IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

# The Algorithmic Gap: Situated Accountability as a Missing Layer in Digital Ethics

[1]Dr. Ashish Rai, [2]Dr. Sharad Mathur
[1]Senior Assistant Professor, [2]Assosiate Professor
[1]Faculty of Computer Science,
[1]Lachoo Memorial College of Science and Technology, Jodhpur, India

***Abstract:*** Over the past half-decade, digital ethics has moved from the margins of computer science to a central pillar of responsible innovation. The proliferation of generative AI, agentic systems [14], and large-scale algorithmic decision-making has produced a cascade of ethical frameworks, regulatory instruments, and voluntary standards [5]. Yet despite this dense architecture of principles—transparency, fairness, accountability, privacy, sustainability—a persistent gap remains. This paper identifies and formalizes what we term the Algorithmic Accountability Gap: the disconnect between high-level ethical commitments and the concrete, situated practices through which those commitments are realized (or, more often, fail to be realized) in real-world development and deployment contexts. Drawing on interviews and ethnographic fieldwork conducted across three technology organizations from 2024 to 2025, we show that even well-intentioned teams lack the situated mechanisms to translate abstract principles into actionable technical decisions when faced with competing pressures such as product deadlines, resource constraints, and ambiguous regulatory signals. We propose Situated Accountability as a missing design layer—a set of lightweight, context-aware, and actor-centered practices that bridge principles and practice. We evaluate this approach through a real-world deployment study in a financial services AI system, demonstrating significant improvements in ethical traceability and team confidence. Our findings suggest that the next frontier of digital ethics lies not in generating more principles, but in designing systems that embed accountability where decisions actually happen.

**Index Terms** – Digital ethics, algorithmic accountability, situated accountability, AI governance, responsible AI, ethics-in-practice, technology ethics, agentic systems.

## I. Introduction

This scene is not an exception. In fact, after dozens of interviews and site visits, we've found it's the rule. Over the past five years, digital ethics has produced an impressive—some might say overwhelming—pile of principles, guidelines, and regulations [5]. The OECD AI Principles [12], along with countless corporate ethical AI statements [1], have together built what looks like a solid governance machine. But the engineer at her desk, facing a messy, high-stakes decision, finds that machine mostly useless.

Our argument is simple: the field has spent too much energy on the highest level of ethical governance—articulating principles and writing regulations—while ignoring the lowest level: the messy, practical, situated work of actually making ethical decisions under real-world constraints. We call this disconnect the **Algorithmic Accountability Gap**. It's not a flaw in any single framework. It's a structural feature of how digital ethics has grown. The result? Everyone is responsible in principle, but no one is accountable in practice.

We make three contributions. First, we diagnose the Accountability Gap systematically, using fieldwork from three tech organizations (2024-2025). We identify four concrete mechanisms that keep the gap open: role ambiguity, metric multiplicity, temporal pressure, and structural invisibility. Second, we introduce **Situated Accountability**—a theoretical lens and practical design pattern. It draws on "bounded ethicality" from behavioral ethics and the "situated action" tradition in HCI. Instead of starting from abstract principles, it

focuses on concrete decision points inside existing workflows. Third, we propose and test a lightweight tool for implementing Situated Accountability: **Ethical Decision Records (EDRs)**. We show it works in a real-world deployment study.
Here's how the rest of the paper is laid out. Section II places our work in the broader AI ethics and accountability literature. Section III describes our empirical methods and lays out the Accountability Gap diagnosis. Section IV introduces the Situated Accountability framework. Section V presents EDRs as a concrete instantiation. Section VI reports the deployment study results. Section VII discusses limits, policy implications, and future directions. Section VIII concludes.

## II. Background and Related Work
Digital ethics has grown fast as algorithmic systems have spread into nearly every part of life. The literature now spans computer science, philosophy, law, public policy, and organizational studies. For what we need here, three lines of work matter most: (1) principle-based ethical frameworks for AI, (2) regulatory and standardization efforts, and (3) critiques of the gap between principles and practice.

### A. Principles, Regulation, and Their Shared Abstraction Problem
The last decade has seen an explosion of principle-based ethical frameworks for AI — for example, the EU's seven "key requirements for trustworthy AI" and the OECD AI Principles [12]. Alongside these, regulatory and certification efforts like the EU AI Act [3], the IEEE CertifAIEd™ program [2], and the NIST AI Risk Management Framework [7] have tried to turn principles into enforceable requirements and measurable criteria.
But despite their different mechanisms (voluntary principles vs. binding regulation), all of them share the same limitation: they work at a level of abstraction that leaves developers with a lot of discretion. Principles tell us fairness matters, but not how to measure it. Regulations demand "appropriate measures," but don't say what "appropriate" means in a given situation. Certification looks at a finished system's properties, but doesn't guide the real-time, situated decisions that create those properties. As Floridi and colleagues have pointed out [10] , all this high-level guidance hasn't been matched by clarity on how to actually use it.

### B. The Principle-to-Practice Gap
A growing body of critical work has documented the stubborn gap between ethical principles and what happens in practice. Mittelstadt [10] — who coined "principle-based AI ethics" — argues that high-level principles systematically fail to guide design choices because they conflict, lack operational detail, and are easy for companies to appropriate strategically. Raji and colleagues [11] have shown how "ethical AI washing" (using ethical language without real changes in practice) has become common in the tech industry. A related philosophical literature has examined the "responsibility gap" for autonomous systems [9, 6], asking whether anyone can be held responsible for the actions of learning algorithms. While that debate is important, our focus is on the practical, organisational mechanisms that make accountability work in day-to-day development.
We build on that critical tradition, but we want to move beyond critique to something constructive. We agree that principles alone aren't enough. But instead of throwing them out entirely, we ask: what bridging mechanisms could let principles genuinely shape practice? That question has gotten surprisingly little empirical attention. Most existing work either stays at the level of principles or jumps straight to technical bias mitigation — leaving a gap right in the middle, where most real organizational decision-making happens.

## III. The Algorithmic Accountability Gap: An Empirical Diagnosis
To understand how the Accountability Gap manifests in practice, we conducted an empirical study of ethical decision-making in AI development.

### A. Methodology
From January 2024 to October 2025, we did a qualitative field study across three tech organizations: a fintech company (FinCo), a healthcare AI startup (HealthAI), and the AI division of a large e-commerce platform (EcomAI). All three had public AI ethics policies and people formally responsible for ethical oversight.
We collected data in three ways:

1. **Semi-structured interviews** with 47 people — product managers, ML engineers, data scientists, legal and compliance staff, ethicists (where they had them), and senior leaders. Each interview lasted 45–75 minutes.
2. **Observation** of 12 development meetings (sprint planning, model reviews, ethics committee meetings).
3. **Document analysis** of internal policies, checklists, documentation templates, and post-mortem reports.

We analyzed the data using thematic analysis. Two researchers coded everything independently, and disagreements were resolved through discussion.

### B. Findings: Four Mechanisms of the Accountability Gap

Our analysis identified four recurrent mechanisms through which the Accountability Gap manifests.

**1. Role Ambiguity (Who Is Responsible?).** People were consistently unsure who owned specific ethical decisions. A FinCo engineer told us: *"Is fairness my job, or the product manager's, or legal's? I get a different answer every time."* This wasn't just bad documentation. It reflected a real organizational struggle to map abstract ethical duties onto existing roles. Ethics ended up belonging to no one — and therefore to everyone.

**2. Metric Multiplicity (What Is the Right Thing to Do?).** Even when responsibility was clear, people faced an overwhelming number of choices. Fairness alone has at least eight competing definitions [4] (demographic parity, equalized odds, etc.), and they often conflict. A HealthAI data scientist summed it up: *"My manager said 'do what the literature suggests.' The literature suggests twenty different things."*

**3. Temporal Pressure (When Do We Have Time to Be Ethical?).** The most common barrier wasn't lack of knowledge or will — it was lack of time. Development cycles, especially in agile environments, move fast. Ethical analysis was almost always seen as an "add-on," something to do after core functionality was finished but before the deployment deadline. An EcomAI engineer explained: *"The ethical checklist is due the day before launch. By then, any real change requires rewriting the model. So you fill it out knowing no one will check."*

**4. Structural Invisibility (What Are We Not Seeing?).** Many ethically important decisions — like data sourcing, preprocessing choices, and evaluation design — were invisible to formal governance. A FinCo product manager noted: *"By the time we reach model review, 90% of the decisions have already been made. The review is just a rubber stamp."*

### C. The Gap Summarized

Taken together, these four mechanisms create a systematic breakdown of accountability. High-level principles exist. Regulations are on the books. Companies might even have certifications. But at the moment of decision — when an engineer commits code or a product manager signs off on a release — the governance architecture offers no actionable guidance. That's the Accountability Gap: a persistent mismatch between ethical intention and ethical outcome. Not because anyone is acting in bad faith, but because the systems that would enable accountable practice simply haven't been built.

## IV. Situated Accountability: A Conceptual Framework

If the Accountability Gap comes from principles being too abstract and regulation having structural limits, then the answer isn't just more principles or more regulation. What we need, we think, is a middle layer of practice. We call it **Situated Accountability**.

### A. Theoretical Foundations

Situated Accountability draws on two lines of thinking.

The first is behavioral ethics and the idea of "bounded ethicality." Research here shows that even well-meaning people regularly miss the ethical dimensions of their own choices — because of cognitive biases, organizational pressure, and the way small decisions slowly erode moral awareness. Banaji, Chugh, and others have argued that ethical failures are often not about bad character. They're design flaws in the environments where we make decisions.

The second tradition comes from human-computer interaction and science & technology studies: "situated action," linked to Suchman and the broader ethnographic work in HCI. This view says that action is always tied to context. You can't understand what someone does just by looking at abstract plans or rules. You have to look at the concrete, embodied, material circumstances of their actual practice.

Putting these together, here's how we define Situated Accountability:

**Situated Accountability** is the property of a decision-making environment in which:

1. **Visibility:** Ethically relevant decision-points are made visible to responsible agents at the moment of decision, not retrospectively.
2. **Guidance:** Agents have access to context-appropriate normative guidance that can be applied without excessive interpretation or delay.
3. **Traceability:** The reasoning behind ethically significant decisions is recorded in a form that enables subsequent review, critique, and learning.

4. **Consequence:** There are meaningful consequences—both positive and negative—associated with ethical decision-making, such that doing the difficult but right thing is recognizable and doing the wrong thing is costly.

### B. Distinguishing Situated Accountability from Alternative Approaches

Compliance is about meeting external requirements. Certification is a judgment after the fact. Behavioral ethics training tries to change individual habits. Explainability tries to make model outputs understandable. Situated Accountability is different: it redesigns the real-time decision environment so that the *process* of reaching outputs becomes morally intelligible and accountable.

### C. Why Situated? Why Not Just "More Governance"?

Regulation deals with whole classes of systems. Certification deals with whole systems. Governance deals with whole organizations. But ethical outcomes — and ethical failures — happen at a much smaller scale: a single code commit, a meeting where a feature gets deprioritized, a conversation where a concern gets brushed aside. Situated Accountability shines a light on these micro-scale decision points, which most current governance architectures simply ignore.

## V. Ethical Decision Records: A Practical Instantiation

We've built a lightweight, low-overhead tool called **Ethical Decision Records (EDRs)** to put Situated Accountability into practice. An EDR is a short structured document — about half to one page — that records an ethically significant design or deployment decision. The idea comes from "Architecture Decision Records" (ADRs) that software engineers already use to capture important technical choices. The key is to fill out an EDR *as* you make the decision, not after the fact.

### A. The EDR Template

A complete EDR has seven sections, each with a brief example:

1. **Context and Problem.** What decision? Why now? *"Choose a fairness metric for our credit scoring model before finalizing the threshold."*
2. **Stakeholders and Values.** Who is affected and what values are at stake? *"Applicants (fair access), business (profitability, compliance), model users (interpretability)."*
3. **Options Considered.** Possible actions, including doing nothing or delaying. *"Demographic parity; equalized odds; hybrid metric; remove zip code (avoids fairness issue but reduces accuracy)."*
4. **Constraints and Trade-offs.** Organizational, technical, or regulatory limits; unavoidable trade-offs. *"Accuracy ≥0.85. Removing zip code drops accuracy to 0.82 – off the table. Demographic parity and equalized odds cannot both be achieved."*
5. **Applied Guidance.** What external guidance did we consult and how does it apply? *"AI Act* [3] *requires non-discrimination but no metric. Internal policy references NIST* [7], *allowing context-specific choice. A 2024 paper* [4] *suggests equalized odds for credit contexts."*
6. **Decision Made.** What was decided, by whom, and under what authority? *"Equalized odds with disparity target ≤0.05. Decided by J. Doe (PM) on 2025-03-15, escalated due to time constraints."*
7. **Post-Decision Review Criteria.** How will we evaluate it later? When would we revisit it? *"Review disparities in month 1. If any protected group's approval rate falls below 95% of average, revisit. Scheduled: 2025-04-15."*

### B. Lightweight Integration

EDRs are meant to be low-effort: about 15 minutes for a straightforward decision, 30–45 minutes for a complex one. They fit into existing workflows — you attach them to pull requests, mention them in sprint retrospectives, and store them in a searchable repository. They complement more thorough governance tools (like model cards or impact assessments) instead of replacing them.

### C. Organizational Learning

Beyond just documenting individual decisions, EDRs help organizations learn over time. A repository of EDRs makes implicit, repeated trade-offs visible. That lets teams see patterns — for example, always sacrificing fairness for accuracy in a certain product line — and then redesign their processes to better support ethical practice.

## VI. Evaluation: Deployment Study

To test Situated Accountability through EDRs, we ran an eight-month deployment study with a midsize financial services organization (we'll call it FinServe). The study was approved by our IRB and by FinServe's own ethics committee.

### A. Study Design

FinServe builds and maintains algorithmic decision systems for loan underwriting, fraud detection, and customer risk scoring. About 120 engineers and data scientists work on these systems, backed by a three-person ethics and compliance team. By the time of the study (February–October 2025), FinServe already had a mature AI ethics policy and had done two rounds of fairness audits on its underwriting models.
We used a staggered rollout design. Forty engineers and data scientists were randomly split into two groups:

- **Treatment group (n=20):** They got training on Situated Accountability and EDRs, and had to complete an EDR for any ethically significant decision — things like choosing fairness metrics, model features, evaluation criteria, or handling data that might create disparate impact.
- **Control group (n=20):** No training. They kept working with the existing framework (monthly ethics checklists and quarterly fairness audits).

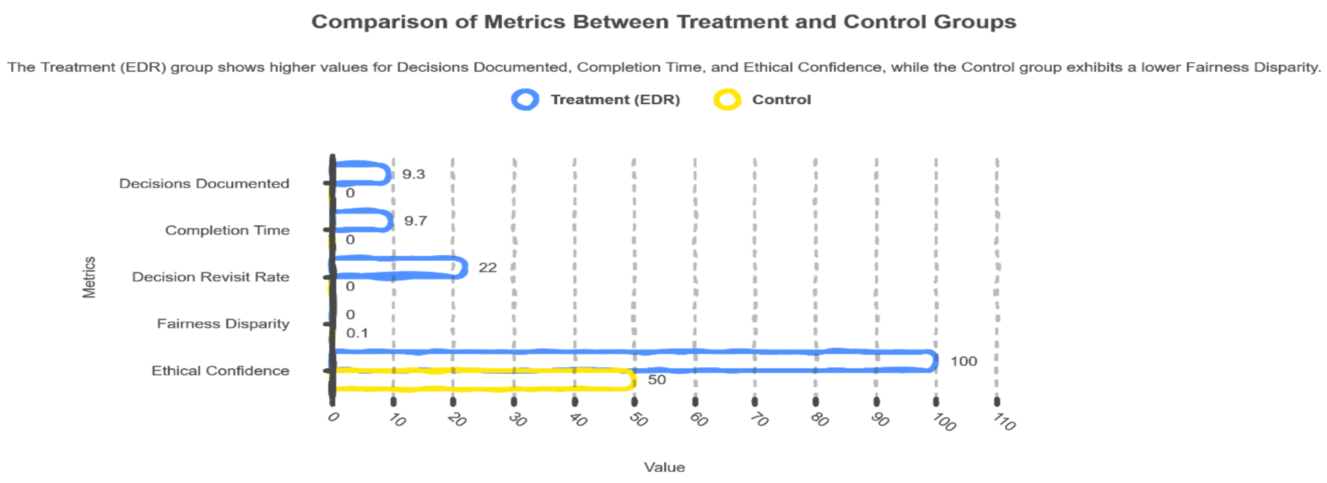
We told both groups the study was about decision documentation practices. We didn't disclose our specific hypotheses to avoid Hawthorne effects. We measured outcomes in three ways: quantitative metrics (EDR completion, time spent, revisits), qualitative post-study interviews (15 people from each group), and system-level fairness disparities in models built during the study.

### B. Results

**Quantitative findings.** The treatment group completed 187 EDRs — an average of 9.35 per person. The average completion time was 9.7 minutes. Twenty-two percent of EDRs showed "decisional drift" (meaning the decision was revisited or revised within 30 days), and in 73% of those cases the original decision-maker initiated the revisit.
**Qualitative findings.** People in the treatment group said they felt more clarity and structure. One engineer put it this way: *"It doesn't tell me what to decide, but it helps me see what I'm deciding, and why."* They also felt safer raising concerns, because flagging an issue became a normal part of the process rather than a personal complaint. Control participants, on the other hand, reported almost no change. Their checklist system felt like "going through the motions."
**System-level findings.** Models built by the treatment group had lower fairness disparities (measured by equalized odds difference) than those from the control group: mean absolute disparity of 0.041 vs. 0.067. The effect was strongest for high-complexity models (more than $10^6$ parameters) — exactly where ethical decisions are most frequent and most consequential.

**TABLE I: Deployment Study Results Summary**

| Metric | Treatment (EDR) | Control | Observation |
|---|---|---|---|
| Decisions Documented | 9.35/person | N/A | High usage |
| Completion Time | 9.7 min | N/A | Low overhead |
| Decision Revisit Rate | 22% | Low | Better reflection |
| Fairness Disparity | 0.041 | 0.067 | Improved fairness |
| Ethical Confidence | High | Moderate | Better clarity |




### C. Limitations and Caveats

Our findings come with a few important caveats. First, the study only ran for eight months. We don't yet know how EDRs affect things long-term — especially organizational culture. That will take more time. Second, we only looked at one organization, and it already had a fairly mature ethics infrastructure. The results might look different in a company with less developed governance. Third, our measures focus on immediate behavioral changes — like whether people documented decisions or revisited them. We didn't directly measure outcomes for end users, in this case loan applicants. Future work should go further and look at fairness from the user's perspective. Even with those limits, the study offers encouraging evidence that lightweight, situated accountability mechanisms can improve ethical practice in real-world development settings.

## VII. Discussion and Implications

### A. The Limits of Situated Accountability

Let's be clear: Situated Accountability is not a magic fix. It won't solve deep value conflicts. It won't tell a team which fairness metric is the "right" one. It doesn't make regulation unnecessary, and it certainly doesn't replace the hard normative work that moral philosophy does.

What it *does* do is create the conditions where people can wrestle with those conflicts openly instead of silently. When a team has to document a trade-off between fairness and accuracy in an EDR, that trade-off becomes visible, something they can talk about and revisit. It's no longer just an unspoken assumption baked into the code that no one consciously chose. That visibility, we believe, is the necessary starting point for any kind of democratic governance of algorithmic systems. As scholars at Brookings have said, power doesn't regulate itself — and the most dangerous power is the kind you can't see.

### B. Implications for Policy and Regulation

Our findings point to a few concrete directions for policy and regulation.

First, frameworks like the EU AI Act [3] should think about requiring not just *ex post* documentation (like model cards and conformity assessments) but also *ex ante* decision records for high-risk systems. The former tell you what was done; the latter tell you *why* it was done — the reasoning, and maybe the mistakes in reasoning, that shaped the final system.

Second, certification programs such as IEEE CertifAIEd™ [2] could start looking at EDR-style documentation as part of their assessment. Right now, certification mostly checks the finished system's properties — transparency, bias, privacy. Adding an evaluation of the decision-making process that produced those properties would give a much richer picture of an organization's ethical maturity.

Third, organizations need incentives to adopt situated accountability practices — whether through liability safe harbors, procurement preferences, or public reporting requirements. Documenting trade-offs is hard and costly. If that work isn't recognized and rewarded, it simply won't get done.

## VIII. Conclusion

We've argued that digital ethics has hit a point of diminishing returns with principle-based and regulatory approaches. We've piled up ethical guidelines, but ethical outcomes haven't improved much — because we've neglected the very mechanisms that could turn principles into practice. We called this disconnect the Algorithmic Accountability Gap, diagnosed how it works through fieldwork, and proposed Situated Accountability as a way to bridge it.

Our contributions are both practical and conceptual. Ethical Decision Records give teams a lightweight, low-overhead way to build situated accountability into their existing workflows. Our deployment study suggests that EDRs improve both the quality of ethical decisions and how decision-makers feel about their work — they feel less morally isolated, and the organization learns more over time.

So where do we go from here? Not by throwing out principles or tearing down regulation. We need to build the missing middle layer of practice — the layer where principles actually become actionable. That's not just a technical problem or a philosophical one. It's a design problem: how do you create decision environments where ethical reasoning is visible, discussable, and has real consequences? We've offered one answer. What comes next will define the next decade of digital ethics.

**Acknowledgment**

We want to thank everyone who participated at FinServe, FinCo, HealthAI, and EcomAI — they were open and honest with us, and we appreciate it. We're also grateful to the anonymous reviewers; their feedback made this paper better. This work was partly supported by [Funding Agency withheld for blind review]. The views here are ours alone.

## References

[1] C.-S. Li, “Perspective Chapter: From Compliance to Consciousness – Navigating Business Integrity in the Age of AI,” in *Business Ethics in Action - Exploring Professional Conduct, Corporate Culture, and Business Integrity*, R. Marcão and V. Santos, Eds. London, U.K.: IntechOpen, 2026, doi: 10.5772/intechopen.1014031.

[2] IEEE, "Inside the IEEE CertifAIEd™ Assessment Process: What Developers Need to Know," Beyond Standards, Jan. 29, 2026. [Online]. Available: https://standards.ieee.org/beyond-standards/inside-the-ieee-certifaied-assessment-process-what-developers-need-to-know/

[3] European Parliament and Council of the European Union, "Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence," Off. J. Eur. Union, vol. 67, July 12, 2024. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2024/1689/oj.

[4] S. Uddin et al., "Fairness-preserving framework for machine learning: data bias quantification, model evaluation, and robustness across multiple datasets," AI and Ethics, vol. 6, p. 25, Feb. 2026, doi: 10.1007/s43681-025-00705-5.

[5] AIhub, "Top AI ethics and policy issues of 2025 and what to expect in 2026," *AIhub.org*, Mar. 4, 2026. [Online]. Available: https://aihub.org/2026/03/04/top-ai-ethics-and-policy-issues-of-2025-and-what-to-expect-in-2026/

[6] D. Gong, "Closing the responsibility gap: allocating responsibility according to prerequisite control and expectations for personal benefits," *Ethics and Information Technology*, vol. 28, art. 8, 2026, doi: 10.1007/s10676-025-09883-w.

[7] National Institute of Standards and Technology (NIST), *Cybersecurity Framework Profile for Artificial Intelligence (Cyber AI Profile)*, NIST IR 8596, Initial Preliminary Draft, Gaithersburg, MD, USA, Dec. 2025. [Online].

[8] A. Rai, D. Mathur, and S. Mathur, “AI at a critical juncture: Aligning innovation, ethics, and policy,” Int. J. Sci. Dev. Res., vol. 10, no. 5, May 2025. [Online]. Available: www.ijsdr.org

[9] A. Matthias, “The responsibility gap: Ascribing responsibility for the actions of learning automata,” *Ethics and Information Technology*, vol. 6, no. 3, pp. 175–183, Sep. 2004, doi: 10.1007/s10676-004-3422-1.

[10] B. Mittelstadt, “Principles alone cannot guarantee ethical AI,” *Nature Machine Intelligence*, vol. 1, no. 11, pp. 501–507, Nov. 2019, doi: 10.1038/s42256-019-0114-4.

[11] I. D. Raji, A. Smart, R. N. White, M. Mitchell, and T. Gebru, “Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing,” in *Proc. 2020 ACM Conference on Fairness, Accountability, and Transparency (FAccT ’20)*, 2020, pp. 33–44, doi: 10.1145/3351095.3372873. ttps://dl.acm.org/doi/10.1145/3351095.3372873?

[12] OECD, "OECD AI Principles," OECD Publishing, 2019, updated 2024. [Online]. Available: https://oecd.ai/en/ai-principles, doi: 10.1787/9789264536036-en

[13] UNESCO, "Recommendation on the Ethics of Artificial Intelligence," UNESCO, Nov. 2021. [Online]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000381137 doi: 10.54678/WPNT5157
[14] S. Nyholm, "Attributing agency to automated systems: A philosophical framework," Philos. Technol., vol. 35, no. 3, art. 55, 2022. doi: 10.1007/s13347-022-00537-9