



Romanticising the algorithm: power dynamics and ideological implications in human–AI collaboration for higher education

Giulia Banfi¹ · Marco Luca Pedroni¹

Received: 17 December 2025 / Accepted: 30 April 2026
© The Author(s) 2026

Abstract

This article examines how undergraduate students appropriate generative AI within higher education, focussing on how their practices both reproduce and contest prevailing algorithmic and neoliberal logics. Drawing on 40 reflexive diaries produced by 186 students enrolled in a journalism course, the study employs thematic coding to analyse how AI tools were integrated into collective production processes such as brainstorming, drafting, and editing, and how students reflected on the benefits, limitations, and ethical implications of these uses. The analysis identifies four distinct profiles of AI use: selective optimisers (targeted, efficiency-oriented use), critical experimenters (reflective and exploratory use with supervision margins), exploratory users (creative and uncertain engagement), and tech functional (minimal and technical use). These profiles reveal how students negotiated AI's potential to enhance efficiency and creativity while simultaneously raising concerns about accuracy, reliability, and academic integrity. The findings suggest that students do not adopt AI uncritically but situate its use within broader negotiations over competence, authorship, and legitimacy. AI is perceived as a supportive but non-substitutive partner, requiring continuous human oversight and reflexive engagement. The results underscore the tension between instrumental benefits, such as time-saving, and deeper pedagogical questions concerning responsibility, agency, and the redistribution of academic labour. By highlighting the ambivalences and situated practices of AI appropriation, the article contributes to critical debates on the role of generative AI in higher education, pointing to the need for pedagogical frameworks that balance efficiency with reflexivity and safeguard human decision-making in contexts increasingly shaped by algorithmic technologies.

Keywords Generative artificial intelligence · AI education · Human–AI collaboration · Higher education

1 Introduction

“Congress failed to meet the moment on social media. Now we have the obligation to do it on AI before the threats and the risks become real.” With these words, delivered during a public hearing at the United States Senate in May 2023, Senator Richard Blumenthal issued a clear warning: artificial intelligence must not be addressed with the same delay that characterised the governance of social media. Present at that same hearing was Sam Altman, CEO of OpenAI, one of the most visible figures in contemporary debates on

automation, AI and computational power (TIME 2023). The senator's intervention captures two key dynamics. On the one hand, it reflects the climate of moral panic that often accompanies technologies perceived as disruptive, particularly when they impinge on sensitive domains such as communication, education or employment. On the other, it highlights a growing recognition that legislators cannot react belatedly to innovation: they must assume responsibility for shaping its introduction within regulatory, ethical and social frameworks, ensuring that such technologies contribute to positive social outcomes rather than exacerbating inequalities or systemic risks.

Unlike earlier technologies designed primarily to increase the efficiency of established processes (e.g. factory automation or the digitisation of bureaucratic procedures), generative AI disrupts the very structure of these processes by reconfiguring them through computational and algorithmic logics. In doing so, it unsettles educational systems, the

✉ Giulia Banfi
giulia.banfi@unife.it
Marco Luca Pedroni
marcoluca.pedroni@unife.it

¹ University of Ferrara, Ferrara, Italy

production of knowledge, and the very notion of authorship for those who have integrated these tools into their everyday practices (Watson et al. 2025). The public accessibility of generative AI, through tools such as ChatGPT, Claude, Gemini, Perplexity and DALL-E, has attracted political and institutional scrutiny. This urgency is fuelled by the memory of past regulatory delays in addressing the systemic effects of social media. The absence of timely intervention in that earlier domain left enduring consequences: privacy erosion, deterioration of mental well-being, disinformation, and weakened democratic resilience. Such inertia stemmed not only from technological lag but also from a broader failure to grasp the sociopolitical implications of platform governance. Today, this collective memory of regulatory failure underpins the accelerated pace of debates on artificial intelligence. This acceleration, however, is driven not only by ethical and regulatory concerns but also by the mounting influence of market interests and neoliberal logics of production. Recent initiatives such as the European Union's AI Act (2024) and Canada's Artificial Intelligence and Data Act (2022), under parliamentary scrutiny at the time of writing) exemplify this trend.

In recent years, across major sectors such as education, healthcare and public services, as well as within cultural industries—including publishing, journalism, film, television, music and advertising—critical reflections have emerged on the integration of generative AI. Within higher education in particular, the rapid uptake of AI tools has prompted an unprecedented proliferation of institutional policies and guidelines (e.g. Harvard University, Columbia University, Duke University, Peking University, and University of Bologna). Notably, such responses were introduced only after these tools had already gained substantial traction among students and staff. This is less an indication of institutional inertia than of the extraordinary speed at which generative AI has been adopted in educational contexts. The ability of large language models to generate coherent texts, synthesise sources, simulate personalised feedback and automate repetitive tasks raises fundamental questions about learning practices, assessment, and the very nature of education. These questions operate both at the epistemological level, where knowledge production and authorship are being redefined, and at the pedagogical level, where established modes of teaching and learning are unsettled.

Generative AI thus enters an already technocratic and performance-oriented educational environment, reshaping the role of teachers and learners and the value attributed to academic knowledge. While promising efficiency, personalisation and accessibility, it simultaneously risks fostering cognitive dependency, pedagogical de-responsibilisation and epistemic delegation to algorithms. It is therefore not a neutral “enabling” technology but a sociotechnical dispositif, in Foucault's (1977) sense of a heterogeneous assemblage of

discourses, institutions, practices and material arrangements that produces and regulates subjects and knowledge, one that embodies worldviews, value hierarchies and power relations. This regulatory acceleration, often presented as neutral and technical, should instead be understood as an attempt to reassert political control over automated processes that increasingly evade traditional legal and institutional categories. Such neutrality is apparent only from a naïve or market-aligned perspective. In this sense, the governance of AI is not solely a matter of policy, but also of imaginaries that generate divergent visions of educational futures, between promises of democratised knowledge and risks of algorithmic standardisation of human intelligence.

It is within this context that the present contribution focuses on the use of generative AI tools by undergraduate journalism students. This setting is not incidental: journalism is a field in which writing, creativity, interpretation and professional ethics are central, making it a particularly fertile site for exploring the implications of human–AI interaction. Moreover, the students involved had encountered and experimented with generative AI exclusively within higher education, without prior training or meaningful exposure during secondary schooling. The aim of this research is to examine how students' engagement with generative AI systems is socially and ideologically structured. Moving beyond instrumental or performance-based assessments of technology, the analysis interrogates the normative assumptions, representations of competence and power asymmetries that shape emerging practices of human–machine ‘collaboration’.

The guiding research question is therefore: in what ways do students' interactions with generative AI tools reflect, reproduce or challenge the ideological logics and power relations underpinning contemporary university pedagogy, specifically those associated with performativity, individual responsibility, output-oriented assessment and the technologisation of learning processes—by which we refer to the progressive incorporation of digital and algorithmic tools into the material and organisational fabric of teaching and learning, reshaping its practices, temporalities and relations? Three subsidiary questions orient the empirical analysis: (RQ1) How do students appropriate generative AI tools within collaborative editorial production, and what patterns of use emerge across different phases of work? (RQ2) How do students articulate the benefits and limitations of AI use, and what do these articulations reveal about the ideological tensions embedded in their learning practices? (RQ3) In what ways do students' collective reflections on human–AI relations signal forms of critical reflexivity capable of contesting dominant logics of automation and technologisation?

To address this question, this study draws on reflexive diaries written by 186 undergraduate students, organised into 40 project teams. Over a 12-week module focussed on digital media production, students were tasked with experimenting

with generative AI tools while developing editorial outputs such as online magazines, podcasts and social media strategies. The diaries, produced in parallel with these projects, captured situated experiences, perceptions and negotiation practices in engaging with AI within a pedagogical framework explicitly designed to support critical reflection. The texts were analysed using thematic analysis (Braun & Clarke 2006) with NVivo 14, to identify recurring patterns, dissonances and ideologically significant structures.

Ultimately, this article argues that students' engagement with generative AI is neither simply adaptive nor passively reproductive of dominant logics: it is a site of contested negotiation in which ideological arrangements are simultaneously enacted and questioned. By grounding this argument in empirical data and a critical sociological framework, the study contributes to debates on the future of higher education in an era of pervasive automation, and to the broader project of developing pedagogical practices capable of cultivating reflexivity, authorship and critical agency.

2 Literature review

2.1 Exploring human–AI collaboration in higher education

The advent of generative artificial intelligence in higher education has prompted numerous studies (Crompton & Burke 2023; Dell'Acqua et al. 2025; van Niekerk et al. 2025). Much of this research has concentrated on the pedagogical effectiveness of AI tools and their potential to enhance teaching and learning, while simultaneously raising ethical and sociological questions concerning knowledge production and dissemination. Yet, the predominantly technical and application-oriented orientation of these studies tends to obscure deeper epistemological, pedagogical and ideological consequences of AI in the relationship between education and social life (Watson & Romic 2025). Such perspectives often grant institutions a sense of control and neutrality, enabling them to avoid more disruptive questions about knowledge, authority and the very purpose of education.

Far from being instrumental aids (Balmer 2023), large language models such as ChatGPT are increasingly recognised as powerful mediators that reshape epistemic practices and redistribute authorship, attention and authority within educational contexts—what sociotechnical approaches would describe as a form of non-human agency in the co-production of knowledge (Latour 2005). Their use raises questions about the nature of human–machine interaction, particularly how student agency is reconfigured within hybrid forms of algorithmic co-agency. When engaged critically and experimentally, AI participates in what Hutchins (1995) would describe as a system of distributed cognition:

rather than extending individual cognitive capacities, it becomes one node within a broader configuration in which cognitive labour is shared across human agents, tools and collaborative workflows. Mollick (2024) has popularised the term *co-intelligence* to describe such configurations, framing AI as a collaborative partner in creative and intellectual work. We retain the term for its descriptive utility while departing from its techno-optimistic framing: rather than treating co-intelligence as a stable affordance of the technology, we approach it as a contested and ideologically situated practice, shaped by the institutional conditions, power asymmetries and normative assumptions within which human–AI interaction unfolds. From this perspective, generative AI is not simply a device for alleviating time-consuming academic tasks, but a sociotechnical force that reconfigures the very terms of knowledge production.

Within this field of inquiry, only a limited number of authors interrogate the metaphor of “intelligence” and related notions such as “hallucinations”, which risk obscuring the fundamentally probabilistic character of these systems, grounded in statistical pattern recognition rather than human-like cognition. Instead, scholarship frequently emphasises the operational conditions that make human–AI encounters efficient: the ability to formulate effective prompts, the personalisation of learning, and the stimulation of creativity (Nguyen et al. 2023; Baidoo-Anu & Owusu Ansah 2023; Mollick 2024; Xia et al. 2024). In this view, AI is often represented as an assistive figure—tutor, partner or creator—able to facilitate routine tasks and support brainstorming, analysis and content production. Such accounts are commonly framed in techno-optimistic terms, oscillating between enthusiasm for AI's emancipatory potential and concern about the risks of unreflective or improper use. The optimism rests on the assumption that, if appropriately managed, AI can enhance human capacities without fundamentally destabilising educational values.

This opportunity/risk dichotomy reduces analysis to a binary logic privileging technological efficiency and usability (Bissessar 2023; Dixon & Cox 2025), while neglecting the ideological and value systems sedimented in human–AI relations (Watson & Romic 2025). Even when AI is recognised as an active contributor to knowledge, particularly when conceived as a co-agent or tutor (Mollick 2024; Dell'Acqua et al. 2025), such recognition often remains confined to the operational level. What remains underexplored are the structural, epistemic, and cultural conditions that legitimise and shape such practices: the institutional norms, ideological arrangements and knowledge hierarchies that configure how interaction is imagined, enacted and evaluated.

Concerns identified in the literature have largely centred on technical fallibility, such as the production of inaccurate or biased information (Romeo & Conti 2025), and the risk

of excessive cognitive delegation (Formosa et al. 2025). Less examined, however, is the process through which such delegation becomes naturalised: repeated and unreflective use can silently restructure learning practices and reshape student subjectivities, while reinforcing norms of productivity, assessment and control embedded in neoliberal models of governance. The risk is not only that “AI gets things wrong”, but that algorithmic knowledge becomes internalised in ways that redefine reflexivity, authorship and intellectual autonomy. Educational practices may thus become increasingly aligned with a neoliberal rationality privileging efficiency over depth, performance over understanding, and automation over deliberation, positioning students less as critical agents than as self-optimising subjects within data-driven systems.

Generative AI also reshapes discursive practices. Automated writing and argumentation can standardise language, flatten expressive diversity and reduce the formative value of dialectical engagement (Mahaut & Franzon 2025). Rapid access to formally correct textual syntheses may encourage cognitive deskilling (Machidon 2025), whereby the act of thinking is partially externalised to the algorithm, diminishing intentionality, authorship and epistemic subjectivity (Formosa et al. 2025; Machidon 2025). This is compounded by the opacity of generative outputs, whose probabilistic logic makes responses difficult to reproduce. The resulting reliance has significant implications for trust and the cultivation of critical autonomy (Floridi 2023; Dixon & Cox 2025).

Within this perspective, generative AI cannot be understood only as a performance-enhancing technology. Rather, it assumes a dual role (Pronzato 2024): as a sociocultural artefact (Seaver 2017) and as a social agent (Airoldi 2021) that restructures educational experience. Human–AI engagement constitutes a relational space in which boundaries between authority, authorship and cognitive agency are renegotiated, unsettling established norms of intellectual ownership and epistemic responsibility (Formosa et al. 2025; Airoldi 2021). Students do not simply “use” AI as a neutral tool; they operate within sociotechnical settings that shape both the content of learning and the formats of output production. In this sense, cognitive autonomy risks being progressively reconfigured, shifting from the centrality of the individual towards a position mediated by algorithmic logic. Yet, a space for critical reflexivity remains: the capacity of students to exercise intentional and reflective control over their learning practices, not only accepting machine output passively but also positioning themselves in dialectical relation to it.

Knowledge, therefore, should not be valued solely in terms of informational efficiency, but also in the possibility for learners to exercise agency and interrogate meaning, assumptions and consequences. Artificial intelligence is thus not simply a cognitive facilitator, but a sociotechnical force

that actively shapes the symbolic organisation of learning, placing the autonomous production of knowledge in tension with processes of cognitive delegation.

2.2 Sociotechnical perspectives on AI in education

In recent years, governments worldwide have launched strategic initiatives to embed artificial intelligence within educational systems, not only as a tool to support teaching but also as a lever for systemic innovation. Such interventions signal a paradigm shift in which generative AI is framed as capable of restructuring practices of teaching, learning and national development. A prominent example is Estonia’s *AI Leap 2025* programme, designed to integrate AI solutions into schools in order to personalise learning, optimise processes and develop advanced digital skills among students and teachers. Italy has followed a similar path, with regional governments allocating funds to support AI-related projects in schools and universities.

In this context, reflexivity emerges not only as an individual competence but also as a situated and contested capacity—crucial for negotiating decision-making autonomy within environments increasingly shaped by automation and algorithmic logics. When educational policies move beyond a narrow focus on efficiency and technical optimisation, they can function as cultural instruments that open symbolic spaces of negotiation around the meaning and use of algorithms in education.

The introduction of generative tools such as ChatGPT, however, does not result in a straightforward reduction of educational labour. Instead, it redistributes academic stress asymmetrically (Kikerpill & Siibak 2023). While students may experience relief through accelerated text production, teachers and academic staff face additional burdens: monitoring plagiarism, exercising heightened vigilance, and rethinking their roles in academic practices. This asymmetry underscores the need to avoid “showcase policies” and instead accompany technological adoption with sustained public investment and systematic evaluation, both *ex ante* and *ex post*. Such evaluations must be grounded not only in criteria of efficiency but also in equity and sustainability, to ensure that AI in education strengthens rather than undermines autonomy and quality in teaching and learning.

This requires grounding in robust theoretical and empirical research on human–AI engagement (Crompton & Burke 2023; Mollick 2024; van Niekerk et al. 2025). Only on this basis can educational initiatives combine innovation with critical awareness, promoting informed and reflective interaction with technologies. From this perspective, the introduction of AI should not be confined to the transfer of technical competences. It must be framed as an educational process aimed at interrogating the conditions of knowledge production and the power relations that sustain them.

Universities are already reshaping curricula in response to these transformations. Both in Italy and internationally, institutions are experimenting with interdisciplinary programmes that approach AI through philosophical, legal, pedagogical and sociological lenses, thereby moving beyond a narrow technical-engineering paradigm. These initiatives respond to the need to train students and future professionals capable not only of interacting with intelligent systems but also of assuming critical and responsible roles in processes of digital knowledge production. This task is especially urgent in an era marked by the proliferation of disinformation, widening digital inequalities and the opacity of algorithmic infrastructures—challenges that demand new forms of digital and cognitive literacy.

As shown in the growing body of literature (Balmer 2023; Bissessar 2023; Crompton & Burke 2023; Sawalha et al. 2024), generative AI is progressively reshaping the boundaries of knowledge, authorship and educational practice. The questions raised extend far beyond instrumental effectiveness to include epistemological, ethical and social dimensions of learning. This calls for sociological approaches capable not only of recognising the transformative impact of AI in education but also of interrogating the broader symbolic, ideological and institutional arrangements within which such transformations unfold.

Within this analytical horizon, the present study seeks to advance a shift in perspective. Rather than approaching human–AI relations solely in terms of cognitive enhancement or technological affordance, it develops an interpretive paradigm grounded in critical sociology, specifically in the interconnected traditions of critical pedagogy (Freire 1970; Giroux 2011), the sociology of algorithms (Airoidi 2021; Seaver 2017), and Foucauldian analyses of subjectivity and power. Whilst sociotechnical approaches such as actor-network theory offer useful descriptive vocabulary for tracing human–machine relations, our primary analytical interest lies not in mapping assemblages but in interrogating the ideological conditions, power asymmetries and normative arrangements within which such relations are enacted and contested. The analysis therefore draws on ANT-inflected concepts selectively and instrumentally, subordinating them to a critical sociological framework.

3 Methods

This study translates its critical sociological perspective into an empirical design aimed at understanding how students engage with generative AI not simply as users of technological tools, but as meaning-makers embedded in sociotechnical and ideological structures. The methodological framework is rooted in a cultural-interpretive approach, treating

students' written reflections as both situated accounts and discursive artefacts.

Data were collected within the undergraduate course *Digital Media and Journalism* at the University of Ferrara (academic year 2024/2025), during a pedagogical intervention entitled *AI-assisted editorial project*. Conceived as a meta-pedagogical space, the intervention invited students to explore AI both as a means of production and as a subject of critique. Its aim was to examine how generative AI reshapes journalistic practices, influences epistemic agency, and reframes authorship and editorial decision-making.

The intervention was organised in two phases. The first phase comprised five preparatory workshops (10 contact hours) introducing students to journalistic applications of AI: from fact-checking and automated writing to podcast production and ethical analysis. These sessions combined technical exercises with reflexive discussions supported by structured prompts. Tools explored included ChatGPT, Gemini, Notebook LM, Perplexity and various AI-based data-extraction platforms. Students were encouraged to reflect on tensions between efficiency and accuracy, creativity and standardisation, authorship and automation.

The second phase involved the development of collaborative editorial projects by 40 student teams (186 students in total). Each team was tasked with producing a digital magazine or equivalent output (e.g. podcast series, thematic blog) on a socially relevant topic such as the climate crisis, digital rights or disinformation. Teams were organised to simulate newsroom dynamics, with defined editorial roles and internal workflows. AI tools were integrated throughout the production process, including brainstorming, drafting, rewriting, visual design and layout generation. Crucially, these uses were not only functional but also framed analytically within the course as moments of negotiation around competence, authorship and legitimacy.

To capture the sociocultural dimensions of these practices, each team designated one member to keep a reflexive AI diary across the 12-week project. The diaries followed a shared template prompting students to record how, when and why AI tools were used; the perceived strengths and limitations of different applications; team discussions on their ethical and epistemic appropriateness; and emotional responses (e.g. relief, discomfort, pride, dependence). Diaries were collected as individual documents and analysed in full.

In parallel, both researchers maintained a reflexive diary throughout the 10-week course, compiled systematically after each session. Entries varied in depth according to the intensity of each project phase and were organised around three distinct registers: observational notes recording classroom interactions, group dynamics and informal conversations; affective notes capturing the researchers' own emotional responses and relational tensions; and interpretive

notes developing analytical reflections on emerging patterns. This approach draws on established practice in qualitative inquiry, whereby the researcher diary functions not simply as a record of events but as an instrument of reflexive knowledge production (Schatzman & Strauss 1973; Ortlipp 2008), adding a complementary layer of situated, process-oriented data to the students' written accounts.

The full corpus, comprising 40 team diaries and the ethnographic journal, was analysed through thematic analysis, following the reflexive approach outlined by Braun and Clarke (2006). Rather than applying pre-established codes, we adopted an inductive, recursive strategy, identifying semantic and latent patterns through iterative readings. Analysis proceeded in six phases: data familiarisation; initial code generation; theme identification; theme review; theme naming and definition; and writing. Coding was conducted independently and in parallel by both authors (one the course instructor, the other a teaching assistant who had been involved throughout the project) and was managed and refined collaboratively using NVivo 14. Divergences were addressed through structured discussion sessions held at each analytical phase, with disagreements resolved through iterative collective review. No formal inter-coder reliability statistic (e.g. Cohen's kappa) was calculated. This decision reflects the epistemological commitments of reflexive thematic analysis: as Braun and Clarke (2019) have argued, such metrics presuppose a positivist logic of coding as objective categorisation, which is at odds with an approach that treats the researcher's interpretive engagement as a constitutive feature of analysis rather than a source of bias to be controlled. Rigour was instead pursued through transparency of procedure, iterative reflexivity, and the sustained dialogue between two researchers with complementary situated knowledge of the data.

The focus was not limited to *what* students reported about AI but extended to *how* they articulated these experiences, tracing discursive framings, affective investments and rhetorical strategies. Particular attention was paid to tensions, ambivalences and contradictions: for example, students describing AI simultaneously as “a shortcut” and “a partner” or expressing pride in their autonomy while deferring key decisions to algorithmic suggestions.

By combining students' accounts with instructor-based observations, the research design foregrounds the relational, affective and ideological dimensions of AI use in educational settings. A note on positionality is warranted here. One of the two authors served as the course instructor, the other as teaching assistant throughout the project; both were therefore simultaneously pedagogical actors and researchers. We acknowledge this dual positioning as a potential source of interpretive proximity, which we sought to manage through the parallel coding procedure and structured inter-researcher dialogue described above. Students were

informed of the research purposes of the diaries at the project briefing and provided formal consent at the point of submission of their work. This approach enables a situated analysis of how algorithmic systems are domesticated, contested or routinised within the pedagogical logics of higher education. It also aligns with the overarching aim of the study: to interrogate how students' interactions with generative AI reflect, reproduce or challenge the power relations and normative assumptions embedded in contemporary university pedagogy.

This research was carried out within the framework of the project *The Challenges of Generative Artificial Intelligence: An Interdisciplinary Approach at the Intersection of Ethics, Culture and Technological Innovation*, funded by the University of [omissis]. The project sought to transform the classroom into a space of collective inquiry, where teaching and research intersect to make visible the symbolic and political work performed by AI in reconfiguring knowledge, authorship and critical agency.

4 Findings

The diaries included (a) a description of the activities undertaken; (b) the AI tools employed and the objectives pursued; (c) critical evaluations of the outputs generated, including perceived advantages and limitations; and (d) accounts of revisions and adaptations introduced to achieve results deemed acceptable by the group. In their concluding sections, students also provided collective reflections that emerged during the project, alongside free-form observations.

The relative uniformity of the template enabled systematic comparative analysis, allowing both differences of approach and recurring patterns in AI use to be identified. Thematic analysis (Braun & Clarke 2006) ensured a consistent organisation and coding of the data. To facilitate comparison across cases, the diaries were further arranged through a process of case classification. Each group was assigned a set of analytical attributes and values (Table 1), designed to capture both the empirical diversity of observed practices and the conceptual dimensions discussed in the literature. This schema allowed us to move beyond purely narrative descriptions of students' experiences and to situate them within categories reflecting variation in modes of appropriation, intensity of use, evaluative orientations, phases of application, and forms of agency attributed to AI.

While most attributes were derived inductively from the data, the classification of AI agency type drew on the taxonomy proposed by Ziglioli et al. (2025). This framework conceptualises agency along two axes: task complexity and degree of relational engagement. It enables us to trace how students positioned AI within their workflows, as a provider

Table 1 Attributes and values of the thematic analysis

Attribute	Values (short description)
Output type	Magazine; magazine + podcast—type of final product realised by the groups
AI usage profile	Selective optimisers (targeted use for efficiency and quality); critical experimenters (experimental use combined with critical reflection); exploratory users (creative and intensive exploration); tech functional (minimal and pragmatic use)
AI intensity	Low; medium; high—level of intensity in the use of AI tools
Stance toward AI	Enthusiastic (unreserved approval); critical (emphasis on limitations); negotiating (pragmatic and situated reflection)
Predominant AI phase	Research; writing; review; visual; audio—main phase of work in which AI was employed
AI agency type	Teaching agent (low task, low relation—AI provides basic or factual information without collaboration); subordinate assistant (high task, low relation—AI executes specific commands under human direction, with little generative autonomy); cognitive partner (low task, high relation—AI contributes suggestions and co-creates knowledge in a collaborative way); substitutive partner (high task, high relation—AI acts as a human-like agent, taking on advanced roles such as reviewer, editor, or scholar)

of factual information, a subordinate assistant, a cognitive partner, or a substitutive agent, and to examine how such positions reflected broader pedagogical logics of performativity, responsibility, assessment and technologisation.

The case classification produced a set of attributes capturing how groups engaged with AI (Table 1). These attributes included the type of editorial output produced, the intensity of AI use, the stance adopted towards AI, the predominant phase of work in which AI was employed, and the form of agency attributed to AI. Taken together, they situate students' practices within broader ideological and pedagogical structures.

From this classification, *four main usage profiles* were identified.

Selective optimisers ($n = 21$). These groups employed AI in a targeted manner, typically at medium intensity. They used it to accelerate specific steps and refine the formal quality of outputs, but without ceding central control. Their approach reflects the pedagogical logic of performativity, in which efficiency and measurable outcomes are prioritised, but also the logic of individual responsibility, since students emphasised their own role in validating and correcting machine-generated text.

Critical experimenters ($n = 11$). These groups engaged more intensively with AI, often at a high level of use, but combined experimentation with reflexive supervision. They tested the limits of AI without delegating the entire production process. This profile illustrates how students negotiated the tension between output-oriented assessment aimed at producing publishable results and a critical stance towards the technologisation of learning.

Exploratory users ($n = 7$). These groups approached AI as a creative and expressive resource, using it to generate ideas and explore stylistic alternatives, even at the cost of uncertainty. Their practices aligned less with performative efficiency and more with resistance to standardisation, yet still entangled with the logic of technologisation, given

their reliance on algorithmic suggestions for inspiration and experimentation.

Tech functional ($n = 1$). This residual profile reduced AI use to minimal, technical functions, such as drafting headlines or correcting text. The group treated AI as a peripheral support rather than a cognitive partner. This posture reflects a logic of individual responsibility and scepticism towards automation but also indicates limited engagement with the broader pedagogical consequences of AI.

Across these profiles, the prevailing stance towards AI was *negotiating* (29 groups), suggesting that students neither embraced AI uncritically nor rejected it outright. Instead, they engaged in pragmatic reflection, continuously adjusting their practices. In terms of agency attribution, most groups positioned AI as a *subordinate assistant* ($n = 22$) or *cognitive partner* ($n = 11$), reflecting a hybrid space in which students retained ultimate authority but recognised AI as more than a neutral tool. Only a minority attributed to AI the roles of *teaching agent* ($n = 2$) or *substitutive partner* ($n = 5$).

This distribution underscores the ambivalent domestication of AI: while students integrated it into their workflow, they simultaneously reasserted their own epistemic authority, often through practices of checking, revising and contextualising. In doing so, they reflected and, in some cases, contested the four logics underpinning contemporary pedagogy: embracing efficiency but questioning reliability; assuming responsibility while resisting over-delegation; working towards outputs while safeguarding creativity; and engaging with technologisation without surrendering autonomy.

In addition to the case classification, which positioned the diaries within predefined profiles and attributes, we conducted thematic analysis to capture more deeply the discursive and experiential dimensions of students' engagement with AI. The codes developed were not purely descriptive but designed to highlight the symbolic meanings students attributed to AI and the tensions within processes of technological appropriation. These codes clustered around four

key nodes: (a) *AI vision*—representations of artificial intelligence, its imaginaries, and its placement within wider social and cultural contexts; (b) *AI–human interaction*—reflections on the relationship between human and technological agency, ranging from AI as ally, subordinate tool, or potential substitute; (c) *challenges and limitations*—narratives concerning difficulties encountered, distinguishing between issues of accuracy and trustworthiness and more technical problems; (d) *perceived benefits*—accounts of positive aspects of AI use, with particular emphasis on creative support and gains in efficiency.

Introducing these codes allowed us to move beyond a typological reading of cases. While the classification mapped recurring usage profiles, the thematic analysis illuminated the interpretive layers of students' practices, showing how they negotiated benefits and risks, and how creative possibilities were weighed against structural constraints. In this way, classification and thematic analysis complemented one another: the former identified patterns of appropriation, while the latter revealed the meanings and ambivalences embedded in students' discourses.

From this dual analytical strategy, three recurrent domains emerged that structure students' practices and discourses around AI: (1) concrete modes of appropriation and use of AI tools; (2) representations of perceived benefits and limitations; and (3) reflective and collective dynamics within group work. These domains should not be understood as separate, but as interwoven dimensions of the same process. Choices around appropriation and experimentation influenced the evaluation of advantages and drawbacks, which in turn shaped collective negotiations and reflections on the role of AI in academic and media production. Taken together, these domains shed light on how students enacted, resisted or reconfigured the pedagogical logics at stake: striving for efficiency while questioning reliability (performativity), assuming responsibility for verification and correction (individual responsibility), working towards outputs under assessment pressures (output-oriented assessment), and grappling with the pervasive presence of automation in their creative and cognitive labour (technologisation).

4.1 Modalities of AI use

The analysis reveals a process of appropriation that was both widespread and selective. In terms of outputs, almost all groups chose to produce a digital magazine, with only a small minority supplementing this with a podcast. This preference suggests that innovation was channelled into a format perceived as reliable and relatively straightforward, rather than into more technically demanding territories. Within this production process, distinct trajectories of AI use emerged. The largest number of groups followed

a pragmatic pattern, integrating AI in targeted ways and at medium intensity, using it to accelerate specific stages and refine formal quality while avoiding full reliance on machine output. Alongside these, a smaller set engaged with AI more intensively and reflexively, experimenting with its possibilities yet retaining significant margins of human supervision. A further, even smaller, group treated AI as a creative lever, embracing it as a source of expressive possibilities and stylistic variation, even at the cost of uncertainty in the final outcome. Finally, one residual case exemplified minimal and purely functional use, confining AI to basic technical tasks.

These usage patterns were accompanied by a predominantly negotiating stance. The majority of groups neither embraced AI with unreserved enthusiasm nor rejected it outright but instead treated it as a resource to be adjusted and recalibrated according to context. Across the diaries, students were unwilling to cede complete agency to AI. Rather, they incorporated it as an artefact that demanded constant decision-making, correction and coordination. This was confirmed by the distribution of AI use across different phases of work: the technology was employed most frequently in writing and research, which were central to production; its presence was more experimental in the visual and audio dimensions; and it was least used in revision, where quality control was consistently regarded as a human responsibility. In these accounts, the act of reviewing and validating information emerges as a key site where students resisted the full technologisation of their work and reasserted their own epistemic authority.

The types of agency attributed to AI followed a similar pattern. Most groups cast the technology in subordinate or collaborative roles, describing it either as an assistant carrying out specific commands under human direction, or as a cognitive partner, contributing suggestions and co-creating content without replacing human oversight. Only a small minority reduced AI to a teaching agent providing factual input or conversely imagined it as a substitutive partner capable of assuming advanced editorial functions. Overall, appropriation was marked by compromise: AI was embedded as infrastructural support or as a partner in moments of design and creativity, yet it remained subordinate in phases of validation and reliability. What emerges is a form of carefully supervised innovation, where technology was genuinely integrated into students' work but only through selective and negotiated practices. This constant calibration reveals how students simultaneously pursued efficiency, protected authorship and maintained responsibility for final outputs, enacting a guarded engagement that both reflected and contested the pedagogical logics of performativity, assessment and technologisation.

4.2 Perceived benefits and limitations

The second domain concerns the ways in which students articulated the benefits and drawbacks of AI use. Across the diaries, two advantages stood out: efficiency and creative stimulation. AI was consistently described as a tool that accelerated production, reduced cognitive load, and provided immediate drafts, summaries or titles that enabled groups to progress more quickly from planning to execution. It was equally valued as a creative support, capable of generating ideas and stylistic alternatives that helped overcome initial blocks and expand expressive possibilities. One group wrote.

The advantages of AI that we experienced in producing the magazine were several: speed and productivity, it provided us with drafts, catchy titles and explanatory summaries. Another benefit was its support in researching information and contextualising topics. Most importantly, it saved us a great deal of time in writing articles, which would have taken much longer manually (Group 13, Selective Optimisers).

Another diary echoed this sentiment:

AI turned out to be an extremely valuable support. It helped us overcome the initial block of the first work phases, speeding up the creative process and providing us with a starting point on which we could build through research, interviews, insights, edits and personal contributions (Group 25, Critical Experimenters).

Such reflections reveal how students partially internalised the logic of performativity, valuing efficiency and output maximisation, while also recognising AI as a source of inspiration that complemented rather than replaced their own creativity. At the same time, these accounts illustrate a determination to preserve authorship and originality, resisting the complete technologisation of learning.

Alongside perceived advantages, the diaries contained sustained reflections on limitations. The most frequent concerns related to accuracy and trustworthiness: groups often reported that texts were repetitive, superficial, or factually unreliable, requiring constant revision. As one group observed,

The text was often banal and repetitive, even when we reformulated the requests differently. It also gave false information. The limitation lies in knowing how to ask the right questions, giving it all the information you need, otherwise the answer is always the same, just rephrased (Group 4, Selective Optimisers).

Others pointed to a lack of depth and originality, stressing the irreplaceable quality of human authorship. Such concerns directly implicate the critiques of neoliberal models of higher education that privilege measurable outputs over intellectual depth, eroding the very qualities—originality, interpretative nuance, authentic voice—that students here defend:

We noticed limitations, especially when the topic required deeper reflection, critical analysis or integration of specific data. While AI was able to draft informative or technical texts with precision, it often lacked interpretative depth and the sensitivity that a human author can provide. Even when correct, the texts seemed mechanical, lacking true originality or an authentic voice (Group 29, Critical Experimenters).

What initially appeared as a gain in efficiency often translated into additional labour, as students were compelled to revise, verify and contextualise outputs. During writing, AI was framed primarily as an accelerator, while in research and revision phases concerns about credibility prevailed, leading to an insistence on human oversight. In more experimental areas, such as visual or audio production, AI was described less as a producer of final outputs and more as a generator of suggestions and inspiration.

Overall, the diaries portray an ambivalent relationship: AI was both an accelerator of efficiency and creativity, and a source of uncertainty that demanded vigilance. Students managed this ambivalence by maintaining a dialogical relation with the technology, acknowledging its affordances while insisting on the necessity of human intervention. Their refusal to relinquish responsibility demonstrates not only a pragmatic adjustment to technological limitations, but also a subtle critique of pedagogical logics that privilege efficiency and automation as unquestioned values.

4.3 Collective reflections on the future of AI and human–AI collaboration

The reflections recorded in the diaries did not remain confined to the operational dimension of AI use but extended to collective evaluations of the technology's future and the possible forms of engagement between humans and artificial intelligence. Students demonstrated the ability to situate their laboratory experiences within broader horizons, connecting the production of magazines with everyday uses and with projections concerning future educational and professional trajectories. The visions articulated were ambivalent: on the one hand, AI was imagined as a resource capable of transforming creative and learning processes; on the other, students expressed anxieties about technological dependency and the erosion of critical control. These were not monolithic representations, but discourses traversed by tension between enthusiasm and caution, between imaginaries of progress and scenarios of vulnerability.

The collective visions mirrored the distribution of stance towards AI, where a negotiating position clearly predominated, followed by more critical attitudes and only rarely by overtly enthusiastic ones. This indicates that students did not frame AI as either a salvific tool or an existential threat, but rather as a technology that must remain under continuous

evaluation, its value contingent on the context of use and on the extent of human oversight.

Nearly all groups, with a single exception, reflected on the degree of human intervention required and the forms of partnership possible with AI. The coding of human–AI relations confirms that the debate extended beyond technical aspects, taking shape as a collective negotiation of the role of technology in knowledge and production processes. Reflections were most evident in groups that attributed to AI the role of Subordinate Assistant or Cognitive Partner, while more radical configurations—substitutive partner or teaching agent—were minority and marginal.

One diary warned explicitly against full reliance on the machine:

Relying on AI for everything, without direct human control, is very risky, both in terms of form and content. Human intervention was necessary to correct and modify sentence structures and to verify that the sources used by AI applications to create content were accurate. We chose to combine AI with our human intervention to see which output would be most credible, coherent and complete (Group 1, Exploratory Users).

Another highlighted the ethical ambivalence of delegating creative tasks:

Some believed that letting AI help us meant giving up creativity and in a sense cheating. Others argued instead that AI should not be seen as a substitute, but as a tool to refine our ideas, broaden perspectives, and stimulate new reflections. The point of the debate, therefore, was not whether to use artificial intelligence, but how to use it (Group 7, Selective Optimisers).

A further group reflected on the broader professional implications:

Given its usefulness, we often discussed how likely it is that AI will in the future take the place of humans in the workplace and beyond. In this way we realised that the first human faculties to disappear could be creativity and imagination. The question that arises is: will we be able to use them appropriately and only in support of our activities? Moreover, we were able to see how effective it is in helping us carry out university tasks (Group 24, Selective Optimisers).

What emerges is an experience of reflexive hybridisation, in which students did not merely describe what AI produced but debated how it should be integrated into future processes. They recognised the utility of AI as both an accelerator of efficiency and a stimulus for creativity yet insisted on the necessity of vigilant use, one that preserves human decision-making at the centre and safeguards the critical dimension of learning. In so doing, students not only reflected on their immediate experiences while also engaging with the ideological logics underpinning pedagogy, simultaneously reproducing the values of efficiency and productivity while

challenging the idea that automation can or should displace reflexive human judgement.

The typology presented in Table 2 should be understood not as a set of rigid categories but as a heuristic mapping of dominant orientations that surfaced in the diaries. The profiles identify prevailing rather than exclusive configurations, since several groups displayed hybrid or shifting practices, moving across different stances and forms of agency depending on the task and phase of work. The table thus highlights recurrent tendencies while acknowledging the fluid and context-dependent nature of human–AI collaboration. While the selective optimisers represent the most common pattern—anchored in pragmatic efficiency and medium-intensity use—critical experimenters reveal a more reflexive engagement, where exploration was accompanied by deliberate human oversight. Exploratory users, though less numerous, illustrate how some students positioned AI as a creative interlocutor, willing to tolerate uncertainty to pursue novel expressive avenues; their oscillation between cognitive partnership and partial substitution underscores the tensions inherent in imagining AI as more than a tool. The tech functional profile, although residual, is instructive in highlighting cautious and minimal engagements, where AI was relegated to peripheral tasks and framed primarily through a critical lens.

What is particularly significant is the way these profiles map onto the broader pedagogical logics under examination. Selective optimisers embody the drive towards performativity and output-oriented assessment, but their insistence on validation shows that individual responsibility remained intact. Critical experimenters highlight the possibility of resisting unreflective technologisation by maintaining a supervisory stance even under conditions of high intensity. Exploratory users complicate the picture by embracing creativity through AI while simultaneously questioning its reliability, thereby illustrating the contradictory pull of innovation and caution. Tech functional cases, finally, make visible the limits of adoption, foregrounding scepticism and minimal reliance as a form of resistance to the assumption that automation should naturally expand into every domain of academic work.

Taken together, these profiles demonstrate that students did not simply adopt AI as a neutral instrument. Rather, they continuously negotiated its role, embedding it in ways that reflected and contested the performative, individualising and technologising imperatives of contemporary pedagogy.

5 Discussion

The findings suggest that students' engagement with generative AI cannot be reduced to the simplistic dichotomy of opportunity versus risk, a polarisation co-opted as a

Table 2 Profiles of AI appropriation and use among student groups

Profile	Usage characteristics	Recognised benefits	Perceived limitations	Prevailing stance towards AI	Prevailing agency type
Selective optimizers	Targeted, pragmatic use focussed on efficiency and quality; medium intensity	Operational efficiency, process acceleration, reduced cognitive load	Issues of accuracy and reliability of outputs	Predominantly negotiating	Subordinate assistant
Critical experimenters	Intensive and reflective use, with controlled experimentation; maintain supervision margins	Creative stimulation, possibility to test the limits of the tool	Concerns about trustworthiness of results, need for constant validation	Critical/negotiating	Cognitive partner
Exploratory users	Creative and exploratory approach; accept margins of uncertainty	Idea generation, stylistic variety, expansion of expressive possibilities	Uncertainties about accuracy and consistency of outputs	Negotiating	Cognitive partner/substitutive partner
Tech functional	Minimal and technical use, limited to basic tasks	Immediate operativity	Lack of added value, low accuracy	Critical/negotiating (minimal engagement)	Subordinate assistant

legitimising device by commercial actors seeking to expand AI markets. Industry players often present themselves simultaneously as champions of innovation and custodians of ethical usage, thus deflecting critique under a veneer of responsibility. At the 2024 World Economic Forum in Davos, for instance, UN Secretary-General António Guterres criticised tech companies for prioritising profit over potential social harms, even as those firms publicly highlighted their commitment to “responsible AI” to protect brand reputation (Elliott 2024). Recent research on corporate reporting practices reinforces this view: a systematic analysis of more than 30,000 SEC filings shows that references to AI risks have surged in recent years, yet disclosures often remain vague and superficial, echoing regulatory concerns about the quality of companies’ risk reporting (Uberti-Bona Marin et al. 2025). In this sense, the opportunity/risk framing functions as rhetorical cover, reassuring policymakers and investors while leaving structural power dynamics and economic imperatives unexamined. By contrast, the diaries show that students did not frame AI in binary terms, as either a threat to academic integrity or a vehicle of creative emancipation, but approached it through ambivalence, negotiation and situated reflexivity. This observation confirms Watson and Romic’s (2025) argument that the most significant consequences of AI are not technological but epistemological and ideological, reshaping the conditions under which knowledge is produced and valued.

At one level, the students’ practices reflect the broader logic of performativity that characterises contemporary higher education. Efficiency, productivity and output maximisation were repeatedly cited as benefits of AI, and the profiles of selective optimisers demonstrate the internalisation of performance-oriented rationalities. Yet, this orientation was consistently tempered by practices of verification and correction, which reasserted the human role in guaranteeing accuracy and depth. The neoliberal imperative of speed and measurability was therefore neither fully embraced nor outright rejected but reconfigured in practice. What appeared to be a straightforward adoption of AI for efficiency often became a more complex choreography, in which students moved rapidly between trust and scepticism, automation and oversight. This oscillation suggests that rather than passively absorbing dominant logics, students actively negotiated them, appropriating AI for pragmatic gains while simultaneously exposing its fragility and re-establishing their own authority as critical agents.

At another level, the diaries illustrate how the technologisation of learning reconfigures student subjectivities. Experimenting with AI as either a cognitive partner or a subordinate assistant was not only a matter of mastering new technical affordances; it reshaped how students conceived of authorship, legitimacy and epistemic agency. As noted in the diaries, students repeatedly highlighted

that automation does not eliminate labour but redistributes responsibility. In this redistribution, the neoliberal logic of self-responsibility was reframed: students did not imagine themselves as solitary producers of knowledge, but as overseers and curators of algorithmic output, responsible for monitoring, correcting and legitimising machine suggestions. Responsibility, in other words, shifted from the act of creation to the task of surveillance, an ideological transformation that alters how the “educated subject” is imagined and enacted.

The ambivalent stance towards AI further illuminates the contradictions of output-oriented assessment. While many groups valued AI as a shortcut to producing polished outputs that matched institutional expectations, they also lamented the loss of depth, originality and what they described as an “authentic voice.” These reflections expose discomfort with the reduction of education to measurable deliverables. The diaries can thus be read as tacit critiques of assessment regimes that privilege standardisation and efficiency over process, dialogue and intellectual autonomy, echoing broader critiques of the neoliberal university as a space increasingly structured by audit and accountability logics (Bucher 2018; Williamson 2017). The tension students expressed around the loss of an “authentic voice” directly parallels the way critical pedagogy has problematised the silencing of individuality and creativity under conditions of standardisation (Freire 1970; Giroux 2011).

Equally revealing are the imaginative projections students articulated concerning future human–AI relations. Their visions oscillated between optimism and anxiety, between seeing AI as a creative accelerator and fearing the erosion of imagination and critical thought. Such ambivalence finds a precise theoretical anchor in Seaver’s (2017) analysis of the cultural work of algorithms: AI here functions not only as a technical artefact but also as an imaginary shaping how futures are envisioned. Students’ reflections can be understood as micro-politics of negotiation, in which imaginaries of automation are appropriated, contested and re-signified in everyday practice.

Taken together, these findings advance three critical insights. First, AI in higher education should be understood less as a neutral instrument than as a *dispositif* that embodies and materialises ideological logics—efficiency, self-responsibility, output orientation, technologization—while also opening spaces for their contestation. Second, student agency under conditions of AI use is best conceptualised as distributed and relational: authority is neither fully surrendered to the machine nor entirely retained by the human, but continuously renegotiated through verification, correction and reflexive evaluation. Third, the integration of AI calls for a reconceptualisation of critical pedagogy. If responsibility shifts from production to oversight, the task of education is not simply to “teach with AI” but to cultivate capacities for

interrogating the ideological arrangements that normalise its adoption.

What is ultimately at stake is not whether AI should be integrated into education—this integration is already underway—but how to prevent its use from consolidating neoliberal pedagogical regimes that privilege efficiency over meaning and outputs over intellectual formation. The challenge for critical pedagogy is to render visible the ideological labour performed by AI, to open space for resisting automatised subjectivities, and to reassert education as a domain of reflexive, dialogical and emancipatory practice.

In this light, the arrival of generative AI in higher education should not be celebrated as a neutral technological upgrade, nor merely feared as an external disruption. What the diaries reveal is more unsettling: AI has already insinuated itself into the micro-politics of learning, silently reorganising what it means to be a student, a writer, a critical subject. The rhetoric of “opportunities and risks” that dominates institutional and corporate discourse operates as a convenient distraction, directing attention towards surface-level trade-offs while leaving untouched the deeper transformations of subjectivity and power. Students’ accounts demonstrate that the technology does not simply augment skills or threaten integrity; it redefines responsibility, authorship and the very grammar of knowledge.

If universities continue to integrate AI uncritically—through guidelines, “responsible use” policies and market-driven incentives—they risk becoming laboratories not of emancipation but of managed compliance, where reflexivity is narrowed to the task of supervising machines and creativity is reduced to the correction of algorithmic drafts. Yet the diaries also offer a counterpoint: in their insistence on verification, their discomfort with standardisation, and their debates on the future of imagination, students displayed forms of reflexivity that resist full automation of thought. These gestures, fragile but significant, suggest that it is possible to cultivate critical practices with and against AI. The responsibility now falls on academia to build the institutional and pedagogical contexts that can amplify these practices, transforming them from individual negotiations into collective resources for resisting automatised subjectivities.

6 Conclusion

The findings presented in this article underline a tension at the heart of generative AI in higher education. On the one hand, the technology permeates everyday academic practices, reorganising notions of responsibility, authorship and knowledge in line with prevailing logics of efficiency and output. On the other, students’ accounts reveal that this reconfiguration is neither total nor uncontested. Their reflexive practices—questioning accuracy, debating creativity, and

reaffirming the need for human verification—signal a latent capacity to resist the automatising of thought. These gestures suggest that AI need not inevitably consolidate neoliberal pedagogies of optimisation; it can also serve as a point of departure for critical work that interrogates and re-purposes the technology.

This article contributes to the literature in two main ways. First, it moves beyond the dominant risk/opportunity binary by demonstrating that students' engagements with AI are characterised by negotiation, ambivalence and situated reflexivity. Rather than treating AI as either a technical instrument or a threat to academic integrity, the analysis foregrounds its role as a sociotechnical *dispositif* that materialises ideological logics while simultaneously opening space for their contestation. Second, it shows how student subjectivities are being reshaped—where subjectivity is understood, in a Foucauldian sense, as the ensemble of dispositions, self-understandings and modes of conduct produced through and regulated by institutional practices and discursive regimes (Foucault 1982): responsibility shifts from creation to oversight, authorship becomes a relational process of verification and curation, and engagement with AI generates both new forms of creativity and new vulnerabilities. In this sense, the article adds an empirically grounded and theoretically informed account of how AI is domesticated in higher education while revealing the micro-politics through which students resist and rearticulate dominant pedagogical imperatives.

There are, of course, limitations. The study is based on a single course in one institutional and disciplinary setting, and the reflexive diaries, while rich, capture students' representations more than their tacit practices. Moreover, the study does not include longitudinal follow-up, leaving open questions about how these reflexive practices evolve once students enter professional fields or confront AI in other domains of their academic lives. These limitations, however, do not undermine the value of the findings; they point instead to the necessity of comparative and longitudinal research across disciplines, institutions and cultural contexts.

Beyond its empirical contributions, this article positions itself politically and culturally within current debates on AI in higher education. Whereas institutional guidelines and policy reports tend to frame AI primarily in terms of compliance, risk management or technical effectiveness, our analysis foregrounds the ideological and cultural work through which AI reconfigures authorship, subjectivity and pedagogy. In this respect, the study offers not only an account of how students negotiate generative AI in practice, but also a critique of the technocratic imagination underpinning much policy discourse. It demonstrates that the stakes of AI in education are not reducible to questions of regulation or instructional design: they concern the very forms of subjectivity and intellectual autonomy that universities cultivate.

By documenting students' ambivalence and reflexivity, the article provides evidence that higher education can become a site where AI is critically interrogated and re-purposed, rather than merely absorbed into existing neoliberal logics of standardisation and optimisation.

Future research should therefore not only broaden the empirical scope but also continue to interrogate the ideological operations of AI in education. As universities rush to integrate “responsible AI” policies, the risk is that critical reflexivity is reduced to a compliance exercise. Our study suggests that students are already developing practices that can resist this reduction. The challenge for higher education is to amplify these gestures, constructing contexts where AI becomes not a tool for efficiency alone but a catalyst for critical engagement with knowledge, power and subjectivity.

What is ultimately at stake is not whether AI will be integrated into education—this process is already underway—but what kind of education will emerge from its integration. If left to corporate logics, AI will entrench regimes of optimisation, measurability and control. If seized as an object of critique and reflexive practice, it might instead open spaces for reimagining education as dialogical, experimental and emancipatory. The choice is not technological but political: to treat AI as a neutral upgrade, to romanticise it as a transformative ally, or to confront it as a site where the future of critical pedagogy is being negotiated.

Author contributions All authors contributed to the study conception and design. Methodology was developed by M.P. Formal analysis and investigation were carried out by G. B. The first draft of the manuscript was written by G.B. and M.P., and both authors contributed to reviewing and editing the manuscript. All authors read and approved the final version.

Funding Open access funding provided by Università degli Studi di Ferrara within the CRUI-CARE Agreement.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Airolidi M (2021) *Machine habitus: toward a sociology of algorithms*. Polity, Cambridge
- Baidoo-Anu D, Owusu Ansah L (2023) Education in the era of generative artificial intelligence (AI): understanding the potential benefits of ChatGPT in promoting teaching and learning. SSRN. <https://doi.org/10.2139/ssrn.4337484>
- Balmer A (2023) A sociological conversation with ChatGPT about AI ethics, affect and reflexivity. *Sociology* 57(5):1249–1258. <https://doi.org/10.1177/00380385231169676>
- Bissessar C (2023) To use or not to use ChatGPT and assistive artificial intelligence tools in higher education institutions? The modern-day conundrum - students' and faculty's perspectives. *Equity Educ Soc* 0(0):1–13. <https://doi.org/10.1177/27526461231215083>
- Braun V, Clarke V (2006) Using thematic analysis in psychology. *Qual Res Psychol* 3(2):77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun V, Clarke V (2019) Reflecting on reflexive thematic analysis. *Qual Res Sport Exerc Health* 11(4):589–597. <https://doi.org/10.1080/2159676X.2019.1628806>
- Bucher T (2018) *If... then: algorithmic power and politics*. Oxford University Press, New York
- Crompton H, Burke D (2023) Artificial intelligence in higher education: the state of the field. *Int J Educ Technol High Educ* 20:22. <https://doi.org/10.1186/s41239-023-00392-8>
- Dell'Acqua F, Ayoubi C, Lifshitz H, Sadun R, Mollick E, Mollick L, Han Y, Goldman J, Nair H, Taub S, Lakhani K (2025) The cybernetic teammate: a field experiment on generative AI reshaping teamwork and expertise (NBER working paper no. w33641). National Bureau of Economic Research. <https://doi.org/10.3386/w33641>
- Dixon N, Cox A (2025) Fiction writing workshops to explore staff perceptions of artificial intelligence (AI) in higher education. *AI Soc* 40:3983–3998. <https://doi.org/10.1007/s00146-024-02157-x>
- Elliott L (2024) Big tech firms recklessly pursuing profits from AI, says UN head. *The Guardian*. 17 Jan 2024. <https://www.theguardian.com/business/2024/jan/17/big-tech-firms-ai-un-antonio-guterres-davos>
- European Union (2024) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations and Directives (Artificial Intelligence Act). *Off J Eur Union L* 2024/1689, 12 Jul 2024. <http://data.europa.eu/eli/reg/2024/1689/oj>
- Floridi D (2023) AI as agency without intelligence: on ChatGPT, large language models, and other generative models. *Philos Technol* 36:15. <https://doi.org/10.1007/s13347-023-00621-y>
- Formosa P, Bankins S, Matulionyte R, Ghasemi O (2025) Can ChatGPT be an author? Generative AI creative writing assistance and perceptions of authorship, creatorship, responsibility, and disclosure. *AI Soc* 40(5):3405–3417. <https://doi.org/10.1007/s00146-024-02081-0>
- Foucault M (1977) *Discipline and punish: the birth of the prison*. Pantheon, New York
- Foucault M (1982) *The subject and power*. *Crit Inq* 8(4):777–795
- Freire P (1970) *Pedagogy of the oppressed*. Herder & Herder, New York
- Giroux HA (2011) *On critical pedagogy*. Bloomsbury Publishing, New York
- Hutchins E (1995) *Cognition in the wild*. MIT Press, Cambridge
- Kikerpill K, Siibak A (2023) App-hazard disruption: an empirical investigation of media discourses on ChatGPT in educational contexts. *Comput Sch* 40(4):334–355. <https://doi.org/10.1080/07380569.2023.2244941>
- Latour B (2005) *Reassembling the social: an introduction to actor-network theory*. Oxford University Press, Oxford
- Machidon O-M (2025) Generative AI and childhood education: lessons from the smartphone generation. *AI Soc* 40(6):4295–4297. <https://doi.org/10.1007/s00146-025-02196-y>
- Mahaut M, Franzon F (2025) Repetitions are not all alike: distinct mechanisms sustain repetition in language models. *arXiv*. <https://doi.org/10.48550/arXiv.2504.01100>
- Mollick E (2024) *Co-intelligence: living and working with AI*. Ebury Publishing, London
- Nguyen A, Ilesanmi F, Dang B, Vuorenmaa E, Järvelä S (2023) Hybrid intelligence in academic writing: examining self-regulated learning patterns in an AI-assisted writing task. In: Lorig F, Tucker J, Dahlgren Lindström A, Dignum F, Murukannaiah P, Theodorou A, Yolum P (eds) *HHAI 2024: hybrid human AI systems for the social good: proceedings of the third international conference on hybrid human-artificial intelligence*. IOS Press, Amsterdam, pp 241–254
- Ortlipp M (2008) Keeping and using reflective journals in the qualitative research process. *Qual Rep* 13(4):695–705
- Pronzato R (2024) *Algoritmi, strutture e agire sociale. Un'analisi sociologica*. FrancoAngeli, Milano
- Romeo G, Conti D (2025) Exploring automation bias in human-AI collaboration: a review and implications for explainable AI. *AI Soc* 41(1):259–278. <https://doi.org/10.1007/s00146-025-02422-7>
- Sawalha G, Taj I, Shoufan A (2024) Analyzing student prompts and their effect on ChatGPT's performance. *Cogent Educ* 11(1):2397200. <https://doi.org/10.1080/2331186X.2024.2397200>
- Schatzman L, Strauss AL (1973) *Field research: strategies for a natural sociology*. Prentice-Hall, Englewood Cliffs, NJ
- Seaver N (2017) Algorithms as culture: some tactics for the ethnography of algorithmic systems. *Big Data Soc* 4(2):2053951717738104. <https://doi.org/10.1177/2053951717738104>
- US Senate (2023) Oversight of A.I.: rules for artificial intelligence. Hearing, 118th Congress, Senate Subcommittee on Privacy, Technology, and the Law. <https://www.congress.gov/event/118th-congress/senate-event/LC71543/text>
- TIME (2023) Sam Altman. *TIME*100 AI. <https://time.com/collection/time100-ai/6309022/sam-altman-ai/>
- Uberti-Bona Marin LG, Rijsbosch B, Spanakis G, Kollnig K (2025) Are companies taking AI risks seriously? A systematic analysis of companies' AI risk disclosures in SEC 10-K forms. *arXiv*. <https://arxiv.org/abs/2508.19313>
- van Niekerk J, Delpont PMJ, Sutherland I (2025) Addressing the use of generative AI in academic writing. *Comput Educ Artif Intell* 8:100342. <https://doi.org/10.1016/j.caeai.2024.100342>
- Watson S, Romic J (2025) ChatGPT and the entangled evolution of society, education, and technology: a systems theory perspective. *Eur Educ Res J*. <https://doi.org/10.1177/14749041231221266>
- Watson S, Brezovec E, Romic J (2025) The role of generative AI in academic and scientific authorship: an autopoietic perspective. *AI Soc* 40(5):3225–3235. <https://doi.org/10.1007/s00146-024-02174-w>
- Williamson B (2017) *Big data in education: the digital future of learning, policy and practice*. Sage, London
- Xia Q, Weng X, Ouyang F, Lin TJ, Chiu TKF (2024) A scoping review on how generative artificial intelligence transforms assessment in higher education. *Int J Educ Technol High Educ* 21(1):40. <https://doi.org/10.1186/s41239-024-00468-z>

Ziglioli G, Banfi G, Alhassan YA (2025) From tool to partner: exploring common patterns of human-AI collaboration among PhD students in social sciences. In: Hacker P (ed) Oxford intersections: AI in society. Oxford University Press, Oxford

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.