Philosophy, Misinformation, and Artificial Intelligence

Friday 7ᵗʰ November

21ˢᵗ Murdoch Colloquium

https://philevents.org/event/show/137605

Programme & Abstracts

8.30	Registration
9.00-9.15	**Welcome** Opening Address Acknowledgement of Country
9.15 – 10.15	**Session 1: Opening Keynote** Thinking Through AI and Misinformation with Arendt and Deleuze: The Life of the AI Mind (Tauel Harper, Murdoch University)
10.15 – 11.00	Morning tea
11.00 – 12.30	**Session 2: Transformations (or just Hallucinations?)** Sticky Atmospheres and Affective-Perceptual Diagnostics: Reframing Misinformation through Dorothea Olkowski and Bernard Stiegler (Brent Jones, University of Notre Dame Australia) From Product to Process: On the Terms of Comparison Between Human Authored and AI Generated Texts (Norma Lam-Saw & Jason Tuckwell, Western Sydney University) Limiting Information for Epistemic Gain: What Zollman's Agent-Based Models Say About Scientific Communities (Edmund Hassell, Murdoch University)
12.30 – 1.30	Lunch
1.30 – 3.00	**Session 3: Applications and Implications** From Obligation to Supererogation and Back: The Transformative Role of AI in Organizational Ethics (Jacqueline Boaks, Curtin University) "An edifice of false pretenses": Revisiting Veblen's *The Higher Learning in America* in the Age of Educational Technology (Andrew Hutcheon, Edith Cowan University) Emotional AI, Consent, and the Right to Oblivion (Adam Andreotta, Curtin University)
3.00 – 3.30	Afternoon tea
3.30 – 4.00	**Session 4: General Philosophy Section** "Empathy with the commodity-soul": Walter Benjamin on the corruption of Husserlian intersubjectivity (Joel Bourland, Murdoch University)
4.00 – 5.00	**Session 5: Closing Keynote** Rational Debates Beyond Facts and Data (Gottfried Vosgerau, Heinrich Heine Universität Düsseldorf)
5.00 – 6.00	**Session 6: Roundtable Discussion** Nardine Alnemr (Murdoch University) Tama Leaver (Curtin University) Celeste Rodriguez Louro (University of Western Australia) Marco Rizzi (University of Western Australia)
6pm	**Society of Applied Philosophy Best Student Paper Award & Reception**

Thinking Through AI and Misinformation with Arendt and Deleuze:
The Life of the AI Mind

(Tauel Harper, Murdoch University)

AI has exacerbated the problem of understanding the truth, not only because it enables the proliferation of false and misleading material, but also because it encourages a particular form of detachment from thinking and the making of meaning in public. In this paper I explore how AI replicates, extends and challenges the activities of thought and critical contemplation by comparing AI processes to these activities as described by Hannah Arendt on one hand and Deleuze and Guattari on the other. While Deleuze and Guttari's philosophy opens up some possibilities for positive uses of AI, Arendt's humanist position outlines why AI presents a real danger to public credulity.

Sticky Atmospheres and Affective-Perceptual Diagnostics: Reframing Misinformation through Dorothea Olkowski and Bernard Stiegler

(Brent Jones, University of Notre Dame Australia)

Misinformation is often cast as a problem of faulty cognition. People are misled by false propositions, fail to fact-check, and adopt misguided beliefs. Yet this framing risks missing how digital environments modulate perception before beliefs are formed. This paper develops a conceptual account of radicalisation and misinformation that foregrounds perceptual structuring rather than cognitive error.

Drawing on Dorothea Olkowski's post-phenomenological synthesis of Deleuze, Bergson, and Merleau-Ponty, and Bernard Stiegler's media pharmacology, I propose a framework of *affective-perceptual diagnostics.* Central to this framework is the concept of sticky atmospheres: affectively charged perceptual milieus that condition how ideological claims come to feel true. Instead of treating memes, livestreams, or algorithmically curated feeds as conveyors of information (true or false). I argue instead that they function as technologies of attunement. These technologies recalibrate attention, mood, and embodied orientation. Stiegler shows how digital media pharmacologically condition memory and affect, while Olkowski demonstrates how perception is world-making, open, and subject to modulation. Taken together, they reveal how misinformation circulates less as discrete content than as atmospheres that reorganise the very conditions of appearance.

This conceptual reframing allows scholars and policy makers to see digital propaganda and misinformation more broadly not as failures of rationality, but as transformations in perceptual life. Moreover, it suggests that countering misinformation requires more than fact-checking: it demands interventions at the level of atmosphere, attention, and embodiment. By situating misinformation within perceptual and ontological registers, this paper highlights how philosophy can illuminate the deep structures through which digital media shape political worlds.

From Product to Process: On the Terms of Comparison Between Human Authored and AI Generated Texts

(Norma Lam-Saw & Jason Tuckwell, Western Sydney University)

The widespread adoption of large language models (LLMs) capable of generating natural language outputs has significantly impacted human writing activities— from integration into communicative and creative practices to critical reassessments of textual production itself. While much consideration of LLMs' impact on literary studies focuses on comparing human and machine-generated textual products as if they were equivalent, such assumptions obscure the radically different generative processes at work in information-computing and organic systems.

Drawing on philosophy, science and technology studies and cognitive science scholarship that shows how technical objects become mistaken for the organic processes they model, we argue that assumptions about the functional equivalence between LLM and human texts emerge from converging theoretical and technical reductions that systematically eliminate the qualitative dimensions of human meaning-making.

Part 1 traces an immediate historical precedent: structuralism's adoption of information theory, which reduced human communication to signal transmission and bracketed out organic mechanisms of signification. This theoretical reduction, conceiving creativity as recombinatory operations within closed systems, has re-emerged as an explanatory frame for comparing human and LLM-generated texts. Through Simondon's alternative theory of information, which preserves signification as the capacity for living beings to be dynamically informable generators of meaning, we recover a model for thinking beyond this reduction. Part 2 examines the technical operations of transformer-based LLMs, demonstrating how tokenization, vector representations, and alignment procedures generate sophisticated statistical outputs without semantic understanding or embodied experience. While producing fluent text that satisfies formal communicative criteria, these operations yet remain entirely dependent on continuous human eco-organological mediation, and are categorically distinct from human processes of meaning-making. Comparisons that treat the textual product as a common measure obscure this difference, since they elide the incommensurability of the generative processes at work. At the same time, what is being compared remains equivocal: sometimes framed as if LLMs were discrete technical objects opposed to humans, and at other times as socio-technical ensembles already entangled with human actions and cultural resources. This equivocation raises the deeper problem of how to trace and ground emergent, agentive, and creative phenomena across human–machine entanglements.

Far from resolving long standing questions about the nature of human authorship, we show how the advent of LLMs constitute a complexification of human-machine entanglement and provoke new questions about their interactions. We conclude by contrasting two universes: one where the textual generation of both LLMs and structuralism operate through purely quantitative reductions, and another, where human authorship emerges through qualitative, organological processes.

Limiting Information for Epistemic Gain:
What Zollman's Agent-Based Models Say About Scientific Communities

(Edmund Hassell, Murdoch University)

It has long been assumed that greater access to information leads both individuals and communities to better understand the subject they are inquiring about. This notion can be traced at least as far back as John Stuart Mill. More recently, philosophers of science have been testing how the composition of the scientific community affects inquiry, with a particular focus on communication structure. In line with this research is philosopher of science Kevin Zollman, who in 2007 developed a computer model simulating scientists investigating the best between two competing options. A surprising result of his study was that, in some cases, the less connected scientific community outperformed the completely connected one. However, this result emerged from an agent-based model, which necessarily simplifies many aspects of scientific research. Since Zollman's original study, several philosophers have tested his model under varying parameter conditions. More often than not, they found that his conclusions do not hold under alternative parameter conditions (Holman and Bruner 2015; Rosenstock et al. 2017; Weatherall and O'Connor 2018; Hahn et al. 2020; Wu 2023). Yet, some studies have confirmed Zollman's findings (Zollman 2013; Rosenstock et al. 2017; Frey and Seselja 2020). This raises two prominent questions: Under what specific parameter conditions does Zollman's conclusion hold? And if a consistent set of conditions can be found that lead the less connected community to outperforming completely connected communities, how seriously should we take that result given that it's the result of a computer simulation? My presentation will address both of these questions.

From Obligation to Supererogation and Back: The Transformative Role of AI in Organizational Ethics

(Jacqueline Boaks, Curtin University)

This paper explores the potential impact of artificial intelligence (AI) systems on the ability and willingness of organizations to engage in supererogatory acts – i.e., to take actions that go beyond an organization's presumed moral duty. The role of supererogation in organizational ethics has been a distinct, if somewhat underdeveloped, theme in recent literature (Mazutis, 2014; Tencati et al., 2020; White et al., 2023), and much of the discussion has centered on the issues of identifying and managing the threshold between acts that fall within an organization's moral duty and acts that could be considered supererogatory.

In this paper, we argue the advent of AI-powered decision support systems can move this threshold in both directions. On one hand, such systems can raise it, increasing the range of actions that could be perceived as falling within the organization's moral duty and thereby reducing the number of supererogatory actions available. On the other hand, AI systems can also expand the available space for supererogation and thus make it easier for organizations to engage in supererogatory actions. In the core part of the paper, we discuss the interaction between AI and moral action with the aim of illuminating the mechanisms that enable organizations to engage in actions that exceed their moral duty. We thus explore the impact of AI, not on core business functions, but on the legal, social and ethical engagement of organisations.

"An edifice of false pretences": Revisiting Veblen's *The Higher Learning in America* in the Age of Educational Technology

(Andrew Hutcheon, Edith Cowan University)

As education is progressively distorted by the twin influences of Chicago School economics and education technology, it has become a vital task to find lines of inquiry that give us tools to understand and ameliorate the problems. One of the Chicago School's most prominent achievements was repositioning people as 'human capital' and the value of education as the cultivation of that 'asset', generating an intensive instrumentalisation of schools and universities. We are therefore given the 'knowledge economy' and, more specifically in Australia, the 'Education Revolution' putting laptops and tablets in every K-12 classroom. These assumptions have been imported wholesale into universities, where we now, thanks to ongoing regulatory guidance, face the prospect of generative AI permeating throughout the curriculum. Against this threat, this presentation turns to an unlikely source of inspiration, Thorstein Veblen's *The Higher Learning in America* (1918), an early 20th century argument against the increasing commercialisation of universities. Veblen's argument is fundamentally an institutional, and therefore anti-neoclassical, economic diagnosis that identifies the causes, and potential cures, of instrumentalised higher education. Critically, he identifies the propensity of university administration for "edifices of false pretences" that mirrors his earlier theories of conspicuous consumption, pecuniary emulation and status competition. This talk argues that Veblen's work is worth revisiting for its analytical clarity and its use of economic arguments against edifices such as the headlong rush towards generative AI and education technology more generally.

Emotional AI, Consent, and the Right to Oblivion

(Adam Andreotta, Curtin University)

"Emotional AI" refers to technologies that use biometric data and machine learning to interpret human emotions. In recent years, these technologies have been applied in education, driver assistance systems, healthcare, marketing, and security. In this paper, I examine how Emotional AI raises several issues for contemporary accounts of privacy, particularly those grounded in individual choice and informed consent.

I begin by critiquing two common ideas about privacy. The first is that privacy is primarily about individual control over personal information, and the second is what Ignacio Cofone calls the "traditionalist" approach, which assumes that "people behave rationally, optimising choices for their own wellbeing based on available information" (2023, p. 33). Although emotional data may seem like just another form of personal information, I argue that it is uniquely sensitive. Emotions are often involuntary, unconsciously expressed, and deeply embedded in our interactions with the world and everyday communications with others. Moreover, since Emotional AI can detect and store our emotions automatically, it can place individuals in situations where managing their emotional privacy becomes very difficult, if not impossible.

My central claim is that the use of Emotional AI warrants a shift towards a more robust model of privacy, one that recognises the importance of emotional opacity. I do so by drawing on Lowry Pressly's recent book *The Right to Oblivion* (2024), which argues that there is moral value in having certain aspects of our lives partially unknown, unrecorded, or uncommodified. I argue that Emotional AI raises urgent ethical questions about which parts of our emotional lives should remain beyond the reach of technological detection, collection, or prediction.

"Empathy with the commodity-soul":
Walter Benjamin on the corruption of Husserlian intersubjectivity

(Joel Bourland, Murdoch University)

Walter Benjamin addressed the extent to which our dreams, desires, and relationships are modeled on mass-marketed images. Central to what he called the "commodification of things" is our daily encouragement to identify with the staged enjoyments and unapproachable icons of films, ad campaigns, and popular culture. Benjamin refers to this tendency of the collective imagination as an "empathy with inorganic things," or "empathy with the commodity-soul." In fact, he reserves the term "empathy" (*Einfühlung*) exclusively in this critical capacity: to account for how personal experience is routed by the heightened solicitations of what we today call the attention-governed economy.

This paper situates Benjamin's project in relation to the phenomenological tradition: specifically, through his engagement with the concept of "empathy" as developed by Scheler, Husserl, and Stein. For these thinkers, empathy is the primordial act that opens onto intersubjective objectivity, i.e., onto a world of agreed-upon objects carrying common-sense meanings. Benjamin, by contrast, employs the concept to account for representations of quite a different sort: images of longing, fantasy, and projection. If 'others' anchor intentionality, then the manners I have of interpreting my world will be refracted through their attitudes, gestures, and intentions: this, without regard to their reality as 'people' or the reality of the desirable contexts and experiences they apparently envelop. Thus, we can nurture imagined intimacies with the lives of celebrities, or feel a sense of 'belonging' through the proxies of brand-names and online collectives. Motivating this encounter with phenomenology, I argue, is Benjamin's consistent rejection of the notion that the 'abstract' nature of concepts stems from their being subtractive generalities.

Rational Debates Beyond Facts and Data

(Gottfried Vosgerau, Heinrich Heine Universität Düsseldorf)

Public debates and science communication often focus on facts and data. At the same time, empirical research shows that more knowledge does not always lead to the expected actions. A case in point is increased cultural polarization, e.g. with respect to the effects of climate change. In general, there is no direct and easy way from knowing facts to acting accordingly. Deliberations about actions take the form of practical syllogisms. I argue that one reason for the gap between knowledge and action is that the normative premises required in practical syllogisms are often not made explicit. In other words, the goals and values that we need in order to act remain implicit such that often too few of our own goals and values are fully transparent to us. However, our goals and values are very complex and hierarchically dependent on one another. In order to make these goals and values explicit, we must rationally reconstruct the practical syllogisms. Although this method of rational reconstruction is well established, we face another problem: From empirical research we know that many biases influence our thinking. Such biases are often defined as irrational. I argue that the principle of charity – at least in the context discussed here – calls for the inclusion of biases in argument reconstruction, irrespective of whether or not they are irrational. I present different ways of integrating biased thinking in charitable rational reconstruction of practical syllogisms with the aim to make goals and values explicit and transparent, thus enabling fair debates about different goals and how to reach them.