Embodied AI at the Margins: Postcolonial Ethics for Intelligent Robotic Systems

Atmadeep Ghoshal, Martim Brandão, Ruba Abu-Salma, Sanjay Modgil

Department of Informatics, King's College London, United Kingdom
atmadeep.ghoshal@kcl.ac.uk, martim.brandao@kcl.ac.uk, ruba.abu-salma@kcl.ac.uk, sanjay.modgil@kcl.ac.uk

Abstract

As artificial intelligence (AI)-powered robots increasingly permeate global societies, critical questions emerge about their ethical governance in diverse cultural contexts. This paper interrogates the adequacy of dominant roboethics frameworks when applied to Global South environments, where unique sociotechnical landscapes demand a reevaluation of Western-centric ethical assumptions. Through thematic analysis of seven major ethical standards for AI and robotics, we uncover systemic limitations that present challenges in non-Western contexts—such as assumptions about standardized testing infrastructures, individualistic notions of autonomy, and universalized ethical principles. The uncritical adoption of these frameworks risks reproducing colonial power dynamics in which technological authority flows from centers of AI production rather than from the communities most affected by deployment. Instead of replacing existing frameworks entirely, we propose augmenting them through four complementary ethical dimensions developed through a postcolonial lens: epistemic non-imposition, onto-contextual consistency, agentic boundaries, and embodied spatial justice. These principles provide conceptual scaffolding for technological governance that respects indigenous knowledge systems, preserves cultural coherence, accounts for communal decision structures, and enhances substantive capabilities for Global South communities. The paper demonstrates practical implementation pathways for these principles across technological life cycles, offering actionable guidance for dataset curation, task design, and deployment protocols that mitigate power asymmetries in cross-cultural robotics implementation. This approach moves beyond surface-level adaptation to re-conceptualize how robotic systems may ethically function within the complex social ecologies of the Global South while fostering genuine technological sovereignty.

Introduction

The emergence of embodied artificial intelligence (AI) agents marks a pivotal shift in human-machine interaction, bringing forth ethical considerations beyond those associated with purely digital systems. Unlike conventional AI applications, embodied agents—including robots, autonomous vehicles, and intelligent physical infrastructures—operate within and manipulate physical environments, thus exercising technological agency in the material space (Brooks 1991; Chrisley 2003). Although scholars have highlighted that the corporeality of embodied agents elicits contextually grounded emotional, behavioral, and cultural responses (Straub 2022), they have also argued that it often causes complex ethical entanglements involving bodily autonomy (Dale and Latham 2014), physical safety (Azeem et al. 2024), and communal norms, such as appropriate physical distance in social interactions (Shiraev and Levy 2015), culturally specific caregiving practices, and taboos regarding technological manipulation of sacred spaces (Hohenthal and Ruuska 2024). Among embodied AI systems, intelligent robots, which exemplify the convergence of computational intelligence and material form, represent what some scholars identify as the most significant ethical frontier in embodied AI (Dautenhahn et al. 2005; Sharkey 2008). Their physical presence—amplified by their autonomous or semi-autonomous capabilities—transforms how they are perceived, experienced, and integrated into diverse environments, creating an intensified sense of agency and presence that significantly influences human psychology and behavior (Dautenhahn 2007; Fong, Nourbakhsh, and Dautenhahn 2003).

Around the world, AI-powered robots are being deployed for diverse purposes, with especially transformative applications emerging across the Global South. In these regions, robots are increasingly being leveraged to address critical development challenges: providing healthcare services to remote communities (Mutongi and Rigava 2024), automating agricultural (Sparrow and Howard 2020) and manufacturing tasks (Henriques et al. 2024), and providing educational resources in areas with severe teacher shortages (Booysen, Rieger, and Ferrein 2011). However, as these robotic technologies proliferate in diverse settings, from urban hospitals to rural farms, the ethical concerns they raise universally, including systemic inequality (Brezis and Rubin 2023), economic precarity (Gomes and Pereira 2019), labor displacement (Cuccu and Royuela 2024), and surveillance (Chun and Papanikolopoulos 2016), manifest with distinctive intensity and characteristics unique to the contexts of the Global South. These amplified ethical challenges come from structural factors that include limited regulatory frameworks that can effectively govern intelligent robotic deployments (Ashwini et al. 2024), public institutions constrained by resource scarcity (Basu 2019), and deeply embedded power

imbalances between communities of the Global South and actors primarily of the Global North who design, manufacture, and control these technologies (Birhane et al. 2022; Mohamed, Png, and Isaac 2020). This epistemic narrowness in ethical frameworks could undermine responsible deployment of AI robots across the Global South, causing fundamental misalignment between their governing principles and the local belief systems of the places where they are deployed. For instance, agricultural robots operating under ethical guidelines inspired by the West that assume private land ownership and individual entrepreneurship could disrupt communal farming systems that have sustained communities for generations in countries like India or Bangladesh, potentially creating new forms of dispossession as algorithmic optimization fails to recognize collective resource management practices. Similarly, a diagnostic AI robot, if deployed in rural health centers, could create barriers to care by rigidly enforcing Western-centric informed consent protocols. Programmed to require standardized written documentation before proceeding with any diagnosis, the robot could reject alternative consent approaches appropriate to communities with different literacy levels, effectively denying timely care to vulnerable populations while adhering to ethical frameworks misaligned with local realities. Therefore, left unaddressed, the uncritical adoption of Western ethical paradigms for intelligent robots, we believe, risks perpetuating technological colonization, where AI robots serve as vectors for imposing external values and extracting data, labor, and resources—ultimately establishing rather than ameliorating global inequality in ways uniquely amplified by the embodied nature of these technologies.

To address the above ethical and epistemic imbalances, our paper critically examines the limitations of the prevailing frameworks that guide the responsible development and deployment of AI-enabled robotic systems in the context of the Global South. Rather than advocating for their wholesale rejection, we demonstrate where these frameworks fall short and propose supplementary considerations to enhance their effectiveness. Using a postcolonial theoretical lens, we interrogate the normative assumptions embedded within existing ethical paradigms while drawing on established discourses in AI safety and roboethics. Through this analysis, we propose four complementary ethical considerations that augment current frameworks to create more contextually appropriate guidelines for responsible intelligent robotics. In the paper, we also demonstrate how these principles can be operationalized at the task, dataset, and deployment levels to ensure that AI-powered robots respect diverse knowledge systems, preserve cultural coherence, support communal decision-making, and enhance substantive capabilities of communities throughout the Global South region.

Related Work

In this section, we examine scholarly contributions in three intersecting domains that inform our investigation of responsible AI for robotics in the Global South. We first review the ethics of AI agents, tracing the evolution from theoretical principles to practical governance frameworks and their limitations. We then explore responsible robotics approaches, highlighting their distinctive considerations beyond software-based AI ethics. Finally, we analyze AI ethics specifically within Global South contexts, examining how decolonial and postcolonial perspectives challenge dominant frameworks and offer alternative conceptual foundations for our proposed approach.

Ethics of AI Agents

The field of AI ethics has expanded significantly to address the unique challenges posed by autonomous agents: entities capable of perceiving environments and performing goal-directed actions independently (Anderson and Anderson 2007; Moor 2006). This has led to the emergence of machine ethics, a subfield focused on embedding ethical reasoning within agents. Moor (2006) distinguishes between implicit ethical agents, which operate within predefined constraints, and explicit ethical agents, which can reason about ethical principles in novel contexts (Moor 2006). These approaches commonly draw on Western philosophical traditions: consequentialism evaluates actions based on outcomes, deontology based on adherence to rules, and virtue ethics based on moral character (Floridi et al. 2018).

Contemporary governance frameworks generally emphasize fairness, accountability, transparency, and safety, although how these principles are applied varies greatly (Hagendorff 2020). A key challenge remains the "principle-to-practice gap"—the difficulty of operationalizing ethical ideals in technical systems (Mittelstadt 2019). The proposed solutions include value-sensitive design (Friedman and Hendry 2019), impact assessments (Reisman et al., 2018), and technical approaches such as explainable AI and fairness-aware machine learning (Doshi-Velez and Kim 2017; Mehrabi et al. 2021). Critical perspectives have increasingly questioned the sufficiency of mainstream approaches. Benjamin (2019) shows how autonomous systems can reproduce structural inequalities (Benjamin 2019), while Crawford (2022) calls for attention to power asymmetries in system design (Crawford 2022). Mohamed et al. (2020) emphasize the role of historical and cultural context in shaping ethical implications (Mohamed, Png, and Isaac 2020). At the technical level, challenges persist around value alignment and moral uncertainty issues, such as how agents should behave when human values are in conflict or when unexpected behaviors arise in dynamic environments, which require sustained interdisciplinary inquiry (Gabriel 2020; Russell 2020).

Responsible Robotics

Responsible robotics has emerged as a critical subfield that engages in the ethical, legal, and social dimensions of embodied AI systems, particularly those capable of autonomous action in physical environments (Lin, Abney, and Bekey 2014). The physical instantiation of such systems introduces distinct ethical challenges throughout the robotic life cycle, including safety in human-robot interaction (Haddadin 2015) and the attribution of moral and legal responsibility as autonomy increases, which Matthias (2004) terms the "responsibility gap" (Matthias 2004). This concern is amplified in high-stakes domains where meaningful human

control remains contested (Sharkey 2011) and where issues of identity, language, and perceived agency may influence moral attributions (Winkle et al. 2021). Operationalizing responsibility requires multi-pronged strategies. These include formal safety and verification protocols (Guiochet, Machin, and Waeselynck 2017), ethical design processes grounded in Responsible Research and Innovation (Stilgoe, Owen, and Macnaghten 2013), and participatory design methods that emphasize inclusivity and co-construction (Datey and Zytko 2024). However, implementation challenges persist at both the technical and sociopolitical levels. Concerns include labor displacement (Acemoglu and Restrepo 2019; Birhane 2021), privacy violations through pervasive sensing (Finn and Wright 2012), and systemic bias embedded in decision systems (Howard and Borenstein 2017). Robots driven by LLMs further complicate these concerns, posing the risks of discriminatory or unlawful behavior if not adequately constrained (Azeem et al. 2024). Recent scholarship has drawn attention to long-term human-robot relationships, focusing on reciprocity, trust, and social integration (Dautenhahn 2014; van Wynsberghe 2021). These concerns intersect with broader normative agendas, such as aligning robotics with social justice imperatives (Zhu, Wen, and Williams 2024), particularly in under-resourced regions (van Wynsberghe 2021). In these contexts, scholars have highlighted the need for cultural and regulatory re-contextualization, advocating for ethical frameworks grounded in human rights (Lin, Abney, and Bekey 2014), epistemic diversity (Coeckelbergh 2022), and inclusive standardization practices (Fosch-Villaronga and Giraudo 2022).

AI Ethics for the Global South

A growing body of research has critiqued dominant AI ethics frameworks for their limited applicability to Global South contexts, emphasizing how these models—largely shaped in the Global North—often overlook local sociocultural specificities and risk reinforcing structural inequalities (Mohamed, Png, and Isaac 2020). Central to this critique is the concept of data colonialism, which exposes how contemporary data practices replicate extractive logics by capturing value from Global South populations while offering minimal reciprocal benefit, thus mirroring colonial power dynamics (Couldry and Mejias 2020). This critique is part of a broader narrative of technological extraction in the digital ecosystem. Empirical studies have further demonstrated technical disparities in AI performance across demographic lines. Buolamwini and Gebru (2018), for example, revealed substantial accuracy gaps in facial recognition systems when applied to darker-skinned individuals, illustrating how algorithmic bias can reflect and reinforce racial and geographic inequities (Buolamwini and Gebru 2018). In response, scholars have proposed context-sensitive alternatives aimed at aligning AI systems with local values and needs. For example, Varshney (2024) has advocated for the incorporation of indigenous knowledge into value alignment processes (Varshney 2025), while Okolo et al. (2021) have demonstrated the design of culturally-tuned explainable AI systems to be effective for healthcare applications in the Global South (Okolo et al. 2021). Mhlambi and Tiribelli (2024) similarly draw on relational autonomy theory to propose more culturally-grounded models of decolonial AI governance. Complementing these efforts, Ofosu-Asare (2024) emphasizes the integration of indigenous epistemologies to counter cognitive imperialism in AI development (Ofosu-Asare 2024); Widder (2024) calls for the recognition of lived experiences as epistemically valid within AI ethics labor (Widder 2024); and Helm et al. (2023) highlight how techno-linguistic bias in language technologies risks perpetuating epistemic injustice in underrepresented communities (Helm et al. 2023).

Approach

We followed a three-step methodological procedure for our inquiry. First, we selected seven globally influential roboethics and AI governance frameworks based on their institutional legitimacy, policy relevance, and explicit engagement with embodied AI systems. Second, we conducted an abductive thematic analysis (Timmermans and Tavory 2012; Braun and Clarke 2006) of these documents through a postcolonial lens, iteratively coding them to identify recurring normative patterns and epistemic assumptions. This process surfaced four categories that captured key tensions between global ethical standards and localized sociotechnical realities. Third, based on these categories, we developed four corresponding ethical principles designed to enhance the seven selected frameworks and better align the governance of AI robots with the cultural, social, and political contexts of the Global South. We elaborate on each of these steps below.

- **Step 1**: We analyzed seven globally recognized roboethics and AI governance frameworks (IEEE P7007, ISO 13482, EURON Roboethics Roadmap, EPSRC Principles of Robotics, UNESCO's COMEST Report, RoboLaw Guidelines, and AI HLEG Ethics Guidelines) selected for their institutional legitimacy and influence in shaping robotics governance across jurisdictions.
- **Step 2**: We conducted abductive thematic analysis (Timmermans and Tavory 2012; Braun and Clarke 2006) of these frameworks using concepts of postcolonial theory. Two researchers independently coded all frameworks, identifying four dominant themes that reveal shortcomings in Global South contexts: epistemic justice, ontological security, organic autonomy and agency, and social freedom.
- **Step 3**: Based on these themes, we developed complementary ethical principles to address identified gaps. Rather than wholesale replacement, these principles provide conceptual scaffolding for more inclusive and contextually-grounded AI robot governance in the Global South, with practical recommendations for implementation throughout the AI life cycle.

Roboethics Frameworks: A Postcolonial Critique and New Considerations

This section is divided into four parts. The first part details our strategy for selecting ethical frameworks regarding intelligent robotics. Next, we outline our thematic analysis

process, detailing how we developed codes and synthesized them into four distinct themes grounded in postcolonial theory. We then apply these themes to critique the frameworks and, finally, propose alternative ethical principles that respond to the identified gaps.

Selection of Frameworks

To examine current ethical governance frameworks for robotics and embodied AI, we adopted a purposive selection strategy guided by relevance, institutional diversity and conceptual influence. Following practices in previous research (Prem 2023; Hagendorff 2020; Narayanan and Schoeberl 2023), we selected frameworks that met three primary criteria: (1) Institutional legitimacy—published or endorsed by internationally recognized standard setting bodies, intergovernmental organizations, or publicly-funded research initiatives; (2) Relevance to embodied AI - engaging with the algorithmic, autonomous, and socio-technical dimensions of robotics; and (3) diversity of normative scope, representing different frameworks for ethics across legal, philosophical, technical, and policy domains. This led us to identify seven frameworks that collectively shape contemporary roboethics discourse: IEEE P7007 Ethical Standards for Robot Design[1], ISO 13482 Safety Standards for Personal Care Robots[2], EURON Roboethics Roadmap[3], EPSRC Principles of Robotics[4], UNESCO COMEST Report on Robotics Ethics[5], RoboLaw Guidelines on Regulating Robotics[6], and Ethics Guidelines for Trustworthy AI by the European Commission's High-Level Expert Group[7]. Rather than aiming for exhaustiveness, our goal was to assemble a conceptually and institutionally diverse corpus of high-level frameworks that are actively shaping ethical standards in robotic and AI governance. This reflects established approaches in critical and interpretive research, where selection is guided by analytical relevance and the inclusion of diverse normative perspectives, rather than comprehensive enumeration (Clarke 2005; Seale 1999).

Postcolonial Lens: Theoretical Foundations for Our Critique

A postcolonial orientation toward the ethical frameworks for AI robots necessitates a rigorous examination of global power asymmetries and culturally-specific epistemologies that fundamentally determine what constitutes ethical conduct. This critical lens illuminates whose values and interests become codified within these frameworks and which epistemological traditions face systematic marginalization (Mbembe 2001; Morris 2009). Despite claims of universality, we contend that dominant ethical frameworks in AI-enabled robotics emerge predominantly from Euro-American intellectual contexts, encoding assumptions regarding autonomy, governance, and technological progress that may lack resonance or applicability across diverse sociotechnical landscapes. Contemporary scholarship has elucidated that these assumptions frequently obscure colonial legacies embedded within technological practices. For example, Irani and Philip (2018) demonstrate how the prevailing standards of technological development inherently embed Western-centric narratives of progress (Irani and Philip 2018). Arora (2019) illustrates how AI implementations in Global South contexts frequently reproduce colonial hierarchies despite rhetorical commitments to neutrality and inclusion (Arora 2019). Additionally, Benjamin's (2019) conceptualization of the "New Jim Code" provides a penetrating critique of how algorithmic systems perpetuate entrenched racial biases (Benjamin 2019), while Couldry and Mejias (2020) theorize "data colonialism" as an emerging form of extraction that commodifies human experience and knowledge (Couldry and Mejias 2020).

Thematic Analysis of Frameworks

Guided by the body of literature discussed above, we proceeded to conduct an abductive thematic analysis of the seven selected frameworks. Two researchers independently coded the documents using a set of initial sensitizing concepts drawn from postcolonial theory—such as epistemic marginalization (Harding 1998), coloniality of power (Dourish and Mainwaring 2012), relational autonomy (Mackenzie 2021), and techno-governance (Broussard 2018). Through several rounds of collaborative discussion, comparison, and refinement, we iteratively developed a set of inductively enriched codes that captured both explicit ethical framings and latent normative assumptions embedded in the documents. This analytic process allowed us to synthesize recurring patterns across diverse institutional texts and surface the unspoken ethical architectures underpinning them. The resulting codes were clustered into four high-level themes that reflect both postcolonial theoretical commitments and thematic coherence across the corpus. These are **epistemic justice**, **ontological security**, **organic autonomy and agency**, and **social freedom**. We find our themes also echo and expand upon existing scholarship in AI ethics: epistemic justice resonates with Sambasivan et al.'s call to re-imagine algorithmic fairness by recognizing diverse knowledge systems and challenging western hegemony; ontological security connects to Giddens' (1991) foundational work that is further developed by Bilgic and Pilcher (2022), which attends to the destabilization of collective identity and social meaning under technocratic governance (Giddens 1991; Bilgic and Pilcher 2022); organic autonomy/agency builds on Asaro's (2006) critique of liberal individualism in robotics ethics, which calls for relational, interdependent models of agency (Asaro 2006); and social freedom draws on Sen's capabilities approach and Verhulst et al.'s work on digital self-determination, which foregrounds political voice, collective flourishing, and representational equity in sociotechnical systems (Crocker and Robeyns 2009; Verhulst 2023).

[1]https://standards.ieee.org/ieee/7007/7070/

[2]https://dig.watch/resource/iso-134822014-robots-and-robotic-devices-safety-requirements-personal-care-robots

[3]https://ieeexplore.ieee.org/document/4115667

[4]https://doi.org/10.1080/09540091.2017.1313817

[5]https://unesdoc.unesco.org/ark:/48223/pf0000245532

[6]https://cordis.europa.eu/project/id/289092/reporting

[7]https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai?

Framework	Key Principles	Shortcomings in Global South Contexts
IEEE P7007 Ethical Standards for Robot Design	Transparency in decision-making, user safety, robot identification	Presupposes technical literacy among users; limited consideration of collective impacts; high implementation barriers for local manufacturers
ISO 13482 (Safety Standards for Personal Care Robots)	Technical safety requirements, risk assessment procedures, performance standards	Focuses on individual safety without addressing communal impacts; assumes standardized testing environments unavailable in many regions; overlooks informal repair economies
EURON Roboethics Roadmap	Human dignity, user autonomy, robot sociability, environmental sustainability	Based primarily on Western ethical traditions; limited consideration of diverse cultural interactions with robots; assumes consistent regulatory oversight
EPSRC Principles of Robotics	Robots as tools serving humans, accountability, transparency, legal compliance	Inadequately addresses contexts with limited legal frameworks; assumes clear human-robot boundaries that may conflict with diverse cultural perspectives
UNESCO COMEST Report on Robotics Ethics	Human dignity preservation, protection from harms, privacy safeguards	Insufficient guidance for applications in resource-constrained environments; limited attention to economic displacement concerns most relevant to developing regions
RoboLaw Guidelines on Regulating Robotics	Proportionality, precaution, procedural fairness, user-centered design	Assumes robust regulatory frameworks; disproportionate compliance burdens; limited attention to contextual variations in risk perception
Ethics Guidelines for Trustworthy AI (AI HLEG)	Human agency, robustness, privacy, transparency, diversity	Implementation guidance lacks contextual specificity; assumes institutional capacity for oversight; limited attention to diverse cultural interactions with embodied AI

Table 1: Robotics Ethics Frameworks and Their Limitations in Global South Contexts

Limitations of Current Frameworks Through a Postcolonial Lens

We now unpack the four themes that emerged through our thematic analysis of the selected ethical frameworks and apply them as interpretive lenses to critique how dominant approaches construct ethical governance for AI-enabled robotics. We find that most frameworks exhibit tendencies toward universalism, abstraction, and epistemic centralization, which limit their ability to engage with situated perspectives and alternative moral imaginaries. We elaborate on each theme and its implications in the sections that follow, while we summarize our analysis in Table 1

Epistemic Justice: Challenging Knowledge Hierarchies The first theme, epistemic justice, concerns the fair consideration and inclusion of diverse knowledge systems (Fricker 2007) in technological development. Following that, we find that current frameworks such as IEEE P7007 and the AI HLEG Guidelines privilege taxonomic precision and Western scientific rationality while rendering indigenous understanding of human-machine relationships invisible. This could be understood by how IEEE P7007 specifies that *"users of this standard need to have minimal knowledge of formal logics to understand the axiomatization expressed in the Common Logic Interchange Format"*, presupposing an educational background and epistemological orientations aligned with analytical traditions predominant in WEIRD societies. Further examination of the IEEE P7007 standard reveals its underlying assumptions through its reliance on formal ontologies, where it states that *"ethical behavior will be implemented through the use of formal ontologies which will help robots act according to human ethics and cultural beliefs"*. We believe that this formulation problematically suggests that diverse ethical systems can be adequately captured within computational logic frameworks, a presumption that Birhane (2021) has challenged as fundamentally reductionist (Birhane 2021). Similarly, the AI HLEG Guidelines (2019) state that *"trustworthy AI should be lawful, ethical, and robust"*, while defining ethical requirements exclusively through Western values of *"respect for human autonomy, prevention of harm, fairness, and explicability"* that may not encompass ethical priorities in non-Western contexts (Sambasivan et al. 2021).

Ontological Security: Cultural Coherence in Human-Robot Interaction The second theme, ontological security, addresses how robotic technologies can destabilize or reinforce established patterns of social life and cultural understanding. Safety standards like ISO 13482 focus primarily on individual physical safety without consider-

ing how robotic presence and algorithmic decision-making might disrupt community coherence and collective meaning-making systems. The standard specifies that *"the manufacturer determines the required tests and appropriate testing parameters based on a risk assessment of the robot's design and usage"*, assuming standardized testing environments and risk assessment protocols that may be unavailable or culturally incongruent in many Global South contexts. For example, in some indigenous Andean communities, land cultivation is not simply an economic activity, but a deeply spiritual and communal practice governed by reciprocal rituals with the earth. The introduction of an autonomous agricultural robot that optimizes sowing or irrigation without considering these rituals could violate local ontologies of care and stewardship, not only disrupting tradition but also eroding the relational meaning of agricultural labor. A closer read of ISO 13482 reveals that it defines safety as *"freedom from risk which is not tolerable"*, but determines tolerability through quantitative risk assessment methodologies that prioritize physical harm to individuals over collective social harms. The standard instructs manufacturers to *"evaluate each hazardous situation by taking into account the severity of the injury and the probability of its occurrence"*, without provisions to assess disruptions to social fabric or communal practices. Similarly, the EURON Roboethics Roadmap states that *"applications of robotics should (...) comply with existing social rules"*, without acknowledging the plurality of social systems globally or providing mechanisms to identify potential disruptions to them. If such protocols are followed, a robot that can be deployed in a communal fishing village in Kerala could bypass collective decision making by providing catch predictions to individual fishers, disrupting systems of equitable distribution historically maintained through community deliberation. Such standards thus, we argue, fail to account for how AI-enabled robots might fundamentally alter social bonds and communal practices that provide ontological stability in non-Western societies.

Organic Agency and Autonomy The third theme, organic agency and autonomy, moves beyond Western individualistic conceptions to acknowledge diverse modes of decision-making and self-determination prevalent in many Global South contexts. The EPSRC Principles of Robotics characterize robots as *"manufactured artifacts"* subordinate to individual human masters, embedding liberal individualist conceptions of agency that may conflict with more collectivist understandings of autonomy and responsibility (Mhlambi and Tiribelli 2024). The EPSRC framework explicitly states that *"robots are manufactured artifacts: they should not be designed in a deceptive way to exploit vulnerable users; instead, their machine nature should be transparent"*. This principle assumes a clear distinction between humans and machines that may not align with the animistic or relational dynamics present in many indigenous knowledge systems (Kohn 2015). Furthermore, the RoboLaw Guidelines assert that *"autonomy traditionally refers to the capacity for self-governance (...) of human beings"*, subsequently extending this individualistic conception to robotic systems without acknowledging alternative frameworks of agency (Palmerini et al. 2016).

Social Freedom and Substantive Capabilities The fourth theme, social freedom, examines whether AI-driven technologies enhance or diminish substantive opportunities for flourishing across diverse contexts. Drawing on Sen's (1999) capability approach (Crocker and Robeyns 2009), this dimension asks whether robotic systems actively expand the capabilities of historically marginalized communities to live lives they have reason to value. The UNESCO COMEST Report acknowledges international inequality, but frames it simplistically as a question of whether robots might *"bring about a new divide between developing and developed countries"*. The report fails to explore how robotics might interact with existing structural inequalities specific to Global South contexts such as India where caste inequities still persist largely, or how these technologies might be inappropriately adapted across diverse cultural settings, especially those where vulnerable communities might be under greater risk due to such adaptations (Kalyanakrishnan et al. 2018). Similarly, the Ethics Guidelines for Trustworthy AI mention that AI systems should foster fundamental rights and ensure inclusion and diversity, but fail to specify how these objectives might be pursued in contexts where basic infrastructure for technological deployment is lacking. When addressing fairness, these frameworks focus primarily on algorithmic bias while neglecting structural inequalities that determine who can access robotic technologies in the first place, a limitation that Ricaurte (2019) identifies as endemic to Global North perspectives on digital justice (Ricaurte 2019).

Proposed Ethical Principles

Building on our postcolonial critique, we propose four principles that substantively augment existing ethical frameworks for intelligent robots to address their limitations when applied in diverse global contexts. These principles emerge directly from our analysis of the epistemic, ontological, agentic, and distributive gaps in current frameworks, offering a pathway toward more inclusive and contextually appropriate governance of AI-embodied robots. Where frameworks like IEEE P7007 encode Western epistemologies through requirements for *formal logics* and *axiomatization*, we propose a principle of **Epistemic Non-Imposition**, which creates space for indigenous knowledge systems to inform robotic perception and decision making. Similarly, while ISO 13482 and the EURON Roadmap focus narrowly on physical safety without acknowledging potential cultural disruption, our principle of **Onto-Contextual Consistency** ensures robots support rather than undermine the coherent *world-sense* that provides ontological security across diverse cultural contexts. The EPSRC Principles' characterization of robots as *"manufactured artifacts"* subordinate to individual human masters fails to account for communal decision-making prevalent in many Global South societies; our principle of **Respect for Agentic Boundaries** addresses this by ensuring that robots calibrate their interventions to preserve both individual and collective agency. Finally, while the UNESCO COMEST Report and AI HLEG Guidelines

speak of fairness in abstract terms without addressing material inequality, our principle of **Embodied Spatial Justice** recognizes how robots' physical operations can either challenge or reinforce patterns of spatial marginalization related to caste, class, gender, and land ownership. Together, these principles do not replace, but significantly expand existing frameworks, addressing the distinct ethical challenges that arise when AI-powered robots operate within the complex social ecologies of the Global South. In the following subsections, we discuss our principles in detail.

Epistemic Non-Imposition Our analysis of ethical frameworks such as IEEE P7007 and AI HLEG Guidelines reveals a recurring pattern of epistemic injustice. These frameworks encode Western analytic epistemologies—privileging formal logic, ontological precision, and taxonomic universality—while marginalizing indigenous and relational knowledge systems. In response, we propose the principle of Epistemic Non-Imposition in Robotic Cognition: **robots must not impose dominant epistemic frameworks when interpreting, learning from, or interacting with the world. Instead, they must be algorithmically designed to accommodate and engage with diverse epistemologies, allowing, for example, oral histories, spiritual reasoning, or seasonal taxonomies to inform their perception and decision-making processes**. For example, a conservational robot deployed in Northeast India should be able to learn from and reason with indigenous ecological markers rather than override them with euro-centric botanical classifications. This principle affirms epistemic justice as the foundational factor in the ethical deployment of AI robots and calls for systems that foster epistemic plurality rather than reinforce knowledge hierarchies.

Onto-Contextual Consistency Our critique of standards such as ISO 13482 and the EURON Roboethics Roadmap highlights a significant neglect of ontological security, the stability of individuals' and communities' sense of continuity in their social and cultural life-worlds. These frameworks prioritize physical risk reduction and individual safety, yet fail to account for how robots can disrupt collective meaning making, narrative coherence, or culturally embedded understandings of space and time. Thus, we propose the principle of Onto-Contextual Consistency in Robotic Operations: **robotic systems must preserve and reinforce users' culturally grounded *world-sense*, avoiding abrupt or opaque changes in behavior, environment, or narrative interaction that could disrupt identity or continuity**. For example, a care robot in an Afro-Caribbean eldercare facility should not update its memory systems in ways that erase inter-generational stories or disrupt the rhythm of spiritual observance. Instead, it should maintain continuity with the user's life-world and adapt its operations in line with local ontologies. This principle ensures that robots do not become agents of cultural rupture, but rather respect and support the ontological stability of those they assist.

Respect for Agentic Boundaries Our examination of the principles of roboethics, including those of EPSRC and RoboLaw, reveals a deep reliance on liberal individualist models of agency that do not account for communal and distributed autonomy, modes of self-determination that are central to many contexts in the Global South (Sambasivan et al. 2021). These frameworks assume that agency operates through individual human-machine relationships, neglecting how decisions and actions emerge in many Global South contexts through collective deliberation, shared responsibility, and interdependent relationships within communities, where these values are greatly prioritized (Escobar 2018; Awad et al. 2018). By defining robots as subordinate *artifacts* and emphasizing transparency primarily to avoid deception, these frameworks neglect how robotic systems might subtly undermine collective agency by automating decisions that traditionally involve community input, anticipating actions that disrupt social processes, or intervening in ways that bypass established communal decision-making structures. Such techno-paternalism risks reenacting colonial patterns of imposed governance, particularly in contexts where collective human agency has historically been denied or minimized (Weckert 2019). To address this, we propose the principle of Respect for Agentic Boundaries: **robots must be designed to support communal decision-making processes without preempting or overriding collective deliberation**. Robotic intervention should recognize and preserve existing social structures of shared agency, allowing communities to collectively determine when and how robotic support is engaged. For example, consider the introduction of an agricultural robot in an indigenous farming community in central India where decisions about crop rotation, seed selection, and land use are made collectively by elders, women's cooperatives, and youth groups through seasonal gatherings. Instead of using AI to independently optimize planting schedules or recommend high-yield commercial crops based on data, a robot designed with respect for agentic boundaries would function as a facilitator in these gatherings. It would surface environmental data, historical rainfall patterns, or pest forecasts only when invited by the community and in ways intelligible within local ecological and cultural knowledge systems. Crucially, the robot would not rank or suggest actions unless the community explicitly seeks that input. For example, if the group is debating whether to plant traditional millet versus a commercial rice variant, the robot would not intervene to "optimize" for yield. Instead, it would facilitate discussions by presenting data in the context of the long-term goals of the community, such as food sovereignty, soil health, and cultural continuity, reinforcing the community's role as the main agent in the shaping of technological participation.

Embodied Spatial Justice Our review of robotics ethics documents, including the UNESCO COMEST Report and AI HLEG Guidelines, indicates a persistent disregard for spatial justice—that is, how robotic systems redistribute access, movement, and control across physical environments. These frameworks often frame inequality in abstract terms (e.g., the *AI divide*) or focus narrowly on algorithmic fairness, failing to consider how robots, by virtue of their embodiment, can materially reproduce spatial hierarchies. Whether through logistics, surveillance, or automation of

infrastructure, robots can reinforce exclusion by bypassing, surveilling, or displacing already marginalized populations. We, therefore, propose the principle of Embodied Spatial Justice: **robotic systems must be designed with an explicit awareness of how their physical operations affect spatial equity. Their deployment should avoid reinforcing patterns of dispossession related to caste, class, gender, or land ownership, especially in Global South contexts**. For example, if delivery robots in a South Asian smart city are routed to avoid informal settlements, they contribute to the spatial marginalization of slum residents. A justice-oriented spatial ethic calls for inclusive route planning, participatory design, and mechanisms for redress when robotic presence reinforces inequality. Robots, as embodied agents, must not just navigate space—they must help reshape it toward greater justice.

Operationalizing Ethical Dimensions for AI Robots in the Global South

In this section, we identify a comprehensive framework for AI-driven robotics deployment in the Global South, examining implementation at three critical levels: task, dataset, and deployment. The analysis focuses on our proposed four core ethical principles—epistemic justice, ontological security, agency and autonomy, and social freedom—and how they can be operationalized at each level to ensure equitable and culturally-responsive robotic AI systems. Drawing on interdisciplinary research spanning human-computer interaction, critical data studies, and development ethics, this section demonstrates how intelligent technology-based implementations that neglect local knowledge structures, cultural meanings, community governance, and substantive freedoms often fail, while providing evidence-based alternatives that center community participation, cultural congruence, and empowerment throughout the AI development life cycle.

Task-Level Operationalization

To uphold epistemic non-imposition, AI robots must be designed to recognize and incorporate local knowledge practices at the task design level. Rather than imposing universal logic models, AI reasoning should allow situated knowledge to shape how tasks are defined and executed. For example, a generative AI-enabled robot used in rural irrigation planning must accommodate community-driven seasonal knowledge and water-sharing customs instead of relying solely on climate optimization models drawn from Euro-American agronomic data. Similarly, onto-contextual consistency requires that task flows remain culturally legible: robots assisting with domestic caregiving should maintain interaction scripts aligned with local kinship roles and respect practices around eldercare or rituals. In the domain of agentic boundaries, robots must scaffold rather than override human skill: agricultural robots, for instance, should offer just-in-time prompts to help farmers diagnose crop conditions, instead of making automatic decisions that bypass farmer judgment. Finally, embodied spatial justice at the task level involves ensuring that robots do not structure workflows in ways that displace or de-prioritize marginalized users. For example, a sanitation robot in an urban township should not be tasked with operating only in commercial zones, but should also service informal settlements through community-prioritized task allocation.

We may also consider here the example of an **Autonomous Underwater Vehicle (AUV)** for biodiversity and overfishing monitoring, used in nautical space reserved for artisanal crafts such as territorial waters of Kerala in India (Government of India 1976). To uphold our four post-colonial principles, *epistemic non-imposition* would require that risk markers (e.g., species, seasons, catch limits) be *co-defined* with local fishers and marine scientists rather than imposed from other contexts such as the EU Common Fisheries Policy (Council of the European Union 2016). *Onto-contextual consistency* would demand that the vehicle's GIS layers preserve indigenous place names, migration calendars, and taboo stretches, so the data map aligns with the community's own spatial ontology. *Respect for agentic boundaries* could be achieved by first sending each provisional catch alert to democratically elected village co-operatives, allowing them to decide which summaries, if any, should be forwarded to state fishery officers. Finally, *embodied spatial justice* would necessitate directing monitoring and enforcement efforts toward industrial trawlers, which are required to stay beyond the twelve-mile limit during the monsoon trawl ban (The New Indian Express 2023; Gain 2025), rather than toward subsistence canoes operating within their customary waters—thereby protecting rather than surveilling marginalized communities.

Dataset-Level Operationalization

At the dataset level, epistemic non-imposition entails curating training data that includes diverse cultural objects, gestures, languages, and interaction norms, moving beyond the narrow visual and behavioral taxonomies common in Global North datasets. For example, an AI robot designed for community education should be trained not only on textbook diagrams but also on local pedagogical resources, visual metaphors, and oral storytelling traditions. Onto-contextual consistency in data requires preserving the cultural significance of images, objects, and spaces. In robotics navigation, a path identified as "empty" by conventional mapping could have religious or ancestral significance locally; therefore, training data must encode such contextual metadata to avoid symbolic erasure. Respect for agentic boundaries in data practices involves participatory annotation processes in which communities can decide which data is collected and how it should guide robotic learning. For example, a healthcare robot trained through generative feedback loops should be co-supervised by community health workers, who can define the limits of acceptable diagnostic phrasing. Lastly, spatial justice demands datasets reflect geographic diversity—not only by including informal or rural landscapes, but also by ensuring that spatial labels and segmentations are created collaboratively to avoid urban-centric bias. For example, a delivery robot trained exclusively on Western road systems is likely to struggle with navigating the alleyways and footpaths commonly used in many parts of the Global South. A practical implementation of this could be traced

to Wadhwani AI's CottonAce[8], an AI-powered pest advisory tool for cotton farmers in India. Rather than automating pesticide spraying, the system helps farmers detect pest infestations through pheromone trap photos, delivering localized alerts and recommendations in regional languages. It respects farmer knowledge (epistemic non-imposition), supports timely and dignified decision-making (agentic boundaries), and empowers smallholder farmers often excluded from precision agriculture (spatial justice).

Deployment-Level Operationalization

During deployment, epistemic non-imposition requires AI robots to be open to ongoing learning from local users and institutions. This includes feedback mechanisms in which community knowledge continuously informs robot updates—for example, voice interfaces that adapt to local dialects or reasoning engines that recalibrate logic based on indigenous ecological observations. Onto-contextual consistency mandates that AI robots be deployed in ways that reinforce rather than disrupt cultural and spatial routines. For example, a mobile clinic robot that navigates a pastoralist community must adapt to mobile social organization and sacred land use patterns rather than insisting on static GPS routing. In terms of agentic boundaries, robots should be equipped with explainable decision pathways and customizable autonomy levels, letting users decide how much control to retain in domains such as health, education, or local governance. For example, a chatbot-robot hybrid in village legal aid should enable users to co-construct arguments rather than merely offering automated recommendations that foreclose deliberation. Finally, embodied spatial justice demands that deployment plans avoid reinforcing infrastructural exclusion. Delivery robots serving the markets of a smart city must also serve peri-urban slums—not as charity extras but as co-designed service zones. Spatial coverage must reflect community priorities and avoid the tendency of AI infrastructure to "skip over" marginalized geographies. The Flying Labs Network[9], which deploys drones in more than 40 countries to serve community-defined goals—such as malaria prevention, flood mapping, or infrastructure planning—has fulfilled several of these ethical modalities. Each lab is locally run, ensuring control over how, where, and why robotics are used; decentralizing power; and centering local priorities—principles aligned with epistemic non-imposition and spatial justice—thus offering a compelling model for ethical deployment in the Global South.

Discussion

While our ethical considerations were developed in response to the deployment of AI robots in the Global South, their relevance extends far beyond this context. As AI systems increasingly inhabit physical space and interact with human routines, it is necessary to assess how these principles apply to other forms of embodied intelligence. Equally important is the question of how these ethical commitments can be evaluated in practice. Moving beyond high-level abstraction requires the development of concrete, context-sensitive methods that can assess whether these principles are meaningfully upheld throughout the AI life cycle. We discuss the same in this section.

Extending Ethical Principles to Other Forms of Embodied AI

Although the proposed ethical considerations—epistemic non-imposition, onto-contextual consistency, respect for agentic boundaries, and embodied spatial justice—emerged from our analysis of limitations of roboethics frameworks in Global South contexts, they offer a foundational basis for engaging a broader class of embodied-AI systems. Although we have focused mostly on ground robots in service and agriculture, our principles apply beyond these, extending to autonomous vehicles, underwater robots, smart buoys, wearable AI, ambient intelligent environments, and even generative AI agents that operate through embodied interfaces (e.g., voice assistants in smart homes, AI-driven prosthetics, socially assistive avatars). In these systems, embodiment does not always imply a wheeled or humanoid form or physical locomotion, but rather an embeddedness in material, spatial, and cultural worlds through sensors, actuators, and human-facing interfaces. For example, epistemic non-imposition applies directly to AI wearables in healthcare that standardize bodily norms based on Global North biometric data.

Decolonial ethics would require these systems to accommodate diverse embodied experiences and diagnostic frameworks. Similarly, onto-contextual consistency is important for smart home assistants that organize social routines through reminders and automated scheduling. These systems need to align with the rhythms and cultural meanings of households that do not adhere to Eurocentric temporalities or nuclear family norms. Similarly, respect for user agency can guide the design of AI-driven prosthetics that respond to intent without compromising movement autonomy or established rehabilitation practices. Finally, embodied spatial justice applies to autonomous vehicles or drone-based logistics systems whose routing decisions risk reproducing infrastructural inequities: skipping informal settlements or inaccessible terrains. The portability of these principles across AI embodiment types highlights their value not as narrow prescriptions but as critical orientations rooted in decolonial theory, participatory design, and justice-centered computing. Their strength lies in shifting focus from universal standards to situated ethics that acknowledge who defines the task, whose knowledge is valued, and which bodies and spaces are prioritized. Extending these principles allows for an ethics of AI that is not only "responsible" in the abstract, but also accountable to the historical, political, and material conditions in which AI systems are deployed.

Evaluating Ethical Commitments in AI Robotics

Translating ethical principles into measurable outcomes remains a key challenge for AI ethics, especially in robotics, where social, spatial, and cultural impacts unfold over time.

[8]https://tinyurl.com/wadhwani-ai

[9]https://tinyurl.com/2w2vnces

To evaluate commitments like epistemic non-imposition or embodied spatial justice, we must move beyond checklist-style audits or technical validation and instead adopt contextual, process-oriented, and participatory evaluation frameworks (Madaio et al. 2020; Watkins et al. 2021). These assessments should consider not only the robot's actions but also its design process, the knowledge that shaped it, and how users and communities experience its presence over time.

For epistemic non-imposition, evaluation might include qualitative audits examining whose knowledge informed the design and how knowledge hierarchies were addressed. Madaio et al. (2020) propose co-design checklists that trace the inclusion of diverse stakeholders and values across design phases (Madaio et al. 2020), while Sloane et al. (2022) highlight the importance of documenting data provenance and surfacing epistemic exclusions embedded in development workflows (Sloane et al. 2022). Participatory design documentation, local co-design interviews, and contextual inquiry (Suchman 2006) can determine whether dominant epistemologies were privileged or challenged in practice.

In terms of onto-contextual consistency, ethnographic methods, such as longitudinal field observations, workplace studies, and user diaries, are well suited to capture how AI robots affect social rhythms, narrative coherence, and spatial routines (Lee et al. 2012). Measures of cultural fit or disruptions to everyday life can help surface ontological harms that may be invisible to conventional system-level metrics.

Respect for agentic boundaries can be evaluated through task-level interaction studies that assess how often users override or defer to robotic recommendations, whether they feel their decision-making is respected, and how much interpretive transparency is offered. System log analyses, user feedback, and transparency tools like model cards (Mitchell et al. 2019) can support this.

For embodied spatial justice, deployment-level data such as coverage maps, access logs, and demographic reach analyses, are critical. Participatory GIS methods and spatial ethnographies can reveal whether marginalized spaces are included or excluded in robotic service networks (Benjamin 2019; McQuillan 2022).

Importantly, these evaluation strategies require collaboration between designers, researchers, and affected communities. Just as our ethical considerations resist universalism, our evaluation frameworks must likewise be locally grounded, historically informed, and methodologically diverse. Building such tools requires embedding ethics throughout the design and deployment life cycle - not as a retrospective audit, but as a continuous negotiation of what just, inclusive, and meaningful AI systems should be (Watkins et al. 2021).

Conclusion

Our paper has advanced a postcolonial intervention in the ethics of embodied-AI systems by critically examining the normative, epistemic, and sociopolitical assumptions embedded in prevailing governance approaches of these technologies. Through interpretive and speculative analyses, we demonstrated how these frameworks, though often positioned as universal, fall short when applied in Global South contexts, where histories of colonialism, infrastructural inequality, and epistemic exclusion demand greater sensitivity to local values, knowledge, and social configurations. To address these limitations, we proposed four alternative ethical principles: epistemic non-imposition, onto-contextual consistency, respect for agentic boundaries, and embodied spatial justice. These principles not only augment existing ethical paradigms but also offer a scaffold for more inclusive and equitable governance of AI-powered robots. Our operationalization of these principles across task, dataset, and deployment levels illustrated how such values can be made actionable, while real-world examples showcased their feasibility and relevance in practice. Crucially, we argued that these ethical considerations are not limited to traditional robots alone. Their relevance extends across a broader ecosystem of embodied AI systems—from wearables and smart assistants to drones and generative tools—whose sociotechnical impact is profoundly shaped by the environments they inhabit and the communities they engage with. We also highlighted the need for situated participatory evaluation methods that move beyond checklist-style audits to embrace ethical pluralism and cultural embeddedness.

Limitations and Future Work

Our analysis is concentrated on a carefully selected set of globally recognized ethical frameworks for robotics and AI. This selection was informed by institutional legitimacy, conceptual influence, and relevance to embodied AI. However, we acknowledge that our scope was not comprehensive. Many regionally-grounded or community-developed ethical models remain unexplored in this study. Future work should examine such frameworks to broaden the normative basis for ethical robotics. Moreover, our critique was primarily analytical and interpretive in nature, focusing on documented principles and their epistemic assumptions. Although this allowed us to identify systemic limitations and propose alternative ethical considerations, our study did not involve direct engagement with robotics practitioners, policymakers, or affected communities in the Global South. Future research would benefit from empirical studies that involve these stakeholders through interviews, workshops, or participatory design methods. Such engagement would help to contextualize our principles further and assess their applicability in practice.

Positionality Statement

We acknowledge our positions as researchers working at the intersection of robotics, AI, and human-computer interaction (HCI), with experience ranging from early-career to decades-long engagements across these fields. All four authors are currently based in Western Europe. Atmadeep Ghoshal is of South Asian origin, with lived experience in South Asia, and brings regionally grounded perspectives to his work in robotics, AI, and HCI. Martim Brandão is from Western Europe with no lived experience in the Global South nor research expertise in the Global South. He is an

expert in AI with extensive research on evaluations and implementations of fairness, explainability, and other aspects of Responsible AI. Ruba Abu-Salma is from the Middle East (West Asia) and specializes in HCI and human-centered privacy research, with a focus on at-risk populations. Sanjay Modgil is of South Asian ethnicity and has some lived experience in South Asia. He has expertise in the logical foundations of AI and in the philosophy and ethics of AI.

Acknowledgments

This work has been supported by the AI Security Institute (grant number: UKRI841), United Kingdom.

References

Acemoglu, D.; and Restrepo, P. 2019. Automation and New Tasks: How Technology Displaces and Reinstates Labor. *Journal of Economic Perspectives*, 33(2): 3–30.

Anderson, M.; and Anderson, S. L. 2007. Machine Ethics: Creating an Ethical Intelligent Agent. *AI Mag.*, 28(4): 15–26.

Arora, P. 2019. *The Next Billion Users*. London, England: Harvard University Press.

Asaro, P. M. 2006. What Should We Want From a Robot Ethic? *International Review of Information Ethics*, 6: 9–16.

Ashwini, B.; Ghoshal, A.; Suri, V. R.; Achary, K.; and Shukla, J. 2024. "It looks useful, works just fine, but will it replace me?" Understanding Special Educators' Perception of Social Robots for Autism Care in India. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24. New York, NY, USA: Association for Computing Machinery. ISBN 9798400703300.

Awad, E.; Dsouza, S.; Kim, R.; Schulz, J.; Henrich, J.; Shariff, A.; Bonnefon, J.-F.; and Rahwan, I. 2018. The Moral Machine experiment. *Nature*, 563(7729): 59–64.

Azeem, R.; Hundt, A.; Mansouri, M.; and Brandão, M. 2024. LLM-Driven Robots Risk Enacting Discrimination, Violence, and Unlawful Actions. arXiv:2406.08824.

Basu, R. 2019. *Public Administration in the 21st Century: A Global South Perspective*. Routledge India. ISBN 9780367140106.

Benjamin, R. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. Oxford, England: Polity Press.

Bilgic, A.; and Pilcher, J. 2022. Desires, Fantasies and Hierarchies: Postcolonial Status Anxiety through Ontological Security. *Alternatives: Global, Local, Political*, 48(1): 3–19.

Birhane, A. 2021. Algorithmic injustice: a relational ethics approach. *Patterns*, 2(2): 100205.

Birhane, A.; Ruane, E.; Laurent, T.; S. Brown, M.; Flowers, J.; Ventresque, A.; and L. Dancy, C. 2022. The Forgotten Margins of AI Ethics. In *2022 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '22, 948–958. ACM.

Booysen, T.; Rieger, M.; and Ferrein, A. 2011. Towards Inexpensive Robots for Science & Technology Teaching and Education in Africa. In *IEEE Africon '11*, 1–6.

Braun, V.; and Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2): 77–101.

Brezis, E. S.; and Rubin, A. 2023. Will automation and robotics lead to more inequality? *The Manchester School*, 92(3): 209–230.

Brooks, R. A. 1991. Intelligence without representation. *Artificial Intelligence*, 47(1–3): 139–159.

Broussard, M. 2018. *Artificial Unintelligence: How Computers Misunderstand the World.* The MIT Press.

Buolamwini, J.; and Gebru, T. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In Friedler, S. A.; and Wilson, C., eds., *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, volume 81 of *Proceedings of Machine Learning Research*, 77–91. PMLR.

Chrisley, R. 2003. Embodied artificial intelligence. *Artificial Intelligence*, 149(1): 131–150.

Chun, W. H.; and Papanikolopoulos, N. 2016. *Robot Surveillance and Security*, 1605–1626. Springer International Publishing. ISBN 9783319325521.

Clarke, A. 2005. *Situational Analysis*. SAGE Publications, Inc. ISBN 9781412985833.

Coeckelbergh, M. 2022. Democracy, epistemic agency, and AI: political epistemology in times of artificial intelligence. *AI and Ethics*, 3(4): 1341–1350.

Couldry, N.; and Mejias, U. A. 2020. The Costs of Connection: How Data Are Colonizing Human Life and Appropriating It for Capitalism. *Social Forces*, 99(1): e6–e6.

Council of the European Union. 2016. Council Regulation (EU) 2016/72 of 22 January 2016 fixing the fishing opportunities for certain fish stocks. *Official Journal of the European Union*. Pp. 1–161.

Crawford, K. 2022. *Atlas of AI*. New Haven, CT: Yale University Press.

Crocker, D. A.; and Robeyns, I. 2009. *Capability and Agency*, 60–90. Cambridge University Press. ISBN 9780511800511.

Cuccu, L.; and Royuela, V. 2024. Just reallocated? Robots displacement, and job quality. *British Journal of Industrial Relations*, 62(4): 705–731.

Dale, K.; and Latham, Y. 2014. Ethics and entangled embodiment: Bodies-materialities-organization. *Organization*, 22(2): 166–182.

Datey, I.; and Zytko, D. 2024. "Just Like, Risking Your Life Here": Participatory Design of User Interactions with Risk Detection AI to Prevent Online-to-Offline Harm Through Dating Apps. *Proceedings of the ACM on Human-Computer Interaction*, 8(CSCW2): 1–41.

Dautenhahn, K. 2007. Socially intelligent robots: dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480): 679–704.

Dautenhahn, K. 2014. Human-Robot Interaction. In *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* Interaction Design Foundation. Accessed: 2025-05-10.

Dautenhahn, K.; Woods, S.; Kaouri, C.; Walters, M.; Koay, K. L.; and Werry, I. 2005. What is a robot companion - friend, assistant or butler? In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1192–1197. IEEE.

Doshi-Velez, F.; and Kim, B. 2017. Towards A Rigorous Science of Interpretable Machine Learning. arXiv:1702.08608.

Dourish, P.; and Mainwaring, S. D. 2012. Ubicomp's colonial impulse. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, 133–142. New York, NY, USA: Association for Computing Machinery. ISBN 9781450312240.

Escobar, A. 2018. *Designs for the Pluriverse: Radical Interdependence, Autonomy, and the Making of Worlds*. Duke University Press. ISBN 9780822371816.

Finn, R. L.; and Wright, D. 2012. Unmanned aircraft systems: Surveillance, ethics and privacy in civil applications. *Computer Law & Security Review*, 28(2): 184–194.

Floridi, L.; Cowls, J.; Beltrametti, M.; Chatila, R.; Chazerand, P.; Dignum, V.; Luetge, C.; Madelin, R.; Pagallo, U.; Rossi, F.; Schafer, B.; Valcke, P.; and Vayena, E. 2018. AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, 28(4): 689–707.

Fong, T.; Nourbakhsh, I.; and Dautenhahn, K. 2003. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3–4): 143–166.

Fosch-Villaronga, E.; and Giraudo, M. 2022. Innovation Letter: Experimenting with Competing Techno-Legal Standards for Robotics. *Journal of Law*, Market & Innovation: Vol. 1 No. 3 (2022): 3/2022.

Fricker, M. 2007. *Epistemic Injustice*. Oxford, England: Clarendon Press.

Friedman, B.; and Hendry, D. G. 2019. *Value Sensitive Design*. The MIT Press. London, England: MIT Press.

Gabriel, I. 2020. Artificial Intelligence, Values, and Alignment. *Minds and Machines*, 30(3): 411–437.

Gain, E. 2025. Kerala's 52-day trawling ban to begin on June 9. *LiveMint*. 5 June 2025.

Giddens, A. 1991. *Modernity and Self-Identity*. Palo Alto, CA: Stanford University Press.

Gomes, O.; and Pereira, S. 2019. On the economic consequences of automation and robotics. *Journal of Economic and Administrative Sciences*, 36(2): 135–154.

Government of India. 1976. Territorial Waters, Continental Shelf, Exclusive Economic Zone and Other Maritime Zones Act, 1976. Act No. 80 of 1976.

Guiochet, J.; Machin, M.; and Waeselynck, H. 2017. Safety-critical advanced robots: A survey. *Robotics and Autonomous Systems*, 94: 43–52.

Haddadin, S. 2015. *Physical Safety in Robotics*, 249–271. Springer Fachmedien Wiesbaden. ISBN 9783658099947.

Hagendorff, T. 2020. The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines*, 30(1): 99–120.

Harding, S. 1998. *Is Science Multicultural?* Race, Gender, and Science. Bloomington, MN: Indiana University Press.

Helm, P.; Bella, G.; Koch, G.; and Giunchiglia, F. 2023. Diversity and Language Technology: How Techno-Linguistic Bias Can Cause Epistemic Injustice. arXiv:2307.13714.

Henriques, G. C.; Passaro, A. M.; Garcia Alvarado, R.; González Böhme, L. F.; Quitral Zapata, F. J.; and Araya, S. 2024. Southern creative robotics – A collaborative framework to implement computational design, fabrication and robotics. *International Journal of Architectural Computing*, 23(1): 153–173.

Hohenthal, J.; and Ruuska, T. 2024. Disclosing the sacred in technological practices for sustainability. *Acta Borealia*, 41: 31 – 43.

Howard, A.; and Borenstein, J. 2017. The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity. *Science and Engineering Ethics*, 24(5): 1521–1536.

Irani, L.; and Philip, K. 2018. Negotiating Engines of Difference. *Catalyst: Feminism, Theory, Technoscience*, 4(2): 1–11.

Kalyanakrishnan, S.; Panicker, R. A.; Natarajan, S.; and Rao, S. 2018. Opportunities and Challenges for Artificial Intelligence in India. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '18, 164–170. New York, NY, USA: Association for Computing Machinery. ISBN 9781450360128.

Kohn, E. 2015. Anthropology of Ontologies. *Annu. Rev. Anthropol.*, 44(1): 311–327.

Lee, M. K.; Kiesler, S.; Forlizzi, J.; and Rybski, P. 2012. Ripple Effects of an Embedded Social Agent : A Field Study of a Social Robot in the Workplace. In *Proceedings of SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*, 695 – 704.

Lin, P.; Abney, K.; and Bekey, G. A., eds. 2014. *Robot ethics: the ethical and social implications of robotics*. Intelligent Robotics and Autonomous Agents series. London, England: MIT Press.

Mackenzie, C. 2021. *Relational Autonomy*, 374–384. Oxford University Press. ISBN 9780190628956.

Madaio, M. A.; Stark, L.; Wortman Vaughan, J.; and Wallach, H. 2020. Co-Designing Checklists to Understand Organizational Challenges and Opportunities around Fairness in AI. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, 1–14. ACM.

Matthias, A. 2004. The Responsibility Gap. Ascribing Responsibility for the Actions of Learning Automata. *Ethics and Information Technology*, 6(3): 175–183.

Mbembe, A. 2001. *On the Postcolony*. Studies on the History of Society and Culture. Berkeley, CA: University of California Press.

McQuillan, D. 2022. *Resisting AI*. Bristol, England: Bristol University Press.

Mehrabi, N.; Morstatter, F.; Saxena, N.; Lerman, K.; and Galstyan, A. 2021. A Survey on Bias and Fairness in Machine Learning. *ACM Comput. Surv.*, 54(6).

Mhlambi, S.; and Tiribelli, S. 2024. Correction: Decolonizing AI Ethics: Relational Autonomy as a Means to Counter AI Harms. *Topoi*.

Mitchell, M.; Wu, S.; Zaldivar, A.; Barnes, P.; Vasserman, L.; Hutchinson, B.; Spitzer, E.; Raji, I. D.; and Gebru, T. 2019. Model Cards for Model Reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, 220–229. New York, NY, USA: Association for Computing Machinery. ISBN 9781450361255.

Mittelstadt, B. 2019. Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11): 501–507.

Mohamed, S.; Png, M.-T.; and Isaac, W. 2020. Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. *Philosophy & Technology*, 33(4): 659–684.

Moor, J. 2006. The Nature, Importance, and Difficulty of Machine Ethics. *IEEE Intelligent Systems*, 21(4): 18–21.

Morris, R., ed. 2009. *Can the Subaltern Speak?* New York, NY: Columbia University Press.

Mutongi, C.; and Rigava, B. 2024. Robotics in Healthcare: The African Perspective. In *2024 IEEE 3rd International Conference on AI in Cybersecurity (ICAIC)*, 1–10.

Narayanan, M.; and Schoeberl, C. 2023. A Matrix for Selecting Responsible AI Frameworks. Technical report, Georgetown University.

Ofosu-Asare, Y. 2024. Cognitive imperialism in artificial intelligence: counteracting bias with indigenous epistemologies. *AI & SOCIETY*.

Okolo, C. T.; Kamath, S.; Dell, N.; and Vashistha, A. 2021. "It cannot do all of my work": Community Health Worker Perceptions of AI-Enabled Mobile Health Applications in Rural India. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21. New York, NY, USA: Association for Computing Machinery. ISBN 9781450380966.

Palmerini, E.; Bertolini, A.; Battaglia, F.; Koops, B.-J.; Carnevale, A.; and Salvini, P. 2016. RoboLaw: Towards a European framework for robotics regulation. *Robotics and Autonomous Systems*, 86: 78–85.

Prem, E. 2023. From ethical AI frameworks to tools: a review of approaches. *AI and Ethics*, 3(3): 699–716.

Ricaurte, P. 2019. Data Epistemologies, The Coloniality of Power, and Resistance. *Television & New Media*, 20(4): 350–365.

Russell, S. 2020. *Human Compatible*. Harlow, England: Penguin Books.

Sambasivan, N.; Arnesen, E.; Hutchinson, B.; Doshi, T.; and Prabhakaran, V. 2021. Re-imagining Algorithmic Fairness in India and Beyond. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21, 315–328. New York, NY, USA: Association for Computing Machinery. ISBN 9781450383097.

Seale, C. 1999. *The Quality of Qualitative Research*. SAGE Publications Ltd. ISBN 9780857020093.

Sharkey, N. 2008. The Ethical Frontiers of Robotics. *Science*, 322(5909): 1800–1801.

Sharkey, N. E. 2011. The Automation and Proliferation of Military Drones and the Protection of Civilians. *Law, Innovation and Technology*, 3: 229 – 240.

Shiraev, E. B.; and Levy, D. A. 2015. *Cross-Cultural Psychology: Critical Thinking and Contemporary Applications, Fifth Edition*. Routledge. ISBN 9781315664439.

Sloane, M.; Moss, E.; Awomolo, O.; and Forlano, L. 2022. Participation Is not a Design Fix for Machine Learning. In *Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, EAAMO '22. New York, NY, USA: Association for Computing Machinery. ISBN 9781450394772.

Sparrow, R.; and Howard, M. 2020. Robots in agriculture: prospects, impacts, ethics, and policy. *Precision Agriculture*, 22(3): 818–833.

Stilgoe, J.; Owen, R.; and Macnaghten, P. 2013. Developing a framework for responsible innovation. *Research Policy*, 42(9): 1568–1580.

Straub, I. 2022. Embodied Intelligence in Physical, Social and Technological Environments. *IOP Conference Series: Materials Science and Engineering*, 1261(1): 012024.

Suchman, L. 2006. *Human-Machine Reconfigurations: Plans and Situated Actions*. Cambridge University Press. ISBN 9780511808418.

The New Indian Express. 2023. After trawl ban, fishing operations prohibited beyond territorial waters. *The New Indian Express*. 14 July 2023.

Timmermans, S.; and Tavory, I. 2012. Theory Construction in Qualitative Research: From Grounded Theory to Abductive Analysis. *Sociological Theory*, 30(3): 167–186.

van Wynsberghe, A. 2021. *Responsible Robotics and Responsibility Attribution*, 239–249. Springer International Publishing. ISBN 9783030541736.

Varshney, K. R. 2025. *Decolonial AI Alignment: Openness, Visesa-Dharma, and Including Excluded Knowledges*, 1467–1481. AAAI Press.

Verhulst, S. G. 2023. Operationalizing digital self-determination. *Data & Policy*, 5.

Watkins, E. A.; Moss, E.; Metcalf, J.; Singh, R.; and Elish, M. C. 2021. Governing Algorithmic Systems with Impact Assessments: Six Observations. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '21, 1010–1022. New York, NY, USA: Association for Computing Machinery. ISBN 9781450384735.

Weckert, J., ed. 2019. *Computer Ethics*. The International Library of Essays in Public and Professional Ethics. London, England: Routledge.

Widder, D. G. 2024. Epistemic Power in AI Ethics Labor: Legitimizing Located Complaints. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '24, 1295–1304. ACM.

Winkle, K.; Jackson, R. B.; Bejarano, A.; and Williams, T. 2021. On the Flexibility of Robot Social Identity Performance: Benefits, Ethical Risks and Open Research Questions for HRI. In *Workshop on Robo-Identity: Artificial Identity and Multi-Embodiment at HRI 2021*.

Zhu, Y.; Wen, R.; and Williams, T. 2024. Robots for Social Justice (R4SJ): Toward a More Equitable Practice of Human-Robot Interaction. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '24, 850–859. New York, NY, USA: Association for Computing Machinery. ISBN 9798400703225.