

Unsettling AI literacies: ethics, power and the pedagogies of critical research

Desestabilizando os letramentos em IA: ética, poder e as pedagogias de pesquisa crítica

Cuestionando las alfabetizaciones en IA: ética, poder y pedagogías de la investigación crítica

Giota Alevizou

King's College London (KCL), London, England

<https://orcid.org/0000-0001-6615-5142>

giota.alevizou@kcl.ac.uk

Abstract

As GenAI systems have moved from novelty to everyday use in research and learning practices, higher education institutions are increasingly adopting them in the realms of “literacy”, “responsible use”, and “ethical awareness”. Such framings often assume strategic simplifications: imagining ethics as a checklist of principles and downplaying the infrastructural forces that shape what counts as legitimate knowledge. This paper argues that approaching AI as a neutral “tool” obscures the value regimes and power asymmetries embedded in its very architecture. Drawing on critical & education technology studies, it proposes a reorientation from AI literacy to critical AI socio-technical practices. Such practices have informed the curricular design, pedagogical implementation, and teaching reflections of two core modules within an online MA distance-learning programme. These are pedagogical practices that aim to foreground situated accountability, epistemic plurality, and alternative imaginaries. By centring observations from curriculum design & learning engagements, this paper advocates for an approach to AI in educational research that foregrounds critical AI literacies: pedagogies attentive to power asymmetries, and that inspire the emergent ways learners may reimagine (and sometimes resist) AI’s role in academic inquiry. This allows for reframing ethics as a politics of value production and infrastructural power and informs how educational research can move beyond the friend-foe binary and toward more contestable engagements with GenAI in knowledge-making.

Keywords: critical AI literacies; socio-technical practices; postgraduate research training; digital futures pedagogy; reflective practitioner inquiry.

Resumo

À medida que os sistemas de IA generativa (GenAI) deixaram de ser uma novidade para se tornarem parte do cotidiano nas práticas de pesquisa e aprendizagem, as instituições de ensino superior vêm adotando essas tecnologias sob as perspectivas de “letramento”, “uso responsável” e “consciência ética”. Tais abordagens frequentemente pressupõem simplificações estratégicas: conceber a ética como uma *checklist* de princípios e minimizar as forças infraestruturais que moldam o que é considerado conhecimento legítimo. Este artigo argumenta que tratar a IA como uma “ferramenta” neutra obscurece os regimes de valor e as assimetrias de poder embutidos em sua própria arquitetura. Com base em estudos críticos e da tecnologia educacional, propõe-se uma reorientação do letramento em IA para práticas sociotécnicas críticas relacionadas à IA. Essas práticas embasaram a concepção de currículo, a implementação pedagógica e as reflexões docentes de dois módulos centrais em um programa de mestrado a distância. Trata-se de práticas pedagógicas que visam destacar a responsabilidade situada, a pluralidade epistêmica e imaginários alternativos. Ao centrar-se em observações sobre a concepção curricular e as dinâmicas de aprendizagem, este artigo defende uma abordagem da IA na pesquisa educacional que privilegie letramentos críticos em IA — pedagogias atentas às assimetrias de poder e que inspirem formas emergentes de os estudantes reimaginarem (e, por vezes, resistirem a) o papel da IA na pesquisa científica. Isso permite ressignificar a ética como uma política de produção de valor e de poder infraestrutural, orientando a pesquisa educacional para além da dicotomia amigo-inimigo, em direção a interações mais passíveis de contestação com a IA generativa na construção do conhecimento.

Palavras-chaves: letramentos críticos em IA; práticas sociotécnicas; formação de pesquisadores; pedagogia de futuros digitais; pesquisa reflexiva da prática

Resumen

A medida que los sistemas de IA generativa han pasado de ser una novedad a formar parte de la vida cotidiana en las prácticas de investigación y aprendizaje, las instituciones de educación superior los están adoptando cada vez más bajo los marcos de la «alfabetización», el «uso responsable» y la «conciencia ética». Estos enfoques suelen partir de simplificaciones estratégicas: concebir la ética como una lista de verificación de principios y restar importancia a las fuerzas infraestructurales que determinan qué se considera conocimiento legítimo. Este artículo sostiene que considerar la IA como una «herramienta» neutral oculta los regímenes de valor y las asimetrías de poder inherentes a su propia arquitectura. Basándose en los estudios críticos y de tecnología educativa, propone pasar de la alfabetización en IA a unas prácticas sociotécnicas críticas en torno a la IA. Dichas prácticas han orientado el diseño curricular, la implementación pedagógica y la reflexión docente de dos módulos troncales en un programa de máster en línea. Se trata de prácticas pedagógicas que buscan poner en primer plano la responsabilidad situada, la pluralidad epistémica y los imaginarios alternativos. A partir de las observaciones derivadas del diseño curricular y de las experiencias de aprendizaje, este trabajo aboga por un enfoque de la IA en la investigación educativa que priorice las alfabetizaciones críticas sobre la IA: pedagogías atentas a las asimetrías de poder que inspiren a los estudiantes a reimaginar (y, en ocasiones, cuestionar o resistirse a) el papel de la IA en la indagación académica. Esto permite replantear la ética como una política de producción de valor y de poder infraestructural, y orienta la investigación educativa para trascender la dicotomía amigo-enemigo, avanzando hacia formas de interacción con la IA generativa — en la construcción del conocimiento — que permitan el debate y el cuestionamiento.

Palabras-clave: alfabetizaciones críticas en IA; prácticas sociotécnicas; formación en investigación de posgrado; pedagogía de los futuros digitales; investigación del profesional reflexivo.

1 Introduction

Generative AI (GenAI) has entered higher education not through debate but as “a fait accompli”. Institutions have responded by reaching for the familiar language of literacy, responsible use, and ethical competence. While these frameworks construct AI as a powerful and increasingly familiar companion, ethics is often presented as a checklist of principles to be applied before research begins. This paper argues that, besides the apparent instrumentalism of such frameworks, they also imply a political intention: they naturalise particular configurations of epistemic power, render invisible the extractive infrastructures on which AI systems depend, and foreclose the more generative, contested, and democratic engagements with AI in knowledge-making that postgraduate teaching and research training urgently need.

In this paper, I propose a theoretically grounded account of how students navigate the promises and frictions of AI in learning and research practices, one that considers the structural conditions shaping their encounters with these systems and the situated, often inventive ways they respond to them. Existing scholarship attends to competencies, risks, and integrity, yet has far less to say about value formation, epistemic plurality, and the micro-political dimensions of increasingly expansive AI-mediated knowledge work.

Drawing on critical technology studies, feminist and decolonial STS, and reflective practitioner observations from postgraduate teaching in two core modules of an online distance-learning programme at a UK Russell Group university, the paper proposes a reorientation from AI literacy to critical AI socio-technical practices. After reviewing relevant literature in the following section, I propose three interlocking concepts that structure this reorientation: AI as socio-technical infrastructure; situated accountability as the epistemological and ethical core of critical AI engagement; and critical AI socio-technical practices as a form of democratic knowledge-making. These concepts offer a framework for educational research that moves beyond the friend-foe, and towards a pedagogy that is reflexive, structurally informed, and open to alternative AI futures. To contextualise these concepts, I draw on insights from curricular design and reflections on patterns of student engagement with pedagogical materials.

2 Literature review

2.1 The domestication of AI in higher education: literacy, responsibility, and the limits of framing

The emergence of large language models (LLMs) as everyday research & pedagogy tools has generated a substantial institutional response; one largely organised around the concept of “AI literacy”. Frameworks produced by major educational organisations have proliferated rapidly. Initiatives such as the EDUCAUSE AI Literacy in Teaching and Learning framework champion a “holistic and durable approach” that balances technical skills, critical evaluation, and ethical considerations (Kassorla *et al.*, 2024), while the ED-AI Lit framework similarly identifies six components (knowledge, evaluation, collaboration, contextualisation, autonomy, and ethics) stressing the importance of critically evaluating AI's implications (Allen; Kendeou, 2024). Long and Magerko's (2020) foundational competency model, widely cited across this literature, defines AI literacy as the capacity to evaluate, communicate with, and use AI as a tool across social and professional settings; a definition that, as will be argued below, both inaugurates and constrains the field's imagination.

What unites these frameworks, and, possibly, what constitutes a productive site of critical scrutiny, is their shared epistemic structure. Empirical reviews of AI literacy curricula show that since 2023, studies positioning AI primarily as a tool have increased rapidly, *dwarfing* perspectives that engage with AI as a sociotechnical system (Gu; Ericson, 2025). Frameworks tend to converge on four areas: technical understanding, evaluative skills, practical application, and ethical considerations (Mihram, n.d.). Such areas, while not without value, structure engagement in ways that consistently foreground individual competence over structural critique. As Ng *et al.*'s frequently cited four-dimensional model illustrates, being “ethical and socially responsible” (Ng *et al.*, 2024, p. 1087), implies that ethics is rendered procedural: a set of principles to be internalised¹ and applied rather than a terrain of value production to be contested. The use of AI is positioned as a discrete competency, separate from the technical and practical ones.

This procedural, checklist-like view of ethics has consequences; it tends to deviate from the conditions under which AI systems are built and deployed. Looking beyond fields within education and educational technology, emerging research shows how institutional

¹ Based on these framework the authors propose: Affective Learning: Intrinsic motivation, confidence, and self-efficacy toward AI.; Behavioural Learning: Behavioural commitment, collaboration, and usage of AI tools.; Cognitive Learning: Understanding, applying, evaluating, and creating AI systems; Ethical Learning: Awareness of ethical principles, privacy, bias, and responsible use of AI.

discourse frames GenAI not only as a tool but as a marker of technological literacy, contributing to new social boundaries around competence, epistemic authority, and inclusion (Rughiniş *et al.*, 2025). The integrity literature has grown rapidly alongside literacy frameworks, with studies mapping generative AI research in higher education between 2022 and 2024 indicating that teaching and learning practices (36%) and ethical considerations (27%) together dominate the field, with academic integrity concerns particularly prominent (Frumin *et al.*, 2026). Nonetheless, this kind of ethical attention remains largely focused on authorship and disclosure rather than on what AI systems do to knowledge.

Systematic reviews of generative AI in postgraduate research have identified “digital labour” and “the blurred lines between authorship and automation” as emerging concerns (Mabirizi *et al.*, 2025), but these are, too, largely framed in terms of disclosure obligations rather than as structural questions about who produces, owns, and profits from research labour. The dominant framing across policy, institutional guidance, and curriculum design thus constructs AI as a powerful but tractable companion: something to be used responsibly rather than critically inhabited. It is this construction that this paper sets out to unsettle.

2.2 Critical technology studies: AI as infrastructure of power

Another body of scholarship — spanning critical technology studies, feminist and decolonial STS, political economy, and critical educational research — argues against instrumentalising frameworks of GenAI use, but they obscure what is most politically consequential about AI systems and infrastructures. Such infrastructures cannot be understood apart from the material, institutional, and discursive conditions of their production, and any pedagogy built on a GenAI-as-a-tool metaphor will reproduce rather than contest the power asymmetries embedded in those conditions.

The infrastructural argument has been presented by several critical researchers, who (Couldry; Mejias, 2019; Crawford, 2021; Hao, 2025) demonstrate that AI is a technology of extraction (drawing on minerals, low-wage labour, and data harvested from every sphere of social activity) and that its planetary network concentrates rather than distributes power. Hao, for example, drawing on the viewpoints of Silicon Valley engineers, Kenyan data labourers, and Chilean water activists, traces how the wheels of extraction (in terms of compute power, data volume, and human cleaning labour paid at sweatshop wages) underpin the polished outputs that reach research environments in the Global North, while Couldry and Mejias (2019) locate these dynamics within the *longue durée* of colonial extraction, arguing that data colonialism constitutes a new phase of a five-hundred-year process in which appropriation

presents itself as civilisation. Srnicek (2025) brings forward current developments perspectives through a sharp political economy analysis: far from being disruptive, contemporary AI looks set to consolidate the power and position of Big Tech, whose strategies of value-chain capture and compute monopoly are positioning a handful of firms as the infrastructural gatekeepers of global knowledge production. These accounts implicitly draw on Bowker and Star's (1999) concept of infrastructure in relation to knowledge production: that the most powerful systems are those that become invisible in use, shaping what can be known and enacted from below the threshold of conscious choice.

The decolonial tradition of tech researchers sharpens the political stakes of these epistemological claims. Safiya Umoja Noble (2018) demonstrates that algorithmic systems are neither neutral nor equal-opportunity, reproducing racial and gender hierarchies through design decisions that present themselves as technical rather than political. Birhane *et al.* (2022) provide rich insights into the ways in which ML research systematically privileges the values of its Big Tech-affiliated producers (Birhane *et al.*, 2022) disproportionately benefiting those already powerful. For educational research, this matters because learners who uncritically adopt LLMs may also reproduce specific configurations of epistemic authority while remaining structurally unaware of it.

These critical and decolonial insights find their discursive and political articulation in work on sociotechnical imaginaries and the politics of digital futuring. Alevizou and Gallagher (2026) develop the concept of the autonomy-automation paradox to analyse how digital futures discourse forecloses democratic imagination, showing how metaphors (from “frontier” and “cloud” to “smartness” and “responsible AI”) function as floating signifiers (Laclau, 2005): stabilising uncertainty, naturalising automation as the fulfilment of human agency, and rendering the tension between autonomy and democratic participation politically inert. Their call for forms of futuring grounded in embodied collective intelligence (acknowledging entanglement and cultivating alternative imaginaries rather than accommodating pre-given technological trajectories) has inspired the pedagogical ambition this paper pursues.

Within educational research, specifically, an important cluster of critical scholarship converges on similar concerns, providing the educational grounding that the critical technology studies tradition alone cannot supply. Shapiro, Souto-Otero and Watermeyer (2026) offer perhaps the most direct challenge to functional AI literacy from within the field. Arguing that conventional frameworks remain “anchored in functional skills acquisition and fail to address the fundamental epistemological challenges posed by probabilistic, opaque

algorithmic systems” (p. 1), they propose reconceptualising AI literacy as a metacognitive social practice: one that transcends individual competencies to encompass collective capacity for critical engagement with AI design, deployment, and governance. Their four-dimensional framework, which extends Selber’s functional, critical, and rhetorical literacies (Selber, 2010) by adding a metacognitive dimension, argues for a central competence in an era of generative AI shifts from rhetorical production to reflexive oversight: evaluating when to trust, question, or recalibrate AI-generated content, and participating actively in shaping the governance conditions under which AI operates. Crucially, they argue that lack of clarity about AI literacy “is not merely a pedagogical problem but a governance problem that redistributes risks downward to individuals while leaving wider structural issues insufficiently examined” (Shapiro, Souto-Otero and Watermeyer, 2026, p. 2). I aim to map this framing directly onto this paper’s argument that the procedural ethics of AI literacy frameworks actively foreclose structural critique. Their analysis of the “platform university” — in which teaching and research have been re-infrastructured by platforms, normalising surveillance-evaluation hybrids that crowd out critical and participatory pedagogies (Komljenovic *et al.*, 2025; Williamson *et al.*, 2023) — provides the institutional context within which the pedagogical observations in this paper are situated.

This institutional critique is deepened by Costa and Murphy (2025), who argue that current approaches to generative AI in education reproduce a “tyranny of freedom” in which the language of empowerment and individual choice masks technocratic governance, echoing Freirean critiques of literacy as solutionism. Their analysis resonates with the broader argument advanced by recent critical scholarship in the *Learning, Media and Technology* journal — notably the observation that AI literacy risks becoming synonymous with how to use AI rather than with questioning the social reality it ushers in, precisely the political and emancipatory approach that a Freirean literacy tradition would demand (Costa; Murphy, 2025; Williamson *et al.*, 2023). Selwyn’s (2024) sustained critique of digital and AI literacy agendas as overly instrumental — tied to employability and adoption while neglecting critical digital citizenship, power asymmetries, and the political economy of technology — provides a longer lineage in critical education studies. The decolonial thread in this literature is developed directly by Zembylas (2023) and related work, applying decolonial AI frameworks to higher education teaching and arguing that ethical AI pedagogy must interrogate the coloniality of AI systems as well as their encoding of Global North epistemological defaults, and their extraction of data from communities without consent or reciprocity. Swist *et al.*’s (2023) socio-technically just pedagogies framework for curriculum-making in higher

education similarly calls for pedagogical design that foregrounds relational accountability and structural critique rather than technical proficiency.

These traditions share an emerging argument that runs directly counter to the literacy discourse that opened this review: what students need is not better competence in using AI but a different relationship to it altogether, one grounded in structural analysis, epistemic accountability, and the capacity to contest and reimagine the conditions under which AI enters knowledge-making. The conceptual framework that follows in section 3 is built on this.

2.3 AI in research training: tensions and emergent practices

The third, emerging body of literature this paper engages reflects the speed at which GenAI has entered postgraduate research environments. The growing empirical scholarship on students and LLMs in academic work documents shows a consistent pattern: instrumental adoption accompanied by persistent epistemic and affective discomfort. Research attests that GenAI enhances writing across multiple dimensions (*e.g.* from ideation and language refinement to cross-lingual support) while simultaneously producing overdependence, authorship anxiety, and what students describe as a creeping estrangement from their own intellectual voice (Kosmyna *et al.*, 2025; Zhai *et al.*, 2024). These affective responses may indicate that something structurally important is at stake in the relationship between writing, thinking, and selfhood.

Some work on academic writing as a social and material practice provides a productive theoretical anchor here: writing is not the transcription of pre-formed thoughts but a constitutive practice through which researchers come to know. When LLMs are inserted into this process, the question is not whether the product is authentic in an integrity sense, but what happens to the epistemic work of research when key stages of it are delegated to a system generating statistically plausible text from aggregated patterns. Boud and Lee (2009) and Kamler and Thomson (2014), who have researched doctoral writing, centre the developmental and identity-forming dimension of research training: the slow construction of scholarly voice, which sits in productive tension with the LLM's capacity to produce instantly what researchers take years to develop. Likewise, the concept of self-authorship (Magolda, 2023) is relevant here for exploring this impact: the development of an internal locus of epistemic authority that AI, when uncritically adopted, threatens to foreclose before it is fully formed.

Though very preliminary at this stage, what I aim to offer in this paper is an account of how postgraduate students navigate some of these tensions in practice: using AI to interrogate its own value regimes, resisting or refusing automated suggestions, or exposing how model outputs flatten conceptual nuances. These practices are neither expressions of mastery nor failures of compliance. They may point towards a different conception of what AI pedagogy might be for. The conceptual framework developed in the next section offers some direction towards this.

3 Conceptual framework: from AI literacy to critical AI literacies

The literature reviewed above maps a productive tension in how higher education is responding to generative AI: between a well-resourced discourse of literacy, responsibility, and ethical competence, and a body of critical scholarship that identifies the political assumptions those framings carry (their naturalisation of particular configurations of power, their rendering invisible of the extractive conditions on which AI systems depend). The conceptual framework proposed here does not resolve this tension but argues that inhabiting it critically (making it pedagogically generative rather than procedurally managed) is precisely the work of a critical AI education. Three interlocking concepts structure this argument: AI as socio-technical infrastructure, which unpacks what the “tool” metaphor obscures; situated accountability as epistemic practice, which is the framework’s theoretical core and most original contribution; and critical AI socio-technical practices as democratic knowledge-making, which translates both into a pedagogical programme that is principled without being prescriptive.

3.1 AI as socio-technical infrastructure

The first move within the framework is to reframe AI not as a tool but as infrastructure. The tool metaphor carries assumptions (about neutrality, instrumentality, and the separability of technology from its uses) that limit the kinds of critical engagement this paper is concerned with. The alternative draws on Bowker and Star’s (1999) concept of infrastructure: a technical substrate as well as a system embedded in social relations, invisible to habitual users, and profoundly shaping which actions and knowledge forms are possible. Infrastructures encode decisions about whose needs matter, whose knowledge counts, and whose labour is legible.

For postgraduate research training, this reframing is consequential. To understand AI as infrastructure is to ask not only “how do I use this effectively?”, but “who built this, on whose data, through whose labour, toward whose ends?”. It is to recognise that the generic academic voice LLMs produce is not a stylistic quirk but an epistemological default: the materialisation of training decisions and fine-tuning choices that encode certain knowledge hierarchies over others. If AI is infrastructure, the pedagogical question is not how to teach its use but how to teach its politics.

3.2 Situated accountability as epistemic practice

If the first concept names the problem, the second proposes a pedagogical response. Situated accountability, as proposed here, is neither a principle to be adopted nor a competency to be achieved. It is an ongoing practice of epistemic positioning: a commitment to asking, in any knowledge-making encounter, from where am I knowing, through whose systems, and to whom am I answerable?

The concept extends this tradition in three directions. The first, draws on Markham’s (2021) argument that critical engagement with AI futures requires working at the limits of what current discourse makes thinkable: cultivating the capacity to intervene in the speculative architectures through which algorithmic futures are constructed. Situated accountability is therefore not only historically informed (accounting for how knowledge was produced) but also forward-looking: it involves the reflexive exercise of speculative agency toward knowledge-making practices that are not yet available but are not, therefore, impossible. This connects to Alevizou and Gallagher’s (2026) call for a kind of futuring grounded in embodied collective intelligence, which informs the design logic of both modules this paper draws on: a shared commitment that learners develop not only critical analysis of existing digital systems or platforms, but also the imaginative and practical capacity to propose alternatives to them.

The second extension draws on Birhane *et al.* (2022) and Shapiro, Souto-Otero and Watermeyer (2026) to argue that “situated accountability” means treating the AI system itself as an object of accountability: asking what values it encodes, whose epistemological defaults it reproduces, and what forms of knowledge it makes harder to think. This is particularly salient in the context of the second module, whose sustained engagement with digital colonialism, linguistic justice, and data sovereignty (including the question of whose languages, knowledge systems, and cultural practices are legible in LLM training corpora)

gives students the structural vocabulary to ask these questions not only of AI in the abstract but of the specific tools they use in their own research.

The third extension draws on decolonial frames to give situated accountability a structural rather than purely individual dimension. Accountability here is not simply a matter of individual researchers reflecting on their positionality; it requires recognising that those positions are structured by relations of extraction and asymmetric power that exceed any individual's control.

Situated accountability, so understood, is a disposition rather than a procedure: cultivated through practice, deepened through encounter, and necessarily incomplete. It requires the willingness to refuse the comfortable authority LLM outputs offer, to sit with the discomfort of partial and contested knowledge, and to maintain one's claim to intellectual authorship in the face of systems that offer to relieve that burden. This connects to Baxter Magolda's (2004) concept of self-authorship (the development of an internal locus of epistemic authority), which an LLM, if uncritically adopted, threatens to foreclose before it is fully formed in the postgraduate researcher.

3.3 Critical AI socio-technical practices as knowledge-making

The third concept synthesises the diagnostic work of the first and the epistemological core of the second into a positive account of what pedagogical practice might look like. The shift from "AI literacy" to "critical AI socio-technical practices" is deliberate and carries theoretical weight at every level.

"Literacy" implies a stable competence to be acquired: a destination. "Practices", by contrast, are always in process: plural, contested, revised through encounter, and constitutively relational. This difference matters pedagogically because the relationship between researchers and AI is not stable but emergent, actively made through every interaction in every research context. The "socio-technical" qualifier signals, in the tradition of STS, that AI systems are not technical objects with social consequences but assemblages in which the social and technical are mutually constitutive from the outset. To engage AI socio-technically is to resist the separation between the algorithm and the social relations it embeds, between the output and the labour that produced it.

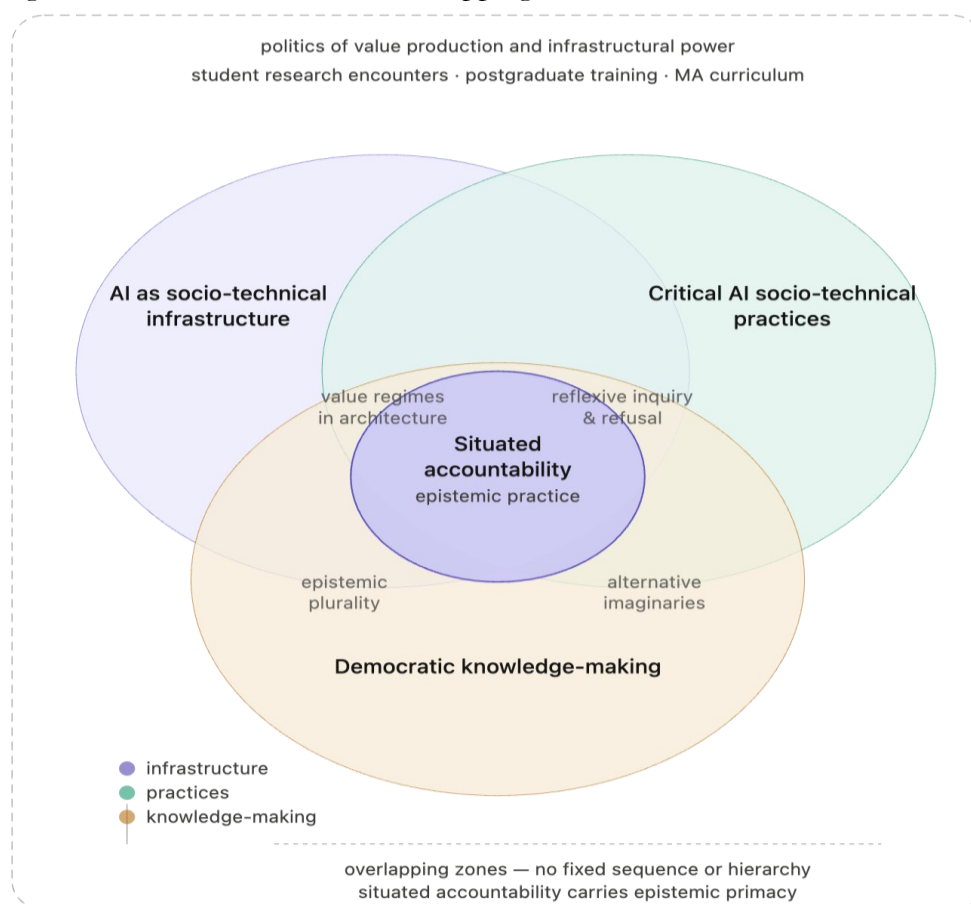
Critical here means oriented toward transformation, towards developing consciousness about the conditions of one's situation and cultivating agency to act differently within and against them. Shapiro, Souto-Otero and Watermeyer (2026) frame this metacognitive AI literacy as civic practice; this paper aims to advance this, arguing that such

practices can be grounded in structural analysis of AI as infrastructure rather than in the development of individual reflective competence alone.

Concretely, such practices include treating AI outputs as objects of analysis rather than sources of authority; using AI to expose rather than conceal the assumptions embedded in research framings; attending to the affective and epistemological textures of AI-assisted work (*e.g.* opportunities for creative invention but also but also anxieties, estrangements, moments of resistance) as data about what is at stake rather than problems to be resolved; and experimenting collectively with modification, and reimagination of AI's role. These are orientations to be developed, not techniques to be applied: necessarily context-specific, disciplinarily situated, and relationally produced.

The three concepts interlock in a fluid sequence. The infrastructural analysis establishes what AI is and does so at a structural level; situated accountability establishes the epistemological and ethical stance from which critical engagement becomes possible; and critical AI socio-technical practices translate both into a pedagogical programme that is actionable without being prescriptive.

Figure 1 – Critical AI literacies overlapping dimensions



Source: Created by the author.

The three concepts are not offered as a finished system but are reliant on tutor/student voices and interaction. The framework anticipates that those voices will complicate, extend, and in places resist it.

4 Methodological Note: critical AI literacies in the virtual classroom

The observations that follow arise from sustained pedagogical practice across two postgraduate online distance learning modules designed by the author and a collaborator as core components (modules) of an MA programme at a UK University. The first module examines current and emerging trends in digital culture and media through a historically informed sociotechnical lens (moving from platformisation, appification, and algorithmic governance through to creative AI and AI ethics); the second deepens this critical vocabulary through sustained engagement with a range of topics including sustainability, digital colonialism, linguistic plurality/justice, data sovereignty, copyright, digital labour, and the design of sustainable digital futures. These modules were designed so that students encounter AI not as an isolated topic but as a formation within broader digital landscapes, which they learn to analyse structurally, politically, and ethically over twenty weeks of study.

The observations that follow draw primarily on the first module, which provides the main arc of the observational/reflective account across all four cohort rounds. From the second module, two to three specific components are primarily drawn upon, pertaining to digital colonialism, linguistic plurality, and copyright /authorial anxiety. This selective scope reflects the nature of reflective practitioner inquiry: the observations reported are those where patterns were sufficiently consistent across cohorts to warrant analytical attention, not a comprehensive account of either module's pedagogical arc. Both modules share similar underlying design commitments (critical engagement with digital systems, reflective practice, and movement from structural analysis toward alternative imaginaries) but differ in emphasis: the first builds the sociotechnical analytical vocabulary; the second applies and extends it toward sustainability, decolonial practice, and intervention design.

The student cohort has been diverse in ways that matter analytically: it spans recent graduates and established professionals across multiple sectors and geographical contexts, bringing to the modules a range of relationships to digital systems (as users, workers, designers, and citizens) that enrich discussion and complicate easy generalisations. The observations reported here are drawn from four successive cohort rounds, giving the patterns described a degree of consistency across time and across different student compositions.

The mode of delivery shapes what can and cannot be observed. Student encounters with AI happen largely outside the direct purview of teaching: in the space between a discussion prompt and a submitted response, between a reading and a forum contribution. What becomes visible are traces: the analytical moves students make in asynchronous discussions, the conceptual frameworks they reach for when asked to reflect, the moments of discomfort or resistance that surface in their writing. These traces are interpreted here through the conceptual framework developed in the preceding sections, and offered in the tradition of reflective practitioner inquiry (Schön, [1983] 2009): a mode of educational knowledge production that is necessarily partial, positioned, and accountable to the pedagogical relationship from which it emerges. No formal ethics approval was required for this work. The observations draw on curricular design decisions and reflective notes on patterns of interaction across cohorts. What follows is an account of tendencies and recurring moments interpreted by a teacher who is also one of the modules' designers: a location that is itself a form of situated knowledge, with all the productive partiality that implies.

More specifically, the observations reported here were produced through three overlapping practices sustained across four cohort rounds. First, facilitation notes: brief written records made prior to and immediately after each week's discussion moderation, capturing recurring analytical moves, points of friction, and moments where students appeared to engage AI use reflexively rather than instrumentally. Second, pattern recognition across student-generated texts: systematic re-reading of asynchronous activity contributions with attention to the presence or absence of structural analysis, the conceptual frameworks students reached for, and the affective registers in which AI encounters were described. Third, retrospective synthesis: at the end of each cohort round, a reflective summary of the most consistent and analytically significant patterns was made and cross-checked against the preceding rounds to identify what was cohort-specific and what recurred across different student compositions. No individual student data was collected or retained.

This positioning expresses the situated accountability the framework proposes. To claim to observe students is already to occupy a particular epistemic location, shaped by institutional authority, disciplinary convention, and the asymmetries of the teaching relationship. As Shapiro, Souto-Otero and Watermeyer (2026) argue, metacognitive AI literacy is not merely an individual achievement but an organisational and institutional condition, shaped by what pedagogical architectures make possible and what modes of delivery make visible or invisible. The observations that follow are doubly situated: emerging from a specific pedagogical encounter with AI in postgraduate research training and

interpreted through a framework that insists all knowledge is partial, located, and accountable.

5 Critical AI literacies in practice

5.1 Frictionless absorption and the invisibility of infrastructure

The four subsections that follow report patterns observed across both modules, though with different weights. The patterns are organised to mirror the conceptual framework: from the diagnostic (AI as invisible infrastructure) through the epistemological (the limits of procedural ethics; the emergence of situated accountability) to the generative (alternative imaginaries and critical AI socio-technical practices).

A consistent pattern across postgraduate encounters with GenAI is neither resistance nor enthusiasm but, I would argue, something slightly more ambivalent: a kind of frictionless absorption, almost immediately shadowed by unease. Students deploy LLMs with little hesitation in the practical work of research (generating literature overviews, structuring arguments, smoothing prose, crossing language barriers) while simultaneously expressing a discomfort they struggle to articulate and that standard responsible-use guidance is poorly equipped to address. This coexistence is a design effect: AI tools present themselves as neutral, weightless, and serviceable (by now embedded in search engines) in ways that earlier platform architectures did not, making the political economy of their operation structurally harder to distinguish.

The pattern becomes visible across both modules. Within the first weeks, students demonstrate sophisticated critical instincts about the platforms that structure their everyday digital lives. Reflecting on the interfaces they inhabit daily, students notice how design decisions compose user identity: how a platform addresses them personally through pronoun choice, algorithmic curation, and the illusion of individualised space, while simultaneously positioning them within a broader economy of data extraction and behavioural profiling. One student, reflecting on a streaming platform's character icon system, observes that self-expression on the app is mediated by corporate network constraints: what feels like individuality is, in fact, a choice within a closed set of options designed to sustain the platform's own coherence. Another notes the asymmetry between the comfort of algorithmic personalisation in a shopping context and the discomfort of what feels like self-surveillance on a social media platform. A third traces how a single day's digital activity moves across multiple roles (consumer, citizen, learner, content creator, listener) and recognises that the

transition between these roles is seamless precisely because the platforms themselves script it, making each identity feel natural within its own logic. The pedagogical facilitation at this stage aims not to supply the analysis but to name and develop what students are already doing: making explicit that the leakiness they observe in platform architectures [following Chun's (2017) account of networked habit, which was part of the core readings] is a structural feature rather than an incidental one.

What is notable is that this critical awareness does not automatically extend to AI systems as a category of infrastructure. The conversational interface of an LLM presents a partner rather than a system, an interlocutor rather than a political economy (a design effect that the module's attention to appification helps to name: the fetishisation of the app as a self-contained object concealing the infrastructure, governance, and data extraction beneath applies with equal force to the LLM interface). What gets obscured are both "labour" and the relational and infrastructural dimensions at which power operates.

This invisibility is particularly consequential in curricula committed to futures thinking and alternative imaginaries. When students engage with Zylinska's (2020) analysis of AI art (asking not whether machines can be creative but how creativity is distributed across humans, machines, and infrastructures) the analytical move they are invited to make is precisely the one that unreflective LLM use makes harder: the refusal of the anthropocentric question in favour of the relational and political one. A model trained to produce plausible elaborations of existing discourse will reproduce the anthropocentric framing as the natural starting point. A student who has genuinely grasped Zylinska's provocation (also a core reading) and then uses a model to develop the argument may find that the provocation has been translated back into the question it was trying to displace. Similarly, when students in the second module encounter questions of linguistic justice, machine translation, and whose languages and knowledge systems are legible in LLM training corpora, the tools for asking structural questions about AI are present (whose data, on whose terms, to whose benefit), but the transfer to their own research practice requires explicit pedagogical invitation.

The implication is not that AI assistance should be prohibited but that its epistemological defaults must become an explicit object of analysis. Students who have learned to ask whose interests a platform's design serves have all the tools needed to ask the same questions of the AI systems they use in their academic work. The pedagogical challenge is to make that transfer explicit rather than leave it as an inference most students will not reach without prompting.

5.2 From procedural ethics to political accountability

A consistent pattern across both modules reveals that even sustained engagement with platform political economy, algorithmic governance, and digital colonialism does not automatically produce what the conceptual framework calls political ethical reasoning. Students can identify bias, name fairness and transparency as values, and articulate concerns about corporate accountability. What they find structurally difficult (until explicitly prompted) is connecting these ethical categories to the material and labour conditions of AI's production: the move from use-ethics to production-ethics that situated accountability proposes as its core.

This difficulty is both conceptual and affective. Applying structural critique to one's own tools requires locating oneself within the conditions one is critiquing: recognising that one is not simply an analyst of AI's political economy but a participant in it. This is a more uncomfortable move than critiquing structures from the outside, and one that standard responsible use guidance forecloses by offering a procedural resolution (acknowledge, attribute, check for accuracy) that makes the structural question feel already answered. This is precisely the checklist ethics that Ricaurte (2019) identifies as operating at the micro level (individual behaviour, responsible use, bias mitigation), while leaving the macro and meso conditions² of AI production unexamined.

Several theoretical bridges make structural self-implication possible. The discussion of vectorialism (Wark, 2004) — the insight that controlling the conditions under which information becomes discoverable is more totalising than controlling content itself — gives students the conceptual tools to analyse LLMs as vectoralist infrastructure: systems that shape not merely what knowledge is available but what arguments feel thinkable, legitimate, and expressible in academic form. The privacy debates I prompted for (structured around whether privacy is a “luxury” rather than a “right”) generate some interesting and politically engaged student reasoning across the modules: arguments about structural asymmetry, the non-consensual character of the platform social contract, and privacy as a collective condition rather than individual preference. Students who can make these arguments about data privacy in social media contexts already possess the political vocabulary for asking

² Macro conditions refer to geopolitical and planetary dimensions: military infrastructures, extractive resource economies, environmental degradation, while meso conditions are identified as institutional and social governance: how platform monopolies organise public life, distribute power, and shape what counts as legitimate knowledge.

equivalent questions about AI training data (whose data, under what conditions of consent, to whose benefit, at whose cost).

The second module deepens this further. Its sustained engagement with digital colonialism, linguistic justice, and data sovereignty gives students structural vocabulary that the first module's platform political economy opens toward but does not fully develop. Students who have encountered the argument that Global English and monolingual LLM training corpora reproduce particular epistemic hierarchies (rendering some knowledge systems legible and others invisible) are equipped to ask not only what AI systems do in the world but what the specific tools they use in their research are doing, and on whose terms. The pedagogical challenge across both modules is making that continuity legible: between students' critiques of surveillance capitalism and digital colonialism in their discussions, and the AI infrastructure they use in their own knowledge-making.

5.3 Situated accountability in practice: micro-resistances, affective knowledges

A third and more generative pattern emerges across both modules: the development, among some students, of nascent situated accountability (forms of engagement with AI and digital systems that are neither instrumental adoption nor principled refusal, but something more inventive and politically alert). These practices are not the product of explicit instruction; they arise from the encounter between the modules' theoretical resources and students' own experience of navigating digital systems as researchers.

The most visible form involves attending critically to dominant system logics rather than accepting them as given. One student, reflecting on search engine behaviour, describes preferring Reddit, Quora, and Facebook groups to Google because they feel more manual and organic. The facilitation response names this as a counter-algorithmic instinct while immediately complicating it: these platforms are not outside algorithmic enclosure but differently governed by it (vectoralist infrastructure nonetheless, as Wark (2004) would have it). What begins as personal preference is developed into situated awareness: not "I prefer Reddit to Google" but "I am learning to ask whose interests each vector serves". This move (taking the student's experience seriously as epistemically significant while inviting structural analysis) models what situated accountability looks like pedagogically: attending to one's own navigations of digital systems with both critical rigour and genuine curiosity.

A creative exercise that asks students to generate AI images of human concepts (memory, grief, labour, utopia), document their prompts, and reflect on where creativity occurred is precisely structured to make the encounter with AI an object of analysis rather than a transparent service. Students confront, in real time, the gap between their own imaginative intention and the model's probabilistic output: aesthetic defaults and knowledge hierarchies they had not anticipated and had not chosen. One student connects creativity to emotional renewal and self-making, articulating a conception of creative agency that is irreducibly personal and embodied (precisely what a probabilistic model cannot replicate). Another identifies creativity as collaborative and socially embedded, using this to think through what human creative decisions in AI-assisted work consist of. These are not reflections on AI creativity in the abstract; they are students developing resources for understanding their own epistemic agency in relation to AI systems (the ground on which situated accountability is built).

One thread emerges with richness across both modules: a student whose research engages with fan culture and emotional labour. The modules' theoretical attention to affect as a dimension of digital life (how users' emotional investments in platforms and content are simultaneously genuine and extractable) gives this student vocabulary for thinking about the affective dimensions of their own research encounter with AI. What does it mean to use a model trained on vast quantities of human emotional expression to theorise a research object whose central feature is precisely the affective investment of human participants? The student's attentiveness to what the model can and cannot do in relation to affect exemplifies what Baxter Magolda (2004) calls self-authorship: the maintenance of an internal locus of epistemic authority in the face of systems that offer to relieve that burden. The second module deepens this thread further: its engagement with digital labour, copyright, and the extractive conditions of AI production gives students additional structural vocabulary for asking whose emotional and creative labour underwrites the outputs they are using.

These affective textures reveal something that maybe epistemically significant: evidence that students are encountering something structurally important in their AI use, even when they lack the vocabulary to name it. A critical AI pedagogy that addresses only the cognitive and procedural dimensions of AI use will miss the most generative material: the moments of discomfort, resistance, and creative friction that signal emerging situated accountability about the conditions of AI-mediated knowledge production.

5.4 Reimagining AI: from critique to alternative practice

Both modules are designed around the conviction that critique must exceed mere criticism and generate strategic, generative thinking. This commitment is embedded in the assessment structures (which ask students to go beyond identifying problems and towards proposing alternatives) and in the facilitation practice that characterises weekly interactions: the consistent move from description to critique, and from critique to the question of what else might be possible.

The most significant student responses are those that apply this futures-orientation reflexively: turning the invitation to imagine alternative digital landscapes back onto the question of AI in research practice itself, and onto the challenges that AI as infrastructure poses for economic, environmental, and sociocultural sustainability. The theoretical vocabulary that makes this reflexive move possible is produced through the modules' own design logic: their sustained argument that dominant AI futures are shaped by particular sociotechnical imaginaries and that those imaginaries are contestable rather than inevitable. This is the same argument that Alevizou and Gallagher (2026) develop in their account of the autonomy-automation paradox (a paper that grows directly from the intellectual architecture of this programme rather than forming part of the student reading list). Students equipped with this argument can ask not “how do I use this responsibly?” but “what kind of knowledge-making practice do I want to pursue, and what role, if any, does this system play in it?”

The Te Hiku Media case study (Brown *et al.*, 2024; Hao, 2022), introduced in the second module's digital colonialism week, captures situated reimagination with particular clarity. Students encountering a Māori community's insistence that language AI be built within, by, and for community governance (rather than extracted by external actors) encounter a concrete instantiation of what data sovereignty as collective practice looks like. One student's observation (that digital self-sufficiency may not be possible, but digital sovereignty remains an imperative) articulates precisely the disposition the framework calls critical AI socio-technical practices: not opting out of digital infrastructure entirely, but insisting on determining, collectively and on one's own terms, the conditions under which that infrastructure operates. Another student's analysis of how Te Hiku Media inverts the dominant political economy of AI (treating language as relational, situated, and cared for rather than abstractable and scalable) demonstrates the capacity for alternative imaginary thinking that both modules' futures orientation is designed to cultivate.

What the collective arc of these observations suggests is that the conditions for critical AI socio-technical practices are not exceptional. They are producible through theoretical frameworks that give students the vocabulary to name what they are already, often partially, beginning to see. The capacity to inhabit the tension between AI's promises and its dangers reflexively and contestably is a consequence of the analytical frameworks they are given.

6 Conclusion

This paper has argued that prevailing responses to generative AI in higher education (literacy frameworks, responsible use guidance, and procedural ethics) reflect important institutional intentions but rest on assumptions that warrant closer examination. By framing AI primarily as a tool to be mastered and ethics as a set of principles to be applied, these frameworks tend to sideline questions of infrastructural power, render less visible the material and labour conditions on which AI systems depend, and leave underexplored the less generative and contested engagements with AI in knowledge-making that postgraduate research training might otherwise cultivate.

The conceptual framework proposed here offers an alternative orientation. Drawing on critical technology studies, feminist and decolonial STS, and the political economy of AI infrastructure, it reframes the pedagogical question from how we may teach students to use AI responsibly to how we may cultivate the conditions in which students can engage with AI reflexively, structurally, and democratically. The three interlocking concepts do overlap: the infrastructural analysis names what the tool metaphor obscures; situated accountability provides the epistemological and ethical stance from which critical engagement becomes possible; and critical AI socio-technical practices translate both into a pedagogical programme that is actionable without being prescriptive, and principled without being procedural.

Observations from reflective practitioner inquiry in two postgraduate online distance-learning modules illuminate what this reorientation looks like in practice. Students who encounter AI systems or infrastructures within curricula already committed to sociotechnical analysis, platform political economy, and alternative digital futures bring considerably richer critical resources to their GenAI encounters than literacy frameworks assume. The gap between students' structural critique of digital systems in general and their under-reflective engagement with AI tools in particular persists and remains instructive. It is not a failure of intelligence but a design effect: the conversational interface of the LLM

performs neutrality in ways that earlier platform architectures do not, making the conditions of its production structurally harder to see. The pedagogical work of making that infrastructure visible (drawing the thread from vectorialism to LLM epistemology, from data colonialism to training data consent, from platform political economy to the generic academic voice, among others) is demanding, iterative, and never fully complete. The observations reported here suggest it may be possible, and that when it happens, something qualitatively different may indeed emerge: students who treat AI not as a given to be used or refused but as a contested sociotechnical object to be analysed, interrogated, and reimaged.

Several limitations shape what can be claimed from these observations. The reflective practitioner mode of inquiry, while theoretically consistent with the paper's situated accountability framework, necessarily limits generalisability. What surfaces in asynchronous forum contributions, synchronous discussions, assignment design, and facilitation exchanges is partial; it reflects what online distance learning makes visible rather than the full texture of student-AI encounters. The observations draw on a single programme context, within a specific institutional setting, with a cohort that (while diverse in professional background, national context, and career stage, spanning recent graduates through established practitioners) is nonetheless self-selected toward critical digital inquiry. Whether the pedagogical conditions described here are replicable across different disciplinary, institutional, and cultural contexts remains an open and important question, given the decolonial commitments of the framework itself.

These limitations point toward several directions for further study. Comparative research across disciplinary and institutional contexts would test the transferability of the critical AI socio-technical practices framework. Longitudinal research tracking how students' relationships with AI evolve over a programme (and across programmes), rather than within individual modules, would shed light on the developmental dimensions of situated accountability. Participatory research designs that give students a direct role in generating and interpreting the empirical account would enact, rather than merely propose, the kind of bottom-up knowledge-making the framework calls for.

The conditions for critical AI socio-technical practices are not exceptional. They are producible through deliberate pedagogical design: through curricula that treat digital systems as politically contestable rather than technically inevitable; through facilitation that takes students' affective and critical responses seriously as data about what is at stake; and through theoretical frameworks that give students the vocabulary to name what they are already, often partially, beginning to see. The task for educational research is not to resolve

the tension between AI's promises and its dangers, but to create the pedagogical conditions in which that tension can be inhabited critically, contestably, and with genuine democratic agency.

References

- ALEVIZOU, G.; GALLAGHER, R. Futures at the Threshold: Between Autonomy and Automation. *Philosophy & Technology*, v. 39, n. 96, 2026. DOI: <https://doi.org/10.1007/s13347-026-01105-5>. Accessed: 10 Jun. 2026.
- ALLEN, L. K.; KENDEOU, P. ED-AI Lit: An Interdisciplinary Framework for AI Literacy in Education. *Policy Insights from the Behavioral and Brain Sciences*, v. 11, n. 1, p. 3-10, 2024. DOI: <https://doi.org/10.1177/23727322231220339>. Accessed: 10 Jun. 2026.
- BIRHANE, A.; KALLURI, P.; CARD, D.; AGNEW, W.; DOTAN, R.; BAO, M. The Values Encoded in Machine Learning Research. *2022 ACM Conference on Fairness, Accountability and Transparency*, p. 173-184, 2022. DOI: <https://doi.org/10.1145/3531146.3533083>. Accessed: 10 Jun. 2026.
- BOUD, D.; LEE, A. (eds). *Changing Practices of Doctoral Education*. London: Routledge, 2009. DOI: <https://doi.org/10.4324/9780203870488>. Accessed: 10 Jun. 2026.
- BOWKER, G. C.; STAR, S. L. *Sorting things out: Classification and its consequences*. Cambridge: The MIT Press, 1999.
- BROWN, P. T.; WILSON, D.; WEST, K.; ESCOTT, K.-R.; BASABAS, K.; RITCHIE, B.; LUCAS, D.; TAIA, I.; KUSABS, N.; TAKA KEEGAN, T. Māori Algorithmic Sovereignty: Idea, Principles, and Use. *Data Science Journal*, v. 23, n. 15, 2024. DOI: <https://doi.org/10.5334/dsj-2024-015>. Accessed: 10 Jun. 2026.
- CHUN, W. H. K. *Updating to remain the same: Habitual new media*. Cambridge: The MIT Press, 2017.
- COSTA, C.; MURPHY, M. Critical education, generative artificial intelligence and the tyranny of freedom: A critique of modern 'technocracy'. *Technology, Pedagogy and Education*, v. 35, n. 3, p. 1-17, 2025. DOI: <https://doi.org/10.1080/1475939X.2025.2547728>. Accessed: 10 Jun. 2026.
- COULDRY, N., & MEJIAS, U. A. *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Redwood City: Stanford University Press, 2019.
- CRAWFORD, K. *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. New Haven, London: Yale University Press, 2021.
- FRUMIN, I.; VOROCHKOV, A.; KIRYUSHINA, M.; PLATONOVA, D.; TEREENTIEV, E. Mapping the Generative AI Research in Higher Education: 2022-2024 Insights. *Higher Education Quarterly*, v. 80, n. 1, e70075, 2026. DOI: <https://doi.org/10.1111/hequ.70075>. Accessed: 10 Jun. 2026.

GU, X.; ERICSON, B. J. AI Literacy in K-12 and Higher Education in the Wake of Generative AI: An Integrative Review. *Proceedings of the 2025 ACM Conference on International Computing Education Research*, v.1, p. 125-140, 2025. DOI: <https://doi.org/10.1145/3702652.3744217>. Accessed: 10 Jun. 2026.

HAO, K. *A new vision of artificial intelligence for the people* [online]. MIT Technology Review, apr. 2022. Available at: <https://www.technologyreview.com/2022/04/22/1050394/artificial-intelligence-for-the-people/>. Accessed: 10 Jun. 2026.

HAO, K. *Empire of AI: Dreams and nightmares in Sam Altman's OpenAI*. New York: Penguin Press, 2025.

KAMLER, B.; THOMSON, P. *Helping doctoral students write: Pedagogies for supervision*. 2nd. ed. New York: Routledge, 2014. DOI: <https://doi.org/10.4324/9781315813639>. Accessed: 10 Jun. 2026.

KASSORLA, M.; GEORGIEVA, M.; PAPINI, A. *AI Literacy in Teaching and Learning: A Durable Framework for Higher Education* [online], oct. 2024. Available at: <https://www.educause.edu/content/2024/ai-literacy-in-teaching-and-learning/introduction>. Accessed: 10 Jun. 2026.

KOMLJENOVIC, J.; BIRCH, K.; SELLAR, S.; BERGVIKEN RENSFELDT, A.; DEVILLE, J.; EATON, C.; GOURLAY, L.; HANSEN, M.; KERSSSENS, N.; KOVALAINEN, A.; NAPPERT, P.-L.; NOTEBOOM, J.; PARCERISA, L.; PARDO-GUERRA, J. P.; POUTANEN, S.; ROBERTSON, S.; TYFIELD, D.; WILLIAMSON, B. Digitalised higher education: key developments, questions, and concerns. *Discourse: Studies in the Cultural Politics of Education*, v. 46, n. 2, p. 276-292, 2025. DOI: <https://doi.org/10.1080/01596306.2024.2408397>. Accessed: 10 Jun. 2026.

KOSMYNA, N.; HAUPTMANN, E.; YUAN, Y. T.; SITU, J.; LIAO, X.-H.; BERESNITZKY, A. V.; BRAUNSTEIN, I.; MAES, P. Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task. *arXiv*, v. 1, 2025. DOI: <https://doi.org/10.48550/arXiv.2506.08872>. Accessed: 10 Jun. 2026.

LACLAU, E. *On populist reason*. London, New York: Verso, 2005.

LONG, D.; MAGERKO, B. What is AI Literacy? Competencies and Design Considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, p. 1-16, 2020. DOI: <https://doi.org/10.1145/3313831.3376727>. Accessed: 10 Jun. 2026.

MABIRIZI, V.; KATUSHABE, C.; MUHOZA, G.; RUGASIRA, J. A systematic review of the impact of generative AI on postgraduate research: Opportunities, challenges, and ethical implications. *Discover Artificial Intelligence*, 5, article number 238, 2025. DOI: <https://doi.org/10.1007/s44163-025-00495-3>. Accessed: 10 Jun. 2026.

MAGOLDA, M. B. B. *Making Their Own Way: Narratives for Transforming Higher Education to Promote Self-Development*. New York: Routledge, 2023. DOI: <https://doi.org/10.4324/9781003445883>. Accessed: 10 Jun. 2026.

- MARKHAM, A. The limits of the imaginary: Challenges to intervening in future speculations of memory, data, and algorithms. *New Media & Society*, v. 23, n. 2, p. 382-405, 2021. DOI: <https://doi.org/10.1177/1461444820929322>. Accessed: 10 Jun. 2026.
- MIHRAM, D. *Research Guides. AI and Teaching and Learning*: examples of AI literacy frameworks [online]. n.d. Available at: <https://libguides.usc.edu/c.php?g=1457247&p=10905199>. Accessed: 10 Jun. 2026.
- NG, D. T. K.; WU, W.; LEUNG, J. K. L.; CHIU, T. K. F.; CHU, S. K. W. Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology*, v. 55, n. 3, p. 1082-1104, 2024. DOI: <https://doi.org/10.1111/bjet.13411>. Accessed: 10 Jun. 2026.
- NOBLE, S. U. *Algorithms of oppression*: how search engines reinforce racism. New York: New York University Press, 2018.
- RICAURTE, P. Data Epistemologies, The Coloniality of Power, and Resistance. *Television & New Media*, v. 20, n. 4, p. 350-365, 2019. DOI: <https://doi.org/10.1177/1527476419831640>. Accessed: 10 Jun. 2026.
- RUGHINIȘ, C.; VULPE, S.N.; ȚURCANU, D.; RUGHINIȘ, R. AI at the knowledge gates: Institutional policies and hybrid configurations in universities and publishers. *Frontiers in Computer Science*, v. 7, e1608276, 2025. DOI: <https://doi.org/10.3389/fcomp.2025.1608276>. Accessed: 10 Jun. 2026.
- SCHÖN, D. A. *The reflective practitioner*: How professionals think in action. New York: Basic Books, 2009.
- SELBER, S. A. *Multiliteracies for a digital age*. Carbondale: Southern Illinois University Press, 2010.
- SELWYN, N. On the Limits of Artificial Intelligence (AI) in Education. *Nordisk Tidsskrift for Pædagogikk Og Kritik*, v. 10, n. 1, 2024. DOI: <https://doi.org/10.23865/ntpk.v10.6062>. Accessed: 10 Jun. 2026.
- SHAPIRO, H.; SOUTO-OTERO, M.; WATERMEYER, R. Metacognitive AI literacy: Going beyond the AI skills gap agenda. *Learning, Media and Technology*, p. 1-15, 2026. DOI: <https://doi.org/10.1080/17439884.2026.2652638>. Accessed: 10 Jun. 2026.
- SRNICEK, N. *Silicon empires*: The fight for the future of AI. Cambridge: Polity Press, 2025.
- SWIST, T.; MALLAWA ARACHCHI, T.; CONDIE, J.; HANCKEL, B. Socio-technically just pedagogies: A framework for curriculum-making in higher education. *Learning, Media and Technology*, v. 49, n. 3, p. 492-505, 2023. DOI: <https://doi.org/10.1080/17439884.2023.2237884>. Accessed: 10 Jun. 2026.
- WARK, M. *A hacker manifesto*. Cambridge: Harvard University Press, 2004.
- WILLIAMSON, B., MACGILCHRIST, F., & POTTER, J. Re-examining AI, automation and datafication in education. *Learning, Media and Technology*, v. 48, n. 1, p. 1-5, 2023. DOI: <https://doi.org/10.1080/17439884.2023.2167830>. Accessed: 10 Jun. 2026.

ZEMBYLAS, M. A decolonial approach to AI in higher education teaching and learning: Strategies for undoing the ethics of digital neocolonialism. *Learning, Media and Technology*, v. 48, n. 1, p. 25-37, 2023. DOI: <https://doi.org/10.1080/17439884.2021.2010094>. Accessed: 10 Jun. 2026.

ZHAI, C.; WIBOWO, S.; LI, L. D. The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments*, v. 11, article number 28, 2024. DOI: <https://doi.org/10.1186/s40561-024-00316-7>. Accessed: 10 Jun. 2026.

ZYLINSKA, J. *AI art: Machine visions and warped dreams*. London: Open Humanities Press, 2020.

Language review: Ananda Missailidis

Submission date: 21/04/2026

Approval date: 28/05/2026