https://doi.org/10.17163/soph.n40.2026.02

Digital Ghosts, Moral Algorithms, and the Challenge of Teaching Ethics in the Posthuman Age

Fantasmas digitales, algoritmos morales y el desafío de enseñar ética en la era poshumana

Amitabh Vikram Dwivedi[*]
amitabh.vikram@smvdu.ac.in
https://orcid.org/0000-0002-7779-2654
Shri Mata Vaishno Devi University, UT of Jammu & Kashmir, India
https://ror.org/036x6w630

Suggested citation: Dwivedi, Amitabh Vikram. (2026). Digital Ghosts, Moral Algorithms, and the Challenge of Teaching Ethics in the Posthuman Age. *Sophia, Colección de Filosofía de la Educación*, (40), pp. 81-117.

Abstract

The article tracks how AI sorting tools and profit-seeking metrics quietly rewrite the ethical rulebook of the university. Introducing the concept of "digital ghosts" to describe the artefactual traces of authorship, intent, and authenticity produced by technology-mediated work, it contends that current educational models encourage a posthuman form of responsibility characterized by weakened accountability. Through empirical vignettes—including algorithm-driven citation inflation, dissertations constructed with large language models, and a July 2025 incident in which institutional affiliations were openly sold on LinkedIn—the article documents an academy where ethical judgment is displaced by platform logic and metrics, and where growing dependence on dashboards increases integrity risks for students and faculty alike. It explains how algorithmic infrastructures reshape authorship and evaluation, examines the consequences across academic roles, and proposes a program to restore accountability that pairs provenance and audit-ready practices with pedagogy integrating technical literacy and ethical reasoning. The article concludes that academic integrity must be deliberately redesigned for digitally mediated scholarship rather than assumed to remain intact.

Keywords

Digital Ethics, Academic Integrity, Posthuman Education, Moral Algorithms, Surveillance capitalism, Virtue Ethics.

[*] He is Dean of Humanities and Social Sciences and Professor in the Faculty of Languages and Literature at Shri Mata Vaishno Devi University (India), as well as a Harvard-certified expert in Writing and Public Speaking. His research focuses on linguistics, linguistic documentation, and the preservation of endangered South Asian languages, literature, and culture. Google Scholar: https://scholar.google.com/citations?user=R1bv5L4AAAAJ&hl=en
h-index: 9

Resumen

Este artículo analiza las implicaciones éticas derivadas de las tecnologías digitales (TIC) y las infraestructuras algorítmicas en la educación contemporánea, especialmente la inteligencia artificial (IA), los sistemas de clasificación basados en datos y los incentivos orientados al beneficio económico. Examina cómo estos factores transforman la conducta moral de estudiantes y docentes. A partir del concepto de "fantasmas digitales", entendido como las huellas de autoría, intención y autenticidad que persisten en obras mediadas o generadas por IA, se sostiene que los modelos educativos actuales propician una forma de responsabilidad poshumana marcada por la erosión de la rendición de cuentas tradicional. Mediante viñetas empíricas, se exploran ejemplos de decadencia ética, como la inflación algorítmica de citas, las tesis redactadas con modelos lingüísticos de gran escala y un escándalo ocurrido en julio de 2025: la compraventa de filiaciones institucionales a través de LinkedIn. Estos casos evidencian el rumbo preocupante que sigue la educación, donde la ética se diluye en los intersticios de un sistema tecnológicamente integrado. Cuanto más dependen las instituciones de plataformas y métricas digitales, mayor es el riesgo para la integridad y la responsabilidad académica. En este contexto, urge reconsiderar el significado práctico de la integridad y redefinir los marcos de responsabilidad ante las transformaciones tecnológicas que reconfiguran la idea misma de lo humano.

Palabras clave

Ética digital, integridad académica, educación poshumana, algoritmos morales, capitalismo de vigilancia, ética de la virtud.

Introduction

Recent scholarships have called for greater attention to the algorithmic infrastructures that inform and produce scholarship in a contemporary, digitally mediated academy (Recio Sastre, 2025). Artificial intelligence (AI), data-driven ranking systems, and profit-motiveness have overwhelmed universities post-AI, transforming them from mere venues for faculty and student interactions. This change raises crucial questions about the responsibility to protect everyone who is affected, students, faculty and institutions. This troubling rise of 'digital ghosts' undercuts the conventional understanding of authorship and pays us to re-think what we do and do not steward, are accountable for, and have faith in.

This article explores the ethical aspects surrounding these digital technologies in modern education. Through studying how changes in these factors affect ethical conduct in both students and faculty, the study aims to show how present-day educational paradigms have played a major role in engendering an attenuated mode of normative responsibility characterized by that old post-human chestnut, moral accountability.

Although educational institutions are becoming paperless, the core of academic integrity seems to be losing its texture. All of this raises important questions regarding the veracity of knowledge generation in an environment increasingly characterized by algorithmic publications and

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

sequestering those human spirits involved in the production of knowledge. At the root of this dilemma is the need to understand how a posthuman information system challenges ethical practices and shifts the burden of accountability from human stakeholders.

The study proposes that a proliferation of digital ghosts—products of AIs and algometrics—causes epistemic irresponsibility in academia. Such a change in scale calls for us to rethink integrity and accountability, to rematerialize ethical obligation according to the worlds put into being by our digital technologies.

As our educational practices become more and more intertwined with digital innovations, it is critical that academics, policymakers, and society at large engage in discussions about what the ethical implications of this transformation are. This leads to reflections on the discussions about academic integrity and ethical routines, ultimately revealing what we could expect when digital algorithms increasingly shape educational activities as part of algorithmic governmentality.

Considering the recent scandals involving institutional affiliations, digital technologies, and academic fraud, this topic is particularly relevant. With the backdrop of a rapidly changing pedagogical environment, knowledge of this topic has become essential, largely because existing ethical frameworks need to change to accommodate new realities facilitated by technologies. As more stakeholders are committed to visibility and quantitative metrics instead of real engagement, addressing this topic is crucial for education and the future of scholarship.

A qualitative research approach has been applied in this paper, importing empirical vignettes to represent these forms of decline that become emergent due to the algorithmic influences and AI-generated outputs. The analysis is embedded within a comprehensive literature review and uses case studies techniques together with applicable theoretical frameworks. Approaches range from a detailed revisiting of major corruption scandals to algorithmic and computational assemblages, and from considerations of neoliberal transformations in authorship and citation practices to both an analysis of the moral economy within which such revelations arise and reflection on their aftermath. Data collection consists of a review of the literature and reports related to Digital Technology (ICTS) in education combined with the integration of opinions being expressed in current discussions about digital ethics and other matters.

To provide a conceptual framework that presents definitions along with an exploration into these phenomena. It then presents case studies and real-life examples illustrating ethical degradation in contemporary

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

academia. Subsequent sections expand on conversation and discourse analysis, examining their effects and their relationship to academic honesty and moral responsibility. Finally, the conclusion offers insights and recommendations for recalibrating accountability in posthuman educational contexts, along with suggestions for future research.

The world of digital ghosts

This section defines "digital ghosts" and locates their ontological stakes in contemporary scholarship, then shows how AI systems, metric platforms, and citation engines erode epistemic responsibility. The text isolates four entrenched modes of academic ghosting— large language models (LLM) generated manuscripts, recursive citation rings, traded authorship slots, and rentable institutional affiliations—showing how each flips the prestige economy from veracity to visibility. It then tracks the moral fallout once human deliberation is ceded to opaque evaluation algorithms: responsibility evaporates into the circuitry, and "good" becomes whatever the dashboard can count. A concise analysis of the Pune "affiliation-rental" scandal—conducted in full public view—illustrates the mechanics of this erosion. The section ends by flagging two forthcoming graphics: a risk-assessment matrix of algorithmic harms and a timeline that reconstructs the Pune incident step-by-step.

The present-day modern universities are haunted by the so-called *digital ghosts*—phantom agents whose presence is manifested in automatically generated prose, algorithmically generated reading lists, and self-replicating citation graphs, silently infusing scholarly networks with nonhuman actors that escape any responsible human authorship. But these echoes of innovation are not innocent; as Pestre (2021) notes with insight, they mark a worrying "erosion of epistemic responsibility", the fundamental duty to be held accountable for what we say, observe, measure, and interpret. With each occasion when a language model rolls out a literature review in seconds, or when metrics services fabricate an h-index in a matter of hours, human judgment is ceded to black-box computational doppelgängers (Pasquale, 2015; Zuboff, 2019). In Hans (2015) so-called "transparency society", the chilling irony is that as outputs have become more visible, the moral agency behind them has vanished from sight.

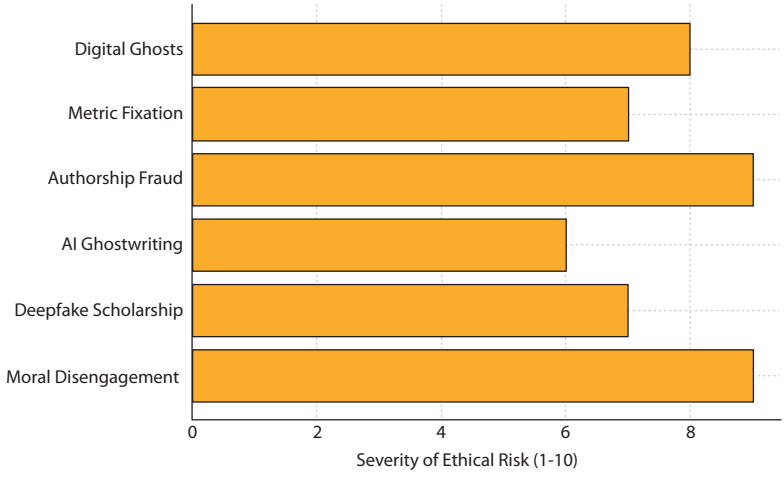
The issue is exacerbated by the fact that digital ghosts are not limited to simple plagiarism, which is the easiest target for detection software.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

They signal a deeper, ontological shift in what scholarship is and how it enters the world of academia, a transition Hayles (1999) foresaw while discussing the "posthuman". Knowledge objects now circulate detached from embodied labor and situated accountability, assuming at least four recurrent forms. For example:

- Complete manuscripts written by LLM are adapted to nominal human authors, which is called automation rhetoric.
- Citation algorithms trigger recursive feedback loops that magnify the apparent authority of works that few have actually read (Fong & Wilhite, 2017).
- black markets that resemble less a collegial collaboration and more a commerce (Hosseini *et al.,* 2018).
- Meanwhile, institutional ties became profitable assets that were sold to improve global ranking metrics, reversing Merton's (1973) norms of the common good and disinterested scrutiny.

Figure 1
Classification of the severity of ethical risks in the digital age on a scale of 1 to 10



Taken together, these expressions help indicate the form of a scholarly ecosystem in which visibility and calculability, and even taxation, outweigh the kind of accountability that used to be embodied in the person occupying the desk down the hall and in which the spectral weight of digital ghosts increasingly writes the story of academic life. Collectively,

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

these manifestations point to a posthuman academic economy whose principal currency is not truth but transmissible appearance. Scholarship is rewarded not for the depth of inquiry it embodies but for its algorithmic resonance, its capacity to register as a quantifiable event in ranking dashboards, citation indices, and social media analytics. In such an environment, the *digital ghost* is not an anomaly to be exorcised; it is the structural norm that reveals how profoundly academic value has been recoded in the age of moral algorithms.

Posthuman Decision-Making and Moral Erosion

Braidotti's (2013) notion of posthuman decision-making captures a historical juncture in which the locus of agency migrates from embodied subjects to machinic assemblages—software agents, predictive models, and data architectures that follow their own internally consistent, but socially uninterpretable, logics. Within higher education, this migration is concretized in an interlocking infrastructure of algorithmic ranking dashboards, auto-updating citation indices, plagiarism-detection *suites*, grant-scoring algorithms, and recommendation engines that advise search committees whom to hire and libraries what to buy. The cumulative effect is precisely what Pasquale (2015) describes as the "black-box society": consequential judgments about scholarly merit, career advancement, and institutional prestige are rendered by opaque computational procedures whose assumptions, weightings, and training data are seldom disclosed or debated.

This reconfiguration of agency changes the moral terrain of the academy. Arendt's (1963) portrait of the "banality of evil" cautions that moral catastrophes frequently result, not from diabolical intent, but from the routine failure to exercise reflective judgment—thoughtlessness grounded in procedural routine. Algorithmic governance intensifies this hazard by providing a ready-made *alibi*: responsibility can be diffused into code, outsourced to datasets, or attributed to "the system". Mittelstadt *et al.* (2016) note that most machine-learning pipelines are designed for optimization and prediction, not for normative deliberation; they lack any intrinsic capacity to weigh justice, fairness, or academic integrity. When promotion committees defer to auto-generated h-index cutoffs that update every night, or journal editors reflexively trust in citation-impact forecasts, they make decisions with as yet unexamined moral implications, ceding ethical agency to statistical correlation. The consequence of this evolution has been a pseudo-voluntaristic iteration of Arendtian thoughtlessness: people are pushing the buttons, but all moral inquiry

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

which ought to direct their behavior is tacitly extradited to algorithms all but unencipherable.

The Pune Incident: A Case Study of Academic Ghosting

On 7 July 2025, a routine LinkedIn scroll revealed an extraordinary post from a mid-career economics professor at a private management institute in Pune. The message was bluntly transactional: "40,000 for sole or first-author papers; 25,000 for middle-author slots—just add our institute's name. DM for details." In a matter of hours, the offer had accrued thousands of views, dozens of "likes," and a comment thread that read like an ethnography of contemporary scholarly culture before the author hastily deleted the evidence. Screenshots, however, had already begun circulating on Twitter, ResearchGate, and departmental WhatsApp groups, ensuring the post's afterlife as a case study in what might be called the *financialization of affiliation*—treating institutional branding as a strip-mall add-on to any manuscript nearing submission.

Figure 2
Censored post on LinkedIn referring to questionable research ethics

What made the incident especially revealing was not the rogue solicitation itself—similar deals have long circulated quietly through conference corridors—but the breadth and tone of the reactions it elicited. Marketing scholar observed with weary pragmatism, "The educa-

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

tion industry has new business verticals. Many assistant professors have quit teaching to do this full-time" (LinkedIn comment, 9 July 2025). Others expressed open disgust or dark humor, yet an undercurrent of resignation ran through the thread: this is simply how the metrics game is now played. The spectrum of responses maps neatly onto Bandura's (2016) taxonomy of moral disengagement. Some commenters employed advantageous comparison ("At least he's not fabricating data"), others displaced responsibility ("NIRF rankings[1] leave us no choice"), and still others minimized consequences ("A harmless branding exercise"). The most sweeping abdication came from Prof. M. R. Saeed: "Who cares for ethics? Do they exist in politics, in corporate life, or—worst of all—in academia? Faculties are forced to do this or perish!" (LinkedIn comment, 9 July 2025).

In a single afternoon the Pune post rendered visible an entire shadow economy of authorship swapping and affiliation rental that usually operates off-platform, crystallizing the broader crisis of integrity now confronting digital academia. It showed how quickly moral norms can be reframed as market opportunities once scholarly reputation is mediated by algorithmic rankings and citation dashboards, and how effortlessly individual actors rationalize participation in those markets when systemic pressures make ethical deviation appear not only normal but necessary.

The Crisis of Truth in Academia

This section maps how metric fixation recasts academic value through rankings and dashboards (NIRF, Shanghai, REF, SJR, etc), reading them with Foucault's disciplinary power to show how quantitative visibility displaces ethical purpose. It then tracks the monetization of authorship—paper mills, citation cartels, and byline trading—through the LinkedIn solicitation as a concrete hinge between incentive structures and everyday practice. Then, brief global scan follows (Clarivate delistings, Hindawi retractions, Chinese paper-mill pipelines, Brazilian citation rings) to illustrate that the integrity problem is systemic rather than local. The section culminates in an account of surveillance capitalism in academia, where impact metrics, downloads, and alt-metrics become extractive resources, alienating intellectual labor from inquiry and normalizing data manipulation over discovery.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

Institutional Pressures in Rankings and Metrics

Over the last decade higher education has been recast as what Muller (2018) names a regime of "metric fixation," an environment in which the worth of a person, program, or institution is inferred almost exclusively from numerically tractable surrogates—impact factors, h-indices, placement percentages, and citation tallies—while contextual judgment and ethical purpose recede from view. Nowhere is this logic more visible than in India's NIRF, whose annual scorecards determine government funding, student demand, and media prestige. By rewarding sheer publication volume and citation counts, the framework fabricates what one business-school faculty member, A.S. Chandel, calls "a game played just to climb the NIRF ladder and pull in admissions" (LinkedIn comment, 9 July 2025). Similar scoreboards—from the Shanghai Ranking to the UK's REF—propagate the same incentives globally, encouraging scholars to slice studies into multiple "least publishable units," court citation cartels, or purchase authorship outright.

A useful lens is Foucault's (1977) genealogy of disciplinary power which sees rankings as panoptic technologies that watch, compare and normalize academic bodies through metrics. Under their eye, researchers become what Foucault referred to as "docile bodies": self-regulating agents who match behavior to indicator floors and ceilings, transforming their intellectual labor into outputs amenable to spreadsheets rather than principled inquiry. "Institutions and disciplines wail the ranking's obsession with the number of publications to the detriment of the very teaching-learning mission they report to" (LinkedIn comment, 9 Jul 2025). The effect is a drift from the epistemic curiosity to the structural necessity of metric congruence, which prioritizes quantitative visibility over intellectual honesty and cultural depth.

Academic Moonlighting and Monetization of Authorship

The now-infamous LinkedIn solicitation exposes a larger market logic in which academic authorship itself functions like a task on an "intellectual Uber", a gig economy of scholarship where visibility can be bought, sold, or rented by the click. Far from an isolated lapse, the post belongs to a mature ecosystem of monetized shortcuts that operates through several well-defined channels:

- *Paper Mills.* These unscrupulous entities (i.e., tutoring services, editing agencies, and infamous paper mills) produce plug-and-

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

play manuscripts with falsified data sets and tailored figures—and even pre-printed journal titles in exchange for payment. For instance, a recent analysis of publishers' databases uncovered thousands of these articles, many featuring nearly identical peer reviews and cut-and-paste graphs, thereby raising doubts about the authenticity of academic publishing (Else & Van Noorden, 2021).

- *Citation Cartels.* In the dark underbelly of academia, researchers gather in Slack groups, private Telegram channels, and exclusive guest editorships in high-impact special issues to build "citation rings." They use underground networks to help each other recycle articles, the recycling in the form of a reference list that bumps up impact metrics while not providing meaningful scholarship engagement (Fong & Wilhite, 2017). These activities pervert the actual worth of academic labor and erode the very foundations of rigorous scholarship.
- *Authorship Trading.* This is an incredibly dangerous trend, but it appears that authorships can be traded similar to political favors or securitizations of "visibility swaps." Hosseini *et al.* (2018) reveal rate cards, where the price of bylines is established: first author credits are high, acknowledgments are lower, and a ghostwriting service token is generated for one ghostwritten article. Authorship has emerged as a commodity—one that devalues research by introducing broad for-profit interests into the heart of the academic intellectual consortium.

The practical ingenuity of this marketplace is captured by quantitative-methods lecturer Pratik B., who noted during the LinkedIn thread that scholars are "hijacking Google Scholar suggestions from namesakes to pad their citation count, while others buy citations at 1,000 apiece like it's an online sale" (personal communication, 9 July 2025). From a Marxian viewpoint, these practices exemplify the alienation of intellectual labor: the scholarly object remains separated from its producer and returns into circulation as a fungible exchange value, evaluable by rankings, impact factors, and tuition money. The writer's trade, long associated with inquiry and evidence, is fit to be traded in data markets.

Global Examples of Citation Cartels and Paper Mills

As mentioned, the integrity problem is not geographically contained; it is structurally global. In 2023, Clarivate Analytics[2] removed 82 titles from

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

the Web of Science after forensic analysts uncovered "anomalous citation patterns"—a polite euphemism for networks of articles whose references were engineered to pump impact-factor scores (Oransky, 2023). Similar irregularities keep surfacing in the case files of the Committee on Publication Ethics (COPE), whose incident reports read like a catalogue of industrial-scale deception:

- *Chinese paper-mill pipelines*. Text mining by independent sleuths traced more than 400 fabricated articles to commercial ghostwriting shops in China. All were published between 2019 and 2021 in Elsevier journals, complete with bogus e-mail reviewers and copy-pasted microscopy images (Schneider, 2021).
- *Hindawi mass retractions*. In 2023 the open-access publisher Hindawi pulled more than 300 papers after auditors found template peer-review reports, recycled figures, and author lists that changed between acceptance and print—a pattern the company itself labelled "systematic manipulation of the publication process" (Retraction Watch, 2023).
- *Brazilian citation rings*. As early as 2013, bibliometric mapping exposed clusters of Brazilian researchers who funneled hundreds of reciprocal references through special issues and conference proceedings to vault their journals up national ranking tables (Van Noorden, 2013).

Taken together, these events reveal an academic variant of what Zuboff (2019) calls "surveillance capitalism". Metrics generated from citations, downloads, and alt-metric signals serve as the primary resources for value extraction, while the knowledge these metrics are meant to reflect fades into the background. Journals pursue impact factors to secure APC;[3] revenue streams; universities harvest citation dashboards to climb global ranking lists; individual scholars treat authorship slots as currency. In this economic circuitry, data manipulation—not discovery—becomes the chief mode of production, further alienating intellectual labor from its epistemic purpose and entrenching a marketplace where attention and credibility can be algorithmically manufactured.

Moral Delegation to Machines

This section charts the reconfiguration of authorship and responsibility in scholarly practice under generative AI. Beginning with a fo-

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

cused treatment of "AI as Author," it interrogates how LLM produce citation-dense texts that confound traditional markers of originality, contribution, and proprietorship. It proceeds to document emergent vectors of academic fraud—deepfakes, synthetic corpora, and fabricated identities—and argues that these phenomena precipitate a verification crisis, reallocating labor from knowledge production toward authentication. Drawing on Bruno Latour's (2005) "distributed agency", the analysis then disaggregates responsibility among students, platform developers, institutions, and evaluative infrastructures, represented via a network map of an AI-generated manuscript. The section ends by foregrounding the normative stakes: when moral reasoning is folded into optimization routines, rightness collapses into score-keeping. It then sketches counter-measures—mandatory provenance trails and tamper-evident authorship verification—that could reinstate accountability inside the algorithmic loop.

AI as Author in ChatGPT

The arrival of LLM—ChatGPT, Claude, Gemini, etc., —and their rapidly proliferating cousins—has redrawn the cartography of scholarly authorship. Having learned from terabytes of digital text, these systems can churn out a slick citation-laden prose that mimics the writing style, argumentative approaches, and disciplinary conventions of peer-reviewed literature. One prompt might produce an abstract (with namedropping from Derrida on Latour), or a literature review laced with imagined DOIs, or a results section adorned with fabricated tables. But this feat of technology comes at an epistemic cost: It destabilizes categories of originality and contribution on which academic credit has traditionally relied (Lund *et al*., 2023).

The resulting condition might be termed *authorship ambiguity*. Whereas plagiarism once involved a clear human act of appropriation, we now face outputs whose provenance is algorithmic, collective, and probabilistic. Did the doctoral candidate merely polish the AI draft, or did the model supply the conceptual core of the argument? Who— if anyone— owns sentences statistically assembled from millions of copyrighted sources? Early empirical work suggests the practice is no longer marginal: Several Scientific, Technical and Medical (STM)[4] publishers report that between 15% and 20% of new submissions exhibit textual fingerprints consistent with LLM: unusual lexical bursts, compression artifacts in references, or stylistic homogenization across multi-author manuscripts.

These developments sharpen the distinction proposed by Floridi *et al.* (2018) between "the ethics of AI" (how we, as moral agents, design

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

and deploy such systems) and "AI ethics" (the misguided expectation that machines can shoulder normative reasoning for us). When ChatGPT drafts a discussion section, the ethical responsibility to verify sources, check logical coherence, and declare contributions still belongs to the human researcher; delegating those tasks to the model confuses computational fluency with moral agency. Yet the speed and fluency of LLMs invite precisely that abdication, tempting scholars to treat ethical deliberation as another routine function to be outsourced. Unless academic fields develop clear standards about provenance, recommendations of disclosure thresholds, and pedagogical recommendations, the line between intellectual hand-craftedness and automated imitation will erode further, creating a marketplace for ideas where accountability itself might be just as manufactured as the prose on the page.

Deepfakes and Citation Hijacking

Text-based ghostwriting is only the opening act. A new suite of generative tools now makes it possible to counterfeit virtually every element of the research record, transforming fraud from a labor-intensive craft into a point-and-click service. The most salient developments include:

- *Deepfake conference talks and lectures.* With little more than a head-and-shoulders photograph and a voice sample scraped from a webinar, commercial "avatar studios" can synthesize high-definition videos in which a scholar appears to outline experimental protocols, walk through PowerPoint slides, and field audience questions—all for a paper that never existed. Lip-sync, eye-tracking, and ambient-noise engines make the footage difficult to distinguish from a genuine keynote, allowing fraudulent presenters to impress hiring committees or secure speaking-fee honoraria without leaving their desks.
- *Algorithmic data factories.* Specialized GANs (generative adversarial networks) and diffusion models can now produce datasets—MRI scans, microscopy images, genomic sequences, or social-survey spreadsheets—that satisfy expected statistical distributions and pass most plausibility checks. A researcher can prompt, "Generate a CSV[5] of 10,000 respondents showing a significant positive correlation between mindfulness training and GPA ($p < .01$)," and receive perfectly formatted numbers that will survive routine peer review unless raw-data audits or code replication are required.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

- *Synthetic scholarly identities.* Profile-generation services stitch together AI portraits, fabricated co-author networks, and curated publication lists populated with either real DOIs (cleverly faked ones) that redirect to paywalled dead ends. These phantoms register ORCID iDs, open Google Scholar accounts, and apply for editorial-board seats, thereby injecting nonexistent experts into the citation economy. Once embedded, they can be called upon to boost the metrics of paying customers, or to generate positive peer-review reports on-demand.

Combined, these capabilities are the simulacrum of Baudrillard (1994): a copy encompassed with no relationship to an original that circulates as though it were real, replacing reality with a hyperreality that only refers to itself. In the academic sphere, the result is an epistemic verification crisis. Traditional gatekeeping—peer review, conference Q&A, institutional affiliation checks—was designed for malicious actors who might embellish, not for algorithms that can fabricate entire worlds at scale. When a compelling deepfake lecture, a statistically impeccable dataset, and a scholar with an impressive h-index can all be conjured *ex nihilo*, the ontological basis of scholarship—its anchoring in observable procedures, accountable individuals, and reproducible evidence—begins to dissolve. The central scholarly task thus shifts from producing knowledge to authenticating reality, a reversal that threatens to divert intellectual energy away from discovery and into an endless game of forensic catch-up.

Who is responsible? The Problem of Distributed Agency

Handing off ever-larger portions of reading, synthesizing, and writing to generative algorithms reconfigures the moral map of scholarship in precisely the way Latour (2005) foresaw when he spoke of "distributed agency." The final act—uploading a term paper to the course portal—now rests on a sociotechnical relay in which multiple actors leave fingerprints:

- The student who frames the prompt and decides to hit "submit".
- The engineering team whose transformer architecture translates that prompt into prose.
- The venture-capital investors who press for frictionless user growth.
- The university that provides no clear guidance beyond plagiarism boilerplate.

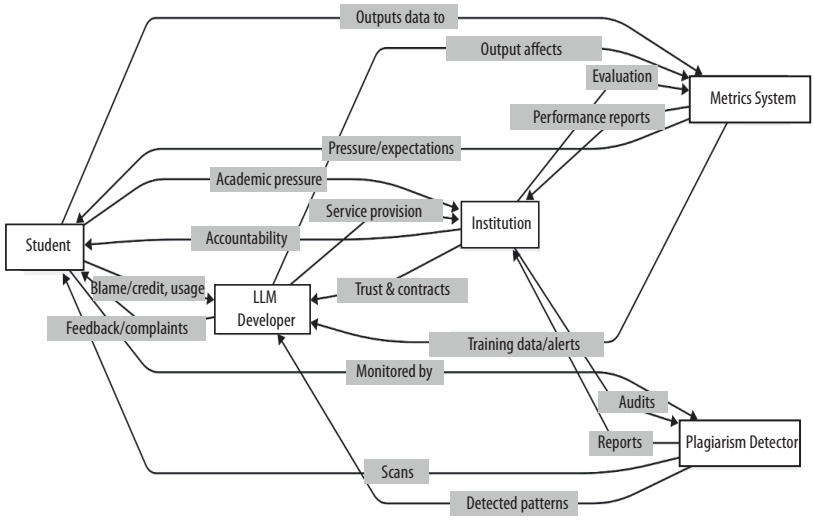
Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

- The global ranking systems that reward throughput and citation visibility over reflective learning.

When the essay is later flagged as "AI-assisted," culpability scatters across this network. Is the student a cheat or merely an opportunistic user of a campus-licensed tool? Are developers liable for releasing a product that can outwrite undergraduates? Should the lecturer have required oral defenses to verify authorship? No single node can credibly lay full claim to the outcome, demonstrating the Latourian observation (Latour, 2005) that action belongs to an *assemblage* and never to a singular subject.

This dispersion creates the moral atmosphere which Bauman (2001) would call "liquid modernity." As it enters the brittle ecosystem of update cycles, policy memos, and dashboard metrics, norms that once seemed solid—authorship, originality, academic integrity—melt into fluid streams. Anticipating the outcome perfectly, behavioral ecologist M. Sharma captured its essence in a much-commented revulsion on LinkedIn: "All is fair in love, war and academic ratings" (personal communication, 9 July 2025). In such liquidity, responsibility behaves like mercury on a tabletop—every time an institution tries to pin it down, it breaks into smaller beads that roll elsewhere, leaving no single actor readily accountable.

Figure 3
Map of agency distributed in an AI-generated job



Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

The Axiological Gap and Vanishing Values in Digital Labor

This section situates a crisis of authenticity in academic production and shows how digital tools erode values across three layers: epistemic authenticity (where metric incentives displace inquiry), authentic relations (where co-authorship and collaboration become transactional), and institutional authenticity (where rankings rewire the university's mission). It then tracks the decline of responsibility and justice: drawing on Levinas to frame responsibility to the Other, and on Fraser to show how "abnormal justice" emerges when algorithmic management masks power and reproduces inequality. Finally, it analyzes moral disengagement as a systemic feature of the digitized academy, using Bandura's mechanisms (moral justification, euphemistic labeling, advantageous comparison, displacement, minimization), a comparison table to align classical and digital-academic forms, and Zuboff's instrumentarian power to explain how platform design and real-time metrics normalize these violations.

Authenticity in Academic Production

The rapid digitization of scholarly life has thrust universities into what Taylor (1991) would describe as a full-blown "crisis of the ethics of authenticity." Taylor's concept hinges on the ideal that individuals and communities should act in ways that express their true convictions and values. Digital tools that can fabricate prose, inflate metrics, or monetize prestige now make that ideal difficult to realize—or even to recognize. The result is a multi-layered erosion of authenticity that affects every plane of academic practice:

- *Epistemic authenticity. The integrity of knowledge-making itself:*
  a. Ideal: Research is pursued for the intrinsic purpose of understanding the world, guided by curiosity, rigor, and methodological transparency.
  b. Erosion: Bibliometric dashboards, tenure clocks, and algorithmic recommendation systems reward volume as well as sheer number of citations. They are teaching a generation of scholars how to port one study into multiple "papers," or cite friends in citation rings, or outsource the literature review to a generic large language model. The craving for metrics

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

displaces the craving for the truth; the paper is turned into an instrument of strategy rather than evidence of inquiry.

- *Authentic relations. The nature of academic relations:*
  a. Ideal: Partnership is based on shared intellectual respect and common cause, with credit given where the real contribution has been made.
  b. Erosion: Co-authorship now functions more and more like a transaction—the adding of authors for their branding, their deletion to avoid a conflict of interest, and finally, expressing fair reciprocity by inclusion in a co-authored writing orgy to accrue citations. Communication is limited to template emails and automatically generated project boards, where one once found the deep, dialogic engagement that used to characterize mentorship and joint discovery.
- *Institutional authenticity. The mission and self-conception of the university:*
  a. Ideal: Universities act as guardians of critical thought and public enlightenment, structuring incentives around teaching excellence, open debate, and socially relevant research.
  b. Erosion: Those rankings convert that mission into a numbers game in which grant income, high-impact publications, and social-media reach dominate. Administrative offices establish "ranking task forces," funnel funds toward quick-win research areas, and pressure faculty to publish in any venue that boosts institutional metrics, even if the scholarship is tangential to local needs or student learning.

Across these three layers the same dynamic repeats: digital affordances that promise efficiency and reach simultaneously undermine the conditions for sincere expression and honest contribution. What looks like productivity on a spreadsheet often masks intellectual shortcuts; what looks like collaboration on a CV can conceal transactional alliances; what looks like institutional success in a ranking table may signal a drift away from education in the public interest. In Taylor's terms (Taylor, 1991), the academy risks *trading authenticity*—the alignment of actions with deeper moral purpose—for a shallow, performative version of itself optimized for digital visibility.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

The Decline of Responsibility and Justice

Lévinas (1961) frames ethics not as a set of rules that follow from being, but as an "infinite responsibility" toward the *Other* that calls to us even before we formulate concepts of self, knowledge, or community. The moment that face-to-face relation is re-routed through screens, dashboards, and automated recommendation engines, its moral intensity thins out. Within today's digital academy—email negotiations, manuscript-tracking portals, LinkedIn talent hunts—the *Other* appears less as a concrete person and more as a data point or a profile to be leveraged. The recent LinkedIn episode, in which departments openly bartered for co-affiliation slots to boost their ranking scores, lays bare that erosion. Institutional labels were swapped like tradeable assets, with little reflection on how such transactions affect the real students, staff, and local publics whose identities those labels are meant to signify.

Justice inside the university has traditionally meant three intertwined commitments: (i) Accurate recognition of who actually contributed intellectual labor; (ii) fair distribution of resources such as grants, mentoring, and publication space; (iii) vigilant protection of the evidentiary record on which further scholarship depends.

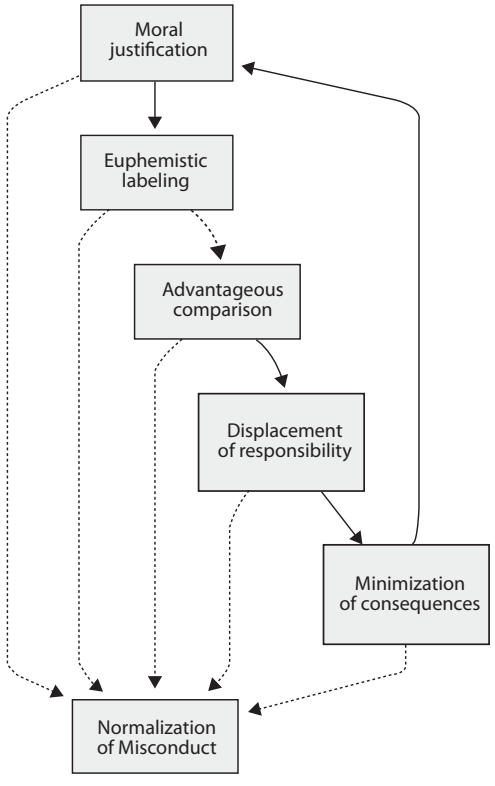
Algorithmic management tools—citation indices, funding dashboards, hiring filters—now promise to safeguard these ideals through objective measurement. Nevertheless, as Fraser (2008) warns in her discussion of "abnormal justice," systems that look impartial from the outside can smuggle in distorted decision rules and opaque power relations. Ranking algorithms amplify advantages enjoyed by well-funded institutions; predictive analytics prioritize grant applicants whose past success already signals privilege; plagiarism detectors flag the prose of ESL write more often than that of native speakers. The decision pathway hides behind code or proprietary metrics, causing new hierarchies to materialize precisely at the moment they become harder to see. In effect, the machinery that was supposed to democratize recognition, distribution, and integrity installs a fresh regime of inequality while masking the tracks of injustice under a veneer of computational neutrality.

Moral Disengagement as a Systemic Feature

Bandura's later work on moral disengagement (Bandura, 2016) describes the cognitive levers people pull when they want to behave unethically without suffering self-reproach.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

Figure 4
The moral disengagement loop in academic publishing



The digitized academy has woven these levers into everyday routines, institutional policies, and even the user interface design of manuscript systems. Five mechanisms are particularly visible:

Moral justification

- Classic form: "I cheated because the stakes were high and my family depends on me."
- Digital-academic variant: "Every lab in our field pads citations or pays paper mills—if we don't, we won't make the next ranking cutoff."

Listen for crisis language and claims of inevitability. Require precommitted standards before review cycles so reasons cannot be retrofitted later. Use cohort-level audits to compare behavior against the field

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

median, not against rumors. Pair amnesty windows with forward-looking fixes, for example retract and disclose, then institute caps on countable outputs. Teach the "least permissible shortcut" test: would you endorse this reason if your name were public and the rule applied to rivals as well.

A single glance at league-table metrics or funding-agency success rates is enough to frame questionable practices as collective self-defense rather than wrongdoing.

Euphemistic labelling

- Classic form: "collateral damage" instead of "civilian deaths."
- Digital-academic variant: An example of a digital-academic variant includes phrases like "strategic collaboration," "guest authorship," or "impact-factor optimization," which are used instead of the more straightforward term "authorship fraud."

Build a plain-speech glossary that maps euphemisms to what happened. Add form fields that force concrete actions, for example "contributed methods section" rather than "collaborator." Run a jargon lint tool on policies and call-for-papers. During reviews, ask for evidence that matches the verb, such as commit history, analysis notes, or correspondence. Penalize mislabeling even when the underlying work passes, since language shapes repeat behavior.

Journal submission portals and consultancy brochures normalize these phrases, allowing actors to hide moral stains behind managerial jargon.

Advantageous comparison

- Classic form: "Yes, I lied, but I did not steal."
- Digital-academic edition: "We are just tacking on honorary co-authors; the true villains are those who cook up entire data sets or resort to deep-fake images."

Evaluate actions against a fixed norm, not against worse examples. Use sanction grids that scale with impact and intent so minor harms are still addressed. In class, run "ladder of violations" drills where each rung is judged on its own merits. Watch for moral licensing after small good deeds. Require a forward remedy, for example remove unearned authorship and correct records, rather than settling for comparisons.

By juxtaposing more minor violations of professional norms with high-profile cases of data fabrication, researchers shield themselves from the brunt of ethical critique.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

Displacement of responsibility

- Classic form: "I was just following orders".
- Digital-academic variant: "The tenure system, grant panels, and ranking exercises make it so it will be published this way; we have little choice in the matter."

Map who decides what with a simple RACI table (responsible, accountable, consulted, informed). Insert human checkpoints where numbers are advisory, not binding. Require signed authorship and methods attestations from named individuals. Publish the decision trail for promotions and special issues so reasons can be inspected. Reform incentives, but do not let structure erase agency. If a rule made the outcome likely, change the rule and still address the choice that was made.

Responsibility is shuffled onto faceless structures: algorithmic dashboards, KPI scorecards, and automated citation trackers. Individuals understand that they are merely participating in a game.

Distortion or minimization of consequences

- Classic form: "It was only a harmless prank."
- Digital-academic variant: "Adding one extra name to the byline doesn't hurt anyone; it's just a line of text in a PDF."

Quantify downstream effects. Show how a single false credit alters h-index curves, grant scores, and shortlists. Use impact statements in retraction or correction notices that trace who lost opportunities. Pair minimization with case reflections from affected peers. Require remedial steps, for example corrected metadata, letters to committees, and redistribution of awards where possible. Make consequence tracking part of annual reports, not a one-off apology.

The tangible harms—misallocated credit, skewed hiring decisions, inflated institutional prestige—are rendered invisible behind a screen of abstraction.

What makes these mechanisms especially potent today is that they are no longer confined to private rationalizations inside a researcher's head. First, they are codified in institutional policies (e.g., revenue-sharing models that reward journals purely by article volume); secondly, they are baked into platform architecture (e.g., "suggested citations" that nudge authors toward reciprocity networks); and finally they are reinforced by metric dashboards that refresh in real time, turning moral shortcuts into instantly visible performance enhancements.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

Table 1
Comparison of Classical vs. Digital-Academic Moral Violations

Mechanism	Classical Form	Digital-Academic Form
Moral justification	"I cheated for my family."	"Everyone buys citations to stay competitive."
Euphemistic labeling	"Collateral damage"	"Impact optimization" instead of authorship fraud
Advantageous comparison	"At least I didn't steal."	"Not as bad as fabricating data."
Displacement	"I was just following orders."	"Ranking systems leave no other choice."
Minimization	"It was just a prank."	"One extra name doesn't hurt anyone."

Source: Own elaboration.

This would be an example of what Zuboff (2019) has called "instrumentarian power": power that operates not by explicit command and control, but by nudging the shaping of behavior through technological infrastructures backed by data feedback loops. Scholars alter their conduct—often preconsciously—to match what the platform rewards, while the platforms themselves remain largely outside the field of ethical visibility. The result is a self-reinforcing ecosystem where moral disengagement is no longer an individual psychological trick; it is the ambient operating system of digital scholarship.

Teaching in a Ghosted World

This section diagnoses how dubious research practices erode faculty role-modeling and produce a hidden curriculum that prizes metric manipulation over integrity. It then presents a pedagogy that fuses narrative imagination with algorithmic literacy through three classroom methods—case analysis, multi-perspective role play, and "digital archaeology." The paper proposes a contemporary virtue-ethics toolkit (digital prudence, algorithmic temperance, cyber-courage, techno-justice) and lists concrete habits to instantiate those virtues (studio-style learning, rotating stewardship roles, structured reflective debriefs, and institutional safeguards). The goal is to help students and faculty see how systems sha-

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

pe behavior and to develop enduring practices that center inquiry, fairness, and accountability.

The Crisis of Role Models

When dubious research practices become routine, universities face an acute pedagogical paradox: how can instructors nurture ethical judgment in students while simultaneously condoning—or quietly benefiting from—the very misconduct they are supposed to condemn? The dilemma was captured with disarming candor by professor B.K Panigrahy in reference to the Pune faculty member who brokers "affiliate positions" for ranking points: "He is an honest economics professor, driving the economy of his university under instructions from top management" (personal communication, 9 July 2025). In other words, dishonest behavior is recast as institutional loyalty; moral transgression is applauded as strategic compliance.

Such cognitive contortions leave educators in a fog of normative uncertainty. Lecture hall homilies about plagiarism and scholarly virtue ring hollow when students can observe their mentors swapping authorship slots, massaging citations, or submitting AI-written abstracts to meet quarterly KPI targets. The classroom becomes a site of ethical doublethink: the syllabus prescribes integrity, while departmental practice rewards metric gaming.

Ortega and Gasset foresaw a similar danger nearly a century ago. In *The Revolt of the Masses* (1930), he deplored the "barbarism of specialization," a condition in which professionals achieve razor-sharp technical proficiency yet remain ethically and culturally anemic. Today's academic exemplars often fit the description: they master h-index optimization, open-access fee waivers, and social media amplification strategies but struggle to embody the intellectual honesty that higher education claims to cultivate.

The upshot is a generation of faculty who:

- Train doctoral candidates to slice a single dataset into multiple "minimum publishable units,"
- Regularly accept or offer "honorary" co-authorships to curry favor.
- Treat algorithmic similarity scores as a box-ticking substitute for genuine citation ethics.
- Gublicly endorse research-integrity policies while privately teaching students how to bypass them.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

In such an environment, the subtle message that students learn is the one provided by the hidden curriculum: Success is not a result of exploring rigorously a question or a mystery, nor of reporting objectively what one finds, but of writing in ways that make the loopholes easy to find. The moral imagination that universities exist to develop effectively shrinks, limited to a strategic calculus of which rules can be broken without the risk of life-altering social-media shaming

Developing Critical Moral Imagination

Nussbaum (1997) maintains that democratic citizenship hinges on "narrative imagination"—the cultivated habit of picturing life from vantage points other than one's own and of empathizing with hopes, fears, and constraints that differ from our own biographies. In a university system increasingly mediated by code, dashboards, and recommender engines, the traditional narrative lens is no longer sufficient. Students and faculty must also acquire what we might call an algorithmic imagination: an intellectual capacity to visualize how technological infrastructures allocate visibility, distribute rewards, and encode bias—and to grasp the human consequences that flow from those hidden design choices. Where narrative imagination inquires, "What would it feel like to be that person?" we can say, "What goods does the system make possible or impossible for that person?" and algorithmic imagination adds, "What realities does the system make possible or impossible for that person?"

This double imagination can be animated through three pedagogical strategies:

Case-based analysis

- *Method:* professors present a case with published incidents, like the Pune affiliation-trading scandal or mass retractions from image-manipulation rings, and ask students to map participants, motivations, and decision points.
- Learning outcome: By dissecting real-life cases, students come to see in action where general ideas (honesty, justice, and responsibility) clash with ranking tables, publication quotas, and venture-funded platforms. The exercise roots moral theory in the specific institutional ecologies.

Use a timeline with evidence tags so students distinguish allegation from proof. Require a "fork in the road" memo that outlines two plausible

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

actions at each decision point and the risks to different stakeholders. Add a counterfactual: What if the ranking rule or journal policy were different. This draws a line from policy to behavior. Score the post-mortem on how cleanly it maps cause to effect—clear boxes, straight arrows, testable links. Deduct for theatrics: if the slide ends with one villain's face in red ink, you've missed the system and mistaken the assignment. Close with a policy brief that names one structural fix and its unintended effects.

Multi-perspective role-play

- Method: teams are assigned stakeholder roles (graduate student, journal editor, university administrator, AI start-up executive, whistleblower) in seminars. They navigate a conundrum, such as whether to retract a paper based on synthetic data.
- Learning outcome: role-play makes participants take on competing imperatives (career survival, fiduciary duty, public trust), which will hone both narrative and algorithmic imagination. Students identify how ethical results depend not only on goodwill but also on technical limitations and power asymmetries.

Hand every participant a sealed brief: private data, red-line constraints, and a mandate that guarantees clash. Half-way through, e-mail a disruptive "new regulation" or "funder ultimatum" that rewrites the game board. Final deliverable: one joint decision document and, stapled to it, a compulsory dissent paragraph that immortalizes the minority voice in the official file. Assess outcomes by the quality of justification and the transparency of the decision-making process, and penalize performative behavior that lacks clear accountability (named owners, deadlines, measurable commitments). Finish with a post-role audit: participants step out of character and rate their team's decision for fairness and feasibility.

Digital archaeology

- Method: students "source dig" into the simple tools of everyday academic work: They reverse engineer the weighting of a plagiarism detector, track the venture capital lineage of a citation index company, and map the geographic bias of a search algorithm.
- Learning outcome: this inquiry uncovers the sedimented histories, economic interests, and political choices written into ostensibly neutral software. It helps students learn to question default settings and to recognize hidden feedback loops so that

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

they can avoid constructing things that lead to exploitation and inequality.

Pair every finding with a reproducible test. For example, run multilingual texts through the detector and report false positive rates by register. Build a "tool dossier" that includes ownership, data sources, known issues, and user costs. Require an exit strategy: a configuration change, a mitigation protocol, or a viable alternative, with trade-offs stated plainly. Evaluate on evidence quality and practicality. The risk is cynicism. Counter it by measuring improvements after applying the mitigation and reflecting on what remained unsolved.

Together, these practices move ethical education beyond rule memorization toward an embodied literacy in how stories and systems coproduce academic life. By pairing Nussbaum's narrative imagination with a robust algorithmic imagination, universities can prepare scholars who not only empathize across human differences but also scrutinize and reshape the technological architectures that increasingly govern those differences.

Virtue Ethics for the Algorithmic Age

MacIntyre's *After Virtue* (1981) laments the collapse of shared moral narratives and proposes a return to an Aristotelian view in which stable communities cultivate character-forming habits. Translating that program into a digitized university means naming the particular excellences a scholar now needs and then creating routines that let those excellences take root. Four stand out:

Digital prudence (phronesis)

- Definition: the practical wisdom to sense how learning-management systems, citation dashboards, or generative AI will affect the integrity of one's work and relationships.
- Example: Choosing to slow down an assignment cycle when a new large-language model update makes "push-button essays" temptingly simple, or deciding to publish code and data so that automation does not obscure reproducibility.

Prudence is situational, so teach it as timed checkpoints: pre-design (anticipate tool effects), mid-semester (audit unintended shortcuts), and post-mortem (record what to change next term). Pair every tech adoption with a "guardrail" (e.g., oral defenses, code notebooks, provenance statements). Evaluate prudence by evidence of foresight (risk logs,

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

contingency plans), not just outcomes. The failure mode is overreaction (blanket bans). Calibrate instead: allow assistive use with declared boundaries and verification steps.

Algorithmic temperance

- Definition: a disciplined wariness toward the seductions of numbers (impact factors, Altmetric donuts, h-index milestones) that can hijack scholarly motives.
- Example: Setting departmental caps on the number of papers counted toward annual reviews or deliberately valuing mentoring and community outreach alongside citation counts.

Temperance needs structural levers, not pep talks. Convert values into rubrics: set ceiling weights for any single metric; require a "narrative of contribution" and a "replication/negative-results" credit. Run sensitivity tests on your own promotion criteria to show how small weight shifts distort behavior. Track a "shadow ledger" of uncounted labor (mentoring, dataset curation) and fold it into annual reviews. The pitfall is metric substitution (new badges for old). Keep any quantitative indicator explicitly provisional.

Cyber-courage

- Definition: the fortitude to call out questionable practices (paper mills, ghostwritten reviews, fake-author bots) even when such critique threatens rankings, grants, or collegial goodwill.
- Example: junior researchers signing a collective letter requesting an investigation into a senior colleague's suspicious images, or an editor refusing fast-track submissions that bypass proper peer review.

Courage scales with protection. Build safe channels (confidential research-integrity ombud, anti-retaliation policy, legal counsel briefings). Normalize evidence-first reporting templates (image checks, data PRISMA, timeline). Run "courage drills": give juniors a scrubbed dataset with a planted ethical trip-wire and five minutes to speak up while seniors answer, "Good catch—what do we do next?" Put "raised the alarm" in performance reviews, but pair it with clear due-process lanes so no one plays Twitter cop. When protection is codified, dissent stops feeling heroic and starts feeling normal.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

Techno-justice

- Definition: a standing commitment to deploy digital tools in ways that widen, rather than narrow, access and recognition.
- Example: using open-source software and preprint servers to level resource disparities or auditing plagiarism-detection algorithms for bias against non-native English writers.

Treat tools as policy decisions. Run equity impact assessments before adoption (costs, bandwidth, language effects, accessibility). Maintain an "inclusion budget": dedicate funds/time to translation, low-resource access, and assistive tech. Demand vendors open the hood, release full training provenance and subgroup error rates or lose their license to sell. Sunset any model that flunks a bias audit. Gauge justice by who gains and who gets sidelined, tracked through distributional dashboards and lived-account videos, not by how many clicks the dashboard collects. The risk is performative openness; tie openness to real stewardship—maintainers, documentation, and community governance.

MacIntyre insists these virtues cannot be memorized like policy clauses; they must be "habituated" through repeated action inside practices that have clear internal goods and communal standards. In concrete terms:

- *Studio pedagogy.* Embed students in research collectives where data transparency and code review are daily rituals, not special events.
- *Rotating stewardship roles.* Assign lab members to manage the shared Git repository, the pre-registration log, or the peer-feedback channel so that prudence and justice become muscle memory.
- Reflective debriefs. After each publication cycle, hold a postmortem asking which digital shortcuts were resisted, which were indulged, and how the group's character was shaped in the process.
- Institutional guardrails. Craft promotion criteria that reward the public release of negative results or replication data as much as splashy citation counts, reinforcing temperance and techno-justice at the structural level.

Habituation of this sort resurrects what MacIntyre calls a "practice-dependent" conception of virtue (MacIntyre, 1981): scholars learn to want the right things by repeatedly doing the right things within communities that honor those efforts. Only then can digital prudence, algorithmic temperance, cyber-courage, and techno-justice solidify into the shared moral grammar the modern academy too often lacks.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

Toward a New Pedagogy of Truth

This section outlines an ethics-infused digital-literacy curriculum that integrates technical competence with moral reflection. It proposes three interlocking strands—critical code studies, data-ethics workshops, and platform pedagogy—so students learn to read code as argument, foresee harms in data practices, and critique the commercial logics of ranking and recommendation systems. Rooted in Freirean critical pedagogy, the model turns dialogue and praxis into classroom practices that redesign evaluative metrics and challenge default platform configurations. Highlighting a new algorithmic reflection.

Integrating Ethics and Digital Literacy

Conventional digital-literacy courses rarely move beyond "how-to" tutorials—how to format a spreadsheet, verify a URL, or toggle a privacy setting. That skills-first approach leaves the deepest question unanswered: What kinds of human relations, power dynamics, and moral consequences are baked into the tools we are learning to master? We therefore propose an expanded, ethics-saturated model of digital literacy that treats technical proficiency and moral reflection as inseparable. The framework rests on three mutually reinforcing strands:

Critical code studies

- Purpose: to teach students to read code the way humanists read novels, looking for assumptions, exclusions, and value judgments.
- Practice: to annotate the source code of a plagiarism detector (or, if proprietary, its patent filing and white papers) to expose how similarity thresholds are selected, which linguistic registers are penalized, and whose cultural allusions are favored.
- Outcome: students come to understand that an algorithm is not a neutral machine but a polemical text, one that can—and in some cases should—be debated, revised, or rejected.

Begin with a single, auditable Lego brick—a ten-line tokenizer or cosine-similarity stub—then snap the rest of the castle around it; the first provable click reassures everyone that the bigger machine can also be opened up and counted. Require "diff diaries": students record how a one-line change (stop-word list, n-gram size) alters outcomes on multilingual samples. This exposes hidden bias quickly. Common failure: treating do-

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

cumentation as neutral. Make students compare docs to actual behavior with unit tests on code-switched and ESL prose. Score the work on what can be run and measured (unit tests, error diffs, performance traces) not on adjectives in the abstract. Equity check: feed the model text that never went near an Ivy-League press (Swahili medical forums, Quechua tweets, Tamil zines) so the benchmark is not just English in a three-piece suit.

Data-ethics workshops

- Purpose: to confront the moral ripple effects of data acquisition, cleaning, modeling, and sharing.
- Practice: conduct a mock Institutional Revision Board (IRB) review of a project that scrapes conference abstracts for trend analysis, requiring participants to identify potential harms, hidden biases, and uncontrollable downstream reuse.
- Outcome: students develop habits of anticipatory reflection, learning to ask, "Who benefits? Who is exposed to risk? Whose voice is missing?" before a single byte is collected.

Move past checklist IRBs. Stage a five-minute "harm rehearsal": each team member writes a one-sentence refusal from the viewpoint of the people most likely to be hurt ("Dear Professors, we decline because…"). Swapping lab coats for lived experience turns the checklist into a consent conversation. Require a "provenance manifest" with every dataset: origin, transformations, uncertainties, and access rules. The usual gap is post-release accountability; add a sunset review where teams revisit the dataset after four weeks and log any unintended uses. Grade on clarity of mitigation plans, not just identification of risks. Keep a bias budget: for every cleaning step, students must state which signals are lost and why that loss is acceptable.

Platform pedagogy

- Purpose: illuminate how learning-management systems, reference managers, recommendation engines, and social-media channels mold what counts as legitimate knowledge.
- Practice: map the revenue model of a citation database, examine how its ranking algorithm weights various fields, and simulate how small changes in weightings redistribute scholarly visibility.
- Outcome: Users become capable of repositioning themselves from passive consumers of platform outputs to active negotiators—or critics—of the epistemic terrain those platforms construct.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

Treat platforms as institutions with incentives. Have students sketch a simple objective function ("maximize time-on-platform × APC revenue") and test policy tweaks against it. A 12% tweak to "international collaboration" or "open-access premium" often flips leaderboards, making fragility tangible. Add an "exit and voice" exercise: propose either a configuration change (voice) or a migration path to an alternative tool (exit), including costs to under-resourced users. Pitfall: moralizing without alternatives. Require a working prototype or settings guide (e.g., disabling "related articles" that bias toward the same publishers). Evaluate on feasibility and user impact, not just critique.

Pedagogical Philosophy

This triad extends Freire (1970) critical pedagogy—rooted in dialogue, reflexivity, and transformative action—into digital space. Freire concept of conscientização (critical consciousness) is translated here into "critical digital pedagogy," where first the dialogue means not only speaking with peers but also "listening" to what code and data reveal about institutional imperatives; second the reflection is coupled with praxis: after analyzing a ranking algorithm, students might propose alternative metrics or draft an open letter to the vendor; and finally empowerment shifts from merely operating technology to shaping, contesting, and redesigning it in the service of democratic and equitable knowledge production.

By embedding ethical inquiry at every technical touchpoint, the revised curriculum cultivates graduates who can (i) debug both the software and the social biases it embeds, (ii) approach data sets with the same moral seriousness they would apply to human subjects, and (iii) recognize platforms as political actors rather than invisible infrastructure.

In short, critical digital pedagogy turns digital literacy from a checklist of competencies into a sustained practice of ethical, intellectual, and civic engagement with the technologies that increasingly govern academic life.

Algorithmic Reflexivity

The study argues that twenty-first-century scholarship demands a new form of self-awareness: "algorithmic reflexivity." Where traditional critical thinking questions arguments and evidence, algorithmic reflexivity asks students to interrogate the very computational currents that steer what they see, read, and value. It is the discipline of pausing mid-scroll to ask, "Why is this paper, this metric, this ad the one that appears before me, and

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

what is it silently teaching me to desire or ignore?" To cultivate that habit of mind, courses can weave three iterative exercises into the curriculum:

Algorithmic audits

- Task: each student keeps a two-week log of how search engine autocomplete, citation manager recommendations, or social media feeds shape the sources they consult. Then, they reverse-engineer the signals (language, geography, co-citation networks) that the system used to reach those suggestions.
- Goal: render visible the invisible nudges that route literature reviews, topic selection, and even the phrasing of research questions.

Logs usually reveal patterned drift: students converge on the same few "safe" papers, often from well-indexed venues, while non-Anglophone or field-adjacent work disappears. The audit also surfaces how phrasing a query in different registers changes what appears. Have students run the same query from a campus IP, a mobile network, and a VPN to see localization effects. Assessment should reward the clarity of inference, not the volume of screenshots. The pitfall is fatalism. Close with a design move: "What three settings or disclosures would make this feed fairer?" That turns critique into capability.

Metric deconstruction projects

- Task: work in small groups on something that can be measured or ranked in an academic context (journal impact factors, Altmetric badges, departmental league tables) and trace the data inputs, weighting models, and ownership/sponsorship model. They demonstrate how tiny changes to parameters redistribute prestige and resources across fields, regions, or demographic cohorts.
- Goal: demystify quantification; show that metrics are not natural facts but rather contestable artifacts, and encourage students to imagine fairer replacements.

Sensitivity tests are the turning point. A one-point weight change on "international collaboration" often reorders an entire table, which teaches students how fragile "excellence" can be. Require two deliverables: a short brief that states what the metric claims to measure and a counter-metric that aligns with a value the class names in advance, such as mentoring or data transparency. Common failure modes are treating ownership and incentives as afterthoughts. Make them central: who pro-

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

fits when this metric rises, and who pays. End with a public annotation of the metric's webpage to practice accountable critique.

Digital-ethics portfolios

- Task: students maintain a "living document" all semester to document ethical dilemmas encountered in their digital workflows: scraping (posting data without an obfuscation step), using generative AI for prose or copy, and citing from the recommendations of opaque algorithms. They append to each declaration the relevant norms they followed in choosing their solutions, which are inspired by existing ethical guidelines, along with their own verdict and a subsequent evaluation of the consequences.
- Goal: establish a longitudinal practice of reflection that mines fragmented insights for a personalized working code for responsible digital scholarly conduct.

Portfolios work when they are specific. Push students to paste prompts, commit hashes, and version notes, not summaries. The follow-up entry is crucial: "What changed after I chose X?" That temporal layer turns rule-following into judgment. Include at least one "reversal" exercise where the student revisits a decision after learning a new standard and writes what they would now do differently. Grading should privilege candor and course-connected reasoning over spotless behavior. The risk is performative confession. Counter scepticism by attaching bite-sized receipts: lock your search plan into a public registry, list every AI prompt in an appendix, or paste the consent email for that scraped dataset—small, checkable acts that turn transparency from slogan into footnote.

Conclusion

The uproar over the Pune-LinkedIn scheme is not a stray bit of gossip but a flare exposing a deeper shift in the academy's moral scaffolding, where digital avatars, phantom co-authors, and KPI algorithms now co-write the scripts by which universities, departments, and individual scholars live, and where the priority cannot be a perpetual game of catching cheaters but a comprehensive reimagining of ethical life in a data-saturated university. Read across the spectrum—from Arendt's warning about thoughtless conformity, through MacIntyre's account of fragmented virtue, to Zuboff's analysis of instrumentarian control (Zuboff, 2019), and one lesson re-

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

peats: durable reform must address both the machinery that shapes behavior and the character of those who operate it. The path forward requires institutional courage to break addiction to scoreboards and to re-anchor evaluation in quality, openness, replication, and public value; pedagogical innovation that embeds ethics in design studios, case simulations, and algorithmic audits; technological wisdom that meets AI and analytics with disciplined critique, transparency demands, and open alternatives; and moral imagination that dares administrators, instructors, students, and funders to picture universities where integrity overtakes efficiency. The weary aside "Why am I not surprised that this is from India?" mislocates a global pathology: metric worship, authorship peddling, and AI shortcuts are symptoms of a universal digital condition. Levinas's call for ethics as first philosophy reminds us that responsibility for the other must precede calculations of profit or speed (Levinas, 1961). Recentering on that principle is how we dispel today's digital ghosts and choose an academy where technology serves rather than supplants our shared moral life.

Notes

1 The National Institutional Ranking Framework (NIRF) is a methodology adopted by the Government of India to rank higher education institutions in the country based on various parameters, including teaching, learning, and resources; research and professional practices; outreach and inclusivity; and perception.
2 A global company known for providing trusted insights and analytics to accelerate the pace of innovation. It offers various services, including the Web of Science, Journal Citation Reports, and InCites.
3 Article processing charge (APC) is a fee that authors typically pay to publishers for the processing, editing, and publication services they provide for their academic articles, especially in open-access journals.
4 These are publishers that specialize in producing academic journals, books, and other content in the fields of science, technology, and medicine.
5 Comma-separated values (CSV) is a simple file format used to store tabular data (spreadsheets or databases). In a CSV file, each line represents a data record, and each field within that record is separated by a comma.

Bibliography

ARENDT, Hannah
1963 *Eichmann in Jerusalem: A report on the banality of evil*. Viking Press.
BANDURA, Albert
2016 *Moral disengagement: How people do harm and live with themselves*. Worth Publishers.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

BAUDRILLARD, Jean
1994 *Simulacra and simulation*. University of Michigan Press.
BAUMAN, Zygmunt
2000 *Liquid modernity*. Polity Press.
BRAIDOTTI, Rosi
2013 *The posthuman*. Polity Press.
ELSE, Holly & VAN NOORDEN, Richard
2021 The fight against fake-paper factories that churn out sham science. *Nature*, *591*(7851), 516-519. https://doi.org/10.1038/d41586-021-00733-5
FLORIDI, Luciano, COWLS, Josh, BELTRAMETTI, Monica, CHATILA, Raja, CHAZERAND, Patrice, DIGNUM, Virginia, LUETGE, Christoph, MADELIN, Robert, PAGALLO, Ugo, ROSSI, Francesca, SCHÄFER, Burkhard, VALCKE, Peggy & VAYENA, Effy
2018 AI4People-An ethical framework for a good AI society. *Minds and Machines*, *28*(4), 689-707. https://doi.org/10.1007/s11023-018-9482-5
FONG, Eric & WILHITE, Allen
2017 Authorship and citation manipulation in academic research. *PLoS One*, *12*(12), e0187394. https://doi.org/10.1371/journal.pone.0187394
FOUCAULT, Michel
1977 *Discipline and punish: The birth of the prison*. Vintage Books.
FRASER, Nancy
2008 *Scales of justice: Reimagining political space in a globalizing world*. Columbia University Press.
FREIRE, Paulo
1970 *Pedagogy of the oppressed*. Continuum.
HAN, Byung-Chul
2015 *The transparency society*. Stanford University Press.
HAYLES, Katherine
1999 *How we Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. University of Chicago Press.
HOSSEINI, Mohammad, HILHORST, Mirjam, DE BEAUFORT, Ineke & FANELLI, Daniele
2018 Doing the right thing: A qualitative investigation of retractions due to unintentional error. *Science and Engineering Ethics*, *24*, 189-206. https://doi.org/10.1007/s11948-017-9894-2
LATOUR, Bruno
2005 *Reassembling the social*. Oxford University Press.
LÉVINAS, Emmanuel
1961 *Totality and infinity*. Duquesne University Press.
LUND, Brady, WANG, Ting, MANNURU, N. R., NIE, Bin, SHIMRAY, S. & WANG, Zhe
2023 ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, *74*(5), 570-581. https://doi.org/10.1002/asi.24750
MACINTYRE, Alasdair
1981 *After virtue*. University of Notre Dame Press.
MERTON, Robert
1973 *The sociology of science: Theoretical and empirical investigations* (N. W. Storer, ed.). University of Chicago Press.

Sophia 40: 2026.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

MITTELSTADT, Brent Daniel, ALLO, Patrick, TADDEO, Mariarosaria, WACHTER, Sandra & FLORIDI, Luciano
2016 The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2). https://doi.org/10.1177/2053951716679679
MULLER, Jerry
2018 *The tyranny of metrics*. Princeton University Press.
NUSSBAUM, Martha
1997 *Cultivating humanity*. Harvard University Press.
ORANSKY, Ivan
2023 "Truly devastating": Four journals won't get new Impact Factors this year because of citation shenanigans. *Retraction Watch*. https://bit.ly/4iEutAD
ORTEGA Y GASSET, José
1930 *The revolt of the masses*. W. W. Norton.
PASQUALE, Frank
2015 *The black box society*. Harvard University Press.
RECIO SASTRE, Alejandro
2025 The Question about Ethics Teaching as Boundary of Artificial Intelligence. *Sophia, Colección de Filosofía de la Educación*, (39), 53-81. https://doi.org/10.17163/soph.n39.2025.01
RETRACTION WATCH
2023 *Hindawi retracts more than 300 papers*. https://bit.ly/4p6LjdL
SCHNEIDER, Leonid
2021 The full-service paper mill and its Chinese customers. *For Better Science*. https://bit.ly/4otcU86
TAYLOR, Charles
1991 *The ethics of authenticity*. Harvard University Press.
VAN NOORDEN, Richard
2013 Brazilian citation scheme outed. *Nature*, *500*(7464), 510-511. https://doi.org/10.1038/500510a
ZUBOFF, Shoshana
2019 *The age of surveillance capitalism*. Public Affairs.

Data Availability Statement

The social media comments referenced in this article were collected from a public LinkedIn thread posted on July 9, 2025. The original post has been deleted, but the author has retained screenshots of the discussion, which are available upon request for verification purposes.

Sophia 40: 2026.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.

Authorship declaration	
Author/s	**Contribution**
Amitabh Vikram Dwivedi	Conceptualization, data processing, etc.

Declaration on the use of Artificial Intelligence
The author Amitabh Vikram Dwivedi of the article entitled: "Digital ghosts, moral algorithms, and the challenge of teaching ethics in the posthuman era," STATES that the document was created with the support of Artificial Intelligence (AI) at a rate of 1%.

Date of receipt: July 13, 2025
Date of review: September 20, 2025
Date of approval: November 10, 2025
Date of publication: January 15, 2026

Sophia 40: 2026.

Print ISSN:1390-3861 / Electronic ISSN: 1390-8626, pp. 81-117.