

Black-box creativity and generative artificial intelligence

[Luke Tredinnick](#) and [Claire Laybats](#)

The rapid rise of generative artificial intelligence applications poses a number of significant challenges for the ways in which we organise our social, cultural and working lives. As we have explored in *Business Information Review* over recent years, new technologies promise to transform existing workplace routines, encroaching on professional roles just as it has already encroached on routine occupations ([Kirkwood, 2018](#); [Richardson, 2020](#); [Tredinnick, 2017](#)). Artificial Intelligence (AI) poses new problems in the evaluation of information and news, and in understanding the role of disinformation in social discourse ([Tredinnick and Laybats, 2023](#)). Artificial Intelligence facilitates new forms of cybercrime ([Tredinnick and Laybats, 2023](#)) and poses new ethical dilemmas about the use of technology ([Carter, 2018](#); [2020](#)). But while we have tended to focus on the ways in which technology may transform the contexts within which we live and work, generative artificial intelligence has also been silently rewiring how we conceptualise the human voice, agency, and creativity. This more subtle challenge is in many ways of greater profundity. Creativity has been widely understood as a distinctly human characteristic ([Carruthers, 2012](#)); other species can be said to engage in creative behaviour but none of this behaviour approaches the complexity and productivity of human creativity ([Gabora and Kaufman, 2010](#)). Yet as we move into an era where creative acts fall ever more comfortably within the province of generative AI, the ways in which we think about creativity and its relationship to the human mind is bound to shift.

It is a truism that emergent technologies are generally met with a mixture of enthusiasm about their transformative potential, and apprehension about their negative social effects. The expansion of printing and literacy in the 18th century and steam press of the 19th were met by widespread misgivings about the impact of popular literature on the morals and behaviour of a newly literate mass public. Brantlinger for example has written that for many late-Victorian intellectuals:

mass culture and mass-literacy themselves threaten a sort of cultural entropy or abjection, the swamps or sewers of mediocrity or vulgarity into which, they feared, excellence – high intelligence, literary and artistic aura – was sinking ([1998](#): 180)

[Black, 1996](#) has explored some of the responses to the Public Libraries Act (1850) and their concerns with public morals. Radio broadcasting was met by fears about the “chaos of the ether” arising from “a jumble of signals” and a “blasting and blanketing of rival programmes” ([Briggs, 1995](#): 64). Television was likewise accompanied by fears about the erosion of family life and the degrading of social values; the Pilkington report for example stated that mass appeal programmes were “vapid and puerile, their content often derivative, repetitious and lacking any real substance” ([HMSO, 1960](#)). More recently the World Wide Web was accompanied by a mixture of utopian and dystopian perspectives. The media theorist Henry Jenkins for example celebrated the potential for a new more participatory culture ([2008](#)), and Lawrence Lessig the potential of digital reproduction to re-negotiate our relationship to the creative act ([2001](#)), but the tech anti-evangelist Andrew Keen opined about social media’s

potential to replace expertise and insight with an army of amateurs ([Keen, 2007](#)). Concerns about the negative effects of technological change are invariably rooted in valid questions, nevertheless the observable pattern has been that the social threat of new technology is often overstated.

The response to generative AI has followed a similar pattern, but with an unusually heavy skew towards its potential negative social effects. Indeed there is an apocalyptic undertone to much of the recent coverage. This is in itself worthy of note. It may be that the challenges posed by AI are indeed more profound than those presented by prior technologies, that AI presents an existential risk, or that the unprecedented pace of change will outstrip our ability to legislate and regulate its use. Nevertheless given the wholesale reorganisation of social and economic structures following the wider adoption of network technologies in the 1990s, or the quiet emergence of ubiquitous information in the mobile revolution of the 2000s ([Tredinnick, 2022](#)), the idea of a qualitative difference driving this wider anxiety become more difficult to maintain. Generative AI is in its infancy and its true effects are likely to be relatively slow. Society will adapt to AI in ways that we cannot possibly currently imagine. The paroxysm of anxiety that has accompanied the wider awareness of generative AI may therefore not reflect a qualitative difference between the challenges posted by artificial intelligence compared to those of prior technological transformations, but instead the social context in which AI has emerged.

This latest technological revolution has emerged into a world that has been buffeted by a series of political, social and economic shocks since the turn of the millennium, from international terrorism, through the financial crisis, to the global environmental emergency, and still adjusting to the post-pandemic reality. Given this it is perhaps hardly surprising that the discourse around AI has channelled an apocalyptic bleakness. Indeed, Artificial intelligence has more than any other technological change often been associated with mythic representations. The discourse around AI has for example drawn on ideas of promethean transgression of knowledge stolen from the gods; the films [Transcendence \(2014\)](#) and [2001: A Space Odyssey \(1968\)](#) bookend this representation in US film. Popular representation of AI have also drawn parallels with the Pygmalion myth or the Golem of Jewish folklore, such as Maria in [Metropolis, 1927](#), or the gunslinger of [Westworld \(1973\)](#) each grotesque parodies of religious creation myths. And finally popular representations of AI draws on the myth of Narcissus as we become captivated by machines created in our own image. The apocalyptic sting of the discourse around generative AI today is therefore perhaps a mirror to our wider anxieties as we pass through a period of unusual instability towards a more uncertain future.

The emerging discourse around generative AI is perhaps most interesting as a mirror to our wider social anxieties. AI is on the one hand a series of technologies with currently still largely overstated applications in a variety of fields and contexts that are likely in the long term to profoundly alter the ways in which we process information and aspects of culture. But it is also a reflection of the ways in which we understand, value, and seek to model fundamental qualities of human nature, communication and the human mind. When we stare into the eyes of the machine we are confronted with our own image.

The arresting uncanniness of the current crop of generative AI is located in its ability to create apparently new content that superficially mirrors the kinds of deliberative choices that people make in the creative process. The fact that a machine can write, paint images, compose music, create code or do the many other things that were once characteristic of human creativity, and can do so with at least a passing resemblance to human creative practice, creates the

unnerving impression that there is something more to the machine than rigidly applied logical rules. The machine appears to have developed a ghost of its own. Of course this is an emotional response that we need to approach with a series of carefully considered caveats. What appears to be a step-change in the ways in which software function is in fact an incremental shift; there is nothing fundamentally different about the way that simulated neural networks work compared with traditional algorithmic processes. In fact whatever the remarkable nature of the outputs that they produce and however much that intrinsic nature is hidden behind training regimes, simulated neural networks remain algorithmic processes. Generative AI is as rule-bound as any other algorithmic process. It is simply that those rules are derived from the training data post hoc. The products of generative AI are therefore already implicit in the algorithmic process combined with the data on which it has been trained. The fundamental difference between generative AI and the computer processes of the past are not in the nature of the machines themselves, but in how we interpret its products, including the values, motivations, and meanings that we read into them. Nevertheless if a rigid rule-bound process can reproduce an albeit flawed verisimilitude of deliberative creativity, what does that say about how we understand the creative act itself? Two common responses have emerged in the creative sector to the emerging challenge of AI. The first is that the products of generative artificial intelligence are by definition mediocre, lacking the mark of individual expression that distinguishes the creative mind. This is largely true at the present. Large language models produce readable but uninspiring prose; image generators produce lively but cliched images; music generators produce derivative music. But not only are these apparent flaws sometimes deliberate - ChatGPT is designed to reproduce a flat style to inhibit the kinds of statements it might otherwise produce - but also reflect only the current state of affairs. It is important to consider not only what Generative AI can achieve now, but also the inevitable progress that will take place over the next 5, 10 or 15 years. And the future potential of the technology if it follows its current trajectory is truly breath-taking. A humanity located in the gaps in AI will inevitably prove to occupy an inexorably shrinking and indeed diminished domain.

The second common response is in many ways more interesting. Generative AI it is often suggested does not create anything fundamentally new, it just rearranges elements of human culture into different forms in a kind of simulacrum of creativity. It follows that generative AI lacks true creativity as we normally understand it. In a superficial sense this is true; the products of generative artificial intelligence are re-combinations of those existing cultural elements of culture on which it has been trained. A large language model learns how language works by observing large volumes of written text and deriving what words are most likely to follow on from what has already been written. This is a purely mechanical response that pays no heed to questions of meaning, significance, or style that might be at the forefront of a human writer. An AI image generator works by learning from large image bases what kinds of visual components are associated with particular prompts, or how elements of existing images can be seamlessly blended together to make them look similar to images that are already common. These technologies are indeed not creative in the sense in which we normally use the world. As the creative sphere increasingly relies on generative AI, we are perhaps condemning ourselves to endlessly recycle the culture of the past in an increasingly homogenised culture.

But this begs the question of how we understand creativity and the creative act in the first place. What do we understand the writer to be doing in the writing process, or the artist to

be doing in the composition of a new work. Where is the originality in the original creative act actually located? Is there something fundamentally different in human creative acts that are beyond the emulation of computer programmes?

Alan Turing is most famous for his work at Bletchley Park during the war, but his insights into machine learning in the 1950s have proven to be more prescient than they really had any right to be given the technological context of the age. While the Turing Test ([Turing, 1950](#)) no longer seems sufficient to demarcate human and machine intelligence, there is a philosophical underpinning to his work perhaps reflecting his own difficulty in understanding other people ([Hodges, 1983](#)) that is still relevant. Turing predicted that intelligent machinery would be met by “great opposition from the intellectuals who were afraid of being put out of a job” and added that “once the machine thinking method had started, it would not take long to outstrip our feeble powers” ([1951](#): 475). His imitation game – often known as the Turing Test - is an example of a black box process: what goes on inside the black box is not open to observation, the only thing that we can consider are its inputs and outputs. If the products of machine and human are indistinguishable, there is no reason to search for things to distinguish them; we may as well assume that broadly the same processes apply. John Searle’s Chinese Room thought experiment was by contrast an attempt to break-open Turing’s black box ([1984](#)). For [Searle \(1984\)](#) it very much mattered what goes on inside the box and what goes on inside the box influences our understanding of what is produced. A language processing model that simply applies rigid rules to an input to produce an acceptable and comprehensive output cannot be said to understand Chinese.

Current AI technologies based on artificial neural networks are black box processes. A black box is a system that can be understood only in terms of its inputs and outputs, rather than in terms of its internal processes. After they have been trained we feed into a generative AI engine a input prompt and receive an output text, image, or piece of music. But what has gone on inside the box to generate that apparently new work, or the ways in which it reflects and explicitly draws on the material used in training, remains opaque to inspection. This has generated a number of concerns about the ethics of generative AI: to what extent should the original creators of the training materials be credited for or recompensed for the products of generative ai, and to what extent should technology companies be allowed to profit on the creative work of others? Who do the products of a generative AI programme belong to? Are they even anything more than bricolage pastiches of prior creative work?

But the human mind is also a black box, and while we may identify the products of human creativity as original works, we do not know from where they derive or quite how they have arrived. We can sometimes recognise conscious or unconscious influences on the products of human creativity, and we might even judge some of that output as being wholly or significantly derivative of prior works. Nevertheless even the most startlingly original creative work is still the product of a black box process. The use of magnetic resonance imaging has given us a glimpse of the processes at work inside the human brain but no more than that; although we have an increasingly detailed map of the human brain we still have no very clear idea how cognition works, and in particular how consciousness emerges. The consequences of generative AI may be to cause us to question how we understand the creative process in the first place, and to what degree we identify in human creativity true originality.

If both human creativity and generative artificial intelligence are examples of black box processes, it is perhaps natural to draw a parallel. It is at least reasonable to describe the work of the artist as taking as input the entire artistic tradition combined with their own unique

experiences, and producing as output original creative works in response to particular conscious prompts, without imagining any extra-material process that maintains a sharp distinction between human and computer artistic practice. It is reasonable to think of writers taking as input the entire literary tradition and prior use of language to which they have been exposed in combination with their own unique experiences and producing as output original creative works in response to particular conscious prompts. It is also reasonable to local human creativity in reworking the influences to which an individual has been exposed through the filtering of the individually wired human brain to produce something new. Creativity is perhaps rooted not in originality *per se*, but in original and arresting re-compositions of existing cultural artefacts that are both recognisable as part of an ongoing tradition while contributing something new. This synthetic creativity is not automatically condemned to reproducing endless pastiches. Of course there is no doubt that the human brain is many orders of magnitude more complex than any current generative AI system, it is not altogether clear that this reflects a qualitative difference in the nature of creativity, rather than merely a quantitative difference.

Generative AI is therefore a form of technology that is unlike any that we have previously had to encounter, not because it works in a fundamentally new or different way, but because it challenges us to rethink things we have assumed to be uniquely human. The anxiety in that challenge is that it strips us of our humanity. While generative AI as it currently stands is but a shadow of the human mind, it does challenge us to confront the anthropocentric assumptions that run through culture. It is wrong to assume that human creativity is intrinsically superior by virtue of originating in the human mind, of that is somehow located in creative acts that are more real or authentic. More is different.

In [1989](#), Penrose used the metaphor of the emperor's new mind to describe the apparently absent mind of the computing machine. Penrose argued that consciousness must be non-algorithmic and therefore impossible to simulate in a Turing machine. But perhaps it is not the mind of the machine that is dressed in invisible clothing. As we embark on another new era of technology, we may find that generative AI eventually forces us to shed the emperors clothes of human exceptionalism rather than recognise the nakedness of machine-minds as Penrose implied. If this is a loss, it is a loss of an unearned belief in human distinctiveness, but one accompanied by an almost unlimited potential to unleash new forms of creative practice. The creative acts of the machine are after all still creative acts originating in the human mind; generative AI is an extension of human intelligence not a replacement for it. The subtle challenge of AI will therefore not be in replacing human intellectual labour – although that will certainly happen in many as yet unthought of contexts – but in requiring a renegotiation of the value that we attribute to intellectual labour in the first place.

September issue of Business Information Review

We are decades away from generative AI truly challenging human creativity, although that will come. Nevertheless until then there are more pressing issues presented by emerging technologies. This issue of Business Information Review explores some of these issues in two research papers that explore emerging issues of artificial intelligence. The first if these is written by editorial board member Ayinde Lateef and addresses the role of large language models in management. Entitled "ChatGPT as an Important Tool in Organizational Management: A Review of the Literature" Lateef notes that "LLM AI models such as ChatGPT

are groundbreaking technologies that transform communication, illustration, and creation processes in unprecedented ways, marking a significant departure from traditional approaches” and explores ChatGPT’s impact on organizational management, providing insights to those aiming to navigate the complexities and harness the benefits of this transformative technology. The second paper is entitled “Unleashing the potential of Chatbots in Business: A Bibliometric Analysis” and was contributed by Sivakumar Alur. It highlights the opportunities for research in areas like chatbot design cues, user experience, chatbots in collaboration with human agents, and chatbot ethics and privacy.

In addition September’s issue of *Business Information Review* also sees a number of other papers focussed on more general issues. We are delighted to have Martin White returning to our pages with a highly valuable paper exploring work-arounds and Shadow-IT, focussing on the use of workarounds in the digital workplace and their potential impact on the level of trust in enterprise information. Mostafa Sayyadi and Michael Provitera have contributed an opinion article on becoming a Digital Era Ready Company. And finally this issue’s out-of-the-box column deals with the growing threat of information warfare, particularly issues associated with orchestrated disinformation campaigns, and addresses the role of information professionals in mitigating that risk.

Finally we are delighted to announce the best paper prize winner. This year the panel voted for Dominique Poole-Avery’s paper “Arup’s Knowledge & Information Handbook: A case study in knowledge management good practice” published in June 2022. This was a case study exploring the development of Arup’s Knowledge & Information Handbook, a digital guide to support a global organisation in making the best use of its systems, tools, and knowledge sharing resources to provide access to the right knowledge, information and people at the point of need.

References

2001: *A Space Odyssey* (1968) *Kibrick, S* [Film]. United Kingdom and United States: Metro-Goldwyn-Mayer.

Black A (1996) *A New History of the English Public Library: Social and Intellectual Contexts 1850 - 1914*. London and New York: Leicester University Press.

Bratlinger P (1998) *The Reading Lesson: The Threat of Mass Literacy in 19th Century British Fiction*. Bloomington; Indianapolis: Indiana University Press.

Briggs A (1995) *The History of Broadcasting in the United Kingdom, Volume I: The Birth of Broadcasting*. London & Oxford: Oxford University Press.

Carruthers P (2012) The distinctly-human mind: the many pillars of cumulative culture. In: Hatfield G, Pittman H (eds), *Evolution of Mind, Brain and Culture*. Philadelphia: University of Pennsylvania Press, Inc: 325–345.

Carter D (2018) How real is the impact of artificial intelligence the business information survey 2018. *Business Information Review* 35(3): 99–115.

Carter D (2020) Regulation and ethics in artificial intelligence and machine learning technologies: where are we now? Who is responsible? Can the information professional play a role? *Business Information Review* 37(2): 60–68.

Gabora L, Kaufman SB (2010) Evolutionary perspectives on creativity. In: Kaufman J, Sternberg R (eds), *The Cambridge Handbook of Creativity*. Cambridge, UK: Cambridge University Press, 279–300.

HMSO (1960) *Report of the Committee on Broadcasting, CMND 1753*. London: HMSO.

Hodges A (1983) *Alan Turing: The Enigma*. London: Burnett Books.

Jenkins H (2008) *Convergence Culture: Where Old and New Media Collide*. United Kingdom: NYU Press.

Keen A (2007) *The Cult of the Amateur: How Today's Internet Is Killing Our Culture and Assaulting Our Economy*. United Kingdom: Nicholas Brealey.

Kirkwood H (2018) The current state of artificial intelligence and the information profession. *Business Information Review* 35(1): 9–11.

Lessig L (2001) *The Future of Ideas: The Fate of the Commons in a Connected World*. United Kingdom: Knopf Doubleday Publishing Group.

Metropolis (1927) *Lang*. Germany: Parufamet: F [Film].

Penrose R (1989) *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford: Oxford University Press.

Richardson S (2020) Cognitive automation: a new era of knowledge work? *Business Information Review* 37(4): 182–189.

Searle J (1984) *Minds, Brains and Science: The 1984 Reith Lectures*. Harvard University Press.

Transcendence (2014), *Pfister, W* [Film]. United States: Warner Bros. Pictures.

Tredinnick L (2017) Artificial intelligence and professional roles. *Business Information Review* 34(1): 37–41.

Tredinnick L (2022) Ubiquitous information. *Business Information Reviews* 39(1): 9–11.

Tredinnick L, Laybats C (2023) The dangers of generative artificial intelligence. *Business Information Review* 40(2): 46–48.

Turing A (1950) Computing machinery and intelligence. In: Copeland J (ed), *The Essential Turing*. Oxford: Clarendon Press, 441–464.

Turing A (1951) Intelligent machinery: a heretical theory. In: Copeland J (ed), *The Essential Turing*. Oxford: Clarendon Press, 472–475.

Westworld (1973), *Crichton, M* [Film], United States: Metro-Goldwyn-Mayer.