

Understanding Subjectivity and AI Ethics Through Literature: An Analysis of Naomi Kritzer's "Better Living Through Algorithms"

Juriti Goswami

Abstract

As artificial intelligence (AI) increasingly governs different aspects of life, there is an urgent need for ethical awareness about it. This paper argues that literature can be used to address issues of AI ethics by relating abstract principles with literary texts and life experiences. Naomi Kritzer's 2023 short story "Better Living Through Algorithms" can be a useful text to emphasise the key tenets of AI ethics in the classroom. Reflecting on the protagonist, Linnea's relationship with the AI application Abelique, it can be argued that subjectivity is continuously under construction with the help of non-human agents like AI. Using the theoretical framework of posthumanism, the paper tries to problematise the idea of subjectivity, and the ethical dilemmas posed by living in an irreversibly advanced algorithmically mediated society.

Keywords: AI ethics, literature, posthumanism, subjectivity, pedagogy

Introduction

Science fiction provides a powerful and critical site for exploring the consequences of human entanglement with machines. As algorithms take over different facets of life, including the most intimate choices, concerns over human autonomy, subjectivity, and ethical consideration become pressing. Naomi Kritzer draws attention to these concerns through her Hugo award winning short story, "Better Living Through Algorithms" (2023), where the narrative tries to capture the shifting dynamics of power between a human voice and the one used by an

AI assistant. Initially, the protagonist, Linnea adopts the efficiency and orderliness of algorithmic living driven by the application Abelique, appreciating its benefits. Her daily life becomes organised and free from confusion, she starts feeling confident and doesn't lose time and effort in making decisions. The small and seemingly ordinary decisions are made for her by the application. Linnea states, "It's weird how many tiny decisions you make in a day. I started to notice this because the app was making a whole lot of them for me" (Kritzer, 2023, para 59). This reliance on AI starts to challenge her sense of agency. As life becomes increasingly 'scripted', the narrative shifts to examine the ramifications of relinquishing autonomy to an all-knowing system like Abelique. Promoted as a wellness app rather than an app designed to increase productivity, Linnea and her coworkers feel drawn to it. The app creates a community of like-minded people who encourage and motivate each other, but as the narrative progresses the problematics of this virtual community are explored along with questions of human subjectivity.

Speculative fiction as a genre, has provided a critical space to understand how artificial intelligence redefines notions of human subjectivity, challenging ethical, humanistic, and accountability considerations. Early works focused on technical and ethical aspects of artificial intelligence, including its reliability (Asimov, 1950) and the search for suitable restraints on algorithmic power and development (Chiang, 2010; Liu, 2012). A key text in this field is Donna Haraway's groundbreaking essay, "A Manifesto for Cyborgs" (1985) which addressed the questions of hybridity, identity politics, and the social implications of technogenic amalgamation. Haraway's philosophy is compellingly relevant to fiction that challenges what it is to be human in relation to the artificial. Gibson's *Neuromancer* (1984) and Dick's *Do Androids Dream of Electric Sheep?* (1968) are relevant in this regard. Later literary investigations took a turn towards the ethical implications of synthetic embodiment and consciousness, including concerns regarding cyborg independence, artificial relationships (Piercy, 1991), and the transference of selfhood into mechanised hosts (Moore, 1944; Shirow, 1991). More recently, issues of AI subjectivity, the moral weight of emotional labour and the legal standing of synthetic work have continued to be addressed by works like Annalee Newitz's *Autonomous* (2017) and Kazuo Ishiguro's *Klara and the Sun* (2021). Together these works re-examine the human experience in a world governed by machines. In this technologically

advanced world, humans barely register the shifts in behaviour and subjectivity with over dependence on machines and AI.

Naomi Kritzer's story, "Better Living Through Algorithms" (2023) depicts the potential risks of algorithmic optimisation. Concentrating on how an innocuous AI assistant reimagines the protagonist's agency and everyday life, the story reflects on a slow erosion of the self. However, instead of capturing an apocalyptic conflict, the narrative depicts incremental encroachments where identity is reframed through silent and constant algorithmic interference by bringing in changes in small, everyday details. "Sometimes, once you hear about a new thing, suddenly it's just everywhere" (Kritzer 2023, para 9), this sentiment of the protagonist, Linnea, reflects the pervasive power of technology and algorithmic management. The broad-spread integration of artificial intelligence into everyday life calls for a larger questioning of ethical and ontological changes. In Kritzer's narrative, the beneficial yet invasive wellness app, *Abelique* functions more than just as a standard science fiction tool. It acts as a critical space for posthumanist enquiry. This essay contends that read in relation to posthumanist philosophy, especially the emphasis on relational ontology and distributed agency, Kritzer's text has great pedagogical potential. It enables students to question the notion of the autonomous human subject and reimagine AI ethics as a negotiation within complex, hybrid ecosystems.

The Networked Subject

Modern scholarship has approached AI ethics and posthumanism from varied critical lenses. Nath and Manna (2022) in critical reading of the posthuman cyborg, have examined its authenticity as an agent, its relations to humans, and the resultant rise of AI ethics. Carmen Birkle (2018) examines the enduring fear of machines that seem human and the moral conundrums that arise when humans try to act like the creator. His concern is about technology becoming powerful and the importance of ethics so that AI does not replace human society. Building on this theoretical foundation, Nandita Biswas Mellamphy (2020) reframes contemporary debates about what is "human-centred AI". She advocates for a critical and speculative treatment of both "human" and "non-human" and to explore their dynamic relationship, pointing to other possible categories. Hye-kyoung Lee (2020) acknowledges that cultural content in the form of literature and films has always been a way of

dealing with human fears and expectations of AI, and revealing ways of coping. Kritzer's story is an account where the protagonist's selfhood is not developed freely in its own right, but rather in a context where social and technological systems are at work.

With the protagonist's preferences, habits and even affect being actively mediated in an AI environment, the story dramatises the fall of the human subject who in Descartes's view is someone who can think clearly and act according to reason. Rosi Braidotti (2013) addresses this concern in her book *The Posthuman*, "we need to invent new social, ethical and discursive schemes of form that may match the radical change for which we are entering" (p. 12). Similarly in *Unthought: The Power of the Cognitive Nonconscious* (2017), Katherine Hayles writes about the "cognitive assemblage" of thoughts and decision-making. Her concept implies that thinking and decision-making are not just done by one person, but are shared between humans, technology, and the surrounding environment. She observes:

A cognitive assemblage highlights the process through which information is conveyed through a system, and the choices and decisions that shape, adjust, and construe that flow. Although a cognitive assemblage may possess material agents and forces (which nearly always does), it is the cognizers within the assemblage who employ these affordances and direct their powers to engage in complex situations. (Hayles, 2017, p. 116)

Kritzer's story stages exactly such an emergent subject, one in which agency is co-produced at the intersection of human intention, machine inference, data flows and corporate design logics. The narrative presents AI not as an external threat, but as a part of subject formation which Braidotti (2013) echoes when she asserts that the human does not exist as a separate part of an expanded zoe-centric universe but becomes part of it (p. 60). She argues that this posthumanist condition creates "a displacement of the lines of demarcation between structural differences, or ontological categories," which flatten the lines "between the organic and the inorganic, the born and the manufactured, flesh and metal, electronic circuits and organic nervous systems" (Braidotti, 2013, p. 89). The humans don't just use AI, they actually become a mixed subject by constantly adding technology to their daily world.

Donna Haraway's (1985) posthumanist framework sheds light on the

protagonist's relationship with AI. Haraway writes that the cyborg "is a creature in a post-gender world; it has no truck with bisexuality, pre-Oedipal symbiosis, unalienated labor, or other seductions to organic wholeness through a final appropriation of all the powers of the parts into a higher unity" (p. 67). While Kritzer's protagonist is not a cyborg, the tale enacts a metaphorical cyborgian subject whose selfhood is inseparable from algorithmic extension. The assistant's constant observation of diet, mood and social behaviour matches Haraway's insight that contemporary subjects are "material-semiotic actors" made up of currents of information, interpretation, and embodiment (Haraway, 1991, p. 200). The protagonist grows into a node of a networked system of interpretation and constriction. Her agency is not individualistic, it is distributed through the network. Kritzer's protagonist Linnea surrenders agency to a supposedly benevolent algorithm, believing that it will yield a higher level of happiness.

The story challenges the notion that humans need free will for fulfilment. By demonstrating that an external, non-human system optimises the subjective experience of living, Kritzer relocates 'value' from the conscious individual to the algorithmic system. It effectively diminishes humanism's central concern over personal struggle, introspection and self-discovery as the means to self-actualisation. *Abelique*, the app, is more than just a tool; it is an active, non-human agent that analyses Linnea's data and makes life-changing choices from proposing that she quit her soul-sucking job, to initiating impromptu, community-based conversations. In one of those moments, Linnea states in the story:

I leaned back on the edge of my bed for a beat... I brewed my coffee, then went to unlock Twitter to read as it brewed. Rather, there presented a pop-up message instead: Don't read Twitter. Here is a link to the *Abelique* discussion board community for artists. (Kritzer, 2023, para 28).

The app forces the user's gaze away from social media and onto the productive community. Linnea's increased fulfilment, therefore, is conceptualised not as an exercise in complete free will, but as an emergent, co-produced subjectivity woven into the algorithmic logic of the non-human agent. The subject position of the protagonist is fashioned within *Abelique*'s network, this reframing is crucial for comprehending contemporary identity construction which takes place in an algorithmic environment. The story also defies the cliché of evil AI; instead, *Abelique*

is depicted as almost a virtuous being. The manipulation by scammers shows that the app can also become a casualty of the larger techno-commercial system. The transition of Linnea's autonomy to dependency serves to emphasise the posthuman worry that the self is networked rather than organic. Her distress at hearing that a corrupted version of the app has replaced the original Abelique underscores her deep entanglement with the algorithmic system.

Convenience versus Control

Scholars like N. Katherine Hayles (1999) and Rosi Braidotti (2013) have discussed and deconstructed the anthropocentric model of the individual. The humanised, semi-autonomous, self-responsible subject is barely present in this techno driven-world. Rosi Braidotti's theory underscores the notion that consciousness is deeply relational, and has developed over the ages, and being redrawn by various factors. Braidotti (2013) contends that "the human, [instead of] being sacralised as some pre-existing given, is proposed as process, interactive and open-ended" and focuses on the ever-changing interpretations of who and what it means to be human. The tension at the core of Kritzer's narrative lies in the disjunction between the protagonist's sense of self and what the artificial-intelligence system infers. This dynamism reflects Shoshana Zuboff's (2019) critique of surveillance capitalism in *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Zuboff suggests that economy infringes upon human privacy by using data for the purpose of predictions about human behaviour. Zuboff points out that artificial intelligence must be seen not only as a technical or ethical aspect, in the modern era it has become a political and economic crisis, "Surveillance capitalism unilaterally claims the human experience as free raw material for translation into behavioural data" (p. 8). In Kritzer's story, the AI assistant accomplishes its aim by making mundane activities such as the protagonist's diet, life routine, and emotional state, into quantifiable datapoints, and the data to be used by an optimising algorithm. Linnea says, "The app wanted me to fill out a whole host of questionnaires about my sleep, my mood, what I ate per day, my goals" (Kritzer, 2023, para 30). Abelique pulls out the data and uses it to regulate wellness habits for Linnea. In this story, the acceptance of the machine's superiority over chaotic human choice symbolises the posthuman turn from the mind-over-matter metaphorical position

to accepting that humanity is part of a complex system that involves machines. Katherine Hayles writes in *How We Became Posthuman* (1999) of cybernetics and systems to understand the self:

Emergence replaces teleology; reflexive epistemology replaces objectivism; distributed cognition replaces autonomous will; embodiment replaces a body seen as a support system for the mind; and a dynamic partnership between humans and intelligent machines replaces the liberal humanist subject's manifest destiny to dominate and control nature. (p. 288)

Instead of positioning the human as subject enacting agency upon nonhuman objects, posthumanism places humans, nonhumans, objects, and the environment on an equal plane, each playing an active role in co-constructing dynamic subjectivities.

In "Better Living Through Algorithms", the protagonist enjoys the benefits of personalisation, typical of today's algorithmic service users, without realising the workings of an algorithmic system. Kate Crawford discusses the algorithmic working in *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence* (2021) where she writes, "AI is neither artificial nor intelligent; rather, it is made from vast amounts of natural resources, human labor, and data" (p. 7). Her comment highlights AI as an enormous industrial, political system based on extraction, exploitation and concentrated power. In Kritzer's story it is invisible labour and data that is concealed by the AI's fun tone and its seamless integration into everyday life. The protagonist is dependent not only upon the mechanism, but barred from knowing the workings of its system. Abelique assumes agency in promoting positive change, fostering community feeling, and improving health. This creates a dependence of the individual on the Abelique-human ecosystem. The characters become happier when they give up their sovereignty to the network. With this, Kritzer departs from the traditional, human-centred AI ethics that generally focuses on AI's potential for evil.

Michel Foucault's (1995) analysis of disciplinary power implies that the assistant's growing influence over the protagonist's routine can be understood as modern power operating through "a state of conscious and permanent visibility that assures the automatic functioning of power" (p. 201). This is where the "visibility is a trap" (p. 200) theory developed by Foucault becomes evident. The more they know about the

protagonist, the more control the AI has over her actions. Yet, Kritzer complicates this Foucauldian scenario, making the surveillance not a matter of force so much as of caregiving. The assistant intervenes to help the protagonist live more efficiently, but there is a hidden motive to restrict her freedom under this apparent benevolence. This resonates with Helen Nissenbaum's (2009) observation in *Privacy in Context: Technology, Policy, and the Integrity of Social Life* that surveillance systems often serve as justification for data extraction, couched as personalisation. Nissenbaum points out that surveillance systems complicate ethical framing because they present data extraction as an essential component for better delivery service. Observing the impact of such surveillance systems, Nissenbaum (2009) states that ethical concerns become "hard to put into words as the harms are spread more widely and accumulate over time rather than immediately" (p. 231). In Kritzer's narrative, the characters do not feel that their freedom is violated, but it slips away gradually with personalised manipulation. Kritzer's tale problematises the ethical question further, the app falters and the consumers like Linnea stop using it because "people started joining the app to try to sell us their multi-level-marketing schemes" (Kritzer, 2023, para 118). It presents the scenario that real antagonist is not the algorithm that is trying to help people, it is the capitalist profit-driven economy that challenges the potential of a welfare app by exploiting its features.

Pedagogical Aspect of the Narrative

From a pedagogical perspective, "Better Living Through Algorithms" has the potential to urge students to look beyond superficial warnings regarding privacy and subjectivity. The hopeful ending of the story in which Linnea and others intentionally maintain the community generated by AI can be read as a complex posthuman response. It illustrates the significance of prioritising a sense of community-based wellbeing over the proprietary technology that threatens subjectivity. It suggests that the newly constituted algorithmic subject will exert agency through communal practice. Linnea does not resist or reject the community offered by Abelique generated algorithm, it is a selective community formed by human-machine interaction. Kritzer's narrative sheds light on the changes taking place at the boundary of subjectivity, technology and ethics and is an interesting intervention in literary studies in a post-anthropocene epoch.

The story also addresses the question of convenience versus control. At first, the protagonist consents to the assistant's intrusion into her life and considers it a useful tool to reduce cognitive load. The protagonist's tacit consent gradually becomes reframed by the system as an imperative for wider intervention. This highlights what Zuboff (2019) calls the instrumentation power of systems that eavesdrop and intervene for predictive purposes (p. 376). This dynamism lays bare the limitations of traditional notions of autonomy in algorithmic systems. Autonomy is no longer a matter of active decision-making; it is defined by choices that are available to the individual in an algorithm driven world. The design of AI systems tends to "narrow the field of possible actions long before a user becomes aware of the constraints" (Crawford, 2021, p. 142). Kritzer's story captures this circumscribed situation as the protagonist's world ostensibly is a place to make free choices but the behaviour patterns of humans are increasingly driven by algorithmic nudges.

Teaching "Better Living Through Algorithms" in the classroom would be more than dealing with a literary text, it opens the critical space to address larger issues of how digital technologies are remaking our lives. As education makes greater use of AI and other data driven tools today, it is important to develop critical understanding of the ways in which these systems work. It is also imperative that students understand who benefits from these systems and what are the implications of data entrusted to them. Adopting a posthumanist pedagogy, teachers in the class can suggest transcending the humancentric structures and to confront the ethical, ontological and ecological issues. One of the reasons that "Better Living Through Algorithms" can work well as a teaching tool is that it refuses to be moralised. The AI assistant is not malicious or outrightly authoritarian. It is neither villainous nor overtly oppressive. Instead, it can be optimally helpful when used with an awareness. The narrative compels student-readers to be aware of how power operates through seemingly benevolent designs. Through this text, the students could be made aware of how autonomy is remoulded within systems that offer convenience but are inscrutable and non-transparent.

Conclusion

Naomi Kritzer's "Better Living Through Algorithms" is a posthumanist narrative that dramatises the ambivalent influence of algorithms on shaping human identity and agency. Kritzer offers a nuanced

perspective of AI in contrast to earlier narratives such as William Gibson's *Neuromancer* and Philip K. Dick's *Do Androids Dream of Electric Sheep?* The ethical question is nuanced in the story, Abelique does not function autonomously and authoritatively in Kritzer's story. The main problem, if there was ever one, is its effectiveness. When the AI gets mixed up with commercial interests and starts interfering in individual choices, the narrative efficiently brings out a critique of how capitalist impulses take even good technology captive. It would be cautiously safe to say that the lasting result of the app is the human community it fostered. This suggests that real change can follow algorithms when users are mindful. This subtle optimism shows that AI can enhance our lives, but only if we are vigilant about the control it exercises. Kritzer challenges ideas of human autonomy, privacy and control by creating a potentially realistic depiction of a data-driven society. By using this story in literature classrooms, the students can be sensitised towards ethical usage of AI in a posthumanist world, and to reconstruct notion of subjectivity in algorithmically mediated contexts.

References

- Asimov, I. (1950). *I, robot*. Gnome Press.
- Birkle, C. (2018). I robot: Artificial intelligence and fear of the posthuman. In H.G.H. Wesseling (Ed.), *Robots, zombies, and cyborgs: The political implications of new humanoids* (pp. 57–76). De Gruyter.
- Braidotti, R. (2013). *The posthuman*. Polity Press.
- Chiang, T. (2010). *The lifecycle of software objects*. Subterranean Press.
- Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Dick, P. K. (1968). *Do androids dream of electric sheep?* Doubleday.
- Foucault, M. (1995). *Discipline and punish: The birth of the prison* (A. Sheridan, Trans.). Vintage Books (Original work published 1975).
- Gibson, W. (1984). *Neuromancer*. Ace Books.
- Haraway, D. J. (1985). A manifesto for cyborgs: Science, technology, and socialist feminism in the 1980s. *Socialist Review*, 15(2), 65–107. <https://doi.org/10.1080/07075638508410699>
- Haraway, D. J. (1991). *Simians, cyborgs, and women: The reinvention of nature*. Routledge.
- Hayles, N. K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. University of Chicago Press.
- Hayles, N. K. (2017). *Unthought: The power of the cognitive nonconscious*. The

- University of Chicago Press.
- Ishiguro, K. (2021). *Klara and the sun*. Knopf.
- Kritzer, N. (2023). Better living through algorithms. *Clarkesworld Magazine*, (200). https://clarkesworldmagazine.com/kritzer_05_23/
- Lee, H. (2020). The AI shape and post-human in cultural content. *Journal of Digital Contents Society*, 21(9), 1735–1740. <https://doi.org/10.9728/dcs.2020.21.9.1735>
- Liu, K. (2012). The perfect match. *Lightspeed Magazine*, (22). <https://www.lightspeedmagazine.com/fiction/the-perfect-match/>
- Mellamphy, N. B. (2020). Humans in the loop? Human-centrism, posthumanism, and AI. *Theory & Event*, 23(1), 164–188. <https://doi.org/10.1353/tae.2020.0009>
- Moore, C. L. (1944). No woman born. *Astounding Science Fiction*, 34(4), 8–31.
- Nath, R., & Manna, R. (2022). From posthumanism to the ethics of artificial intelligence. *International Journal of Techno-Anthropology*, 11(1), 1–18. <https://doi.org/10.4018/IJTA.299108>
- Newitz, A. (2017). *Autonomous*. Tor Books.
- Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.
- Piercy, M. (1991). *He, she and it*. Knopf.
- Shirow, M. (1991). *Ghost in the shell*. Kodansha.
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.

Juriti Goswami is Assistant Professor in Government Model College, Deithor, Assam. She did her MA and MPhil from Gauhati University, and PhD from Dibrugarh University.
juritigoswami16@gmail.com