



OPEN ACCESS

EDITED BY

Kelly Merrill Jr.,
University of Cincinnati, United States

REVIEWED BY

Kenzo Seto,
Federal University of Rio de Janeiro, Brazil
Thomas Sommerer,
Johannes Kepler University of Linz, Austria

*CORRESPONDENCE

Oğuz Alyanak
✉ oguz.alyanak@oii.ox.ac.uk

RECEIVED 02 May 2025

ACCEPTED 25 June 2025

PUBLISHED 14 July 2025

CITATION

Graham M, Alyanak O and Valente JCL (2025)
Reflection AI: feeding the machine - the
hidden labour behind AI tools and ethical
implications for higher education.
Front. Commun. 10:1614817.
doi: 10.3389/fcomm.2025.1614817

COPYRIGHT

© 2025 Graham, Alyanak and Valente. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Reflection AI: feeding the machine - the hidden labour behind AI tools and ethical implications for higher education

Mark Graham, Oğuz Alyanak* and Jonas CL Valente

Social Sciences Division, Oxford Internet Institute, University of Oxford, Oxford, United Kingdom

As university instructors integrate AI tools, such as large language models (LLMs) into their pedagogy, they must grapple with the ethical and practical implications of these technologies. This reflection examines the overlooked labour of Cloudworkers and data workers whose contributions make AI systems functional. Drawing on insights obtained from Fairwork's Cloudwork and AI research, it argues for the adoption of the Fairwork scoring system, as a methodology, as well as a heuristic, to guide ethical engagement with AI and urges higher education instructors and students to advocate for improved working conditions in AI supply chains. Additionally, it explores the multifaceted impacts of AI technologies on global labour markets, highlighting pathways to more equitable practises through education, policy, and institutional intervention. By centring the experiences of cloudworkers and data enrichment employees, the article urges various stakeholders to foster a more ethical approach to AI in higher education.

KEYWORDS

artificial intelligence, working conditions, Fairwork, higher education, AI ethics

Introduction

The integration of AI tools like ChatGPT, Grammarly, and image generation systems into higher education classrooms has transformed how we teach and learn (Grassini, 2023). These technologies promise enhanced efficiency, creativity, and access to knowledge (Heaven, 2023; OpenAI, 2023). Yet, beneath their polished interfaces lies an invisible workforce of Cloudworkers and data workers whose labour sustains these systems. These workers often perform monotonous, underpaid, and emotionally taxing tasks, such as moderating harmful content or labelling data for machine learning algorithms (Arsht and Etcovitch, 2018; Fairwork, 2021; Hao and Seetharaman, 2023). This global workforce operates largely out of sight, raising critical ethical questions for educators and institutions that rely on AI tools.

This finding highlights a gap in literature on AI ethics in higher educational settings, where discussions on the ethical use of AI tools, such as applications which draw on large language models (LLMs), omit a critical engagement with the production of such tools, e.g., who are the actual workers that enable the use of AI tools in classrooms, and what are the conditions (from structural, global inequalities to unfair practises in the workplace) under which these tools are produced?¹

This article provides some guidelines for diverse stakeholders, with a particular emphasis on higher-ed, including lecturers, postgraduate students, and institutional staff/administration, to help make

¹ On the contrary, there is extensive literature on working conditions of data workers, including but not limited to Mann and Graham (2016), Miceli and Posada (2022), Muldoon et al. (2024), and Shestakofsky (2024). Literature on working conditions in BPOs that train the AI (e.g., ChatGPT, openAI) to date have been limited to journalistic accounts (see Perrigo, 2023; Rowe, 2023; Hao and Seetharaman, 2023).

informed decisions about the use of AI systems. These recommendations draw primarily on the Fairwork project which evaluates working conditions in companies that train the AI and gives them a Fairwork score that underlines where they align with Fairwork's principles assessing fairness in the workplace, and where they fail to do so. As researchers of the Fairwork project, we additionally propose these stakeholders to familiarise themselves with Fairwork research approach, methodology and output, such as our annual scores evaluating working conditions in AI suppliers, to learn more about the companies and workers building the AI systems and choose service providers which attain a higher Fairwork score.

As educators in higher education, our responsibility extends beyond equipping students with the latest tools. In our respective institutions, we must critically examine the labour practises underpinning these technologies and question whether our pedagogical approaches inadvertently perpetuate exploitation. In this reflection, we argue that universities must recognise the hidden labour behind AI tools and adopt ethical frameworks, such as the Fairwork scoring system, and use it as a methodological framework and a heuristic tool, to ensure that their use aligns with principles of social justice and equity. Moreover, universities must not only educate students about these issues but also take an active role in driving policy changes that demand accountability from AI corporations. It should therefore be our responsibility to be reflexive on ways we bring AI systems into our universities and invite our students to be mindful of the precarious labour that enables such systems to exist.

The hidden workforce behind AI

Education—be it K-12 or higher-ed—has long been considered a site sustained in large part through the invisible or hidden labour provided by instructors (Staudt Willet and He, 2024). Scholars working on this topic have highlighted how teachers, as well as staff working in educational institutions, engage in hidden, invisible, and un-or-underpaid work. In recent years, in light of growing scholarship on AI, scholars have started to question how education has become a site of datafication, and how educators, as well as staff (admin, tech, etc.), sustain datafication by providing hidden and unpaid labour (Selwyn, 2021). In this paper, we focus on labour provided by workers who enable the very AI tools used in classrooms. In contrast to previous studies, the workers we highlight are not located in educational sites, such as schools or universities, but rather at home in front of their computers using Cloudwork platforms or in cubicles in BPOs often located thousands of miles away, feeding the machine day after day (Muldoon et al., 2024; Tubaro et al., 2020).

AI tools such as LLMs are often marketed as autonomous and efficient, but they are anything but independent (Shen et al., 2024). Every intelligent output relies on a foundation of human labour. This workforce, consisting primarily of Cloudworkers² and data enrichment³

workers, plays a pivotal role in data services (such as data labelling), training AI systems, content moderation, and ensuring accuracy (Gray and Suri, 2019). Workers manually annotate data, train datasets, and refine algorithmic outputs. Tasks such as tagging images, labelling text for sentiment, and flagging inappropriate content are essential to building reliable AI models (Muldoon et al., 2024). Workers, moreover, review and filter vast quantities of harmful or explicit content to train AI moderation tools, often exposing themselves to psychologically harmful material without adequate mental health support (Roberts, 2016). They also correct errors, refine algorithmic outputs, and provide feedback to improve system performance. These tasks require precision and attention to detail but are typically undervalued in terms of pay and recognition (Fairwork, 2021).

These jobs are typically outsourced to Cloudwork platforms operating in a planetary labour market (Anwar and Graham, 2020) and business process outsourcing (BPO) firms in low-income countries where labour is cheap and labour protections are weak (Graham et al., 2017). However, it should be noted that in countries like Kenya, despite agendas that prioritise company narratives over worker wellbeing under the rubric of job creation (The Republic of Kenya, Presidency, 2024), there is growing worker activism. In Kenya alone, there are three worker organisations which organise data workers. These are Techworker Community Africa, the Data Labeler's Association, and the African Content Moderators Union. Regardless, workers continue to face precarious employment conditions, including low wages that fail to meet local living standards, inconsistent hours and unpredictable income, and a lack of benefits such as health insurance, paid leave, or mental health resources (Ustek Spilda et al., in press).

Non-compensated worktime is one example of challenges experienced by those workers. Cloudworkers, on average, spent 8.5 h per week on unpaid tasks, such as applying for jobs or managing demanding clients (Fairwork, 2023a). A significant issue for Cloudworkers engaged in data enrichment tasks on microwork platforms is non-payment; a global survey reported that 27% of these workers encountered this problem (Ibid). Additionally, the same report indicated that these workers earned an average of USD 2.15 per hour.

The psychological toll of such work, especially in moderation roles that expose individuals to disturbing content, exacerbates the ethical concerns associated with these practises (Roberts, 2016). Whilst students and educators in wealthier regions benefit from the efficiency of AI tools, the labourers enabling these systems remain largely invisible, their contributions unacknowledged and undervalued. Expanding awareness and advocacy for these workers is vital to building a fairer technological ecosystem.

Implications for university instructors and students

The ethical challenges posed by AI tools are not limited to the corporate sector; they extend into higher education, where these technologies are increasingly central to teaching and learning. University instructors must recognise the human labour embedded in AI tools and educate students about this reality. University departments should acknowledge the challenges associated with the use of these services and adopt ethical practises in their procurement, deployment, and use by faculty, instructors and students. These ethical practises should include evaluating the services and systems implemented based on fair standards, and informing students of the precarious conditions

² Cloudwork can be defined as “remotely performed labour mediated by digital labour platforms – companies that connect workers with clients through a digital interface, exert control over and extract value through the labour process” (Howson et al., 2023, p. 733).

³ Data enrichment can be defined as “Data curation for the purposes of machine learning model development that requires human judgment and intelligence. This can include data preparation, cleaning, labelling, and human review of algorithmic outputs, sometimes performed in real time” (Partnership on AI, 2021, p. 9).

that underscore the labour which enables such tools to exist. Here is one way of doing so: the human labour feeding AI could be incorporated into digital literacy or ethics components of the curriculum, fostering a deeper understanding of the global economy that sustains these tools (Noble, 2018). For example, when teaching with AI writing assistants like ChatGPT, instructors could include discussions about the data enrichment processes that enable these tools to function. This would provide students with a holistic understanding of the technology and encourage critical thinking about its ethical dimensions.

Encouraging reflective thinking about the production of AI and the labour involved in these educational processes aligns with a critical pedagogy perspective. This approach seeks to promote critical awareness of power imbalances and historically rooted issues, emphasising the necessity to challenge systems and advocate for social change as a “freedom practise” (Freire, 2005). In this context, academic institutions and their staff must illuminate the labour-intensive processes that underlie the AI systems they adopt. They must also address the power asymmetries and challenges faced by vulnerable social groups, including workers within these supply chains.

Teaching and learning are not neutral, as the ideology of traditional educational practises suggests (Giroux, 2024). Whilst some may view AI merely as a powerful tool to support the learning process, a critical perspective takes a different approach. This perspective seeks to expose how various forms of power and inequality—social, cultural, and economic—manifest in both formal and informal education for children and adults (Apple et al., 2009, p. 3). Such a critical take on AI in higher educational settings is essential because the theories and actions used to explain social phenomena “structure the possibilities for knowing, acting, feeling, reflecting, and transforming” (Robertson and Dale, 2015, p. 3). Consequently, the theories and perspectives on AI in the classroom significantly influence how students engage with this impactful technology.

AI tools should not be framed as substitutes for human creativity and critical thinking but as complements to them (Holstein and Alevén, 2020). Assignments could ask students to reflect on their experiences using AI tools, including the ethical considerations of relying on such technologies. Research in recent years has identified the risks and harms associated to AI (Slattery et al., 2024), including its environmental impacts (Valdivia, 2024). For instance, students might be tasked with researching the working conditions of data labellers or proposing ways to make AI supply chains more equitable. In doing so, the Fairwork scoring scheme (explained below) can provide a useful methodological framework, as well as a heuristic tool, to assess whether our day-to-day engagement with AI systems promote practises which are exploitative of the very workers, the so-called “human in the loop,” who enable these systems by constantly training the machine. One does not have to go the full length of the Fairwork research.

A simple desk research on companies that utilise workers to feed the AI machine could serve as a starting point to better inform users to decide whether to continue utilising AI tools. If, for example, ChatGPT is known to engage with firms who pay their workers below minimum wage, or demand unpaid labour, or do little to mitigate workers’ exposure to physical, or mental risk, these should serve as a red flag for the users, including educators, to seek alternatives, such as companies which take the additional measures to protect their workers as they train the AI. This should also serve as a red flag to administrative staff who often serve as decision makers bringing AI systems into universities. Reports published by Fairwork could serve

as a guide to obtain more information on working conditions in BPOs. For example, educators, as well as administration, who encourage the use of translation and transcription platforms or buy institutional subscriptions, can refer to the several reports published by Fairwork to choose or subscribe to service providers/platforms which rank higher in Fairwork evaluations (Fairwork, 2022; Fairwork, 2023b; Fairwork, 2025). Students could also be reminded of the “invisible” or “ghost” workers that power the AI (Altenried, 2020) and be encouraged to learn more about the conditions that shape data workers’ experience at AI-training sites, such as BPOs, by reading reports published by the Fairwork project (e.g., Fairwork, 2023c). We therefore encourage educators to use these reports to further educate students on the risks and harms associated with AI work, and administrative staff to use our findings in making informed decisions in subscribing to AI service providers, to take heed in making ethical decisions regarding the adoption and use of AI tools. Additionally, instructors should model responsible AI use by emphasising transparency. This includes disclosing the AI tools used in teaching and discussing their potential ethical implications openly (Fairwork, 2021).

Moreover, universities have a unique opportunity to amplify these lessons through interdisciplinary collaborations that integrate insights from computer science, sociology, and economics. Hosting guest lectures, workshops, and public forums on AI ethics can provide students with diverse perspectives. Beyond the classroom, these initiatives could spur broader movements towards ethical AI practises within academia and beyond.

The role of the Fairwork framework

Whilst instructors can foster ethical awareness, institutional change is essential to addressing the systemic issues underpinning AI tools. This is where the Fairwork scoring system offers a critical intervention.

Data sources and evaluation process

Fairwork evaluates companies using a robust, data-driven methodology grounded in five core principles: fair pay, fair conditions, fair contracts, fair management, and fair representation (Fairwork, 2021). Its methodology begins with a thorough review of publicly available data, including company policies, terms and conditions, and public statements. This is supplemented by direct communication with the companies being assessed, providing them an opportunity to share additional information and clarify their practises. Worker interviews are a critical component of the evaluation process, offering firsthand insights into working conditions, pay structures, and contractual arrangements. This triangulation of sources ensures that evaluations are comprehensive and grounded in reality.

The data collected is analysed against the five Fairwork principles. For example, under the principle of fair pay, companies must demonstrate that workers earn at least the local minimum wage after accounting for expenses. For fair conditions, platforms are assessed on their ability to provide safe and healthy working environments, which includes protections against physical and psychological harm. The principle of fair contracts examines whether contracts are transparent and accessible, avoiding clauses that disproportionately disadvantage workers. Fair management focuses

on mechanisms for dispute resolution and prevention of discrimination, whilst fair representation evaluates whether workers have a voice in governance and decision-making processes (Fairwork, 2021) (See Tables 1, 2).

The integration of rigorous evaluation methodologies and worker-centred advocacy distinguishes Fairwork as a transformative force in the AI and platform economy. Universities can enhance their engagement with Fairwork principles by developing partnerships that extend the framework's applications to local and regional AI projects.

The ten-point scoring system

Companies are scored on a scale of up to ten points, with each principle contributing a maximum of two points. To achieve a full score (of 2) under a principle, companies must meet both basic and advanced criteria. For instance, to score both points for fair pay, a platform must ensure not only that workers earn above the minimum wage but also that they earn a living wage that accounts for local cost-of-living standards. Similarly, fair contracts require both clarity in terms and active measures to ensure contracts do not exploit workers' lack of legal knowledge or bargaining power.

Scores are published annually, fostering accountability and incentivising continuous improvement. High-scoring companies are celebrated as exemplars, setting benchmarks for ethical practises within the industry. Conversely, lower scores serve as a call to action, urging companies to address deficiencies. The iterative nature of this scoring system ensures that companies remain motivated to enhance their labour practises year after year (Fairwork, 2021; Graham et al., 2025).

By adopting the Fairwork framework, universities can promote transparency and ethical accountability in their selection of AI tools. Institutions can use Fairwork scores to inform procurement decisions, ensuring that the tools they adopt align with their values. Additionally, by collaborating with Fairwork to audit the tools they use, universities can play an active role in advocating for improved working conditions in AI supply chains, setting a precedent for other sectors to follow.

Practical steps for universities

Whilst the scoring exercise is a key component of the Fairwork methodology, it should be noted that it is one way to the means, which is to improve working conditions for data workers. Fairwork is an action-research project (Alyanak et al., in press). What that means is that both our methodology, as well as our research output, is intended to bring change to the future of AI work. The scores are a starting point to a larger discussion with multiple stakeholders, where we show that a fairer future of work is possible in AI. For the companies we engage with, we show how this is possible by changing their policies and practises to increase their score. For the regulators, the scores help us make a statement for them to understand that there are top-level interventions needed to better regulate the digital economy. For workers, the scores help us weave webs of solidarity, where we not only remind them the wrongdoings in this economy, but also provide feedback on ways they can demand rights from the companies they work for. And for the larger public, the scores should serve as a

reminder that behind the codes and algorithms, there are always workers, in flesh and blood, working day and night to perfect the machine.

It is, therefore, imperative that the institutions and educators reflect on the Fairwork framework to start up a conversation within departments and with students in classroom, and during office hours, over why even the most basic labour rights continue to be violated by the companies training the AI, and what action students, as consumers of AI tools, can take to demand more humane conditions. Educators, furthermore, are welcomed to include a text to course syllabi to remind students to be aware of the very workers—as well as the conditions they are subjected to—that enable the tools they use in completing assignments. Students, in short, should be invited to make informed decisions in using AI tools. One way of doing so is by inviting them to read Fairwork reports, which offer comprehensive insights into workers and companies that train the AI.

In addition to reminding students of the hidden aspects of AI labour, such as violations of labour rights, universities should also publicly commit to using Fairwork scores in decision-making processes for AI tools and digital services. This could involve integrating Fairwork scores into procurement policies, ensuring that only tools from companies meeting specific ethical thresholds are considered. Additionally, institutions should require vendors to demonstrate compliance with Fairwork principles during the bidding process, reinforcing the importance of fair labour practises. By establishing clear guidelines and accountability measures, universities can set a standard for ethical engagement with AI technologies (Noble, 2018). Moreover, they should partner with the organisation to audit AI tools used in teaching, research, and administration. Educating faculty and students is essential, with workshops and courses offered on the ethical implications of AI, emphasising the human labour behind these tools.

Another pressing issue related to AI is the increasing adoption of Generative AI systems in society, particularly in educational settings. Many universities and faculties are grappling with how to establish ethical guidelines for the use of these tools. These systems impact the educational process in various ways—serving as support or even substitute for student assignments and offering new methods for augmenting or automating teaching tasks. In light of Generative AI, there is a need for a critical approach that combines reflections on the underlying data work involved in GenAI development with inquiries into the broader impact of GenAI use on work. A team of leading researchers has recommended that institutions prioritise this topic in AI discussions and design social protections for both workers and skills development, which would apply to instructors as well as students (Global Partnership on AI, 2023).

Institutions must use their influence to push for stronger labour protections in AI supply chains, both nationally and globally. Universities can advocate for these protections by partnering with labour rights organisations and conducting independent audits of AI supply chains to identify violations and areas for improvement. For example, institutions could join coalitions like the Fairwork project to amplify their impact and align efforts with international standards. Past successes, such as universities influencing tech companies to adopt greener energy practises, demonstrate that academic institutions can affect significant industry change (Graham et al., 2017). Universities might also host conferences or publish reports to bring attention

TABLE 1 Fairwork AI principles.

Principle	First threshold	Second threshold
Fair pay	<p><i>Pays at least the local minimum wage</i></p> <ul style="list-style-type: none">Workers, regardless of their employment status or contract type, must earn the local minimum wage or the wage set by collective sectoral agreement (whichever is higher) for all hours worked.Workers, regardless of their employment status or contract type, are paid on time and in-full.	<p><i>Pays at least the local living wage</i></p> <ul style="list-style-type: none">Workers, regardless of their employment status or contract type, must earn at least the living wage, or the wage set by collective sectoral agreement (whichever is higher) for all hours worked.
Fair conditions	<p><i>Ensures safe working conditions</i></p> <ul style="list-style-type: none">The employer implements policies and practises that protect workers' safety from task specific risks. This should, at a minimum, account for well-evidenced risks such as:High job strain, which can lead to a range of negative health impacts including cardiovascular disease and mental health disorders. Secondary traumatic stress, which can be associated with repeated exposure to traumatic content. Muscular skeletal injuries, which may emerge as a result of unsuitable equipment, excessive workload or perverse incentivisation in physical jobs.Risks related to a specific job are flagged to workers before they accept the job (such as indicating that they might be exposed to violent content.)The employer places a maximum limit on standard working time that meets either the applicable national regulation or, in cases where there is no applicable national regulation, the ILO standard of 40 h a week.Workers are entitled to take breaks during working time that is defined under the applicable national regulation, or in cases where there is no applicable national regulation, is equivalent to a minimum of 1 h for every 8 h worked.If the work arrangements require workers to work in shifts, workers are given the option to choose their shifts, and reasonable accommodations are made for workers with additional needs due to health, safety and other personal reasons (such as pregnancy, care requirements, disability and other health conditions.)	<p><i>Ensures paid leave, and a safety net</i></p> <ul style="list-style-type: none">Workers have access to paid time-off (such as bereavement, parental, sick and annual leave.)Where core medical treatment is not provided by a public system, such as a national healthcare scheme, the employer makes a meaningful provision to the health care costs of its workers.

(Continued)

TABLE 1 (Continued)

Principle	First threshold	Second threshold
Fair contracts	<p><i>Provides decent contracts</i></p> <ul style="list-style-type: none"> Workers must sign a contract and/or give informed consent to terms of conditions upon signing up, and for each subsequent contract extension. The contract or terms and conditions is presented in full, in clear and comprehensible language that all workers could be expected to understand. The contract or terms and conditions are easily accessible to workers in paper and/or electronic form. If these conditions differ for different contract types, reasonable steps are taken to inform workers about the differences in contract types. The party employing the worker must be identified in the contract or terms and conditions, and subject to the law of the place in which the worker works. Workers working on long-term projects that exceed the probation time are provided with the option to sign an employment contract lasting at a minimum the same length of time as the project. The contracts or terms and conditions do not include clauses that revert prevailing legal frameworks in the countries where workers work. 	<p><i>Provides secure employment</i></p> <ul style="list-style-type: none"> Workers with 3 years or more of consistent short-term employment are provided with the option to move onto permanent contracts if they so desire. The employer should make reasonable adjustments in wages and conditions between both: fixed-term and permanent employees and outsourced workers; and any outsourced or indirectly employed workers and directly employed workers. Workers who are outsourced or indirectly employed should be compensated for additional costs incurred, including visa/work permits and their extensions, insurance, pensions, and other social security premiums. In cases of justified redundancy or contract non-renewal, the employer should provide workers with severance allowance commensurate with tenure at the company and retraining opportunities. In cases where the redundancies are being made because reasons of an economic, technological, structural or similar nature, workers or their representatives are consulted, and steps are taken to minimise the resulting redundancies. If desired, workers should be able to invite worker representatives to their end of contract meetings with the relevant HR departments. In the case of subcontracting arrangements, where part or all of the work is subcontracted to other companies, management implements a reliable mechanism to monitor and ensure that the subcontractor is living up to the standards expected from the company itself regarding working conditions.
Fair management	<p><i>Treats workers fairly</i></p> <ul style="list-style-type: none"> Management should refrain from deploying any form of depersonalised bullying or mobbing in order to ensure organisational goals are met. There is a policy in place which guarantees that any form of harassment in the workplace will not be tolerated. There is a policy in place which guarantees that the employer will not discriminate against persons on the grounds of racial, ethnic, social or minority background, caste, religion or belief, political or any other opinion, language, gender, gender identity, sex, sexual orientation, disability, age, geographical location, or any other status. Workers have the right to appeal dismissals and other disciplinary measures. Workers are not disadvantaged for voicing concerns or appealing disciplinary actions. 	<p><i>Creates clear and effective systems for data management, explanations, and appeals</i></p> <ul style="list-style-type: none"> Where AI systems are involved in work, employers must create explainability mechanisms such as transparency reports or question and answer processes that allow workers to understand both the model behaviour of the system as a whole and specific decisions. Workers must be able to appeal decisions made by AI systems through a multi-stakeholder process that reflects collective worker voice, and successful appeals to lead not only that specific decision being revised but also wider revisions of decision-making process. Management avoids excessive surveillance in the workplace, and avoids use of invasive technologies. Workers must not be subject to excessive data collection practises and should be informed about the data that is being collected about them. Employers must apply the principle of data minimisation (collecting the minimum amount of personal data required to fulfil a legitimate purpose) in their collection processes.
Fair representation	<p><i>Assures freedom of association and the expression of worker voice</i></p> <ul style="list-style-type: none"> There is a documented mechanism for the expression of collective worker voice that allows all workers, regardless of contract type or duration to participate in collective groups without risks. There is a formal, written statement of willingness to recognise, and bargain with, a collective, independent body of workers or trade union, that is clearly communicated to all workers, and available on the company webpage. Freedom of association is not inhibited, and workers are not disadvantaged in any way for communicating their concerns, wishes and demands to the company management, or expressing willingness to form independent collective bodies of representation. 	<p><i>Supports democratic governance</i></p> <ul style="list-style-type: none"> Workers play a meaningful role in governing the company. In a written document available, the company publicly and formally recognises an independent collective body of workers, an elected works council, or trade union, and takes meaningful steps towards signing a collective bargaining agreement. This recognition is not exclusive and, when the legal framework allows, the company should recognise any significant collective body seeking representation.

TABLE 2 Fairwork Cloudwork (online work) principles.

Principle	First threshold	Second threshold
Fair pay	<p><i>Workers are paid on time and for all completed work</i></p> <ul style="list-style-type: none"> Non-payment is not an option for clients and there are mechanisms to ensure workers are paid. Payments are made within an agreed timeframe. Workers can choose to be paid in a recognised national currency. Workers can request funds from their account on a regular basis with reasonable withdrawal thresholds. 	<p><i>Pays at least the local living wage</i></p> <ul style="list-style-type: none"> For hourly-paid work, workers earn at least their local minimum wage after costs. For piece-rate work: (a) the vast majority of workers earn at least their local minimum wage after costs, and (b) a reasonable estimate of the time it takes to complete each task is provided to each worker before they accept the work.
Fair conditions	<p><i>Health and safety risks are mitigated</i></p> <ul style="list-style-type: none"> The allocation of work and/or supply of new workers is managed to promote job availability, and reduce unpaid work and overwork. 	<p><i>Ensures paid leave, and a safety net</i></p> <ul style="list-style-type: none"> There are policies to protect workers from risks that arise from the processes of work. There are processes for job-related health and safety risks (including psychological risks) to be identified and addressed. Risks related to a specific job are flagged to workers before they accept the job (such as indicating that they might be exposed to violent content). There are clear reporting channels and documented penalties for clients who jeopardise workers' health and safety. There are adequate and ethical data privacy and security measures applicable to workers, laid out in a documented policy.
Fair contracts	<p><i>Clear terms and conditions are available</i></p> <ul style="list-style-type: none"> The contract is written in clear and comprehensible language that the worker could be expected to understand. The contract is available for workers to access at all times. Workers are notified of proposed changes in a reasonable timeframe before changes come into effect. Changes should not reverse existing accrued benefits and reasonable expectations on which workers have relied. The contract does not require workers to waive rights to reasonable legal recourse against the platform. 	<p><i>Contracts are consistent with the workers' terms of engagement on the platform</i></p> <ul style="list-style-type: none"> Clients are encouraged to inform workers about how their work will be used. The worker is not subject to non-compete clauses. Except in cases where the worker is in a standard employment relationship the platform makes clear to workers and clients that: Working schedules cannot be imposed upon workers. The worker retains the freedom to choose which tasks to accept or refuse. When workers choose not to accept tasks, this does not punitively impact their rating or reputation.
Fair management	<p><i>There is a due process for decisions affecting workers</i></p> <ul style="list-style-type: none"> There is a channel for workers to communicate with a human representative of the platform. This channel is documented in policies that are easily accessible to workers, and communications are responded to within a reasonable timeframe. There is documentation of things the worker is not allowed to do, and workers receive an explanation for all punitive actions, including reductions in their rating/platform standing, non-payment, work rejections, penalties, account blocks, deactivation and any other disciplinary actions. Explanations for punitive actions and work rejections include information on how they can be appealed. The process for workers to appeal punitive actions and work rejections is non-arduous, documented in the contract, and available to workers who no longer have access to the platform. 	<p><i>There is equity in the management process</i></p> <ul style="list-style-type: none"> There is a policy which guarantees that the platform will not discriminate against persons on the grounds of racial, ethnic, social or minority background, caste, religion or belief, political or any other opinion, language, gender, gender identity, sex, sexual orientation, disability, age, geographical location, or any other status. There are mechanisms to reduce the risk of clients discriminating against workers on any basis listed above. The platform specifies the methods used to manage and allocate work (including when algorithms are used). Substantive changes to methods of managing and allocating work are preceded by a worker consultation.
Fair representation	<p><i>Workers have access to representation, and freedom of association</i></p> <ul style="list-style-type: none"> The platform commits to a dispute resolution process in which workers can access an independent representative freely chosen by them or an unbiased, independent, and accessible dispute resolution system. Freedom of association is not inhibited and groups of workers are not disadvantaged in any way for communicating their concerns, demands and wishes to management. 	<p><i>There is collection governance or bargaining</i></p> <ul style="list-style-type: none"> It is democratically governed by workers. It formally engages with an independent collective body of workers, an elected works council or trade union, and has not refused to engage with collective bodies who seek representation and/or bargaining. New workers are advised of the existence of these bodies. If such a body does not exist, or such bodies have not requested recognition by the platform, the platform has engaged in a formal process of dialogue with local and/or international representative bodies of workers to discuss what structures and processes of representation could look like for platform workers.

to labour issues in AI, thereby pressuring companies to commit to better practises.

Moreover, collaborative initiatives between universities and international organisations could facilitate the development of ethical guidelines for AI tools, ensuring that labour practises are central to global technology standards. By championing these causes, universities can reinforce their roles as leaders in social responsibility and innovation.

These recommendations may face challenges in implementation. Students, for example, may choose to forego ethical concerns amid pending deadlines, and institutions may procure services from providers who resort to ethically dubious labour practises. Teachers can face issues to critically approach AI use in the classroom when institutions do not have a policy or adopt permissive rules about this topic.

However, it is imperative for all users—be they university admin, instructors, or students—to be informed of the ethical debates that envelop the very tools they use, and to be critical of such use. This is why higher education institutions should discuss and implement policies and guidelines that acknowledge the problems in AI supply chains and provide orientations on how to address them in classrooms. Another alternative to addressing these challenges can be for universities to develop their own models with ethical considerations in mind. Whilst the development of large language models requires substantial resources, a consortium of universities promoting the use of AI in education could embark on a joint initiative to that end.

As underlined by Freire (2005), education fosters critical thinking, and educators should strive in their instruction to be critical of use of AI in educational setting. Such an approach would pave way towards challenging systems in place and working collectively towards social change.

Conclusion

The adoption of AI tools in education offers transformative potential but also implicates us in systems of global labour exploitation. As instructors and institutions, we have a moral obligation to acknowledge and address the hidden labour behind these technologies. By centring the experiences of Cloudworkers and data enrichment employees, we can foster a more ethical approach to AI in higher education. The Fairwork framework provides a practical and actionable pathway for achieving this goal. Universities must take the lead in promoting transparency, advocating for fair labour practises, and ensuring that the tools we use align with our values. This is not just about teaching with AI; it is about teaching responsibly, with an unwavering commitment to justice and equity for all workers, visible and invisible alike.

References

Altenried, M. (2020). The platform as factory: Crowdwork and the hidden labour behind artificial intelligence. *Capital Class* 44, 145–158. doi: 10.1177/0309816819899410

Expanding these efforts through global collaborations, interdisciplinary research, and active policy engagement will help ensure that AI tools serve as instruments of equity rather than exploitation. The ethical adoption of AI in education is not merely a challenge; it is an opportunity to model the values of justice and human dignity in a rapidly evolving technological landscape.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The research draws on general findings from the Fairwork project. It does not rely on primary data in its analysis. Requests to access these datasets should be directed to info@fair.work.

Author contributions

MG: Writing – original draft, Writing – review & editing. OA: Writing – original draft, Writing – review & editing. JV: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Alyanak, O., Bertolini, A., Ustek-Spilda, F., Valente, J., Warin, R., and Graham, M. (in press). "Action research: the Fairwork project" in *The handbook of digital labour* (Sage). Eds. E. Bulut, J. Y. Chen, R. Grohmann, K. Jarrett.

- Anwar, M. A., and Graham, M. (2020). Digital labour at economic margins: African workers and the global information economy. *Rev. Afr. Polit. Econ.* 47, 95–105. doi: 10.1080/03056244.2020.1728243
- Apple, M. W., Au, W., and Gandin, L. A. (2009). "Mapping critical education" in *The Routledge international handbook of critical education*. eds. M. W. Apple, W. Au and L. A. Gandin (New York, Oxford: Routledge), 3–19.
- Arshat, A. and Etcovitch, D. (2018). The human cost of online content moderation. *Harvard Journal of Law and Technology*. Available online at: <https://jolt.law.harvard.edu/digest/the-human-cost-of-online-content-moderation> (Accessed June 10, 2025).
- Fairwork. (2021). Fairwork labour standards in the platform economy. Available online at: <https://www.fair.work> (Accessed May 26, 2005).
- Fairwork (2022). *Fairwork translation & transcription platform ratings 2022*. Oxford, United Kingdom: University of Oxford.
- Fairwork (2023a). *Work in the planetary labour market: Fairwork Cloudwork ratings 2023*. Oxford, United Kingdom: University of Oxford.
- Fairwork (2023b). *Fairwork translation & transcription platform ratings 2023*. Oxford, United Kingdom: University of Oxford.
- Fairwork (2023c). *Fairwork AI ratings 2023: The workers behind AI at Sama*. Oxford, United Kingdom: University of Oxford.
- Fairwork. (2025). *Cloudwork Report 2025: Advancing Standards in Digital Labour and AI Supply Chain Governance*. Oxford, United Kingdom.
- Freire, P. (2005). *Pedagogy of the oppressed (30th Anniversary Edition)*. New York: Continuum.
- Giroux, H. A. (2024). Educators as public intellectuals and the challenge of fascism. *Policy Futures Educ.* 22, 1533–1539. doi: 10.1177/14782103241226844
- Global Partnership on AI. (2023). Policy brief: generative AI, jobs and policy response. GPAI, Montreal. Available online at: <https://gpai.ai/projects/future-of-work/policy-brief-generative-ai-jobs-and-policy-response-innovation-workshop-montreal-2023.pdf> (Accessed May 18, 2025).
- Graham, M., Alyanak, O., Bertoini, A., Feuerstein, P., Kuttler, T., Ustek Spilda, F., et al. (2025). Pressure and praise as an action research methodology: the case of Fairwork. *Environ. Plan. A Econ. Space*. doi: 10.1177/0308518X251336893
- Graham, M., Hjorth, I., and Lehdonvirta, V. (2017). Digital labour and development: impacts of global digital labour platforms. *Dev. Stud. Res.* 4, 12–29. doi: 10.1177/1024258916687250
- Grassini, S. (2023). Shaping the future of education: exploring the potential and consequences of AI and ChatGPT in educational settings. *Educ. Sci.* 13:692. doi: 10.3390/educsci13070692
- Gray, M. L., and Suri, S. (2019). *Ghost work: How to stop Silicon Valley from building a new global underclass*. Boston: Houghton Mifflin Harcourt.
- Hao, K., and Seetharaman, D. (2023) Cleaning up ChatGPT takes heavy toll on human workers, *Wall Street J.* Available online at: <https://www.wsj.com/tech/chatgpt-openai-content-abusive-sexually-explicit-harassment-kenya-workers-on-human-workers-cf191483> (Accessed June 10, 2025)
- Heaven, W. D. (2023). "ChatGPT is going to change education, not destroy it." MIT Technology Review. Available online at: <https://www.technologyreview.com/2023/04/06/1071059/chatgpt-change-not-destroy-education-openai/> (Accessed May 18, 2025).
- Holstein, K., and Alevan, V. (2020). Designing for human–AI complementarity in K-12 education. *AI Mag.* 43, 239–248. doi: 10.1002/aaai.12058
- Howson, K., Johnston, H., Cole, M., Ferrari, F., Ustek-Spilda, F., and Graham, M. (2023). Unpaid labour and territorial extraction in digital value networks. *Global Networks*, 23, 732–754. doi: 10.1111/glob.12407
- Mann, L., and Graham, M. (2016). The domestic turn: business process outsourcing and the growing automation of Kenyan organisations. *J. Dev. Stud.* 52, 530–548. doi: 10.1080/00220388.2015.1126251
- Miceli, M., and Posada, J. (2022). The data-production Dispositif. *Proc. ACM Hum. Comput. Interact.* 6, 1–37. doi: 10.1145/3555561
- Muldoon, J., Graham, M., and Cant, C. (2024). *Feeding the machine: The hidden human labour powering AI*. London: Cannongate/New York: Bloomsbury.
- Noble, S. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: New York University Press.
- OpenAI. (2023). "Teaching with AI." Available online at: <https://openai.com/index/teaching-with-ai/> (Accessed May 18, 2025).
- Partnership on AI. (2021). "Responsible sourcing of data enrichment services. Available online at: <https://partnershiponai.org/paper/responsible-sourcing-considerations/> (Accessed March 21, 2025).
- Perrigo, B. (2023). "Exclusive: OpenAI used Kenyan workers on less than \$2 per hour to make ChatGPT less toxic", *Time*. Available online at: <https://time.com/6247678/openai-chatgpt-kenya-workers/> (Accessed June 10, 2025).
- Robertson, S., and Dale, R. (2015). Toward a critical cultural political economy of the globalisation of education. *Globalisation. Soc. Educ.* 13, 149–170. doi: 10.1080/14767724.2014.967502
- Roberts, S. T. (2016). *Commercial content moderation: digital laborers' dirty work*, In *The intersectional Internet: race, sex, class and culture online*, eds Noble, S. U & Tynes, B. New York: Peter Lang Publishing, Inc, 147–160.
- Rowe, N. (2023). "It's destroyed me completely": Kenyan moderators decry toll of training of AI models, *Guardian*. Available online at: <https://www.theguardian.com/technology/2023/aug/02/ai-chatbot-training-human-toll-content-moderator-meta-openai> (Accessed June 10, 2025)
- Selwyn, N. (2021). The human labour of school data: exploring the production of digital data in schools. *Oxf. Rev. Educ.* 47, 353–368. doi: 10.1080/03054985.2020.1835628
- Shen, Y., Shao, J., Zhang, X., Lin, Z., Pan, H., Li, D., et al. (2024). Large language models empowered autonomous edge AI for connected intelligence. *IEEE Commun. Mag.* 62, 140–146. doi: 10.1109/MCOM.001.2300550
- Shestakofsky, B. (2024). Cleaning up data work: negotiating meaning, morality, and inequality in a tech startup. *Big. Data Soc.* 11. doi: 10.1177/20539517241285372
- Slattery, P., Saeri, A. K., Grundy, E. A., Graham, J., Noetel, M., Uuk, R., et al. (2024). *The AI risk repository: a comprehensive meta-review, database, and taxonomy of risks from artificial intelligence*. doi: 10.13140/RG.2.2.28850.00968
- Staudt Willet, K. B., and He, D. (2024). Educators' invisible labour: a systematic review. *Rev. Educ.* 12:e3473. doi: 10.1002/rev3.3473
- The Republic of Kenya, Presidency (2024). "Government's plan to create 1 million jobs." Available online at: <https://www.president.go.ke/governments-plan-to-create-1-million-jobs/>
- Tubaro, P., Casilli, A. A., and Coville, M. (2020). The trainer, the verifier, the imitator: three ways in which human platform workers support artificial intelligence. *Big Data Soc.* 7:205395172091977. doi: 10.1177/2053951720919776
- Ustek Spilda, F., Brittain, L., Alyanak, O., and Graham, M. (in press). "Datafication, surveillance and automation: capturing workers' experience in the digital economy with Fairwork AI principles" in *Job quality in a turbulent era*. eds. A. Piasna and J. Leschke (European Trade Union Institute (ETUI)).
- Valdivia, A. (2024). The supply chain capitalism of AI: a call to (re) think algorithmic harms and resistance through environmental lens. *Inf. Commun. Soc.*, 1–17. doi: 10.1080/1369118X.2024.2420021