

Terry Flew

Ten Provocations on AI, Trust, and the Future of Communication

ABSTRACT: As we enter into a global artificial intelligence (AI) boom, it is important to bring together the disparate array of conversations, provocations, and prophecies regarding AI and its societal impacts. This article addresses such questions from the standpoint of trust as a concept and communications as a disciplinary field. It notes both historical continuities and areas of discontinuity in these debates and the importance of popular culture as a means of framing AI debates. This article also questions the pessimistic scenario on AI's likely impact on education, noting that it could be potentially positive for the humanities.

KEYWORDS: artificial intelligence, trust, communication, materiality, humanism, education.

There is little doubt that we are currently in an artificial intelligence (AI) boom. AI booms are those moments where everyone and everything seems to be focused on AI as challenge, threat, opportunity, or prophecy. We are currently in one of those moments, although share market volatility points to some concerns about whether AI is being “oversold” in both the economic marketplace and in the marketplace of ideas. In this short paper, I want to put forward 10 provocations about AI, trust, and the future of communication, with the aim of stimulating conversation rather than reaching definitive conclusions. It is timely to open up such conversations with the disciplinary field of communication, as they further the conversation about whether communication is in fact a “post-discipline” (Flew, 2024a; Waisbord, 2019), as well as its future relationship to fields such as computing science and the information and communications technology (ICT) sectors.

1. AI Is Not a New Thing

Alan Turing's essay “Computing Machinery and Intelligence” was published in *Mind* in 1950 (Turing, 1950). While Turing does not use the words “artificial intelligence” in that essay, it begins with the question “Can machines think?” and proposes what has since come to be known as “the imitation game,” where a machine can reach the point that it can communicate in such a way that the receiver of the message cannot know whether they are communicating with a person or a machine. The first event explicitly addressed to AI was the Dartmouth College Research Seminar

on Artificial Intelligence, a summer workshop held in 1956 and attended by, among others, John McCarthy, Herbert Simon, Claude Shannon, and Marvin Minsky. Shannon is, of course, considered one of the founders of information theory (the Shannon-Weaver model of communication), and McCarthy, Simon, and Minsky would become major figures at MIT, Stanford, and Carnegie-Mellon University. Any discussion of AI thus needs to acknowledge that there are significant intellectual legacies in these debates that require acknowledgement.

2. AI Is a Material Thing

The AI theorist Kate Crawford, in her 2021 book *Atlas of AI*, made the observation that “AI is neither artificial nor intelligent” (Crawford, 2021, p. 8). What is apparent from the accounts of Crawford and others is that AI is a profoundly material thing. It is not an abstract “cloud.” The materiality of AI is apparent in the billions of human interactions in digital environments that constitute the data harvested as raw material for large language models (LLMs). It is apparent in the enormous amount of energy consumption required to operate the data centers that drive AI-related activities. And it is both apparent and largely invisible in the work undertaken by large and growing numbers of people around the world who are required to “clean” the data, removing illegal, offensive, and toxic content from the data sets fed into LLMs (Brevini, 2022).

3. Claims to “Intelligence” Draw Attention to Both the Sources and Uses of Data

Dominant AI discourses present “intelligence” as both the input and output of AI systems. The relationship between data inputs and AI outputs is becoming an ongoing source of debate in the development of AI. Insofar as there are algorithmic or participatory biases in the means through which data are accessed in order to be harvested, bias is going to continue to be reflected in the outputs produced using AI (Noble, 2018). The question of whether AI outputs demonstrate intelligence takes us to the frontiers of debates about the replicability of human cognitive processes, particularly those where some form of creativity is required (Boden, 2014).

4. Arts and Humanities Have a Long History of Engagement With AI

The arts and humanities have a long history of engagement with computer science in general, and AI in particular. This is most apparent when

we look, not at the most elite universities, where arts and humanities are very well established, but at the more applied, sci-tech universities where arts and humanities have developed in the context of a close—if often problematic—relationship with big computer science initiatives. In the United States, for example, this includes Stanford, MIT, and the University of California–Berkeley more than it does Harvard, Yale, and Princeton.

A famous example of how this has worked in practice is the work of philosopher Hubert Dreyfus, whose critique of AI, *What Computers Can't Do* (Dreyfus, 1978), developed in a dialogue with the leading AI scholars of his era at MIT and UC Berkeley. Dreyfus's critique of the conceptual underpinnings of AI drew upon his work on Husserl, Merleau-Ponty, Heidegger, and Foucault, to argue the limits of AI models based on what is known as *cognitive simulation*, which proposes that the computer can replicate the human mind and that computing processes can be made analogous to human thought processes. In order to develop this critique, it was important that Dreyfus was a philosopher located in the milieu of researchers and labs in which AI was being trialed, such as the RAND Corporation at MIT, which acted as the stimulus to his work.

5. Popular Culture Has Long Foregrounded Today's AI Debates

2001: A Space Odyssey, produced by Stanley Kubrick, was released in 1968 (Kubrick, 1968). While the film never refers to AI, Kubrick was advised during the making of the film by the leading science fiction writers of the time, such as Arthur C. Clarke, as well as Marvin Minsky, who was by this time the professor of Artificial Intelligence at MIT (Kolker & Abrams, 2024). What is both scary and plausible about the supercomputer HAL in *2001*, and which has enabled Kubrick's film to have a life beyond most science fiction cinema, is that it displays both sentience, or the capacity to experience feelings and sensations, and what we now term generative AI capabilities, or the capacity to perform broad cognitive tasks, such as the ability to discern that the human astronauts threaten the mission and thus have to be killed.

In addition to *2001: A Space Odyssey*, Wikipedia lists 184 films that have dealt at some level with AI, either as a protagonist or as an essential part of the film, and as a computer or program or as a robot or android. This is a very diverse range of cinematic representations and scenarios through which we have developed an understanding of AI through popular culture, as well as published fiction, which present scenarios, both utopian and dystopian, of where AI may take humanity. These insights are very

relevant to contemporary policy debates such as what to do about deep-fakes. Faking images has been a feature of cinema for as long as it has existed, and the question of what is real, what is “fake,” and when and how it matters has been a part of our collective popular imaginary for a very long time.

6. AI Presented New—but Recognizable— Challenges to Communication

Claude Shannon was one of the participants in the 1956 Dartmouth College symposium on AI and is also the founder, with William Weaver, of the mathematical model of communication, more commonly referred to as the sender-message receiver (S-M-R) model (Waisbord, 2019, pp. 20–21). While the notion of communication as information as distinct from meaning has been widely critiqued, the underlying assumption about technologies (media) being the passive transmitters of such messages has remained central. Stuart Hall’s encoding/decoding model of media, which has been central to cultural studies, is effectively a reframing of Shannon and Weaver through the prism of critical media sociology (Flew, 2024b). Even where communication scholars have broken with this essentially passive conception of media technologies, such as Marshall McLuhan, the media are still conceived as “extensions of man” (Flew, 2017).

David Gunkel has proposed that, rather than associating digital technologies with computer-mediated communication (CMC), or an extension of the S-M-R framework into the digital world, we start from a different point that communication involves the study of messages that may be between humans but may also be between humans and machines, or indeed between machines and machines (Gunkel, 2020). In other words, the media themselves can be social actors with whom one communicates and interacts, whether the interaction is by human social actors (e.g., interacting with a service chatbot) or is a process that draws on an understanding of human requirements by circumventing the need for prior permission (e.g., the internet of things, where your refrigerator may communicate with a supermarket to advise you on your need for a carton of milk, and the subsequent financial transaction is automated). Such work is emerging under the banner of human–machine communication (HMC) (Guzman, 2018), but AI may point to a larger paradigm shift, as machines can and do communicate with one another. As Gunkel noted, cybernetics theorists such as Norbert Wiener were speculating with such possibilities as early as 1950 (Gunkel, 2020, pp. 58–59).

7. AI Reconfigures the Relationship Between Communication and Trust

AI has important implications for analyzing the relationship of communication to trust. As I have argued elsewhere (Flew, 2021), communication as a discipline has often worked with proxy categories for trust and distrust: ideology, hegemony, mediated populism, framing, the public sphere, and communicative action rest implicitly upon the concept of trust, yet it is typically not explicitly referred to in communication, as distinct from fields such as sociology.¹

One of the more interesting theorizations of the relationship of communication to trust comes from the German sociologist Niklas Luhmann. For Luhmann, what differentiates a social system from a psychic or living system is that a social system is based around networks of communication, with psychic systems based around consciousness and living systems around life (cells, brains, organisms, matter). The social system is stabilized insofar as complexity can be managed, and the networks are formed around four symbolically generalized media of communication: money, power, trust, and love (Borch, 2011). Trust is only possible in a familiar world—where experiences of the past provide guidance for actions in the future—yet the growing complexity of modern social systems renders this *system trust* inherently unstable. As Luhmann puts it:

Although system trust is shown to be more or less absorbed, more or less latent, it is fundamentally different from the “naïve” experience of familiarity with the everyday world. In system trust one is continually conscious that everything that is accomplished is being *produced*, that each action has been *decided* on after comparison with other possibilities. System trust *counts on* explicit processes for the reduction of complexity i.e. on people, not nature. The great civilizing processes of transition to system trust give people a stable attitude towards what is contingent in a complex world, make it possible to live with the realization that everything could be otherwise. (Luhmann, 2017, p. 64)

As Christian Borch has observed, Luhmann differs from his contemporaries—most notably Jürgen Habermas—in the radical implication of his systems theory that “society and the social are nothing but communication and that only communication communicates” (Borch, 2011, p. 135). In the field of trust studies, Luhmann’s uncoupling of system trust from personal

trust, with the former grounded in communication and the latter in consciousness, gets beyond the binary opposition between those who propose that trust relations are dependent upon trustworthy individuals (Hawley, 2019) and those who identify trust as arising from the rational choices of calculating individuals (Hardin, 2006). Luhmann's framework is well placed for considering the implications of AI for trust, as it grounds expectations of trust in symbolically generalized media, which are of course the raw material of LLMs, and constitutes trust in AI as a question of the trustworthiness of both these systems and the people involved in developing and deploying them. It situates social trust in a different space to that of rationality and consciousness.

8. Many Students Have Already Adopted AI

A 2024 survey of Harvard University undergraduates found that 87.5% of students surveyed were using generative AI tools for study at least once a week, and 30% subscribed to premium services such as ChatGPT Plus (Hirabayashi et al., 2024). The most common use of generative AI was to answer general questions, like "How does a 401k work?" For around a third of students, AI is replacing the role of traditional informational sources like Wikipedia and Google search. Other common uses of AI in this survey were for help with writing assignments (e.g., coming up with ideas, drafting, proofreading), writing emails, and helping with programming assignments and data processing.

The Harvard survey is a reminder that the AI genie is well and truly out of the box in higher education. What is apparent is that questions of ethics are not at the forefront of student considerations. Students are using AI partly for advantage in assessment, but also because the new technologies exert a clear fascination and sit so close to the educational mission. There is thus a need to move from debates about whether AI can be used by students to the terms around which it is being used. Access to AI is becoming the new laptop debate in that, 20 years ago, teachers grappled with the question of how reliant their courses should be on the expectation that all students had a laptop they could bring to class and the equity implications of unequal access. That equity issue is emerging again around access to AI, particularly premium services. Unlike laptops, however, there can be a social stigma attached to using AI for fear that it may lead to cheating or, somewhat differently, to inauthentic expression. There is a need for educators to step back and think about that, otherwise there will be an "AI underground" in terms of teaching and assessment practices.

9. The Humanities and Social Sciences Give Meaning and Purpose to the Science of AI

One concern is that the humanities and social sciences perspective on AI will be cast as one where “we” can “save” the technology from the actions of the unethical. In this scenario, the critical scholar can be cast as the superhero in the cape, awaiting a call from the wings to save humanity from the hubris and blinkered thinking of scientists. Such a scenario no doubt gives comfort as it places such scholars potentially at the center of an alternative narrative, where technological progress and ethical self-formation are fused in the march toward a utopian future. However, since the “lonely hour of the last instance” never ultimately arrives, what we are too often left with in this scenario is the figure of the humanities scholar as the lonely secular prophet. This gloomy scenario seems to point to a onetime “Golden Age” that continually recedes on the horizon, as we are left with bad ideas and bad people (scientists, tech bros, populist politicians) in control of institutions and deliberative processes.

In Wendy Brown’s (2023) Tanner Lectures, published as *Nihilistic Times: Thinking with Max Weber*, she returned to Weber’s lectures on “Politics as a Vocation” and “Science as a Vocation” as a way of thinking through what humanistic scholarship offers in the face of rapid technological change and populist politics. Although Brown views Weber’s promise of value-neutral social science as having done great damage to universities as knowledge institutions, she shares with Weber the sense that science requires conversations outside of its own disciplinary structures in order to give it a sense of purpose and steer its endeavors away from nihilism, or practice that is unmoored from underlying values:

As it topples religious and theological accounts of order and meaning, science cannot replace what it destroys. The inclination to do so, more than merely misguided, is itself a dangerous nihilistic effect, the voids opened in a radically desacralized world create a demand, Weber says, for prophets and demagogues everywhere, and for ideas that excite and incite. (Brown, 2023, p. 65)

AI development without a mission and values, in other words, is prone to being shaped by “prophets and demagogues.” To give its development and deployment a sense of social purpose, it will not be enough to simply call for state regulation, although that is important. It is to be prepared to open up conversations about values and the space that learning, as distinct from

politics, provides for the exchange and contestation of competing visions of the good society. And those are humanities and social science questions.

10. AI May Give New Life to Humanistic Education

I will conclude with a recent reflection on AI and the world it is creating from an arts and humanities perspective. Steve Fuller's recent book *Back to the University's Future: The Second Coming of Humboldt*, considers the long-term implications of the decentralization of knowledge and the resulting "demystification of cognitive authority" (Fuller, 2023, p. 9). From this perspective, AI is to be welcomed as it simplifies the process of both acquiring knowledge and demonstrating its utility in a replicable form. As a result, the collective intelligence of a society is advanced, and there is the opportunity to explore new and different questions.

Its implication for universities, in Fuller's view, is that it clarifies the purpose of teaching, which is not to transmit the findings of research (which is increasingly done by those other than the researcher anyway, whether a sessional academic or ChatGPT), but to enable the findings of research to be the starting point for a conversation that "amounts to teachers casting themselves as learners by presenting their field of research as an open horizon with many unexplored possibilities awaiting fellow travelers" (Fuller, 2024, p. 227). AI thus disrupts the business model of the modern mass university, particularly around the long-form essay as default assessment setting, but also has the potential to put universities on a new and different footing, to which the Humboldtian goals of humanistic scholarship are central.

Note

1. We can see this again with how misinformation debates have evolved. Misinformation, disinformation, and "fake news" are perceived from a largely "supply-side" perspective, where bad actors (populist politicians, Russian hackers, malign tech billionaires, etc.) pollute the infosphere, and the task of researchers, policymakers, fact checkers, and the like is to identify and remove the "toxic" content. This quasi-ecological take on misinformation means that the reasons why such information has provenance with particular individuals and groups can be either bypassed or attributed to their particular pathologies

(e.g., those who vote for right-wing populists are nostalgic for the past) (Flew, 2022). For some interestingly different takes, which bring the politics of knowledge to the fore, see Bratich (2020) and Fuller (2020).

References

- Boden, M. (2014). Creativity and artificial intelligence: A contradiction in terms? In E. S. Paul & S. B. Kaufman (Eds.), *The Philosophy of Creativity: New Essays* (pp. 224–244). Oxford University Press.
- Borch, C. (2011). *Niklas Luhmann*. Routledge.

- Bratich, J. (2020). civil society must be defended: Misinformation, moral panics, and wars of restoration. *Communication, Culture & Critique*, 13(3), 311–332.
- Brevini, B. (2022). *Is AI Good for the Planet?* Polity.
- Brown, W. (2023). *Nihilistic times: Thinking with Max Weber*. Harvard University Press.
- Crawford, K. (2021). *The atlas of AI*. Yale University Press.
- Dreyfus, H. L. (1978). *What computers can't do: The limits of artificial intelligence*. Harper Colophon Books.
- Flew, T. (2017). The “theory” in media theory: The “media-centrism” debate. *Media Theory*, 1(1), 43–56.
- Flew, T. (2021). The global trust deficit disorder: A communications perspective on trust in the time of global pandemics. *Journal of Communication*, 71(2), 163–186.
- Flew, T. (2022). Fake news, trust and behaviour in a digital world. In G. Rawnsley, Y. Ma, & K. Pothong (Eds.), *Research Handbook on Political Propaganda* (pp. 28–40). Edward Elgar.
- Flew, T. (2024a). The rise, fall, and rise again of communication as a disciplinary signifier in Australia: After the “cultural turn” and the “digital turn.” *Journal of Technical Writing and Communication*, 54(2), 129–142.
- Flew, T. (2024b, forthcoming). What Future for Media Industries and Media Industry Studies? *Global Journal of Media Studies*.
- Fuller, S. (2020). *A player's guide to the post-truth condition*. Anthem Press.
- Fuller, S. (2023). *Back to the university's future: The second coming of Humboldt*. Springer.
- Fuller, S. (2024). Classroom as crucible in the Humboldtian university: Reply to Collin. *Philosophy of the Social Sciences*, 54(3), 226–230.
- Gunkel, D. (2020). *An introduction to communication and artificial intelligence*. Polity.
- Guzman, A. (2018). What is human-machine communication, anyway? In A. Guzman (Ed.), *Human-machine communication: Rethinking communication, technology, and ourselves* (pp. 1–28). Peter Lang.
- Hardin, R. (2006). *Trust*. Polity.
- Hawley, K. (2019). *How to be trustworthy*. Oxford University Press.
- Hirabayashi, S., Jain, R., Jurkovic, N., & Wu, G. (2024). *Harvard undergraduate survey on generative AI*. Harvard University.
- Kolker, R., & Abrams, N. (2024). *Kubrick: An odyssey*. Faber & Faber.
- Kubrick, S. (Director) (1968) 2001: *A Space Odyssey*. Warner Bros.
- Luhmann, N. (2017). *Trust and power* (2nd ed.). Polity Press.
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59, 433–460.
- Waisbord, S. (2019). *Communications: A post-discipline*. Wiley-Blackwell.

Terry Flew (terry.flew@sydney.edu.au) is professor of Digital Communication and Culture and Australian Research Council (ARC) laureate fellow at The University of Sydney. His books include *The Creative Industries, Culture and Policy* (SAGE, 2012), *Global Creative Industries* (Polity, 2013), *Media Economics* (Palgrave, 2015), *Understanding Global Media* (Palgrave, 2018), *Regulating Platforms* (Polity, 2021), and *Digital Platform Regulation: Global Perspectives on Internet Governance* (Springer, 2022). He was president of the International Communications Association (ICA) from 2019 to 2020. His ARC laureate is on Mediated Trust: Ideas, Interests, Institutions, Futures.