousand times, and presumably seen by most of his 150 million Twitter followers. This
type of engagement typified the early hype and language that surrounded the launch of ChatGPT, with reports that “crypto” had been replaced by generative AI as the “hot tech topic” and hopes that it would be “transformative for business” (Browne). By March 2023, global economic analysts at Goldman Sachs had released a report on the potentially transformative effects of generative AI, saying that it marked the “brink of a rapid acceleration in task automation that will drive labor cost savings and raise productivity” (Hatzius et al.). Further, they concluded that “its ability to generate content that is indistinguishable from human-created output and to break down communication barriers between humans and machines reflects a major advancement with potentially large macroeconomic effects” (Hatzius et al.). Speculation about the potentially transformative power and reach of generative AI technology was reinforced by warnings that it could also lead to “significant disruption” of the labor market, and the potential automation of up to 300 million jobs, with associated job losses for humans (Hatzius et al.). In addition, there was widespread buzz that ChatGPT’s “rationalization process may evidence human-like cognition” (Browne), claims that were supported by the emergent language of ChatGPT. The technology was explained as being “trained” on a “corpus” of datasets, using a “neural network” capable of producing “natural language” (Dsouza), positioning the technology as human-like, and more than ‘artificial’ intelligence. Incorrect responses or errors produced by the tech were termed “hallucinations”, akin to magical thinking, which OpenAI founder Sam Altman insisted wasn’t a word that he associated with sentience (Intelligencer staff). Indeed, Altman asserts that he rejects moves to “anthropomorphize” (Intelligencer staff) the technology; however, arguably the language, hype, and Altman’s well-publicised misgivings about ChatGPT have had the combined effect of shaping our understanding of this generative AI as alive, vast, fast-moving, and potentially lethal to humanity.

Unsurprisingly, the hype around the transformative effects of ChatGPT and its ability to generate ‘human-like’ answers and sophisticated essay-style responses was matched by a concomitant panic throughout educational institutions. The beginning of the 2023 Australian school year was marked by schools and state education ministers meeting to discuss the emerging problem of ChatGPT in the education system (Hiatt). Every state in Australia, bar South Australia, banned the use of the technology in public schools, with a “national expert task force” formed to “guide” schools on how to navigate ChatGPT in the classroom (Hiatt). Globally, schools banned the technology amid fears that students could use it to generate convincing essay responses whose plagiarism would be undetectable with current software (Clarence-Smith). Some schools banned the technology citing concerns that it would have a “negative impact on student learning”, while others cited its “lack of reliable safeguards preventing these tools exposing students to potentially explicit and harmful content” (Cassidy). ChatGPT investor Musk famously tweeted, “It’s a new world. Goodbye homework!”, further fueling the growing alarm about the freely available technology that could “churn out convincing essays which can’t be detected by their existing anti-plagiarism software” (Clarence-Smith). Universities were reported to be moving towards more “in-person supervision and increased paper assessments” (SBS), rather than essay-style assessments, in a bid to “out-manoeuvre ChatGPT’s plagiarism potential. Seven months on, concerns about the technology seem to have been dialled back, with educators more curious about the ways the technology can be integrated into the classroom to good effect (Liu et al.); however, the full implications and impacts of the generative AI are still emerging.

In May 2023, the Writer’s Guild of America (WGA), the union representing screenwriters across the US creative industries, went on strike, and one of their core issues were “regulations on the use of artificial intelligence in writing” (Porter). Early in the negotiations, Chris Keyser, co-chair of the WGA’s negotiating committee, lamented that “no one knows exactly what AI’s going to be, but the fact that the companies won’t talk about it is the best indication we’ve had that we have a reason to fear it” (Grobar). At the same time, the Screen Actors’ Guild (SAG) warned that members were being asked to agree to contracts that stipulated that an actor’s voice could be re-used in future scenarios without that actor’s additional consent, potentially reducing actors to a dataset to be animated by generative AI technologies (Scheibel and Koblin). In a statement issued by SAG, they made their position clear that the creation or (re)animation of any digital likeness of any part of an actor must be recognised as labour and properly paid, also warning that any attempt to legislate around these rights should be strongly resisted (Screen Actors Guild). Unlike the more sensationalised hype, the WGA and SAG responses to generative AI are grounded in labour relations. These unions quite rightly fear the immediate future where human labour could be augmented, reclassified, and exploited by, and in the name of, algorithmic systems. Screenwriters, for example, might be hired at much lower pay rates to edit scripts first generated by ChatGPT, even if those editors would really be doing most of the creative work to turn something clichéd and predictable into something more appealing. Rather than a dystopian world where machines do all the work, the WGA and SAG protests railed against a world where workers would be paid less because executives could pretend generative AI was doing most of the work (Bender).

The Open Letter and Promotion of AI Panic

In an open letter that received enormous press and media uptake, many of the leading figures in AI called for a pause in AI development since “advanced AI could represent a profound change in the history of life on Earth”; they warned early 2023 had already seen “an out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict, or reliably control” (Future of Life Institute). Further, the open letter signatories called on “all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4”, arguing that “labs and independent experts should use this pause to jointly develop and implement a set of shared safety protocols for advanced AI design and development that are rigorously audited and overseen by independent outside experts” (Future of Life Institute). Notably, many of the signatories work for the very companies involved in the “out-of-control race”. Indeed, while this letter could be read as a moment of ethical clarity for the AI industry, a more cynical reading might just be that in warning that their AIs could effectively destroy the world, these companies were positioning their products as seemingly magical—“digital minds that no one – not even their creators – can understand”—making them even more appealing to potential customers and investors. Far from pausing AI development, the open letter actually operates as a neon sign touting the amazing capacities and future brilliance of generative AI systems.

Nirit Weiss-Blatt argues that general reporting on technology industries up to 2017 largely concurred with the public relations stance of those companies, positioning them as savours and amplifiers of human connection, creativity, and participation. After 2017, though, media reporting completely shifted, focussing on the problems, risks, and worst elements of these corporate platforms. In the wake of the open letter, Weiss-Blatt extended her point on Twitter, arguing that media and messaging surrounding generative AI can be broken down into those who are profiting and fuelling the panic at one end of the spectrum, and those who think the form of the panic (which positions AI as dangerously intelligent) is deflecting from the immediate real issues caused by generative AI at the other. Weiss-Blatt characterises the Panic-as-a-Business proponents as arguing “we’re telling you will all die from a Godlike AI... so you must listen to us”, which coheres with the broader positioning narrative of generative AI’s seemingly magical (and thus potentially destructive) capabilities. Yet this rhetoric also positions the companies creating generative AI as the ones who should be making the rules to control it, an argument so effective that in July 2023 the Biden Administration in the US endorsed the biggest AI companies—Amazon, Anthropic, Google, Inflection, Meta, Microsoft, and OpenAI—framing future AI development with voluntary safeguards rather than externally imposed policies (Shear, Kang, and Sanger).
Artificial Intelligences have inhabited popular imaginaries via novels, television, and films far longer than they have been considered even potentially viable technologies, so it is not surprising that popular culture has often framed the way AI is understood (Leaver). Yet as Emily Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell argue, Large Language Models and generative AI are most productively understood as "a stochastic parrot" insomuch as each is a "system for haphazardly stitching together sequences of linguistic forms it has observed in its vast training data, according to probabilistic information about how they combine, but without any reference to meaning" (Bender et al. 617). Generative AI, then, is not creating something genuinely new, but rather remixing existing data in novel ways that the systems themselves do not in any meaningful sense understand. Going further, Simone Natale characterises current AI tools as "deceitful media" insomuch as they are designed to deliberately appear generally intelligent, but this is always a deception. The deception makes these tools more engaging for humans to use but is also fundamental in selling and profiting from the use of AI tools. Rather than accepting claims made by the companies financing and creating contemporary AI, Natale argues for a more pedagogically productive path:

we must resist the normalization of the deceptive mechanisms embedded in contemporary AI and the silent power that digital media companies exercise over us. We should never cease to interrogate how the technology works, even while we are trying to accommodate it in the fabric of everyday life. (Natale 132)

Real Issues

Although even a comprehensive list is beyond the scope of this short article, it is nevertheless vital to note that in looking beyond the promotion of AI Panic and deceptive media, ChatGPT and other generative AI tools create or exacerbate a range of very real and significant ethical problems. The most obvious problem is the lack of transparency in terms of what data different generative AI tools were trained on. Generally, these tools are thought to get better by absorbing ever greater amounts of data, with most AI companies acknowledging that scraping the Web in some form has been part of the training data harvesting for their AI tools. Not knowing what data have been used makes it almost impossible to know which perspectives, presumptions, and biases are baked into these tools. While many forms of bias have plagued technology companies for many years (Noble), for generative AI tools, in “accepting large amounts of web text as ‘representative’ of ‘all’ of humanity we risk perpetuating dominant viewpoints, increasing power imbalances, and further reifying inequality” (Bender et al. 614). Even mitigating and working to correct biases in generative AI tools will be a huge challenge if these companies never share what was in their training data.

As the WGA and SAG strike discussed above emphasises, the question of human labour is a central challenge for generative AI. Beyond Hollywood, more entrenched forms of labour exploitation haunt generative AI. Very low-paid workers have done much of the labour in classifying different forms of data in order to train AI systems; data workers are routinely not acknowledged at all, even sometimes directly performing the tasks that are ascribed to AI, to the extent that “distracted by the specter of nonexistent sentient machines, an army of precarized workers stands behind the supposed accomplishments of artificial intelligence systems today” (Williams, Miceli, and Gebru). It turns out that people are still doing the work so that companies can pretend the machines can think.

In one final but very important example, there is a very direct ecological cost to training, maintaining, and running generative AI tools. In the context of global warming, concerns already existed about the enormous data centres at the heart of the big technology platforms prior to ChatGPT’s release. However, the data and processing power needed to run generative AI tools are even larger, leading to very real questions about how much electricity and water (for cooling) are used by even the most rudimentary ChatGPT queries (Lizarraga and Solon). While not just an AI question, balancing the environmental costs of data centres with the actual utility of AI tools is not one that is routinely asked, or answered, in the hype around generative AI.

Messing Around and Geeking Out

Escaping the hype and hypocrisy deployed by AI companies is vital for repositioning generative AI not as magical, not as a saviour, and not as a destroyer, but rather as a new technology that needs to be critically and ethically understood. In seminal work exploring how young people engage with digital tools and technologies, Mimi Ito and colleagues developed three genres of technology participation: hanging out, where engagement with any technologies is largely driven by friendships and social engagement; messing around, which includes a great deal of experimentation and
play with technological tools; and geeking out, where some young people will find a particular focus on one platform, tool or technology that inspires them to focus enough to develop expertise in using and understanding that tool (Ito et al.). If young people, in particular, are going to be living in a world where generative AI tools are part of their social worlds and workplaces, then messing around with ChatGPT is, indeed, going to be important in testing out how these tools answer questions and synthesise information, what biases are evident in responses, and at what points answers are incorrect. For some young people, they may well move from messing around to completely geeking out with generative AI, a process that will be even more fruitful if these tools are not seen as impenetrable magic, but rather as commercial tools built by for-profit companies. While the idea of digital natives is an unhelpful myth (Bennett, Maton, and Kervin), if young people are going to be the first generation to have generative AI as part of their information, creative, and search landscapes, then safely messing around and geeking out with these tools will be more vital than ever.

We mentioned above that most Australian state education departments initially banned ChatGPT, but a more optimistic sign arrived as we were finishing this article insomuch as the different Australian states agreed in mid-2023 to work together to create “a framework to guide the safe and effective use of artificial intelligence in the nation’s schools” (Clare). Although there is work to be done, moving away from a ban to a setting that should allow students to be part of testing, framing, and critiquing ChatGPT and generative AI is a clear step in repositioning these technologies as tools, not magical systems that could never be understood.

Conclusion

Generative AI is not magic; it is not a saviour or destroyer; it is neither utopian nor dystopian; nor, unless we radically narrow the definition, is it intelligent. The companies and corporations driving AI development have a vested interest in promoting fantastical ideas about generative AI, as it drives their customers, investment, and future viability. When the hype is dominant, responses can be overdetermined, such as banning generative AI in schools. But in taking a less magical and more material approach to ChatGPT and generative AI, we can try and ensure pedagogical opportunities for today’s young people to test out, scrutinise, and critically understand the AI tools they are most likely going to be asked to use today and in the future. The first wave of generative AI hype following the public release of ChatGPT offers an opportunity to reflect on exactly what the best uses of these technologies are, what ethics should drive those uses, and how transparent the workings of generative AI should be before their presence in the digital landscape is so entrenched and mundane that it becomes difficult to see at all.

Acknowledgment

This research was supported by the Australian Research Centre Centre of Excellence for the Digital Child through project number CE200100022.

References


This research was supported by the Australian Research Centre Centre of Excellence for the Digital Child through project number CE200100022.
ChatGPT Isn't Magic | M/C Journal

30/03/2023, 22:29


Author Biographies

Tama Leaver, Curtin University

Tama Leaver is a Professor of Internet Studies at Curtin University in Perth, Western Australia and expert media commentator. He is the Vice-President of the (international) Association of Internet Researchers (AoIR) and a Chief Investigator in the ARC Centre of Excellence for the Digital Child.

His research interests include digital childhood and infancy online, visual social media, social media, digital death, mobile gaming and the changing landscape of media distribution. He has published in a number of journals including Popular Communication, Media International Australia, First Monday, Comparative Literature Studies, Social Media and Society, Communication Research and Practice and the Fibreculture journal. He is the author of Artificial Culture: Identity, Technology and Bodies (Routledge, 2012); co-editor of An Education in Facebook? Higher Education and the World’s Largest Social Network (Routledge, 2014) with Mike Kent; and Social, Casual and Mobile Games: The Changing Gaming Landscape (Bloomsbury Academic, 2016) with Michele Willson; co-author of Instagram: Visual Social Media Cultures (Polity, 2020) with Tim Highfield and Crystal Abiden; and co-editor of The Routledge Companion to Digital Media and Children (Routledge, 2021) with Lelia Green, Donell Holloway, Kylie Stevenson and Leslie Haddon.

He has been awarded teaching awards from the University of Western Australia, Curtin University, and in 2012 received a national Australian Award for Teaching Excellence in the Humanities and the Arts.

Suzanne Srdarov, Curtin University

Dr Suzanne Srdarov teaches in media and digital cultures, and has a particular interest in media cultures and their intersections with gender. She is a research assistant in the ARC Centre of Excellence for the Digital Child.

License

Copyright (c) 2023 Tama Leaver, Suzanne Srdarov

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Authors who publish with this journal agree to the following terms:

1. Authors retain copyright and grant the journal right of first publication with the work simultaneously licenced under a Creative Commons Attribution - Noncommercial - No Derivatives 4.0 Licence that allows others to share the work with an acknowledgement of the work's authorship and initial publication in this journal.
2. Authors are able to enter into separate, additional contractual arrangements for the non-exclusive distribution of the journal's published version of the work (e.g., post it to an institutional repository or publish it in a book), with an acknowledgement of its initial publication in this journal.
3. Authors are permitted and encouraged to post their work online (e.g., in institutional repositories or on their website) prior to and during the submission process, as it can lead to productive exchanges, as well as earlier and greater citation of published work (see The Effect of Open Access).

Supported by

Copyright © M/C, 1998-2023 ISBN 1441-2616

About M/C | Contact M/C | Accessibility