

Grimwood, Tom ORCID: <https://orcid.org/0000-0001-8099-6191> (2025) What is an artifice? The precariousities of Culbertson's two distinctions on generative AI. Journal of Applied Hermeneutics .

Downloaded from: <https://insight.cumbria.ac.uk/id/eprint/8969/>

Usage of any items from the University of Cumbria's institutional repository 'Insight' must conform to the following fair usage guidelines.

Any item and its associated metadata held in the University of Cumbria's institutional repository Insight (unless stated otherwise on the metadata record) may be copied, displayed or performed, and stored in line with the JISC fair dealing guidelines (available [here](#)) for educational and not-for-profit activities

provided that

- the authors, title and full bibliographic details of the item are cited clearly when any part of the work is referred to verbally or in the written form
- a hyperlink/URL to the original Insight record of that item is included in any citations of the work
- the content is not changed in any way
- all files required for usage of the item are kept together with the main item file.

You may not

- sell any part of an item
- refer to any part of an item without citation
- amend any item or contextualise it in a way that will impugn the creator's reputation
- remove or alter the copyright statement on an item.

The full policy can be found [here](#).

Alternatively contact the University of Cumbria Repository Editor by emailing insight@cumbria.ac.uk.

What is an Artifice? The Precarities of Culbertson's Two Distinctions on Generative AI

Tom Grimwood

Abstract

Culbertson's recent paper published in the *Journal of Applied Hermeneutics* offered two distinctions at work in the reading and understanding of Natural Learning Processing. This paper was a significant articulation of a general hermeneutic response to the prospect of generative AI and its challenges for interpretation. But it also raised some nagging questions on whether there is a risk that we settle too quickly on the promotion of close reading and the aspirations of “thinking with others” in dialogical open-ness, and in doing so, also settle a little too quickly on what the object of the hermeneutic encounter is, at the expense of other possible dialogues, or traditions, at work? This paper argues that a dimension at work in the debate over generative AI often missed from hermeneutic discussions is that of the artifice. This paper explores the artifice, as an interpretative element of the “artificial” at work in AI; not to critique Culbertson's two distinctions, but rather to suggest a certain precarity to their resoluteness, a precarity from which further research would benefit.

Keywords

Generative AI; artifice; non-ideal dialogue; education; technology

Corresponding Author:
Tom Grimwood, PhD
University of Cumbria, Lancaster, UK
Email: tom.grimwood@cumbria.ac.uk

It is all but impossible to ignore Artificial Intelligence (AI) in Higher Education today. Not just in terms of the overabundance of apps, tools and technological innovations constantly advertised, but also in the equally constant discussions across academics and educators about its potential uses and abuses. Here, something else is difficult to ignore. For all the special journal issues, conference panels and policy workshops, the discussions already seem rather limited, if not repetitive. Perhaps this is because they so often draw on cemented historical themes. After all, the conflict between moral concerns about automated production, and the (apparent) inevitability of technological progress can be found accompanying innovations throughout history. Perhaps it is because, as Babette Babich (2025) has noted, such discussions seem to merge applied decision-making with the speculative fiction of technological promise. Perhaps it is because the structures of these fictions are inevitably clichéd; much like, as Beguš (2025) argues, Large Language Models (LLMs) faithfully reproduce clichés in their generated writing. Either way, the dreams of how to utilize AI and aggressively market it in new and impressive ways continue at pace (both its speed and its inevitability are necessary components of any discussion), while those conversations *about* its application in research and education can seem doomed to circle around a tension which the rest of history has never quite managed to resolve.

It was, then, a fortuitous coincidence that I read Carolyn Culbertson's "Why Natural Language Processing is Not Reading: Two Philosophical Distinctions and their Educational Import" (2025) after spending the day chairing a malpractice panel at my university, where each case of plagiarism concerned allegations of AI use on a nursing programme. This was only a week after I had attended a conference on UK Postgraduate Research education where several hours had been spent discussing how to ensure PhD candidates were not using AI in a way which jeopardised the principles of the doctorate. The coincidence was fortuitous because, having been caught up in the frustrations of these discussions – frustrations which mixed the urgency of decision-making, anxiety for the imminent future of assessment and boredom with the circularity of the debate – Culbertson's paper provides an argument which moves away from judicial norms around the right and wrong use of generative AI, and centred instead on the role of meaning in a specific aspect of the issue, Natural Language Processing (NLP).

Culbertson's first philosophical distinction is between the formal semantic approach to textual meaning, utilised by NLP, and the phenomenological-hermeneutic approach of "close reading," which involves a self-reflective encounter between the text and the consciousness and imagination of the reader. For Culbertson, this distinction forms the site of a tension surrounding the epistemic virtues of AI. On the one hand, the move towards "fragmented reading" in the twenty-first century – encouraged by new digital media forms and their bitesize doses of information spread over multiple screens and voices – is aided and abetted by NLP's use of language as the distillation of information. On the other hand, if there is more to engaging with a text or artefact than attaining information, then the increasing use of NLP, especially amongst educators, should be give cause for some concern. Culbertson argues that, while the "significant advances in computational semantics and language modelling" combined with the norm of "fragmented reading" may tempt us "to utilize natural language processing programs as shortcuts to distill information from texts quickly and with little cognitive effort," (Culbertson, 2025, p.14) there is nevertheless a value in the understanding produced through close reading which constitutes an epistemic good, above and beyond the utility of the NLP route to knowledge acquisition. The rights and wrongs of using generative AI based on NLP are not the issue: the question is the

more fundamental one of how we learn. Thus, Culbertson's second distinction is between the creation of "true belief" enhanced by information gathering, contrasted with the development of "understanding" involved in close reading. Understanding, they argue, is more than simply the processing of data, because it involves a range of other virtues beyond the transfer of information:

Through close reading, one sharpens one's capacity to listen carefully to others and, when necessary, to rethink prejudices that hinder one's understanding of them. Reading also instills a disposition to relate to meaning as a collective enterprise and to inhabit the world as one whose meaning is disclosed through the process of thinking with others. It instills a sense of community with people past and present. (2025, p.13)

In arguing for these two distinctions, Culbertson raises several constructive questions for my role as an educator. If hermeneutic understanding "instills a sense of community with people past and present," is one way to address the dramatic rise of AI usage in student work to not insist on strict rules of engagement with generative technologies, but rather to orient the values of our education more around an articulation of what this community is and how it manifests in practice, be it in nursing, social work, education or other disciplines? Is the fundamental relationality of understanding which Culbertson points to something that must be prioritised in our teaching and training, when it may otherwise (intentionally or not) so often get lost amid a concern for skills and employability? Are these the questions we should be posing to our students, alongside (or instead of) the refrain of "did you cheat?"

These questions stem from Culbertson's move towards considering our role as interpreters in the context of NLP, rather than the NLP itself – which is both a welcome and necessary move if we are to address the urgency, angst and boredom raised above. Their argument reflects a growing consensus on the relationship between Gadamerian hermeneutics and generative AI (see, for e.g., Wang, 2021; Pinnell 2024). In turn, though, this raises a further question about the *integrity* of this sense of community which the phenomenological-hermeneutic reading promises; an integrity which is sometimes beholden to speculative fictions of its own. In short: is there a risk that we settle too quickly, and perhaps too comfortably, on the promotion of close reading and the aspirations of "thinking with others" in dialogical open-ness? In doing so, do we risk missing some of the more subtle and nuanced displacements at work in the encounter with generative AI? Is there a sense in which this decides a little too quickly on what the object of the hermeneutic encounter is, at the expense of other possible dialogues, or traditions, at work? Is there a risk that we fall into an extended tautology – "in order to understand closely, we must understand closely"?

This is not a question to ask of Culbertson's argument alone. As I will suggest, it is one prompted by many of those responses to generative AI within the hermeneutic tradition. Consider, for example, Iain Thomson's Heideggerian observation on the danger of the artifice – the "fake or phony" – within the discourse of AI:

Besides its obvious meaning, "the age of artificial intelligence" also more subtly suggests a time marked by the triumph of a kind of fake or phony intelligence, a thorough-going calculative rationality, the very dominance of which conceals the fact that our impressive,

interconnected, and ever-growing technological apparatuses for optimal means–end calculating seem to have left us unable to understand the nature of our own historical situation in a way that could help us shed our debilitating anxiety and begin to chart our collective historical course into a future that is more meaningful and appealing. (Thomson, 2025, p. 30)

It is always troubling when a claim that AI “has left us unable to understand” is framed in terms of an articulate diagnosis of the problem (in other words: an able understanding). But this understanding, just as the weight of the “debilitating anxiety” Thomson references, is not only based on the separation of meaningful interpretation from calculative rationality, but also a long tradition of thought and practice whereby meaningful humanistic practice will always triumph through its appeal to meaning. There is nothing surprising here: the work of critical cultivation, with the awareness of “our own historical situation” which it brings, will necessarily eclipse the passive regurgitation of previously established data. The risk, though, is that the security of the traditions invoked is assured only by the repetition of that tradition, obscuring the original displacement which generative AI prompted in the first place.

What I want to add to this discussion, then, is a dimension which current hermeneutic research on AI can be liable to overlook, largely due to the way that it positions the problem to be addressed as an agonistic choice between surface information and deeper understanding. In the rest of this paper, I want to explore the dimension of the artifice, as an interpretative element of the “artificial” at work in AI, not as a critique of Culbertson’s two distinctions, but rather to suggest a certain precarity to their resoluteness, a precarity from which further research would benefit.

“Probably AI”

As always, it is useful to start with some examples.

Scrolling on social media, I come across a picture from a group which shares images of castles from around the world. It is a marvellous looking, somewhat mysterious ruin. It certainly appears to be some kind a fortress, as the group would require: but it is buried in the foliage of an ancient forest, as though the overgrowth was part of the very architectural structure.

The first comment underneath reads: “Probably AI.”

Further down, another post in a different group. This time, it is a short-written post which questions how the pyramids were constructed, loosely echoing the ideas from Graham Hancock’s pseudoscientific series *Ancient Apocalypse*. This time, the first comment presents a long, well-structured and reasoned response, pointing out in patient detail the faults in the post’s assumptions and the leaps in logic it commits.

The comment has a series of replies. The first reads: “Good use of AI, bro.”

Two rather randomly chosen and benign cases, which could easily be substituted for more pressing ones. A student suddenly improving their essay writing, or a particularly splendid

assembling of a group presentation, may prompt the same “probably AI” response. But those cases would lead us into spaces which demand a judgement or assessment. Instead, and at the risk of entertaining what Gadamer refers to as “idle suspicions” (2004, p. 549), I want to consider these as a specific disposition of interpretation emerging alongside the growth of generative software which is perhaps more obvious in these banal cases.

Of course, it would be easy to dismiss this as beyond the remit of an authentic hermeneutic encounter, at least in the sense which Culbertson outlines. A two-word response hardly constitutes a close reading, after all. But this is also where we need to be careful of slipping into the self-confirming tautologies raised earlier. I am reminded of Hans-Herbert Kögler’s critique of the Gadamerian model of dialogue, which he claims may well be suitable for “the successful philosophical seminar discussion of exceedingly difficult texts that leave everyone transformed and elevated.” He continues: “but if such an idealised image of dialogue, however valuable and rewarding, is ontologically promoted to the all-encompassing process of being/understanding, concern is in place. What is missing is the fact that dialogues happen in non-ideal times, places, and situations.” (Kögler, 2014, p. 57)

The conditions of these “probably AI” responses are clearly non-ideal, in Kögler’s sense. Do they constitute a dialogue, or an attempt to understand? The answer to this question hinges on whether imposing an ideal distinction between fragmented and close reading addresses the challenge of the artificial interlocutor; or, conversely, whether this distinction instead obscures a more fundamental non-ideality regarding discourses on AI. Given the case for the first option has already been made excellently by Culbertson, this paper will explore the second.

A picture that is a little too beautiful, or a post that is a little too essayish, are both placed under suspicion. The contexts are different, of course, as are the merit of the accusations. The image is almost certainly fake. The post is probably not (at least its author insists, further down, that it is simply ‘well-researched’). But whether the images and posts are *artificial* creations is not really of interest if we accept the non-ideality of its appearance within dialogue. What is more interesting, I think, is the different ways in which the idea of an *artifice* pervades these fleeting moments of online engagement. First, there is an artifice of the synthetic content: that the castle is not really a castle, that the information posted is not really knowledge. Second, there is an artifice at work in the presentation: the castle that is not really a castle has been posted in a group about real castles; the information has been posted under a name as if they were the author of that information. And there is a third artifice present in the immediacy of the responses, where suspicion serves as a supplement for knowledge, and implication serves as a supplement for understanding. This is the equivalent of a knowing glance, a raised eyebrow or half-smile for those who can recognise it; a statement that “we do not need to see this,” where the collective “we” is formed in the perlocutionary moment. “Seeing means articulating,” Gadamer writes when he discusses encountering “trick” pictures whose object is not always clear. “While we are still trying various ways of organizing what we see or hesitating between them, as with certain trick pictures, we don’t yet see what is there. The trick picture is, as it were, the artificial perpetuation of this hesitation, the ‘agony’ of seeing.” (Gadamer, 2004, p.79) The suspense created by the possibility of the first two artifices prompts the third, which reinforces the boundaries threatened by the first two: it provides a relief from the agony of seeing – or, at least, probably does.

Exploring the non-ideality of encounters with AI involves exploring how the meaning of these responses arise, not from the engagement with AI content, but engagement with the ways in which the notion of the artificial relates to the concept of the artifice. Each form of artifice in play draws attention to the way in which an interpretative community reacts to the intrusion of deceit. This moves the object of hermeneutic concern *away* from the actual technologies of AI, NLPs, and LLMs (along with the excitement and threat of what they can or can't do), as well as its use and abuse as a form of knowledge-creation, and *towards* the relationality of those interpretative communities discussing them.

What is an Artifice?

In this sense, then, what *is* an artifice? This question might call back to Foucault's paper "What is an Author?," where he drew a contrast between the writer, as the producer of a text, and the author, as a *function* of discourse which survives the living writer as an organising principle of text and its relation to others. Can we consider the artifice in a similar way in relation to the discourse around generative AI and its effect on epistemic communities?

Philosophically the term is first seen in Plato. "Artificers" appear as characters in the *Timaeus* as a kind of classical incarnation of today's content creator. The artificer "keeps his gaze fixed on that which is uniform, using a model of this kind, that object, executed in this way, must of necessity be beautiful; but whenever he gazes at that which has come into existence and uses a created model, the object thus executed is not beautiful." (28a-b) The artificers are beholden to the need to constantly create, always failing to quite imitate their object of study.

While this provides a handy analogy for fragmented media today, it remains within the privileged view of Plato's philosopher. In Western culture, the artifice has a far more applied history: not a person or agent, but an ambiguous (yet formative) relationship between the artisan – the practitioner working through applied phronesis – and what they produce. In his introduction to *The Body Impolitic*, an analysis of the economic relationship between artisans and artifice in the context of local traditional practices as they negotiate global capital, Michael Herzfeld notes: "in many cultures of European origin, people associate the idea of manual artistry with cunning and subterfuge [...]. The ingenuity of the artisan is never solely aesthetic; it always contains a component of indirection and guile." (2004, p.1) This, Herzfeld continues, can be seen in the complex folds of the artisan's traditions with myths of national identity, and narratives of continuous common practices created that respond to contemporary challenges. The value of tradition, in the work of the rural craftsmen that Herzfeld studies, is not in an apprentice imitating their master but rather about understanding the work of authority in order to subvert it. The artisan learns how to be polite to rich tourists keen to purchase the authentic wares of a pre-globalised culture, while simultaneously mocking them; they perfect how to perform boredom and indifference at work so that they might steal the techniques of their master (2004, p.152).

The relationship between a skilled production of a craft and a misleading sleight of hand is not just rooted in practice, of course. The Latin *artificium*, rooted in the art (*ars*) of making (*facere*), and could refer to both craftmaking or cunning. No doubt this follows the poetic, rather than philosophical, models provided by the Greeks: the craftiness of Odysseus is described as *'poly-*

technos’ (euphemistically, a man of ‘many arts’). In French, artifice was used to denote skill or cunning since at least the 14th century, and particularly skills learned from training. In English, the word settled by the 16th century as a synonym for workmanship which is created by humans (as opposed to nature). Evolving technical developments in the early modern period word supported the notion of the artifice as a cunning device or work, emphasising the unnatural or contrived aspect of craft. This resulted in the dominant meaning in English from the 19th century onwards, that of creative deception. Nevertheless, the traditions behind the term can be seen in the four meanings which Mirriam-Webster lists: “clever or artful skill: ingenuity”; “an ingenious device or expedient”; “an artful stratagem”; or “false or insincere behavior.”

While the artifice’s ambiguity between art and deception is present in English and French, the German *Kunstfertigkeit* does not have the same connotations of contrivance or cunning; quite the opposite, in fact, and as such holds a central place within Gadamer’s account of the artistry of interpretation in applied practice. In Gadamer’s example of the work of doctors, clinical judgement is “an awareness appropriate to a particular situation, like that in which diagnosis, treatment, dialogue and the participation of the patient all come together.” It is “a form of attentiveness, namely the ability to sense the demands of an individual person at a particular moment and respond to those demands in an appropriate manner.” (Gadamer, 1996, p. 138) This artistry is quite distinct from the artificial (*künstlich*), which is typically presented by Gadamer, at least in *Truth and Method*, as not only a negative term but a restrictive one: the artificial narrowness of methodologies in nineteenth century hermeneutics (2003, pp.16, 234), the artificial narrowing of the concept of the symbol (p. 68), the artificial truncating of aesthetic questioning (p.138), and so on. Perhaps he is most vehement in this critique of the artificial when he ascribes it to technical terms – that is, words “whose meaning is univocally defined” – which destroys the living meaning of language: “a technical term is a word that has become ossified. Using a word as a technical term is an act of violence against language.” (p. 414)

There is an interesting coincidence between this linguistic different and how Gadamerian hermeneutics has developed its responses to generative AI. In its earlier days, the alignment of hermeneutics with machine learning was considered to hold great potential (see Mallory et al., 1986). Despite this initial interest, more recently there has been a broad tendency to frame AI as distinct from understanding, due to its lack of the art of attentiveness required for interpretation. Echoing the early thoughts of Dreyfus (1972) in his self-explanatorily titled *What Computers Can’t Do*, Jing Wang notes that “human understanding is not formalized, and [...] the cause-and-effect calculations of computers are not understanding” but a “cognitive result obtained by a cognitive subject [...] free of any bias.” (Wang, 2021, p. 141) As such, “an important reason why AI is not capable of understanding lies in its lack of an appreciation and conception of historicity.” (Wang, 2021, p.145) Phillip Pinell, likewise, rejects that “AI can, properly speaking, experience the fusion of horizons” because “they lack historicity and community.” (2024, p. 722) And Culbertson, as we have seen, argues that close reading enables a connectivity to past and present which the fragmented assembling of information does not.

Such responses are sustained by a separation between the life of artistry – embedded within the living history of practice – and the ossification of fakery. This separation has an entirely different appearance within the artisanal exploitation of the artifice. The former keeps its lens on the capability of generative AI; the latter, through the act of crafting, the identity of the craftsman,

and the relationship between both and the wider society, all involve managing a web of different dialogical positions operating in traditions.

The Precarity of Understanding, Artifices, and AI

How, then, does the function of this artisanal blurring of the artifice presents a precarity to the hermeneutic responses to generative AI? I will conclude by suggesting three ways in which it might.

The first precarity lies in the association of formal semantic analysis with the fragmentation of information in digital media. This, of course, draws on a fundamental opposition within hermeneutic thinking between calculative rationality and the life of historically affected interpretation. But is Culbertson's rejection of fragmented reading convincing, as the rapid or even atemporal jumping from one information source to another which precludes full understanding by failing to establish meaningful relations between the hermeneutic subject and its object? Perhaps more pedantically than critically, could it be asked: what of a Nietzsche or a Kierkegaard who used fragments as a mode of expression, the fragmented writing approach of Mark C. Taylor or Avita Ronell, or, of course, any of the pre-Socratics whose teachings come only through scraps of text? There are many cases where the disruption of the relations between past and present communities of knowledge serve to highlight or provoke questions on how those communities are themselves preserved. Far from retreat into the scientific plausibility of formal semantic structures, these examples of fragmented thinking highlight the messiness and complexity of interpretation as an applied practice: which forms of thinking are brought into focus and which are not, what the media are which carry meaning as much as the words or language, and how habits and rituals establish relations as much as logical connection. And while one might feel confident in discerning a "classic" text from a series of unrelated TikTok shorts (even if there are many reasons to challenge the passivity of the fragmented reader of digital media, see Grimwood 2023, pp. 82-85), this would overlook the way that fragmentation within writing and reading has historically transformed the act of understanding into a *performance* of interpretation.

The second precarity also lies in the opposition between living practice – which would include both the interpretative contingency and historicity which provides the "sense of community with people past and present" that Culbertson places at the centre of understanding (2025, p.13) – and the ossification of technical fragments. The general agreement found across the thinkers listed above is effectively held in place by a mutual defence of the "living meaning" of language, a life sustained by its historicity. In turn, situating the object of these hermeneutic inquiries as the generative software (hence the questions: can it think? can it converse? can it create?), rather than dialogue with its proponents or its critics (and, more precisely, both the constitutive rhetoric and speculative fiction at work in dialogue with both).

No doubt, this is enhanced by the grammar of how NLP is engaged: after all, we "ask" AI to do things; ChatGPT and its competitors will respond in the first person. When AI makes a factual error, it is described as "hallucinating." (Choi and Mei 2025) Nietzsche's claim that the "I" is an artifice of European grammar conventions is reborn with agential AI. This is an artifice of its own: just as the employment of the word "natural" in the NLP acronym is a curious double

reference to imitating human thought processes and hiding the mechanics of its computations through user-friendly interfaces, both of which conform to an uncanny replication of “naturalness” rather than actual nature.

Viewing this not as a mere artificial decoy but a function artifice suggests more can be said about the relationship between hermeneutics and AI here. Boris Groys, for example, discusses how AI requires “prompts” to assemble answers, suggesting that it does not *generate* so much as *interpret* its command:

Hermeneutics, as the analysis of interpretations, is at the heart of the human sciences. The ability to interpret seems to define the difference between humans and things, including machines. Things obey natural laws — for example, the law of gravitation — but humans interpret social laws before following them in this or that way, or not following them at all. That is why humans are historical and things are not. Stones react to the law of gravitation today as they did in the past. However, our contemporary interpretation of the state and its laws went through many radical historical transformations. Thus, if AI operates by interpretation, this means that it is also historical. (Groys, 2023, on-line)

On this reading, the historicity argument is inversed: rather than AI not understanding because it cannot appreciate historicity, it is the *very historicity* of AI which frames the manner in which it “understands.” Groys terms AI a “zeitgeist-machine,” which “operates by processing a mass of writing that has been accumulated in a fragmentary and chaotic manner” into an illusory logical structure. This zeitgeist-machine takes the place previously held by the literary canon, which is no longer able to function due to the “mass of accumulated writing” which “cannot be explored and processed by the human mind” and “is experienced as a huge garbage pit into which every new text is thrown as merely an additional piece of garbage” (Groys, 2023, on-line).

Of course, it could be argued that Groys is equivocating history and historicity, with an artisanal sleight of hand. But the key argument is not that AI is an interpreter in the sense that a writer is, but rather an interpreter in the sense of Foucault’s author: that is, a function of the media which carries the historicity of meaning. The interpretation Groys describes is one where the boundaries of tradition are determined by algorithms rather than cultivation. If this is the case, then it would not be enough to differentiate hermeneutic understanding from formal understanding on the basis of historicity; rather, it would be a potentially more challenging case of hermeneutic understanding considering its own use of history, its transmission and its curation in the context of contemporary digital culture. The task facing the close reader is not to oppose NLP’s formal semantic approach with hermeneutic approaches, but to deconstruct the ways in which NLP is both historical and orients its interpretative communities to responses.

In doing so, a third precarity emerges. If the previous two precarities reflect how the artifice functions by subverting the living and what Gadamer termed the “ossified,” this also carries an assumption that machine learning, NLP, LLMs, and AI in general is fundamentally, and perhaps even exclusively, both a mimetic and a mechanistic process. This is, after all, one reason it is so readily associated with calculative rationality; as Thomson notes, the move towards artificial intelligence is fed by “the increasingly widespread belief that the human being is merely a complicated computer, hence something that AIs can not only replicate but even surpass” (2025,

p. 30). Perhaps unsurprisingly, the humanistic traditions underlying hermeneutic responses to AI tend to focus on the cultivation of the mind, and as such on the effects of AI on its users, or the educators of those users.

But this opposition between living interpretation and machine learning is itself an artifice, deployed as it is in between the artistry and wonder of technologies capable of producing AI technology, and fundamental deceptions at work in how this can be achieved. When discussions of education and AI talk of the “human cost,” there is a tendency to focus on what jobs are lost and what knowledge is perverted through the fragmentation of machine processing. Yet, beneath the promise of AI as a low-effort, high reward tool are swathes of people in the Global South hired to “train the algorithm” by relentlessly performing the work of the “bot” (Casilli, 2025; Cant et al., 2024). Or, as Casilli summarises: “AI often means that behind one white-collar worker, millions of blue-collar workers exist as well” (2025, p. 3). Another artifice: what is sold as the work of the machine is, in reality, the work of the exploited human. In this way, Cant et al. situate the idea of autonomous AI within a tradition going back at least as far as von Kempelen’s 18th century “Mechanical Turk,” a chess machine capable of playing human opponents – which was, in reality, a small human hidden inside operating levers. “Today’s AI,” they argue, “operates through a similar illusion” (2024, p. 6).

The dependence of AI on human interventions at every stage means, Casilli argues, that the focus should be less on the expertise of the machine, and more on the people necessary for the machine to work as an artifice of understanding: “Where are they located? What is their employment history? Under what conditions do they work? How are they paid? Where are they recruited?” (2025, p. 33) Or, in more general terms, perhaps the question should not be about how someone learns and understands, whether they have cheated or excelled, but instead draw out the full implications of a simply statement that something is “probably AI”: which is to think through what Wellner describes as a “huge and complex material infrastructure” (2020, p. 2159) that sits behind not only the technology, but its accompanying discourse, and its seemingly endless repetitions.

References

- Babich, B. (2025). Nietzsche AI and Poetic Artistry. In H. Bajohr (Ed.), *Thinking with AI: Machine learning the humanities* (pp. 147-181). Open Humanities Press.
- Beguš, N. (2025). *Artificial humanities: A fictional perspective on language in AI*. University of Michigan Press.
- Cant, C., Muldoon, J. & Graham, M. (2024). *Feeding the machine: The hidden human labor powering A.I.* Bloomsbury.
- Casilli, A. (2024). *Waiting for robots: The hired hands of automation* (S. Brown, Trans.). University of Chicago Press.

Choi, A. & Mei, K. (2025). “What are AI hallucinations? Why AIs sometimes make things up.” *The Conversation*, 21st March <https://theconversation.com/what-are-ai-hallucinations-why-ais-sometimes-make-things-up-242896>

Culbertson, C. (2025). Why natural language processing is not reading: Two philosophical distinctions and their educational import.” *Journal of Applied Hermeneutics*, Article 2, 1-17. <https://doi.org/10.55016/ojs/jah.v2025Y2025.80848>

Foucault, M. (1998). What is an author? In J. Faubion (Ed.), *Aesthetics, method and epistemology* (pp. 205-222). New Press.

Gadamer, H.-G. (1996). *The enigma of health* (J. Gaiger & N. Walker, Trans.). Polity Press.

Gadamer, H.-G. (1960/2004). *Truth and method* (J. Weinsheimer & D.G. Marshall, Trans.). Continuum.

Grimwood, T. (2023). *The problem with stupid: Ignorance, intellectuals, post-truth and resistance*. Zer0 Books

Groys, N. (2023). From writing to prompting: AI as zeitgeist-machine.’ *E-flux*. <https://www.e-flux.com/notes/553214/from-writing-to-prompting-ai-as-zeitgeist-machine>

Herzfeld, M. (2004). *The body impolitic: Artisans and artifice in the global hierarchy of value*. University of Chicago Press.

Kögler, H.-H. (2014). A critique of dialogue in philosophical hermeneutics. *Journal of Dialogue Studies*, 2(1), 47-67. <https://doi.org/10.55207/LWRU9245>

Pinell, P. (2024). Does artificial intelligence speak our language?: A Gadamerian assessment of generative language models. *Political Research Quarterly*, 77(3), 713–728. <https://doi.org/10.1177/10659129241243038>

Thomson, I. (2025). *Heidegger on technology’s danger and promise in the age of AI*. Cambridge University Press.

Wang, J. (2021). Is artificial intelligence capable of understanding? An analysis based on philosophical hermeneutics. *Cultures of Science*, 4(3), 135-146 <https://doi.org/10.1177/20966083211056405>

Wellner, G. (2020). Material hermeneutic of digital technologies in the age of AI. *AI & Society*, 38, 2159–2166 <https://doi.org/10.1007/s00146-020-00952-w>