

Operationalizing AI Ethics in the Public Sector: A Cross-Context Replication in Brazil

Edna Dias Canedo
University of Brasília (UnB)
Computer Science Department
Brasília-DF, Brazil
ednacanedo@unb.br

Fabiana Mendes
Aalto University
Espoo, Finland
fabiana.mendes@aalto.fi

Richardson B. da S. Andrade
University of Brasília (UnB)
Brasília-DF, Brazil
jcrbsa@gmail.com

José Siqueira de Cerqueira
Tampere University
Faculty of Information Technology
and Communication Sciences
Tampere, Finland
jose.siqueiradecerqueira@tuni.fi

Pekka Abrahamsson
Tampere University
Faculty of Information Technology
and Communication Sciences
Tampere, Finland
pekka.abrahamsson@tuni.fi

Abstract

Background: AI ethics encompasses principles such as privacy, fairness, transparency, accountability, and safety that guide the responsible design and use of AI systems. Responsible AI (RAI) translates these principles into actionable practices across the AI life-cycle. In Brazil, the General Data Protection Law (LGPD) provides a privacy baseline, but there is no national AI ethics framework; meanwhile, Generative AI (GenAI) introduces new socio-technical and governance risks. **Objective:** This study replicates and extends the research by Pant et al. [25] to examine how Brazilian public organizations perceive, interpret, and implement AI ethics. It aims to identify awareness levels, institutional governance maturity, and capability gaps in the context of GenAI adoption. **Method.** A mixed-method, cross-sectional survey was conducted with 87 civil servants from federal, state, and municipal agencies. The questionnaire adapted Pant et al.'s instrument to the Brazilian context, incorporating LGPD references and GenAI-specific items. Quantitative data were analyzed descriptively, while open-ended responses underwent qualitative content analysis with open coding and constant comparison. **Results.** Awareness of AI ethics is moderate and concentrated on compliance-oriented principles, whereas participatory dimensions such as Fairness and Contestability remain limited. Governance maturity is low: only 18.6% of organizations have dedicated ethics roles or committees, and over half report never conducting training. Perceived GenAI risks are high across ethics, privacy, and data protection. The main barriers include a lack of AI knowledge, the absence of AI-specific regulation for privacy and data protection, and limited tools to apply LGPD principles. A comparison with Pant et al. shows that, while both studies identify an awareness–practice gap, its cause in Brazil lies primarily in

institutional governance immaturity rather than practitioner capability. **Conclusion.** Brazilian public organizations demonstrate growing recognition of AI ethics but face structural barriers to operationalization. Advancing toward Responsible and Trustworthy AI requires institutional scaffolding—ethics roles, policies, and training programs—alongside regulatory clarification to align GenAI governance with LGPD principles. The study contributes cross-context empirical evidence on AI ethics governance in the Global South and outlines practical levers for embedding ethics-by-design in public-sector AI initiatives.

CCS Concepts

• **Security and privacy** → **Human and societal aspects of security and privacy**; *Social aspects of security and privacy*.

Keywords

AI Ethics, Responsible AI, Generative AI, Public Sector Governance, Data Protection

ACM Reference Format:

Edna Dias Canedo, Fabiana Mendes, Richardson B. da S. Andrade, José Siqueira de Cerqueira, and Pekka Abrahamsson. 2026. Operationalizing AI Ethics in the Public Sector: A Cross-Context Replication in Brazil. In *Proceedings of 19th International Conference on Cooperative and Human Aspects of Software Engineering (CHASE '2026)*. ACM, New York, NY, USA, 12 pages. <https://doi.org/XXXXXXX.XXXXXXX>

1 Introduction

Artificial Intelligence (AI) technologies are increasingly embedded in public services and decision-making processes worldwide [2, 22]. As AI systems gain autonomy and influence, concerns about their ethical, social, and legal implications have intensified [12, 29]. AI ethics encompasses a set of principles such as privacy, fairness, transparency, accountability, and safety aimed at ensuring that AI systems are developed and deployed responsibly [20]. Operationalizing these principles throughout the AI lifecycle is often referred to as Responsible AI (RAI). However, translating high-level ethical aspirations into actionable practices remains a persistent challenge for both developers and organizations [1, 5, 7, 15, 16].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CHASE '2026, Rio de Janeiro, Brazil

© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-XXXX-X/2026/06
<https://doi.org/XXXXXXX.XXXXXXX>

Over the past decade, numerous countries and organizations have released AI ethics guidelines [6], including the European Union's AI Act [23] and Australia's AI Ethics Principles.¹ These frameworks articulate expectations for trustworthy and human-centered AI but differ widely in scope, maturity, and enforcement. Despite their proliferation, empirical evidence shows that principles alone rarely translate into operational routines [19, 25, 26]. Developers and organizations often struggle with unclear accountability structures, limited ethical literacy, and a tendency to prioritize technical or business goals over societal concerns [4, 25].

Recent work has shown that ethical awareness and advocacy are not uniformly distributed across the software workforce. Olson et al. [24] found that practitioners' ability to identify and act upon ethical issues varies according to gender, race, socioeconomic background, and geographic location. Marginalized practitioners, particularly women, black, indigenous, people of color, and Global South professionals, reported stronger ethical sensitivity and willingness to intervene when ethical concerns arise, yet lacked institutional support to act effectively. These findings underscore that ethical capacity is socially situated and that advancing Responsible AI requires inclusive organizational structures that value diversity and empower ethical decision-making. This perspective complements prior evidence on the limits of principle-based frameworks by emphasizing the human and contextual dimensions of ethics integration.

Complementing these findings, Hinton [17] examined how AI ethics principles are implemented across eight Estonian public service organizations using the Value Sensitive Design framework. The study revealed that ethical considerations are often indirect and vary according to the maturity of AI systems and the ethical awareness of civil servants. While principles such as privacy, security, and accountability are partially embedded through GDPR compliance, broader ethical dimensions—such as justice, explicability, and human-centeredness—remain weakly institutionalized. Even in highly digitalized governments, embedding AI ethics in practice depends not only on technical expertise but also on governance structures and sustained competence building. This gap highlights the need for comparative studies in other public sector contexts, such as Brazil, where similar governance and capacity challenges persist amid the rapid adoption of Generative AI.

In Brazil, the legal landscape for AI governance is still emerging. The Brazilian General Data Protection Law (LGPD) [21] establishes foundational principles for privacy and data protection, but does not explicitly address AI-specific ethical or operational requirements. Several bills and strategy documents have called for a national AI framework, yet no official ethical principles or institutional governance structures for AI have been enacted. Meanwhile, the rapid diffusion of Generative AI (GenAI) in government and society has amplified ethical and privacy risks, posing new challenges for accountability, transparency, and public trust [10, 11]. This creates a unique context where ethical awareness and governance capacity coexist with regulatory uncertainty.

Understanding how public organizations perceive and implement ethical AI practices is, therefore, critical. Prior research has

primarily examined the views of AI practitioners in industry and academia [25], but far less is known about how public institutions—tasked with upholding citizens' rights and societal values translate ethical principles into governance routines. Public-sector AI projects often operate under additional constraints, including legal compliance, transparency obligations, and limited technical resources, making them a compelling setting for studying ethics in practice.

To address this gap, this study replicates and extends Pant et al. [25], adapting their survey on AI ethics awareness and challenges to the institutional and regulatory environment of the Brazilian public sector. The replication investigates two research questions: (i) the level of institutional awareness of AI ethics and (ii) the main challenges public organizations face in implementing ethical and responsible AI, particularly in the context of GenAI. By analyzing perceptions from 87 public servants across federal, state, and municipal agencies, the study provides empirical insights into how ethics is understood, governed, and operationalized in a country without established AI ethical principles.

The main contribution of this study is the design and execution of a contextual replication that extends Pant et al.'s original work to the Global South. It offers an evidence-based characterization of AI ethics awareness and governance maturity within Brazilian public institutions and highlights emerging risks associated with Generative AI. Building on these findings, the study provides actionable recommendations for policymakers, educators, and public-sector leaders to strengthen ethical governance, foster Responsible AI capacity, and inform future research on ethics operationalization in public organizations.

2 Related Work

Empirical investigations on how AI practitioners understand and apply ethical principles have proliferated in recent years. The study by Pant et al. [25] offered one of the first systematic analyses of AI professionals' awareness of ethical issues across the AI lifecycle. Through interviews with practitioners, they identified challenges such as limited ethical literacy, the absence of clear operational guidelines, and the tendency to prioritize technical or business goals over ethical reflection. Their work established an empirical foundation for understanding how ethics is perceived and practiced in real-world AI contexts.

Building on this foundation, Ibáñez and Olmeda [19] explored how organizations translate AI ethical principles into practice, revealing that ethical initiatives are often reactive and compliance-driven rather than embedded in design and development processes. Similarly, Sanderson et al. [26] interviewed developers and scientists from Australia's national science agency to assess how government-endorsed AI ethics principles were implemented in practice. Their findings demonstrated persistent tensions among ethical values such as privacy, transparency, and accountability, underscoring that formal principles alone are insufficient without institutional mechanisms and governance support.

Further extending this empirical line, Elia et al. [9] proposed a conceptual framework linking Responsible AI to human influence throughout the AI lifecycle, highlighting how cognitive and cultural factors shape ethical outcomes. Fraenkel [13] complemented these

¹<https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles/australias-ai-ethics-principles>

perspectives by proposing a maturity model for institutionalizing AI ethics, mapping organizational evolution from reactive compliance to proactive ethical integration. In parallel, Dotan et al. [8] proposed a maturity model for AI risk management grounded in the NIST AI Risk Management Framework (AI RMF). Their model translates abstract principles of Responsible AI into measurable practices across the lifecycle of AI systems—mapping, measuring, managing, and governing ethical risks. By introducing structured evaluation metrics and evidence-based scoring, the study advances the operationalization of AI ethics from declarative to actionable practices. These studies collectively underscore the importance of organizational maturity and evaluation mechanisms in advancing ethical AI governance.

Sattlegger and Bharosa [27] examined how ethical AI risks can be embedded within existing public sector risk management practices by applying the Three Lines of Defense (TLoD) model. Through a mixed-method approach combining surveys and interviews with professionals from Dutch government agencies, they found that ethical responsibilities are often diffused and poorly defined across organizational levels. While the TLoD model provides a useful lens for distributing accountability, its implementation tends to focus on compliance rather than proactive ethical reflection. Their findings highlight the importance of organizational mechanisms to manage ethical risks—an issue that remains particularly critical in contexts, such as Brazil, where no formal AI ethics guidelines exist. Complementary research has addressed ethical decision-making in managerial contexts: Manda et al. [22] emphasized leadership and governance mechanisms for responsible AI, while Usha et al. [28] examined fairness and accountability dilemmas in algorithmic decision-making contexts.

Bolgouras et al. [3] examined how key European regulatory instruments—such as the AI Act, GDPR, NIS2, Cyber Resilience Act, and Digital Services Act—interact to form a cohesive governance ecosystem for ethical AI. Their comparative analysis categorized security, functional, and non-functional requirements across frameworks, identifying convergence around principles of transparency, accountability, and user empowerment. The authors argue that harmonizing these regulations mitigates fragmentation and sets a global benchmark for trustworthy AI governance. In contrast, Brazil lacks comparable ethical AI regulations, highlighting the importance of empirical research into how practitioners interpret and operationalize ethical principles in practice.

From a national perspective, empirical studies in Brazil have begun to address related challenges in responsible and ethical AI practice. Falcão and Canedo [10] investigated Brazilian software developers' perceptions of data privacy when using Large Language Models (LLMs). Their survey revealed limited knowledge of privacy regulations, lack of formal training, and concerns about data misuse and transparency—issues that overlap with broader ethical challenges in AI practice. Their findings reinforce the need for educational, organizational, and regulatory mechanisms to operationalize ethical principles, aligning with the focus of our replication study.

Gonçalves et al. [14] analyzed the perceptions of Brazilian executives regarding trust in AI within organizations. Their mixed-method study identified that while business leaders recognize AI's potential to improve operational efficiency and decision-making,

concerns about privacy, ethics, and the absence of responsible AI strategies limit their trust in AI. The findings underscore the importance of training and institutional mechanisms to promote ethical and responsible AI adoption—issues that resonate with practitioners' challenges observed in our study.

Siqueira de Cerqueira et al. [5] proposed the RE4AI Ethical Guide, a Design Science Research artifact aimed at supporting AI development teams in eliciting ethical requirements within agile projects. The guide operationalizes eleven ethical principles through a set of 26 interactive cards containing reflective questions, examples, and tool suggestions to help practitioners translate ethical values into user stories. Its empirical evaluation demonstrated increased ethical awareness and practical applicability, bridging the gap between abstract ethical principles and daily software engineering practices. This work is particularly relevant as it addresses the same challenge—operationalizing AI ethics in practice—that motivates our replication study.

Taken together, the literature indicates growing attention to how AI ethics is operationalized across diverse organizational and cultural contexts. However, most prior work has focused on regions—such as Europe, the United States, and Australia—where national frameworks for AI ethics already exist. In contrast, Brazil lacks formally established ethical principles or governance guidelines for AI, despite the rapid adoption of AI technologies in public and private sectors. This replication and extension study helps fill this gap by examining how practitioners in a context without institutionalized ethical frameworks perceive, interpret, and implement ethical principles in their professional practice.

3 Research Methodology

This study analyzes the perceptions and challenges faced by public sector institutions in Brazil regarding Ethics in Artificial Intelligence, with particular emphasis on the governance and use of Generative AI. It constitutes a conceptual replication of Pant et al. [25], originally conducted with AI practitioners from multiple countries, and adapts their survey to the institutional and governance context of the Brazilian Public Administration.

The replication maintains the same conceptual foundation as Pant et al. [25], who explored two central research questions (RQs): (i) the degree of awareness of AI practitioners regarding ethics, and (ii) the challenges they face when integrating ethical principles into AI systems. In our adaptation, the focus is extended to public institutions, examining organizational practices, training, and governance mechanisms related to AI ethics, particularly in the context of generative AI adoption in government. Accordingly, our study addresses the following research questions:

RQ1: How aware are public sector institutions of ethical principles and risks associated with generative artificial intelligence? This question investigates the level of institutional familiarity with ethical principles in AI, including training, internal policies, and perceived preparedness to apply ethics in practice. It also examines the presence of formal structures such as ethics committees, ordinances, or codes of conduct that support responsible AI use.

RQ2: What are the main challenges faced by public institutions in implementing ethical practices for the use and governance of generative AI?

This second question explores barriers related to human, organizational, regulatory, and technical aspects that hinder the effective adoption of ethical and responsible AI principles in public sector environments.

The objective of this replication study is to assess whether there are differences in perceptions of AI ethics between the Brazilian public sector and findings reported in other nations. It also aims to identify specific governance challenges and contextual factors that influence the ethical use of AI within government institutions.

3.1 Survey Design

The survey was designed as a descriptive, cross-sectional study using both closed-ended and open-ended questions to assess institutional awareness (RQ1) and identify implementation challenges (RQ2). The survey instrument was adapted from Pant et al. [25] to reflect the Brazilian legal and organizational context. The adaptation involved rephrasing items originally targeted at individual AI practitioners to institutional language relevant for public agencies, adding references to governance frameworks, codes of ethics, and data protection regulations. Additionally, new items were introduced to capture perceptions related to the ethical, privacy, and data protection implications of Generative AI adoption in the public sector.

The final instrument comprised eight sections and thirty-one questions, and it is available in our Zenodo package. Sections I–VI address institutional characteristics, governance mechanisms, awareness of ethical principles, and perception of risks associated to Generative AI models, thereby responding to RQ1. Sections VII–VIII focus on perceived risks and barriers to ethical AI implementation, responding to RQ2. Closed-ended questions employed single-choice, multiple-choice, or Likert-type scales, allowing descriptive statistical analysis, while open-ended items enabled qualitative exploration of perceptions and experiences.

We conduct a pilot with five information security and data protection officers from federal institutions to ensure clarity, completeness, and contextual relevance. Minor revisions were made to adapt terminology and eliminate redundancies.

The differences between our questionnaire and the one developed by Pant et al. [25] stem from the distinct contextual focus of this replication. While the original study targeted individual AI practitioners to understand personal awareness and challenges, our adaptation was designed for public sector institutions, where ethical, privacy, and data protection issues are inherently shaped by legal and governance frameworks. In the Brazilian context, the LGPD[21] establishes privacy and personal data protection as fundamental rights, making these dimensions inseparable from ethical AI discussions and institutional accountability mechanisms. Therefore, additional sections addressing governance, regulatory compliance (LGPD and Access to Information Law—LAI), and organizational culture were introduced to capture systemic factors influencing ethical AI adoption.

Information about age, gender, and educational level was not collected because the unit of analysis in this replication is the institution rather than the individual, and the questionnaire sought to characterize organizational practices and governance structures rather than personal attributes. Finally, the alternatives for the questions related to challenges and barriers were derived from the original instrument by Pant et al. [25], complemented by elements from international AI ethics frameworks (e.g., OECD, UNESCO, NIST AI RMF) and Brazilian regulatory guidelines.

3.2 Data Collection

Data collection followed a similar procedure to Pant et al. [25], using an online structured questionnaire distributed through official and professional channels. Data were collected between June and October 2025. The target population included civil servants and staff members from federal, state, and municipal public agencies involved in AI-related activities, information security, data protection, or digital governance. Participants were recruited via institutional e-mails obtained from official websites and through social media posts on professional networks such as LinkedIn and X (formerly Twitter). Participation was voluntary and anonymous, and respondents were informed about the study's objectives and data handling procedures in compliance with the Brazilian LGPD (General Data Protection Law) [21].

A total of **87 valid responses** were collected after data cleaning. Responses were obtained from a variety of Brazilian agencies, representing a diverse range of organizational types, including ministries, oversight agencies (e.g., TCU, CGU), federal universities, and public foundations. Although not probabilistic, this sample offers a heterogeneous view of the institutional landscape of AI ethics governance across the country.

3.3 Data Analysis

Following the approach used by Pant et al. [25], data analysis combined quantitative and qualitative techniques to answer RQ1 and RQ2. Closed-ended items were analyzed using descriptive statistics, including frequency distributions and percentage calculations. Results were organized by survey section to provide an overview of institutional awareness levels (RQ1) and challenges encountered (RQ2). Open-ended responses were examined through qualitative content analysis using open coding to identify recurring themes, patterns, and illustrative excerpts [18]. Codes were iteratively grouped into categories reflecting dimensions such as organizational culture, resource limitations, governance maturity, and technical barriers. Triangulation between quantitative trends and qualitative insights strengthened the interpretation of findings and allowed comparison with results from other nations reported in Pant et al. [25].

4 Findings

A total of 87 professionals answered the survey, all of them working in public administration agencies across different levels of government in Brazil. The vast majority of participants (96.5%) work in the Federal Public Administration, followed by a smaller proportion from State (3.5%) and Municipal administrations. This distribution reflects the current concentration of artificial intelligence and data

governance initiatives within federal institutions, which typically lead national digital transformation and AI ethics programs ².

Regarding work areas, nearly half of the respondents (48.8%) are linked to Information Technology units, followed by Information Security (15.1%), Governance and Compliance (14%), and *Personal Data Protection* (9.3%). A smaller portion reported involvement directly with *Artificial Intelligence* (7%) or in *other administrative or technical areas* (5.8%). This distribution suggests that, while AI-related responsibilities are still emerging, most ethical and privacy considerations in the public sector remain managed within IT and governance structures.

The respondents' institutional diversity ensures an overview of how ethics in AI is being discussed and operationalized within the Brazilian Public Administration, particularly at the federal level, where policy-making and regulatory frameworks for responsible AI are more consolidated.

4.1 RQ1 – Institutional Awareness of AI Ethic

We assessed institutional awareness of AI ethics by examining two aspects related to employees' exposure to the topic: their familiarity with it and their experience. Additionally, we investigated the respondents' perceptions of the risks associated with generative AI models. Lastly, we inquired about the actions taken by the organization to promote the ethical use of AI and to address related incidents, which we refer to as AI Structure and Governance. The following sections provide detailed information on each of these four aspects.

Familiarity with AI ethics. In relation to familiarity with the concept of AI ethics, 13.8% of respondents reported being very familiar, while 29.9% considered themselves familiar and 25.3% somewhat familiar. On the other side of the scale, 28.7% reported being slightly familiar, and 2.3% indicated no familiarity at all. These results suggest that while there is some awareness of AI ethics in the Brazilian public sector, this knowledge tends to be superficial. The predominance of intermediate levels of familiarity reflects the overall profile of respondents, who mainly work in IT, information security, and data protection roles, rather than in specialized AI or ethics governance functions.

Regarding how respondents became aware of AI ethics, the majority cited news and media as their source (56.3%), highlighting the significant role that public discourse and journalistic coverage play in shaping perceptions of AI ethics. A substantial number of respondents also indicated that their awareness originated from workplace rules and policies (41.4%), reflecting the importance of organizational governance mechanisms in promoting ethical reflection—often linked to compliance practices, data protection programs, and information security protocols.

Additionally, several participants reported that they became aware of AI ethics through personal interest or self-study (45.3%), indicating a growing individual engagement with the subject beyond institutional mandates. Smaller yet meaningful proportions

mentioned direct professional experience with AI systems (32.2%) or personal encounters as users of digital technologies as triggers for their awareness. A few isolated responses referred to practical events, such as client complaints (1.2%), suggesting that awareness can also develop reactively when ethical implications become evident in day-to-day work.

To further explore the sources of awareness, we asked if respondents had received formal education related to AI ethics. Most participants (44.2%) indicated that they had not received any formal training or education in this area. Among those who did report some form of training, postgraduate studies such as master's, doctoral, or specialization courses were the most common (25.6%), followed by independent short courses or professional certifications (14.9%). A smaller group (8.3%) noted that they had participated in internal training sessions provided by their organizations, reflecting isolated institutional efforts to address ethical issues in AI, often in conjunction with data protection and governance initiatives.

Overall, the results suggest that awareness of AI ethics among public sector professionals is primarily driven by external communication and institutional regulation rather than by formal training or dedicated capacity-building initiatives. This pattern reinforces the understanding that ethical comprehension of AI within Brazilian public administration remains informal, largely media-driven, and compliance-oriented, consistent with prior findings by Pant et al. [25].

Finally, in relation to the familiarity with the Australian AI Ethics Principles, the most widely recognized principle was Privacy protection and security, cited by 73.6% of participants, followed by Reliability and safety (60.9%) and Accountability (51.7%). Intermediate levels of familiarity were observed for Human-centered values (46%) and Transparency and explainability (44.8%), indicating a growing awareness of ethical aspects related to openness and respect for human autonomy. By contrast, principles such as Human, societal, and environmental well-being (32.2%), Fairness (31%), and particularly Contestability (21.8%) were less frequently mentioned. These numbers are shown in Figure 1.

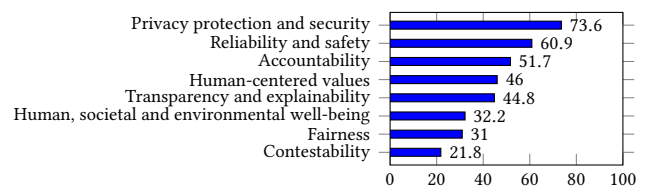


Figure 1: Awareness of AI Ethics Principles.

By analyzing the numbers in Figure 1, we can learn that the ethical awareness of public sector professionals in Brazil is predominantly focused on operational and data-centric concerns, whereas participatory and justice-oriented dimensions of AI ethics remain less consolidated. The distribution closely mirrors findings from Pant et al. [25], where technical and compliance-related principles similarly dominated practitioners' awareness across other national contexts.

Experience with AI. We gathered general data on experiences in AI-related fields and their roles. Furthermore, to enhance our

²Brazilian government policies, guides, and laws about digital transformation and AI ethics (in Brazilian Portuguese language): <https://www.gov.br/governodigital/pt-br/estrategias-e-governanca-digital/transformacao-digital> and <https://www.gov.br/governodigital/pt-br/infraestrutura-nacional-dados/inteligencia-artificial-1/publicacoes/framework-para-a-autoavaliacao-de-impacto-etico-em-inteligencia-artificial-no-setor-publico-federal>

analysis, we also inquired about experience in AI-based software development and the specific roles they held.

A substantial proportion of respondents (29.9%) indicated that they had no professional experience at all in areas related to artificial intelligence, 20.7% reported less than one year, and 32.2% reported between one and two years of experience. The remaining respondents had more than three years of experience, which was distributed across five different experience ranges. Among those with some professional involvement in AI, the most common roles were AI/ML Specialist (7.2%), Data Scientist (7.0%), AI Developer (8.1%), and AI Engineer (3.5%). A smaller share identified as AI/ML Professionals (4.7%) or AI Designers (1.2%).

These results indicate that while a significant portion of respondents possess foundational exposure to AI-related functions, there is still a limited concentration of specialized roles within the Brazilian public sector.

Regarding experience with AI-based software development, most respondents (62.7%) reported having no prior experience in this area. Among those with some background, 14% indicated having between one and two years of experience, while 14% reported less than one year. A smaller proportion had three to five years (7%) or six to ten years (2.3%) of experience.

Among those with practical engagement, the most frequently mentioned tasks were data collection (25.6%), model evaluation (23.3%), model training (17.4%), and software requirements specification (15.1%). Activities associated with data preprocessing—such as data cleaning (11.6%) and data labeling (10.5%)—were also reported, although less frequently. Fewer participants mentioned participating in model deployment (11.6%) or monitoring (9.3%), which suggests that end-to-end AI lifecycle management is still uncommon across public institutions. Activities related to governance and compliance with AI (2.3%) or strategic project management (1.2%) were rarely cited.

These findings indicate that AI development expertise is still emerging among professionals in the Brazilian public sector. Although a minority has accumulated more extensive experience, the overall distribution reveals that most institutions are in the early stages of capacity building for AI-related projects. This limited technical maturity underscores the importance of investing in training, institutional structures, and ethical governance mechanisms to ensure that future AI initiatives are aligned with responsible and transparent practices.

Perception of risks associated with generative AI models. We asked the respondents about the severity level of risks associated lack of control over Generative AI models, focusing on ethics, privacy, and personal data protection. Table 1 summarizes the results we got.

For all risk types, most respondents rated the lack of control in generative AI Models as very high or high. These results indicate that ethical privacy and data protection risks are widely recognized as a serious concern in the context of generative AI. The results emphasize a prevailing sense of institutional alertness regarding the ethical, privacy, and data protection challenges posed by generative AI in both personal and professional contexts.

Table 1: Perceived severity of lack of control in generative AI

Risk Severity	Risk Type		
	Ethics	Privacy	Data Protection
Very High	30.2%	38.4%	43%
High	39.5%	33.7%	38.4%
Moderate	18.6%	19.8%	11.6%
Low	9.3%	5.8%	4.7%
Very Low	2.4%	2.3%	2.3%

AI Structure and Governance. We assessed the AI structure and Governance by asking about the implementation of four elements in the respondent organization: (1) a code of ethics or conduct, (2) a designated person or committee responsible for AI ethics, (3) an internal ordinance or normative act concerning AI ethics, and (4) an internal committee or working group that addresses AI ethics-related issues. Table 2 summarizes the responses we received.

Table 2: Perception on the implementation of elements of AI Structure and Governance

AI Governance Element	Percentage of Answers			
	Yes	In Progress	No	Don't Know
Code of ethics or conduct	64.4%	8%	12.6%	14.9%
Responsible for AI ethics	19.5%	16.1%	42.5%	21.8%
Internal normative on AI ethics	12.6%	23%	40.2%	24.1%
Internal committee for AI ethics	17.2%	16.1%	32.2%	34.5%

Regarding the implementation of a code of ethics or conduct, the results indicate a relatively mature ethical foundation across Brazilian public institutions, especially within the federal administration, where such instruments are mandated by public governance and transparency regulations. However, approximately 15% of respondents reported being unaware of the existence of these documents, suggesting gaps in internal communication or limited distribution of institutional ethical guidelines, even in agencies where they may be formally established.

In contrast, the implementation of governance elements directly related to AI ethics is less encouraging. For all three elements investigated, the percentage of respondents reporting their existence was below 20%. Additionally, over 30% of respondents indicated that these elements are not implemented in their organizations. These findings collectively highlight a fragmented and early-stage institutionalization of ethical governance in AI across the Brazilian public sector, with few organizations exhibiting consolidated structures for oversight, coordination, and ethical deliberation.

We also asked the respondents about receiving internal training on AI ethics. More than half of the participants (55.8%) reported that such training has never been conducted. A smaller group (16.3%) mentioned that training sessions are held occasionally, typically once a year or less, while only 5.8% stated that training takes place regularly, more than once per year. Additionally, 22.1% of respondents were not aware of whether their organization promotes any AI ethics training activities. These findings suggest that institutional capacity-building initiatives related to AI ethics remain incipient across most public organizations.

Finally, we asked questions about incident management related to privacy, personal data protection, and other AI ethics principles over the past five years. The results reveal that most organizations have not yet internally recorded or notified authorities about such incidents, indicating a limited level of institutional maturity in monitoring ethical risks associated with artificial intelligence. For privacy-related incidents, over half of the respondents stated that their institutions had not recorded or reported any occurrences involving the use of AI. Similarly, nearly half were unable to provide this information, reflecting either an absence of structured processes for incident management or a lack of transparency in internal communication channels. Only a few participants confirmed that privacy incidents had been registered or communicated to competent authorities.

The same pattern was observed for incidents concerning personal data protection. While a small number of respondents reported that their organizations had recorded or reported such cases, the majority indicated either the absence of incidents or uncertainty about whether these had occurred. This result suggests that although compliance with Brazil’s General Data Protection Law (LGPD) has become a priority, mechanisms for identifying and documenting AI-specific incidents are still in early development across public institutions.

When asked about incidents related to other AI ethics principles most respondents reported that no such events had been registered or were unaware of their existence. Only a very limited number of respondents replied that their organization has formally reported these occurrences to any competent authority. This highlights a gap between the recognition of ethical principles and their operationalization in practice, suggesting that ethical oversight mechanisms remain largely reactive rather than preventive. Taken together, these findings point to an overall lack of systematic procedures for recording and escalating ethical and data-related incidents in AI contexts within Brazilian public institutions, as shown in Table 3.

Table 3: Organizational Incident management related to privacy, data protection, and other AI-ethics.

Type of Incident	Organizational Action	Percentage of Responses		
		Yes	No	Don't know
Privacy	Internal Record	3.5%	41.9%	45.2%
	Authorities Notification	2.3%	40.7%	48.8%
Data	Internal Record	4.7%	52.3%	45.3%
Protection	Authorities Notification	2.3%	50%	45.3%
AI Ethics Principles	Internal Record	7%	48.8%	52.3%
	Authorities Notification	3.5%	51.2%	57%

RQ.1 Summary: Organizations have the presence of ethical instruments, but awareness of AI ethics remains modest and uneven. Recognition is skewed toward compliance-oriented principles (privacy, reliability, accountability), whereas participatory and justice-oriented dimensions (such as contestability) are less salient. Formal capacity-building is limited, and self-assessed readiness is mostly intermediate, mirroring early-stage professional exposure to AI. Organizational institutionalization is incipient—few designated leads, committees, or formal policies, and training is rare. Findings reveal that the implementation of ethical and privacy practices in AI within Brazilian public agencies remains at an early stage, characterized by limited incident monitoring, weak governance structures, and fragmented accountability mechanisms.

4.2 RQ2 – Challenges to AI Implementation

We investigated the challenges and barriers to implementing AI by using both closed and open-ended questions. Our focus was on three aspects: ethics, privacy, and data protection. Each aspect is discussed in the sections that follow.

Challenges and Barriers related to Ethics in AI. When asked about the main barriers to incorporating ethics into AI. The percentage of responses for each alternative is shown in Table 4.

Table 4: Challenges and barriers to incorporating ethics in AI.

Challenge or Barrier	Percentage of Responses
Lack of AI knowledge/understanding	74.4%
AI complexity	37.2%
Difficulty predicting consequences	36%
Poor data quality	26.7%
Difficulty predicting outputs	24.4%
Lack of awareness of others' work	20.9%
Lack of training data	16.3%
Scope of AI	14%
Other	8.4%

Note in Table 4 that the most chosen alternative is the lack of AI knowledge and understanding (74.4%). Participants emphasized that insufficient conceptual and technical literacy about AI limits the ability to anticipate ethical consequences or evaluate the behavior of models in complex or sensitive contexts. This finding is confirmed by the second and third most chosen alternatives, “complexity of AI systems” and “difficulty in predicting the consequences of AI.” This finding reinforces the need for organizations to invest in training initiatives to implement ethics in AI successfully.

In response to the open-ended questions, the respondents recognized substantial obstacles in translating ethical principles into operational mechanisms. The most recurrent themes included the difficulty of operationalizing ethics in AI practice, particularly the need to embed ethical reflection from the design phase (“ethics by design”), the absence of clear institutional structures or governance

models, and the emergence of new ethical risks such as misinformation, deepfakes, and manipulation of public opinion. Participants also pointed to a general lack of institutional knowledge and awareness, indicating that AI ethics remains a relatively unexplored and fragmented domain within public administration. As respondent BR#42 highlighted:

“Generative AI has brought new challenges, such as deepfakes, fake news, and manipulation of public opinion. In this case, ethics involves not only thinking about what technology can do, but also how to prevent its abusive use in social and political contexts.”

This statement reflects a recurring concern across responses — that the ethical implications of AI are not limited to compliance or governance structures, but extend to the societal impacts of emerging technologies, reinforcing the need for proactive, institutionally supported mechanisms to anticipate and mitigate ethical risks.

Challenges and Barriers related to Privacy in AI. When addressing privacy, respondents highlighted structural and organizational barriers that go beyond technical implementation, as shown in Table 5. The absence of clear guidelines or regulatory frameworks for AI privacy (65.1%) emerged as the most recurrent concern, followed by insufficient staff training (55.8%) and the difficulty of applying anonymization or data minimization techniques in practice (54.7%). Participants also pointed to organizational cultures (53.5%) that fail to prioritize privacy as a central design value, often due to the additional cost and effort required to implement Privacy by Design strategies. Together, these results reveal a gap between legal principles and their operationalization in AI projects, underscoring the need for clearer regulation and institutional capacity building.

Table 5: Challenges and barriers to incorporating privacy in AI.

Challenge or Barrier	Percentage of Responses
Lack of clear privacy/AI regulation	65.5%
Teams lack privacy/data training	55.8%
Technical difficulty (anonymization/minimization)	54.7%
Org. culture not prioritizing privacy	53.5%
Lack of mature privacy-preserving tools	37.2%
High cost/resources (Privacy by Design)	34.9%
Balancing performance vs privacy	26.9%
Other	3.6%

In the open-ended questions, the respondents also expressed concerns about regulatory clarity, institutional readiness, and cultural awareness. Participants frequently pointed to a lack of clear guidelines or legal certainty on how privacy principles apply to AI, especially regarding data processed by cloud services and global technology providers. Several respondents also mentioned gaps in institutional information and weak enforcement by the Brazilian National Data Protection Authority (ANPD), underscoring the need for stronger oversight and more explicit regulatory mechanisms. Another prominent theme concerned cultural and educational aspects, including insufficient public awareness of the risks of sharing personal or business data in AI prompts, and the need for a deeper societal understanding of how AI technologies affect

privacy. From a technical standpoint, participants discussed challenges in data transparency and user control, such as uncertainty about data storage locations, hidden data collection (“passive privacy”), and blurred boundaries between security and surveillance. As respondent (BR#42) observed:

“Many AI systems capture data in the background (voice, image, navigation patterns, biometrics) without users realizing it. This ‘passive privacy’ is an ethical and legal challenge. AI systems used for security often blur the line between protection and invasion of privacy.”

This reflection encapsulates the dual nature of the challenge identified by many respondents: ensuring that AI systems safeguard user privacy not only through compliance with legal requirements but also through transparent, accountable, and culturally informed design practices that address both active and passive data flows.

Challenges and Barriers related to Personal Data Protection in AI. In relation to the barriers to incorporating personal data protection in AI, participants highlighted challenges associated with both legal uncertainty and institutional readiness. Table 6 shows the percentage of responses for each alternative for the closed-ended question.

Table 6: Challenges and barriers to incorporating personal data protection in AI.

Challenge or Barrier	Percentage of Responses
Lack of legal clarity (PD in AI)	66.3%
Low team awareness/capacity	62.8%
Org. culture not prioritizing data protection	50%
Lack of tools to apply LGPD principles	46.5%
Technical limits (anonymization/minimization)	44.2%
Balancing performance vs protection	25.6%
High costs / limited resources	20.9%
Other	2.4%

Observe in Table 6 that the absence of specific regulations or legal clarity on how to apply data protection rules to AI systems emerged as the dominant barrier, followed by low levels of staff awareness and training on data protection. Organizational culture was also mentioned as a limiting factor (50.0%), alongside the lack of practical tools or methods to operationalize LGPD principles such as purpose limitation, necessity, and adequacy (46.5%). Overall, these findings indicate that while public institutions recognize the importance of personal data protection in AI, its effective implementation remains dependent on clearer legal guidance, technical support, and capacity building.

In relation to the related closed-ended question, from analysis of the responses, we identified five thematic categories related to how public institutions perceive the protection of personal data in AI systems. These categories include: Governance and Institutional Readiness, Technical and Legal Safeguards for Data Protection, Institutional Awareness and Capacity Development, Ethical Awareness and Culture, and Risks and Ethical Threats.

Under Governance and Institutional Readiness, respondents highlighted the lack of clear regulatory guidance on how to apply the LGPD and the Access to Information Law (LAI) to AI contexts.

Participants emphasized the need for institutional support mechanisms and cultural change to foster responsible data governance practices within public entities. Several responses also pointed to a persistent lack of institutional information, which limits the implementation of coherent privacy and data protection strategies across agencies. The Technical and Legal Safeguards for Data Protection category focused on ensuring compliance with LGPD principles such as purpose limitation, necessity, and transparency in AI development. Respondents stressed the need to adopt measures like anonymization, pseudonymization, access control, and encryption to reduce risks. Concerns were also raised about data reuse for multiple purposes (e.g., training, validation, model improvement), which can undermine the legitimacy of processing operations and challenge the principle of purpose limitation.

Institutional Awareness and Capacity Development emerged as a third major theme, with participants noting a lack of understanding about data sharing and storage practices, as well as limited professional training opportunities. Respondents recognized that effective data protection in AI depends not only on compliance mechanisms but also on strengthening the knowledge and technical capacity of staff members. In the Ethical Awareness and Culture category, participants reflected on the human and cultural aspects underlying data protection, observing that societal normalization of personal data exposure may weaken privacy-conscious behavior. They underscored the importance of promoting ethical awareness and human oversight in AI systems.

Finally, the Risks and Ethical Threats category captured respondents' concerns about the dangers posed by inadequate data protection in AI, including data leaks and the lack of technical safeguards to prevent unauthorized access or secondary use of personal information. As one respondent (BR#40) noted:

"The protection of personal data in Artificial Intelligence systems requires special attention, as these models generally depend on large volumes of sensitive information. One of the main challenges is ensuring that the principles of the LGPD — such as purpose, necessity, and transparency — are effectively applied in the development and use of AI. In addition, it is essential to adopt technical measures such as anonymization, pseudonymization, and access control, while also promoting an organizational culture focused on data protection."

Overall, these findings show that ensuring personal data protection in AI requires a combination of clear governance frameworks, legal and technical safeguards, and sustained institutional investment in training and ethical culture.

RQ.2 Summary: The key challenges include insufficient AI literacy, technical and data-related constraints, and institutional cultures that fail to prioritize ethics and privacy. Unclear regulatory guidance and limited capacity to operationalize privacy-by-design and data protection principles further hinder compliance with the LGPD. Qualitative evidence highlights emerging risks such as deepfakes and data misuse, emphasizing the urgency of strengthening ethical governance, staff training, and proactive oversight mechanisms. Overall, the results indicate that achieving responsible and trustworthy AI in the public sector requires coordinated legal, organizational, and technical efforts supported by sustained institutional commitment.

5 Discussion

This section interprets our findings (Section 4) in light of prior empirical evidence on AI ethics in practice, with a particular focus on the study by Pant et al. [25]. We discuss convergences and divergences across contexts, highlight institutional drivers behind the awareness–practice gap, examine the distinctive role of generative AI risks in the public sector, and outline implications for governance, capacity building, and future research.

Cross-context comparison: similar gaps, different causes.

Pant et al. [25] showed that practitioners are broadly aware of AI ethics but struggle to operationalize principles due to unclear guidance, competing priorities, and limited formal training. Our results confirm the same structural gap in the Brazilian public sector: awareness levels are predominantly intermediate (Figure 1), with stronger recognition of compliance-oriented principles (privacy, reliability, accountability) and weaker salience of justice-oriented and participatory dimensions (e.g., contestability). However, the causal mechanisms differ. While Pant et al. emphasize individual-level constraints (ethical literacy, tool support), our data reveal predominantly institutional constraints: scarce governance structures (Table 3), rare training programs, lack of formal policies, and fragmented accountability. In other words, where Pant et al. identify a practitioner-centered capability gap, we observe a governance-centered maturity gap.

Institutionalization of ethics: from policy statements to operational routines.

Our findings suggest that ethics in Brazilian public agencies remains largely declarative: codes of ethics are common (65.1%), yet concrete mechanisms—such as designated roles, committees, normative acts, and recurring training—are incipient (Section 4.1). This policy–practice gap echoes prior work on organizational maturity and AI governance [8, 27]. Dotan et al. [8] argue that translating Responsible AI (RAI) into practice requires measurable routines (mapping, measuring, managing, governing). Our data align with this proposition: the absence of routine-building artifacts (e.g., checklists, assurance activities, role definitions) correlates with low institutional readiness to embed ethics by design. Moreover, respondents frequently describe responsibility diffusion, mirroring public-sector risk governance challenges reported by Sattlegger and Bharosa [27].

Regulatory context matters: operating without a national AI ethics framework. Another salient divergence relative to Pant et al. [25] is the regulatory backdrop. Pant et al. aggregated multi-country perceptions where organizations often operate under or alongside explicit guidelines (e.g., sectoral principles or corporate policies). By contrast, our respondents repeatedly indicate lack of clear, AI-specific regulation as a primary barrier for privacy and personal data protection (Tables 5 and 6). This contrasts with the European ecosystem, where converging instruments (AI Act, GDPR, NIS2, CRA, DSA/DMA) offer a harmonized governance baseline [3]. In Brazil, LGPD [21] provides a robust privacy foundation, but respondents perceive legal uncertainty on AI-specific obligations, especially for generative AI and model lifecycle transparency. As a result, organizations default to compliance proxies (privacy, reliability) while justice-oriented principles (fairness, contestability) remain under-institutionalized.

Generative AI as an amplifier of perceived ethical risk. A distinctive contribution of our replication-extension is the explicit measurement of perceived risks in generative AI. Respondents report very high or high severity across ethics (69.7%), privacy (72.1%), and personal data protection (81.4%), indicating stronger risk salience than reported in Pant et al. [25] (which did not separately quantify generative AI risks). Qualitative responses further surface socio-technical threats (deepfakes, misinformation, influence operations) and practical uncertainties (cloud data flows, “passive privacy”, vendor opacity). These concerns support calls for operational artifacts that convert principles into requirements and routines (e.g., RE4AI cards for ethics elicitation [5]) and for competence frameworks and training pipelines tailored to the public sector [10, 14].

5.1 Implications for practice and policy

Embed governance scaffolding. Our results indicate that simply having codes of ethics is insufficient. Agencies should establish minimal scaffolding: (i) clear role assignment (ethics leads/committees), (ii) normative acts that mandate ethical checkpoints in the AI lifecycle, (iii) incident logging and escalation procedures that cover ethical harms beyond data breaches, and (iv) recurrent training with case-based exercises. These align with maturity-oriented guidance (e.g., NIST AI RMF operationalization [8]) and public-sector risk integration [27].

Operationalize ethics by design. To move from declarative to actionable ethics, teams should convert principles into requirements and work items (backlog stories, acceptance criteria, evidence artifacts). Artifacts such as RE4AI [5] can systematize ethics elicitation within agile planning, while checklists and assurance cases support accountability and auditability during deployment and monitoring.

Capacity building and leadership. Both our study and Pant et al. [25] point to training deficits. Public-sector competence frameworks [10, 14] and leadership engagement are essential to institutionalize ethics routines, reduce responsibility gaps, and sustain privacy-by-design/data-protection-by-design in AI projects.

Implications for policy. Given respondents’ emphasis on legal uncertainty, a near-term policy priority is to articulate AI-specific guidance that interfaces with LGPD (e.g., purpose limitation in

relation to model training, risk classification for use cases, documentation and transparency duties). International experiences suggest that harmonized instruments clarify expectations and reduce fragmentation [3]. Even before comprehensive regulation, sectoral guidelines and procurement standards can require ethics checkpoints and evidence logs across the AI lifecycle.

6 Limitations and Threats to Validity

Potential threats to this study’s validity were carefully examined throughout its design and execution. The first potential threat concerns the representativeness of the sample. Although the survey reached 87 respondents from diverse federal, state, and municipal organizations, participation was voluntary and non-probabilistic. Therefore, the results should be interpreted as indicative rather than statistically generalizable. To mitigate this limitation, recruitment targeted a broad range of public agencies with distinct missions and sizes, ensuring heterogeneous perspectives and enabling analytical generalization.

A second threat relates to possible response bias. Self-reported data may reflect social desirability or the overrepresentation of respondents with a preexisting interest in ethics and AI governance. This risk was mitigated through anonymous participation and neutral question phrasing, encouraging honest responses. In addition, triangulation between quantitative trends and qualitative insights reduced the risk of one-dimensional interpretations. The third limitation concerns construct alignment between this replication and the original study by Pant et al. [25]. While the instrument was based on their validated questionnaire, several items were adapted to the institutional and legal context of Brazilian public administration, and new questions were added to address Generative AI. These adaptations may constrain direct one-to-one comparison, but they extend the explanatory scope by incorporating governance, regulatory, and societal dimensions absent from the original study.

A fourth potential threat involves the reliability of analytical inferences. To strengthen validity, the survey instrument was pilot-tested with domain experts, and all procedures—data collection, coding, and aggregation—were transparently documented. Quantitative analysis relied on descriptive statistics, while qualitative coding followed a structured content analysis protocol, with the complete codebook publicly available on Zenodo for reproducibility. Despite these limitations, the replication design ensures strong methodological alignment with Pant et al.’s international study, supporting meaningful cross-context comparison. The combination of rigorous adaptation, transparency, and mixed-method triangulation enhances the credibility and analytical robustness of the findings, providing a reliable foundation for understanding how public organizations perceive and operationalize ethical AI in emerging Generative AI contexts.

7 Conclusion

This study replicated and extended Pant et al. [25] to examine how public institutions in Brazil perceive, interpret, and operationalize ethical principles in Artificial Intelligence, particularly in the context of Generative AI. While the original study explored the awareness and challenges of AI practitioners across countries, our replication adapted the survey to the institutional and governance

reality of the Brazilian Public Administration, where no formal national AI ethics framework yet exists.

Findings reveal that institutional awareness of AI ethics is emerging but uneven, with a focus on compliance-oriented principles such as privacy, reliability, and accountability. Justice- and participation-oriented principles (e.g., fairness and contestability) remain weakly internalized. Governance maturity is low: few organizations have designated roles or committees for AI ethics, formal ordinances are rare, and internal training or incident reporting mechanisms are incipient. These results point to a governance-centered maturity gap, in contrast to the practitioner-centered capability gap observed by Pant et al. [25].

Respondents also perceive high levels of ethical, privacy, and data-protection risk related to Generative AI, with the absence of AI-specific regulation and limited capacity to operationalize the LGPD amplifying these challenges. Together, these findings emphasize that responsible AI in the public sector requires not only ethical literacy but also institutional scaffolding—policies, roles, and accountability routines that translate principles into operational practices.

From a theoretical standpoint, this replication contributes empirical evidence on how AI ethics awareness and governance evolve in a context without national ethical guidelines, extending prior international findings to the Global South. Practically, it offers actionable insights for policymakers and organizations seeking to embed ethics-by-design in AI projects. Priority actions include: (i) creating clear governance structures and normative acts for AI ethics; (ii) integrating ethics checkpoints throughout the AI lifecycle; (iii) promoting capacity building for technical and managerial staff; and (iv) issuing sectoral guidance to align LGPD compliance with emerging AI governance standards.

Future work will deepen this research through longitudinal and comparative analyses to assess institutional learning and the evolution of AI ethics maturity over time. Additional studies could integrate interviews or focus groups with policy leaders to triangulate survey findings and refine practical tools for ethical governance. By addressing institutional, technical, and cultural dimensions simultaneously, Brazil can progress toward a proactive model of Responsible and Trustworthy AI in the public sector.

Data Availability

The material produced during the research, such as the survey form, the 87 responses, and the coding of the open questions are available on Zenodo at <https://doi.org/10.5281/zenodo.17382444>.

Acknowledgements

This work was supported by CONVERGENCE of Humans and Machines (220025) and the EVIL-AI “The identification and the mitigation of the negative effects of Artificial Intelligence Agents” (JAES/2024/EVIL-AI) projects by Jane and Aatos Erkkö Foundation and the “Multifaceted ripple effects and limitations of human-AI interplay at work, business and society (SYNTHETICA)” project (358714) by Research Council of Finland. We thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Grant N° 300883/2025-0.

References

- [1] Mamia Agbese, Rahul Mohanani, Arif Ali Khan, and Pekka Abrahamsson. 2023. Implementing AI Ethics: Making Sense of the Ethical Requirements. In *Proceedings of the 27th International Conference on Evaluation and Assessment in Software Engineering, EASE 2023, Oulu, Finland, June 14-16, 2023*. ACM, <https://doi.org/10.1145/3593434.3593453>, 62–71. doi:10.1145/3593434.3593453
- [2] Najla Abdullah Ahmed Albannai and Muhammad Mustafa Raziq. 2025. Navigating ethical, human-centric leadership in AI-driven organizations: a thematic literature review. *The Service Industries Journal* 45 (2025), 1–28.
- [3] Vaivos Bolgouras, Apostolis Zaras, Christian Leka, Ioannis Stylianou, Aristeidis Farao, and Christos Xenakis. 2025. Eu regulatory ecosystem for ethical AI. *AI and Ethics* 5 (2025), 1–18. doi:10.1007/s43681-025-00749-x
- [4] Nicholas Kluge Corrêa, James William Santos, Camila Galvão, Marcelo Pasetti, Dieine Estela Bernieri Schiavon, Faizah Naqvi, Robayet Hossain, and Nythamar de Oliveira. 2025. Crossing the principle-practice gap in AI ethics with ethical problem-solving. *AI Ethics* 5, 2 (2025), 1271–1288. doi:10.1007/S43681-024-00469-8
- [5] José Antonio Siqueira de Cerqueira, Anayran Pinheiro De Azevedo, Heloíse Acco Tives Leão, and Edna Dias Canedo. 2022. Guide for Artificial Intelligence Ethical Requirements Elicitation - REAAI Ethical Guide. In *55th Hawaii International Conference on System Sciences, HICSS 2022, Virtual Event / Maui, Hawaii, USA, January 4-7, 2022*. ScholarSpace, <http://hdl.handle.net/10125/80015>, 1–10.
- [6] José Antonio Siqueira de Cerqueira, Heloíse Acco Tives Leão, and Edna Dias Canedo. 2021. Ethical Guidelines and Principles in the Context of Artificial Intelligence. In *SBSI 2021: XVII Brazilian Symposium on Information Systems, Uberlândia, Brazil, June 7 - 10, 2021*, Rafael D. Araújo, Fabiano A. Dorça, Renata Mendes de Araujo, Sean W. M. Siqueira, and Awdren L. Fontão (Eds.). ACM, <https://doi.org/10.1145/3466933.3466969>, 361–368. doi:10.1145/3466933.3466969
- [7] Daniel de Paula Porto, Renata De Castro Vianna Prado, Gilmar dos Santos Marques, André Luiz Marques Serrano, Fábio L. L. Mendonça, and Edna Dias Canedo. 2025. Ethical Requirements in the Age of Artificial Intelligence: A Systematic Literature Review. In *Proceedings of the 21st Brazilian Symposium on Information Systems, SBSI 2025, Recife, Brazil, May 19-23, 2025*, Mônica Ximenes Carneiro da Cunha, Davi Viana, Rita Suzana Pitangueira Maciel, and Allysson Allex Araújo (Eds.). SBC, <https://doi.org/10.5753/sbsi.2025.246613>, 663–672. doi:10.5753/SBSI.2025.246613
- [8] Ravit Dotan, Borhane Bili-Hamelin, Ravi Madhavan, Jeanna Matthews, and Joshua Scarpino. 2024. Evolving AI Risk Management: A Maturity Model based on the NIST AI Risk Management Framework. *CoRR abs/2401.15229* (2024), 1–21. arXiv:2401.15229 doi:10.48550/ARXIV.2401.15229
- [9] Miriam Elia, Paula Ziethmann, Julia Krumme, Kerstin Schlögl-Flierl, and Bernhard Bauer. 2025. Responsible AI, ethics, and the AI lifecycle: how to consider the human influence? *AI Ethics* 5, 4 (2025), 4011–4028. doi:10.1007/S43681-025-00666-Z
- [10] Fabiano Damasceno Sousa Falcão and Edna Dias Canedo. 2024. Investigating Software Development Teams Members' Perceptions of Data Privacy in the Use of Large Language Models (LLMs). In *Proceedings of the XXIII Brazilian Symposium on Software Quality, SBQS 2024, Salvador, Bahia, Brazil, November 5-8, 2024*, Ivan Machado, José Carlos Maldonado, Tayana Conte, Edna Dias Canedo, Johnny Marques, Breno Bernard Nicolau de França, Patrícia Matsubara, Davi Viana, Sérgio Soares, Gleison Santos, Larissa Rocha, Bruno Gadelha, Rodrigo Pereira dos Santos, Ana Carolina Oran, and Adolfo Gustavo Serra Seca Neto (Eds.). ACM, <https://doi.org/10.1145/3701625.3701675>, 373–382. doi:10.1145/3701625.3701675
- [11] Mirko Farina, Xiao Yu, and Andrea Lavazza. 2025. Ethical considerations and policy interventions concerning the impact of generative AI tools in the economy and in society. *AI Ethics* 5, 1 (2025), 737–745. doi:10.1007/S43681-023-00405-2
- [12] Anxhela Ferhataj, Fatmir Memaj, Roland Sahatcija, Ariel Ora, and Enkelejda Koka. 2025. Ethical concerns in AI development: analyzing students' perspectives on robotics and society. *J. Inf. Commun. Ethics Soc.* 23, 2 (2025), 165–187. doi:10.1108/JICES-08-2024-0111
- [13] Noah Fraenkel. 2024. Beyond Principles: Virtue Ethics in AI Development—A Developer-Centric. *environments* 2024, 1689 (2024), 46. <https://helda.helsinki.fi/bitstreams/f71f2fee-5875-47fa-adad-867eb55ed749/download>
- [14] Clendson Domingos Gonçalves, Eduardo de Paoli Menescal, Fábio Lúcio Lopes de Mendonça, and Edna Dias Canedo. 2024. Trust in AI: Perspectives of C-Level Executives in Brazilian Organizations. In *Proceedings of the XXIII Brazilian Symposium on Software Quality, SBQS 2024, Salvador, Bahia, Brazil, November 5-8, 2024*, Ivan Machado, José Carlos Maldonado, Tayana Conte, Edna Dias Canedo, Johnny Marques, Breno Bernard Nicolau de França, Patrícia Matsubara, Davi Viana, Sérgio Soares, Gleison Santos, Larissa Rocha, Bruno Gadelha, Rodrigo Pereira dos Santos, Ana Carolina Oran, and Adolfo Gustavo Serra Seca Neto (Eds.). ACM, <https://doi.org/10.1145/3701625.3701654>, 147–157. doi:10.1145/3701625.3701654
- [15] Gabriel M. C. Guimarães, Geraldo Pereira Rocha Filho, Gilmar dos Santos Marques, and Edna Dias Canedo. 2025. May I speak? Perceptions on ethical concerns and power while developing software in AI teams. In *Proceedings of the 24th Brazilian Symposium on Software Quality, SBQS 2025, São José dos Campos, SP, Brazil, November 4-7, 2025*, Gleison Santos, Sheila S. Reinehr, Ivaldir

- de Farias Júnior, Bruno Gadelha, Monalessa Barcellos, Sávio Freire, Breno Bernard Nicolau de França, Edna Dias Canedo, Ana Carolina Oran, Patrícia Matsubara, and Rafael Parizi (Eds.). SBC, <https://doi.org/10.5753/sbqs.2025.13576>, 44–54. doi:10.5753/SBQS.2025.13576
- [16] Erika Halme, Marianna Jantunen, Ville Vakkuri, Kai-Kristian Kemell, and Pekka Abrahamsson. 2024. Making ethics practical: User stories as a way of implementing ethical consideration in Software Engineering. *Inf. Softw. Technol.* 167 (2024), 107379. doi:10.1016/J.INFSOF.2023.107379
- [17] Charlene Hinton. 2023. The State of Ethical AI in Practice: A Multiple Case Study of Estonian Public Service Organizations. *Int. J. Technoethics* 14, 1 (2023), 1–15. doi:10.4018/IJT.322017
- [18] Rashina Hoda. 2022. Socio-Technical Grounded Theory for Software Engineering. *IEEE Trans. Software Eng.* 48, 10 (2022), 3808–3832. doi:10.1109/TSE.2021.3106280
- [19] Javier Camacho Ibáñez and Mónica Villas Olmeda. 2022. Operationalising AI ethics: how are companies bridging the gap between practice and principles? An exploratory study. *AI Soc.* 37, 4 (2022), 1663–1687. doi:10.1007/S00146-021-01267-0
- [20] Arif Ali Khan, Muhammad Azeem Akbar, Mahdi Fahmideh, Peng Liang, Muhammad Waseem, Aakash Ahmad, Mahmood Niazi, and Pekka Abrahamsson. 2023. AI Ethics: An Empirical Study on the Views of Practitioners and Lawmakers. *IEEE Trans. Comput. Soc. Syst.* 10, 6 (2023), 2971–2984. doi:10.1109/TCSS.2023.3251729
- [21] Pereira Neto Macedo. 2025. Brazilian General Data Protection Law (LGPD). *Nartional Congress*, accessed in September 15, 2025 1 (2025), 1–31. <https://www.pnm.adv.br/wp-content/uploads/2018/08/Brazilian-General-Data-Protection-Law.pdf>
- [22] Vijaya Kittu Manda, Veena Christy, and Mallikharjuna Rao Jitta. 2025. Ethical AI and decision-making in management leadership. In *Ethical dimensions of AI development*. IGI Global, <https://www.igi-global.com/chapter/ethical-ai-and-decision-making-in-management-leadership/359644>, 197–226.
- [23] Josh Meltzer and Aaron Tielemans. 2022. The European Union AI Act. *Bruselj: Brookings Institution* 1 (2022), 1–8. https://www.brookings.edu/wp-content/uploads/2022/05/FCAI-Policy-Brief_Final_060122.pdf
- [24] Lauren Olson, Ricarda Anna-Lena Fischer, Florian Kunneman, and Emitzá Guzmán. 2025. Who Speaks for Ethics? How Demographics Shape Ethical Advocacy in Software Development. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency, FAccT 2025, Athens, Greece, June 23-26, 2025*. ACM, <https://doi.org/10.1145/3715275.3732183>, 2847–2862. doi:10.1145/3715275.3732183
- [25] Aastha Pant, Rashina Hoda, Simone V. Spiegler, Chakkrit Tantithamthavorn, and Burak Turhan. 2024. Ethics in the Age of AI: An Analysis of AI Practitioners' Awareness and Challenges. *ACM Trans. Softw. Eng. Methodol.* 33, 3 (2024), 80:1–80:35. doi:10.1145/3635715
- [26] Conrad Sanderson, David Douglas, Qinghua Lu, Emma Schleiger, Jon Whittle, Justine Lacey, Glenn J. Newnham, Stefan Hajkiewicz, Cathy Robinson, and David Hansen. 2021. AI Ethics Principles in Practice: Perspectives of Designers and Developers. *CoRR* abs/2112.07467 (2021), 1–17. arXiv:2112.07467
- [27] Antonia Sattlegger and Nitesh Bharosa. 2024. Beyond principles: Embedding ethical AI risks in public sector risk management practice. In *Proceedings of the 25th Annual International Conference on Digital Government Research, DGO 2024, Taipei, Taiwan, June 11-14, 2024*, Hsin-Chung Liao, David Duenas-Cid, Marie Anne Macadar, and Flavia Bernardini (Eds.). ACM, <https://doi.org/10.1145/3657054.3657063>, 70–80. doi:10.1145/3657054.3657063
- [28] Sruthy Murugan Usha, Chandan Medatwal, Vutti Purendra Prasad, Muralidhar LB, Anurag Rai, and Amit Verma. 2025. The Ethical Implications of AI in Management: Navigating Challenges in Decision-Making. In *2025 International Conference on Pervasive Computational Technologies (ICPCT)*. IEEE, <https://ieeexplore.ieee.org/document/10940240>, 392–396.
- [29] Heidi Vainio-Pekka, Mamia Ori-otse Agbese, Marianna Jantunen, Ville Vakkuri, Tommi Mikkonen, Rebekah Rousi, and Pekka Abrahamsson. 2023. The Role of Explainable AI in the Research Field of AI Ethics. *ACM Trans. Interact. Intell. Syst.* 13, 4 (2023), 26:1–26:39. doi:10.1145/3599974