Government Information Quarterly 42 (2025) 102003

Contents lists available at ScienceDirect

Government Information Quarterly

journal homepage: www.elsevier.com/locate/govinf

Artificial intelligence governance: Understanding how public organizations implement it

Patricia Gomes Rêgo de Almeida [a,b,*], Carlos Denner dos Santos Júnior [a]

[a] University of Brasilia, Department of Business Management-PPGA, Campus Darcy Ribeiro, Prédio da FACE, Asa Norte, Brasília, DF 70910-900, Brazil
[b] Chamber of Deputies of Brazil, Palácio do Congresso Nacional - Praça dos Três Poderes, Brasília, DF 70160-900, Brazil

ARTICLE INFO

Keywords:
AI governance
AI regulation
Responsible AI
AI ethics
AI in public sector
AI in government
Fuzzy QCA

ABSTRACT

While observing the race for Artificial Intelligence (AI) regulation and global governance, public organizations are faced with the need to structure themselves so that their AI systems consider ethical principles. This research aimed to investigate how public organizations have incorporated the guidelines presented by academia, legislation, and international standards into their governance, management, and AI system development processes, focusing on ethical principles. Propositions were elaborated on the processes and practices recommended by literature specialized in AI governance. This entailed a comprehensive search that reached out to 28 public organizations across five continents that have AI systems in operation. Through an exploratory and descriptive aim, based on a qualitative and quantitative approach, the empirical analysis was carried out by means of proposition analysis using the Qualitative Comparative Analysis (QCA) method in crisp-set and fuzzy modes, based on questionnaire responses, combined with an interview and document content analysis. The analyses identified how processes and practices, across multiple layers and directed at the application of ethical principles in AI system production, have been combined and internalized in those public institutions. Organizations that trained decision-makers, AI system developers, and users showed a more advanced stage of AI governance; on the other hand, low scores were found on actions towards AI governance when those professionals did not receive any training. The results also revealed how governments can boost AI governance in public organizations by designing AI strategy, AI policy, AI ethical principles and publishing standards for that purpose to government agencies. The results also ground the design of the AIGov4Gov framework for public organizations to implement their own AI governance.

1. Introduction

After being coined "Artificial Intelligence" in 1956 (Cerka et al., 2017), a variety of research on AI has been developed, initially as knowledge-driven, later as data-driven, or combining them. The growing scope of AI in society has been observed through the combination of this technology with other emerging ones, generating the expression AI Plus (AI+) (Shao et al., 2022), boosting productivity (Mezgár & Váncza, 2022) and its influence on social transformation (Boyd & Holton, 2018).

The benefits offered by AI have reached the Government (Alhosani & Alhashmi, 2024). While Artificial Intelligence (AI) is seen as an enabler of digital transformation for organizations (Holmström, 2022; Kitsios & Kamariotou, 2021), in the public sector, governments' development strategies coincide with their AI strategies (Wirtz et al., 2018). However, at the same time, concerns grow about ethical impacts on AI-dependent decisions when ethical principles are not considered (Ashok et al., 2022; Bonsón et al., 2021; Hopster, 2021; Kazim & Koshiyama, 2021; Stahl et al., 2022; Wirtz et al., 2022). Immersed in this scenario, the movement for a responsible AI (Eke et al., 2023) using AI regulation and governance has involved governments, academia, and international standardization bodies (Carter, 2020; Gianni et al., 2022; Gutierrez & Marchant, 2021; IEEE, 2019, 2020, 2021a, 2021b, 2021c, 2021d, 2022; ISO, 2021a, 2021b, 2021c, 2022a, 2022b; OECD, 2022c).

Even though there is a significant number of theoretical essays (De Almeida et al., 2021), AI governance is still an underdeveloped area of research (Morley et al., 2020; Taeihagh, 2021), requiring a greater understanding of how organizations have interpreted and incorporated

* Corresponding author at: University of Brasilia, Department of Business Management, Campus Darcy Ribeiro, Prédio da FACE, Asa Norte, Brasília, DF 70910-900, Brazil.

E-mail addresses: patricia.rego.almeida@gmail.com (P.G.R. de Almeida), carlosdenner@unb.br (C.D. dos Santos Júnior).

https://doi.org/10.1016/j.giq.2024.102003
Received 11 June 2023; Received in revised form 25 September 2024; Accepted 20 December 2024
Available online 16 January 2025

ethical principles into their practices, processes and structures when producing AI systems (Mäntymäki et al., 2022; Mikalef, Conboy, et al., 2022). In systematic research on AI in public governance, Zuiderwijk et al. (2021) identified research gaps that adopt multiple methods, combining exploratory empirical research with qualitative-quantitative analyses, in order to delve deeper into practices used by AI governance in the public sector. Considering the important role that government bodies play in AI regulation and governance (Cihon et al., 2020; Stix, 2021), the potential benefits and risks that AI can bring to society when it supports the public sector (Ahn & Chen, 2022; Ojo et al., 2019; Sharma et al., 2020), and the challenge of avoiding loss of confidence in AI-supported government decisions (Zuiderwijk et al., 2021), it becomes crucial to understand how public organizations are following the academia, legislation, and standard recommendations concerning ethical principles in the use and development of AI systems.

2. AI governance

An AI that considers ethical principles brings new obligations to organizations (Hickman & Petrin, 2021; Roorda, 2021; Smuha, 2021). However, traditional governance processes and structures are not sufficient for the challenges posed by AI governance (Taeihagh, 2021). In public institutions, the challenge is greater because, in general, citizens do not choose AI products but are obligated to consume them as they are embedded in public services (Zuiderwijk et al., 2021).

In the context of AI governance in public organizations, ethical principles are applied to maximize the benefits of AI and minimize its risks (Rose et al., 2018; Vial, 2019). Considering the focus on impacts generated in society as a premise for AI governance (Djeffal, 2018), researchers point out the need to integrate the Stakeholder Theory with the Social Contract Theory (Bonsón et al., 2021; Wright & Schultz, 2018) to obtain society's perceptions of the values involved in decisions made by AI systems (Hickman & Petrin, 2021; Rahwan, 2017). In the corporate context, subordinated to IT governance (IT Governance Institute, 2003), an effective AI governance requires a multilayered model, the upper layer of which includes mechanisms for capturing government regulatory requirements and legislation, translating them into the organizational context through internal regulations that establish ethical principles and conditions for their application through processes and practices (Mäntymäki et al., 2022), as illustrated by the conceptual research model in Fig. 1a, whose details are found in the subsections 2.1 up to 2.6.

Considering many initiatives to regulate AI through legislation, government policies, and international standards (Fjeld et al., 2020; Gutierrez & Marchant, 2021; OECD, 2022c), it is expected that public organizations implement their own AI governance aligned with such regulation initiatives.

2.1. AI governance actions at the strategic level of public organizations

The relationship among the different scopes of governance, defined by Mäntymäki et al. (2022), establishes that corporate governance contains IT governance, which, in turn, contains AI governance. Consequently, AI governance inherits characteristics from IT governance.

IT governance relates to IT decision-making at the board of directors and executive management, which involves: creating an organizational structure (unit, committee, board), elaborating a strategy to effectively address the organization's needs through AI (Herremans, 2021), and implementing processes that support the board's decisions (Aasi et al., 2014). To formalize IT governance, organizations establish policies (Mäntymäki et al., 2022). Thus, similar to those actions that demonstrate the existence of IT governance at a higher-level decision-making board (Aasi et al., 2014), high-level decision-making actions for AI governance were considered in this research.

Indeed, in Papagiannidis et al.'s (2023) research, the higher-level decision-makers highlighted the importance of an AI strategy to manage the corporate needs that AI systems will address, as well as the AI governance process in their organization. In the same sense, Agarwal's (2023) research points out that the existence of an AI governance structure at a high level of the organization is crucial for setting the organization's strategic direction in AI-related initiatives. Complementing them, Sigfrids et al. (2023) emphasize that, in the public sector, AI policies should consider AI ethical principles, giving special

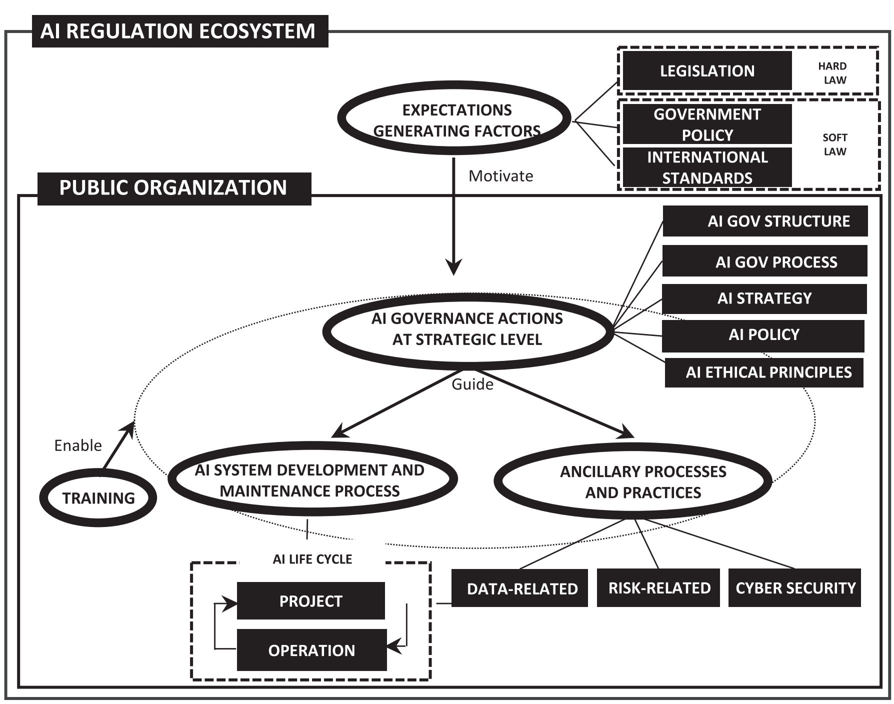

Fig. 1. a: Conceptual research model for AI governance in public organizations. (Source: Self-elaboration.)

attention to a wider socio-technical and political approach instead of merely respecting moral minimums. It requires considering values and gaining citizens' trust. Such an approach imposes establishing an AI ethical principles code at a high level of the organization. Thus, for the context of this research, "creating an AI governance structure," "elaborating an AI strategy," "establishing an AI policy," "implementing an AI governance process," and "establishing an AI ethical principles code" were at the higher level of the conceptual research model and were called "AI governance actions at a strategic level."

Guided by the AI governance actions at a strategic level, ancillary mechanisms are created to establish data-related processes (Janssen et al., 2020; Rhahla et al., 2021), a cybersecurity-related process (Breier et al., 2020; Xue et al., 2020), risk-related processes and practices (NIST, 2022; Wirtz et al., 2022) (Fig. 1a – Ancillary processes and practices) and AI system development-related processes that address AI ethical issues (De Silva & Alahakoon, 2022; Laato et al., 2022) (Fig. 1a – AI system development and maintenance process). Based on the above, the following **Proposition 1** is conjectured: "data-related," "risk-related," "cyber-security," and "AI system development" practices must follow guidelines established at the strategic level of AI governance. Considering that each dimension addressed in Proposition 1 can be decomposed into other actions, for greater accuracy of this study, derived propositions were created for each group of processes and practices.

2.2. *Data-related processes and practices (Fig. 1a – Data-related processes and practices)*

Given AI's reliance on data, a link is established between data quality and the outcome of AI systems (Dwivedi et al., 2021; Kuziemski & Misuraca, 2020). For this reason, focusing on reliable AI systems and data governance becomes crucial (Haneem et al., 2019; Vining et al., 2022).

Sometimes intersecting with AI governance (Alshahrani et al., 2021; Andrews, 2018; Dwivedi et al., 2021; Mäntymäki et al., 2022; Medaglia et al., 2021; Özdemir & Hekim, 2018), a data governance process specifies responsibilities over decisions made about the organization's data, as well as formalizes policies and standards (Abraham et al., 2019; Carretero et al., 2017; Vilminko-Heikkinen & Pekkola, 2019), which requires a great deal of stakeholders' negotiation skills to obtain consensus (Benfeldt et al., 2020; Calzada & Almirall, 2020; Micheli et al., 2020; Ruijer, 2021). Following data governance policies (Carretero et al., 2017), processes are defined for managing data quality (Haneem et al., 2019; Khatri, 2016) and personal data protection (Janssen et al., 2020). For those reasons, **Proposition 1 A** is based on the decomposition of the "data-related" construct into three processes: data governance process, data quality management process, and personal data protection management process, which must follow guidelines established at the strategic level of AI governance.

2.3. *Risk-related processes and practices (Fig. 1a – risk-related processes and practices)*

Proposition 1B comprises actions created to mitigate the risks posed by AI systems (Vetrò et al., 2021; Wirtz et al., 2022). However, traditional risk management processes, supported by quantitative methods (Chen & Deng, 2022; Duijm, 2015), have been criticized, requiring complementary approaches that integrate visions and thus also include a qualitative approach (Aiken, 2021; Fernandes et al., 2021; Gerkensmeier & Ratter, 2018; ISO, 2022a; ISO, 2022b).

Such integrated vision starts with the definition of stakeholders that are impacted, whether directly or indirectly, by the AI system (NIST, 2022; Wirtz et al., 2022). Therefore, identifying stakeholders in the whole AI lifecycle is a requirement for risk management (Wright & Schultz, 2018). In the same direction, audit processes for AI systems are also risk-oriented (De Oliveira, 2019; Erlina et al., 2020), associating them with stakeholders (Zicari et al., 2021). In addition, over time, changes in environmental variables can alter the context for which the AI system was designed, causing behaviors that differ from the desired results. This situation can be avoided by monitoring changes in the environment (rules, social trends, etc.) and feeding the risk management process (González et al., 2020). Thus, **Proposition 1B** was formulated as follows: risk management processes, audit processes, practices for identifying stakeholders, and practices for monitoring changes in the environment, all of which must follow guidelines established at the strategic level of AI governance.

2.4. *Cybersecurity process (Fig. 1a – Cybersecurity)*

Integrated with risk management (Breier et al., 2020; European Union Agency for Cyber Security, 2022), a cybersecurity management process is applied to prevent cyberattacks explicitly designed to exploit vulnerabilities in AI algorithms (Chen et al., 2019; Eggers & Sample, 2020; Gu et al., 2019; McGraw et al., 2020; Xue et al., 2020). To face those threats, organizations implement a security management process (European Union Agency for Cyber Security, 2021) that addresses the entire AI system lifecycle (Jing et al., 2021). Therefore, **Proposition 1C** was thus formulated as follows: the AI system security management process must follow guidelines established at the strategic level of AI governance.

2.5. *AI system development process (Fig. 1a – AI system development and maintenance process)*

Seeking to encompass the entire AI lifecycle, **Proposition 1D** was formulated as follows: practices aimed at ethical principles for AI system development and maintenance processes, both as a project and as an operational product, must follow guidelines established at the strategic level of AI governance. To investigate practices at each phase, Proposition 1D was further broken down into 1D1 and 1D2 to analyze the project phase and the operation phase, respectively.

At the project's starting point (Fig. 2 – Project), the translation of ethical principles into rules on the behavior of the system is deepened (Dennis et al., 2016; IEEE, 2021a), focusing on the definition of groups and attributes to be protected (González et al., 2020; ISO, 2021a; Rajkomar et al., 2018). Rules and ethical dilemmas are identified and analyzed (Anderson & Anderson, 2018; Awad et al., 2020; Bench-Capon & Modgil, 2017; Bonnemains et al., 2018; Locher & Bolander, 2019; Ma et al., 2018; Schrader & Ghosh, 2018; Zicari et al., 2021). Data extraction and preparation tasks demand attention to understand their characteristics and quality, preparing them for pre-processing (De Silva & Alahakoon, 2022). Techniques are applied to identify and minimize data biases (Ashokan & Haas, 2021; Baeza-Yates, 2018; ISO, 2021a; Leavy et al., 2020; Lin et al., 2021; Ntoutsi et al., 2020; Oneto & Chiappa, 2020; Roselli et al., 2019; Silberg & Manyika, 2019), while adjustments are made in the data sample (González et al., 2020). The building and validation of models involve algorithmic research, which requires decisions that also need to be free of bias (Abdollahi & Nasraoui, 2018; Ashokan & Haas, 2021; De Silva & Alahakoon, 2022; Makhlouf et al., 2021). Transparency practices are required to explain the system results (Adadi & Berrada, 2018; Arrieta et al., 2020; Das, 2020; Dazeley et al., 2021; Kale et al., 2022; Phillips et al., 2021). Focusing on the AI system project phase, **Proposition 1D1** was formulated: practices for representing rules and ethical dilemmas, practices for minimizing biases, and practices for providing transparency in the AI system development process must follow guidelines at the strategic level of AI governance.

The sensitivity to changes in the context for which the AI system was created, combined with the fact that AI models are less complex than social realities (Strauß, 2021), impose continuous monitoring after the AI system is in operation (Laato et al., 2022), which implies: a) automatic performance monitoring in charge of the IT staff (De Silva & Alahakoon, 2022; Fjeld et al., 2020; González et al., 2020), b) human oversight of the AI system behavior, usually by someone delegated by

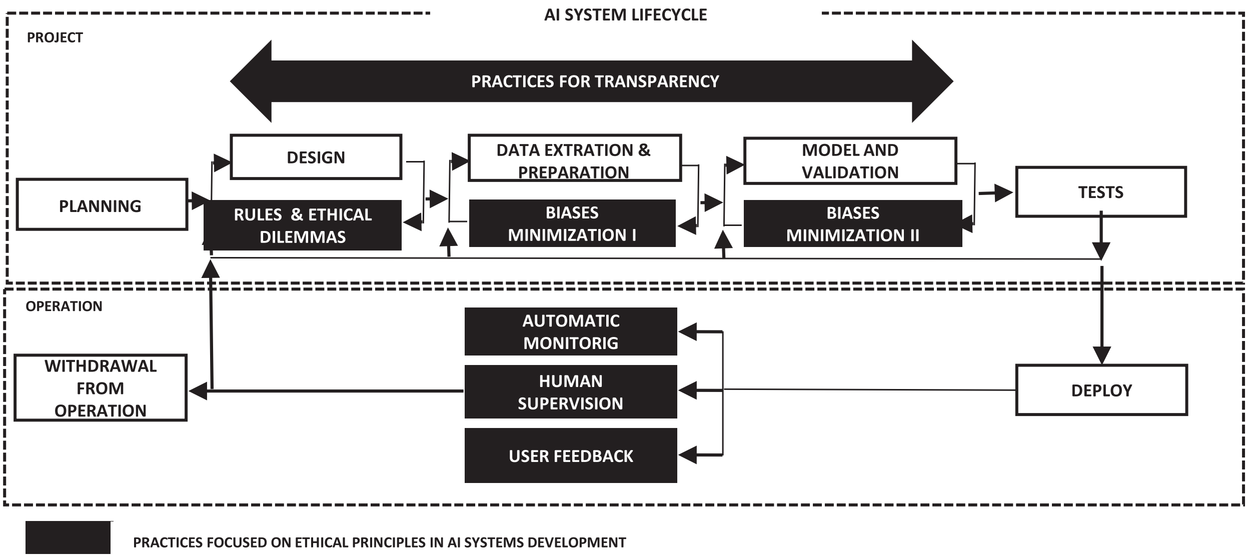

Fig. 2. Conceptual research model for the entire AI system lifecycle.
(Source: Self-elaboration)

the domains' decision-maker (Dignum, 2019; Fjeld et al., 2020; Hickman & Petrin, 2021; Strauß, 2021; Zicari et al., 2021), and c) continuous user feedback, which is often implemented through a feature in the AI system that asks its users' satisfaction with the system outputs and requests tips to refine its performance (Rahwan et al., 2019; Wright & Schultz, 2018) (Fig. 2 – Operation). The combination of those actions motivates the evolution of AI systems in a continuous loop until their withdrawal from operation (De Silva & Alahakoon, 2022; ISO, 2022a; Laato et al., 2022) (Fig. 2). Therefore, with the practices described in the AI system operation phase in mind, **Proposition 1D2** was formulated: practices for automatic monitoring, human oversight, and user feedback must follow guidelines established at the strategic level of AI governance.

2.6. *Training people (Fig. 1 – Training)*

Mikalef, Lemmer, et al. (2022) confirm that the decision-makers' perceptions regarding the potential AI value are drivers for AI adoption in public organizations. In the same sense, educating stakeholders on AI ethics becomes crucial for an effective AI ethical principles implementation in AI system production (Zhou & Chen, 2023), which includes a culture with fewer biases in organizations (Awad et al., 2020; Ma et al., 2018) as well as knowledge about practices required to AI governance (Ligot, 2024). With such purposes, training key stakeholders (decision-makers, developers, system users, and auditors) on AI ethical principles is considered an enabler for AI governance implementation in organizations (Calzada & Almirall, 2020; Herremans, 2021; Makarius et al., 2020; Micheli et al., 2020; Ruijer, 2021). Researchers point to an AI literacy program as a pivotal action to an effective and responsible AI adoption since it is a set of training and awareness actions that reaches staff beyond the IT team, which includes business decision-makers and system users. Decision-makers need it for a wide understanding of how AI can create value for their work processes and the requirements necessary for it (data quality, data protection, data governance, inclusive teams, for example). Users are also considered key stakeholders due to their role in data quality and data protection processes, and there is also the need for an awareness approach regarding ethical implications when AI is used inadequately (Pinski et al., 2024; Schüller, 2022).

Aiming to determine the enabling requirements of training people for AI governance practices, **Proposition 2** was elaborated: training stakeholders on data, the development of AI systems, and ethical principles applied to AI enables the implementation of AI governance in public organizations.

Considering that the complexity of internalizing AI governance in the organizations' processes and structures (Agarwal, 2023) can impact the "make-or-buy" decision on developing AI systems. Gräf et al. (2024) highlight that the opacity of AI systems and the lack of skilled human resources to deal with the whole AI system lifecycle are key factors for such decisions. In general, the decision to buy AI systems is an obstacle to accessing their codes, resulting in a transparency issue (Martin & Parmar, 2024). Adding such concerns to Mikalef, Lemmer, et al. (2022), Ahn and Chen's (2022) and Benfeldt et al.'s (2020) recommendations for training both the managerial and the technical spheres to deal with challenges such as reducing algorithmic opacity, it is possible to formulate **Proposition 3**: over time, in public organizations, training managers and AI system developers on AI ethics sparks interest in obtaining knowledge of AI system codes and contributes positively to AI governance.

3. Research methods and techniques

As an exploratory and descriptive study, the investigation was carried out through empirical research with a qualitative and quantitative approach to fill the gap identified by Mäntymäki et al. (2022) and Zuiderwijk et al. (2021) concerning knowledge of how organizations have interpreted and incorporated ethical principles in AI system production into their practices and processes, as especially demanded by Zuiderwijk et al. (2021) for using data-driven methods with exploratory and multiple approaches to deepen AI governance in the public sector.

3.1. *Sample selection and data collection strategy*

Since AI governance is a global need (Fjeld et al., 2020; OECD, 2022a), an attempt was made to build the sample including populations of public organizations belonging to any branches — Executive, Legislative, or Judiciary (Maluf, 1995) — from five continents. Considering the interest in investigating processes and practices, a search criterion was defined: a public organization should have at least one AI system in operation as part of its official portfolio. As a consequence, organizations that only had AI systems at a prototype stage or projects in an embryonic stage were not included in the sample.

The identification, sample selection, and data collection tasks were performed over three months, consisting of several steps in parallel lines of research involving different actors and sources of information, as shown in Fig. 3 (MRE, 2022; OECD, 2022a); European Commission, 2018). Those communications took place through remote meetings, email, and social media until the contact information from researchers, government and AI system development and research centers, managers in charge of AI or digital transformation strategies, and AI use cases could be obtained. The path described in Fig. 3 culminated in 711 AI systems being developed or used by public organizations (European Commission, 2021b; FCT, 2021; Government of India, 2020; IPS-X, 2021; Misuraca & van Noordt, 2020; OECD/CAF, 2022; Tangi et al., 2022; WEF, 2020, 2020b). After removing redundancies and non-operational systems (prototypes or withdrawn from use), finetuning continued until the contact information for organizations that met the search criteria could be found. Upon invitation, 28 organizations effectively participated in the survey.

3.2. Data collection mechanisms

For the quantitative and qualitative analyses, primary data was used by means of an online questionnaire and semi-structured interviews (Appendix A - Annex 1a and 1b), both of which required basic knowledge of the decisions, processes, and practices related to AI system production. For this reason, the questionnaire was sent out only after the organization indicated a person in charge of the AI system portfolio. The interviews were conducted remotely and, in some cases, two or three individuals represented the organization.

Both the questionnaire and the interview script were submitted to a group of judges for assessment (PhD and Master researchers on ethical AI and data analysis) using methods indicated in the literature for each case (Fig. 4). For the questionnaire, the "Content Validation Coefficient" (CVC) (Aburachid & Greco, 2011; Hernández-Nieto, 2002; Silveira et al., 2018) was used with each question being evaluated on a scale from 1 to 5 to measure levels of clarity and relevance for the research (Appendix A - Annex 2 A). The interview script was evaluated using the "Validation for Qualitative Research Instruments" method (VALI-QUALI) (Torlig et al., 2022), an evolution of MRPQ (Torlig et al., 2019), taking the "Content" and "Semantics" dimensions into consideration. The content evaluation provided a score from 1 to 5 for each question in relation to the "alignment of each question with the research objective" and "adherence of the question to the investigated construct" attributes. The semantic analysis considered the "clarity" and "qualitative expectation of answer for each question" attributes (Appendix A - Annex 2). Pre-tests were carried out to the complete set (questionnaire and interview) using people with a similar profile to the target audience to confirm alignment with the research objectives (Manzini, 2004).

For each question regarding the existence of a practice or process implemented (or being implemented), four scores were considered according to standard answers (100 for "Yes, completely," 67 for "Yes, but only partially," 33 for "No, but a formal decision has been made to implement it," and 0 for "No, and no formal decision has been made to implement it." For the AI governance process and for practices aimed at transparency, adapted answers were made, once they were gathered from the interview (Appendix A - Annex 3).

3.3. Analysis strategy

The analysis of the primary data was carried out using a combination of Qualitative Comparative Analysis - QCA (Ragin, 2008; Rihoux & Ragin, 2008), and content analysis of the interviews and shared documents (Krippendorff, 2013; Saldaña, 2013).

3.3.1. QCA

The QCA is a qualitative research technique that also considers quantitative aspects for samples from 3 to 250 cases (Dias, 2011). Based on Set Theory and Boolean operations to establish logical relationships between sets, the QCA proposes to solve problems whose analysis requires causal inferences in case studies. The method seeks to show which combinations of conditions occurred in a scenario of an expected outcome (Rihoux & Ragin, 2008) to carry out comparative analyses, through associations between certain conditions and the outcome, instead of correlations (Korjani & Mendel, 2012; Ragin, 2008). A condition is required for a given result if the condition is always present when the outcome occurs. A condition is sufficient for a certain outcome

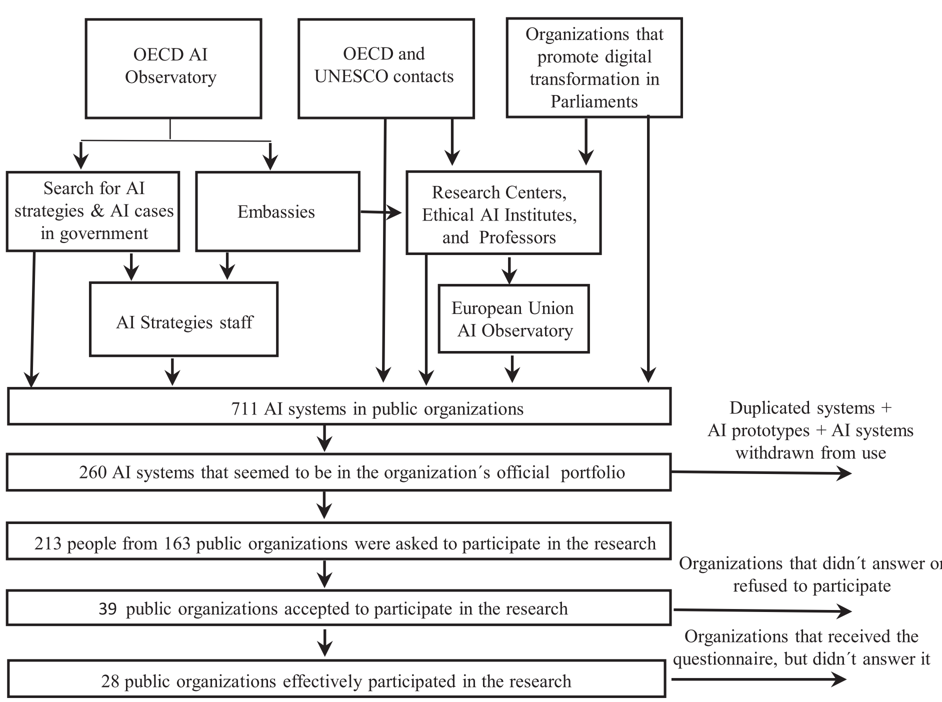

Fig. 3. Sample selection strategy.
(Source: Self elaboration)

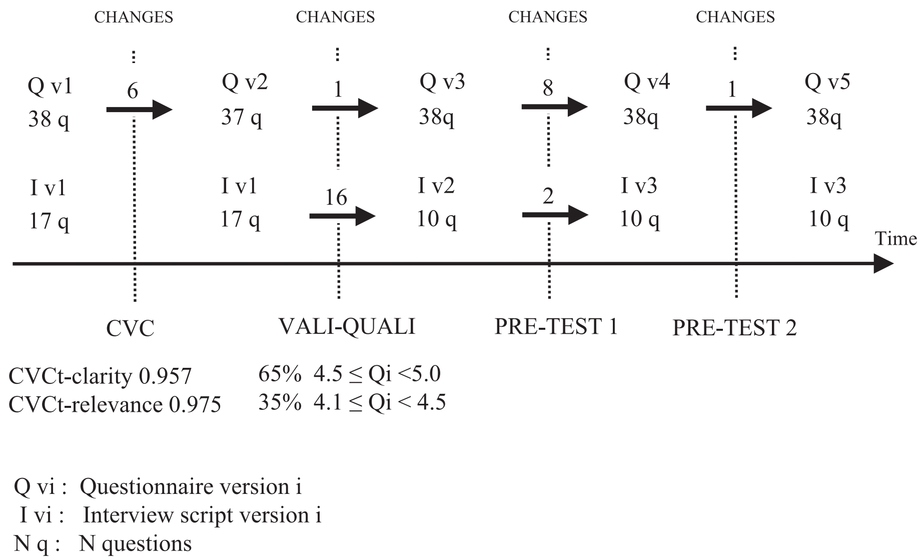

Fig. 4. Evolution of data gathering mechanisms.
(Source: Self-elaboration)

if this result always occurs when the condition is present (Rihoux & Ragin, 2008).

This research used both the crisp-set QCA (values 0 and 1, respectively, for the absence or presence of a relationship between the sets) and the fuzzy QCA, which offers more precision due to the use of a continuous set of values in the interval from 0 (complete absence of membership) to 1 (full membership) (Ragin, 2008; Rihoux & Ragin, 2008). In fuzzy analyses, A is a fuzzy subset of a fuzzy set B if given two fuzzy sets, A = {s1,s2,.., sn} and B = {g1,g2,.., gn}, and si, and gi, being i-th case scores in each set, and si ϵ [0;1] $\subset \mathbb{R}$, gi ϵ [0;1] $\subset \mathbb{R}$, $\forall$i; so, A $\subset$ B if si $\leq$ gi, $\forall$i. A calibration (Freitas & Neto, 2016; Meijerink & Bondarouk, 2018; Ragin, 2008) was made using the original values of the sets, turning them into fuzzy sets, whose values are distributed in the interval between 0 and 1, based on the presence level of the conditions in the outcome set. The main criteria for validating the fuzzy QCA is the consistency indicator to measure the relationship proximity between sets, indicating the degree to which the cases that share a combination of conditions agree with the outcome, with a value between 0 and 1. Consistency $(Xi \leq Yi) = \sum \frac{MIN(Xi,Yi)}{\sum Xi}, \forall i$; where X is the membership score in the causal combination, and Y is the membership score in the outcome (Betarelli-Júnior & Ferreira, 2018). Complementing the outcome interpretation, the coverage indicator offers the quantification of the empirical relevance of a causal combination in the causal combination set: Coverage $(Xi \leq Yi) = \sum \frac{MIN(Xi,Yi)}{\sum Yi}, \forall i$; where X is the membership score in the causal combination and Y is the membership score in the outcome (Rihoux & Ragin, 2008).

3.3.2. *Analysis plan*

A crisp-set QCA was applied for dichotomous variables, and the fuzzy QCA for variables with continuous values (Fig. 5) was associated with the research model constructs. The QCA was applied to analyze the propositions. As for the crisp-set QCA, the TOSMANA software version 1.6.1 (https://www.tosmana.net/) was used, while fsQCA version 3.1 (http://www.socsci.uci.edu/~cragin/fsQCA/software.shtml) was utilized for the fuzzy QCA. Another piece of information was extracted from the analysis of the interview content using MAXQDA 2022 software (https://www.maxqda.com/). Lastly, the union of the two analyses grounded the discussions to reach the conclusions that support the proposed framework.

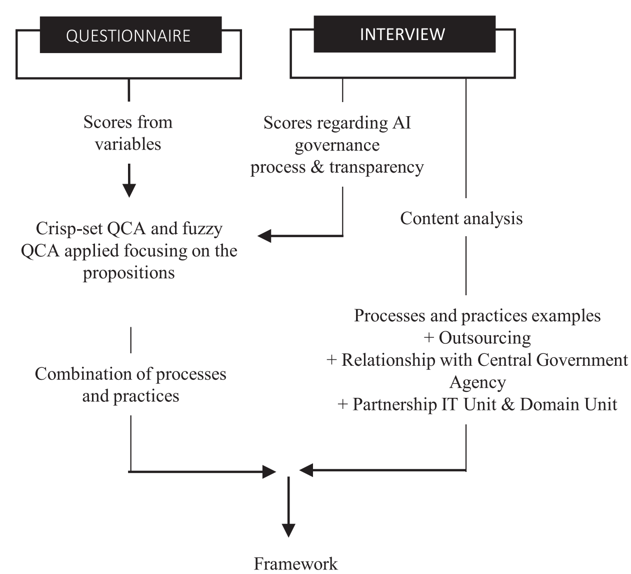

Fig. 5. Analysis plan.
(Source: Self-elaboration)

4. Results and discussions

The participants were people aware of the processes involved in the whole AI lifecycle. They were decision-makers and/or people appointed by them to participate. Table 1a displays the participants' profiles. The sample, resulting from the path taken as shown in Fig. 3, consists of 28 public organizations distributed across five continents, whose characteristics are listed in Tables 1b and 1c. The category of public organizations reveals their importance to citizens. The questionnaire asked the participant to provide information about the organization's five most frequently used AI systems (Appendix A - Annex 5).

The analyses that simultaneously involved dichotomous variables and continuous-value variables were performed by converting variables from continuous into binary and applying the crisp-set QCA (Proposition 3). For analyses in which all variables had continuous values (Propositions 1 A, 1B, 1C, 1D1, 1D2, and 2), the fuzzy QCA was applied with a fuzzy value calibration of 0.05 and 0.95, respectively, for each set's lowest and highest values, as also established by Navarro et al. (2016)

Table 1a
Interviewed profiles.

Organization Unit	N	%	Education	N	%	Position	N	%
Information Technology	17	60.7	PhD	9	32.1	Director/Manager	20	71.4
Innovation	2	7.14	Master	12	42.9	Data Scientist/IT Analyst	6	21.4
Data Science and Statistics*	3	10.7	MBA	4	14.3	Consultant	2	7.14
Corporate Governance	4	14.3	Undergraduate	3	10.7			
Others	2	7.14						

* Cases in which the data science unit is detached from the IT unit.

Tables 1b and 1c
Characteristics of the 28 public organizations in the sample used in the study.

Country	N	%	Coverage	N	%
Angola	1	3.57	National	24	85.71
Argentina	3	10.71	Group of countries	1	3.57
Australia	1	3.57	State/City	3	10.71
Brazil	4	14.29			
Canada	2	7.14	**Branch of Government**	**N**	**%**
Denmark	1	3.57	Executive	17	60.7
Estonia	2	7.14	Legislative	9	32.1
Finland	2	7.14	Judiciary	2	7.14
Germany	2	7.14			
Iceland	1	3.57	**No of Employees**	**N**	**%**
Italy	1	3.57	Up to 100	2	7.14
Japan	1	3.57	101 up to 500	3	10.71
Luxembourg	1	3.57	501 up to 1000	8	28.57
Norway	3	10.71	1001 up to 10,000	12	42.86
Sweden	1	3.57	More than 10,000	3	10.71
Switzerland	1	3.57			

Category	N	%
Parliament	8	28.57
Ministry of Finances\National Agency for Trade and Investment \National Business Authority \National Tax Agency\National Agency for Improving Business	6	21.42
National Agency for Unemployment Insurance Fund\Federal Office for Migration and Refugees\National Agency for Labour and Welfare	3	10.71
National Statistic and Research Institute\National Institute of Research and Technology	3	10.71
Federal Court/Federal Court Dept Specialized in Child Sexual Abuse Crimes	2	7.14
Hospital and Health Research Institute/National Agency for Auditing Businesses Producing Food	2	7.14
National Agency for Account Auditing	1	3.57
State Agency for Innovation	1	3.57
State Agency for Public Transport	1	3.57
Chief of Ministerial President Cabinet	1	3.57
AI systems production time	**N**	**%**
Less than 1 year	1	3.57
1–3 years	8	28.57
3–5 years	11	39.29
More than 5 years	8	28.57

and Codá et al. (2022). The descriptions of the variables are in Appendix A - Annex 6.

Regarding legislation, it is worth mentioning that 100 % of the sample comprises organizations subject to some law that protects personal data. In the sample's scope, only Denmark (Danish Government, 2020) approved a law on data ethics that also addresses AI systems development. Among the bills under discussion, the European AI Act (European Commission, 2021a; OECD, 2022b), despite being in an advanced discussion stage when the data were collected, was not yet approved as a law (European Parliament, 2022a; European Parliament, 2022b).

4.1. Analysis of proposition 1

Proposition 1 was analyzed through its decomposition into dimensions – data, risks, cyber security, and development – in Propositions 1 A, 1B, 1C, and 1D analyses, respectively. In those analyses, to represent AI governance actions at the strategic level of the organization, the following was considered (Appendix A - Annex 4): AI strategy; policy or recommendations directed at AI systems; guide to ethical principles that are applied to AI systems; AI governance processes; and the existence of a structure (unit, position, or board/council/committee to address AI governance). Those analyses carried out the calibration for turning the original values into fuzzy values while considering that the analyses focus on the conditions whose practices and processes have been implemented in any proportion (scoring options 67 or 100 – Appendix A - Annex 3). For those purposes, in all research analyses that used a fuzzy QCA, the crossing point was defined as 60.

4.1.1. Analysis of proposition 1 A

From the decomposition of the practices regarding data into processes for data governance (PDGOV), for data quality management (PDQUA), and for personal data protection management (PDPRO), an analysis of Proposition 1 A was carried out, considering, as an output set, the cases that were in a more advanced stage of implementing AI governance actions at a strategic level (SAIGOV(1)) (Appendix A - Annex 6). The process for managing personal data protection had the highest average (90.5, see Appendix A - Annex 4), which illustrates the most advanced stage of actions to comply with the personal data protection law to which they are subject.

The three preliminary combinations presented by the fuzzy QCA (Table 2) were not valid to the conclusion (combinations 1 and 2 had a consistency below the minimum value accepted by Ragin, 2008 and Scheider and Wagemann, 2012, and combination 3 was not conclusive). However, it should be noted that nineteen cases implemented the three processes – data governance, data quality management, and personal data protection management; and that, of the twelve highest scores for strategic actions towards AI governance, nine cases (75 %, see Appendix A - Annex 4) had implemented, at any stage, all three processes. Due to that, the "Necessary Condition" assessment test was applied to the combination of the three processes, which resulted in a consistency of 0.959862, revealing that high values would have been unlikely to occur in AI governance actions at a strategic level without the simultaneous existence of the three processes tested. Thus, the simultaneous implementation of the three processes is associated with organizations at an advanced stage in those AI governance actions at the strategic level defined in 2.1.

The positive impact when the three data processes are implemented aligns with the researchers' arguments that data governance is crucial for AI governance (Haneem et al., 2019; Vining et al., 2022) since data governance defines stakeholders' roles and responsibilities concerning the data (Sivarajah et al., 2017), which, in turn, requires managers' involvement in many deliberations (Benfeldt et al., 2020) to improve data quality process (Haneem et al., 2019; Rhahla et al., 2021; Vilminko-Heikkinen & Pekkola, 2019) and personal data protection management process. Thus, data governance guides data quality management processes and personal data protection processes (Labadie et al., 2020).

Table 2
Fuzzy QCA applied to proposition 1 A.

Proposition		Combinations	Cases	Results	Coverage and Consistency
1A	1	PDGOV * PDPRO	3,5,9,19,22,27,6,8,13,25	SAIGOV(1)	Cv: 0.825606
			4,20,24,2,7,11,15,16,23	SAIGOV(0)	Cs: 0.71781
	2	PDQUA * PDPRO	8,19,22,25,5,6,9,13,27	SAIGOV(1)	Cv: 0.733564
			20,23,2,4,7,21,15,16,18,28	SAIGOV(0)	Cs: 0.688312
	3	∼ PDGOV* ∼ PDQUA* ∼ PDPRO	21	SAIGOV(1)	Cv: 0.215917
			1	SAIGOV(0)	Cs: 0.75
	"Necessary Conditions" Test	PDGOV * PDQUA * PDPRO		SAIGOV(1)	Cv:0.581795 Cs: 0.959862

4.1.2. Analysis of proposition 1B

Proposition 1B was analyzed based on the decomposition of risk-related practices into a process for risk management (PRISM), a process for auditing AI systems (PAUDIT), a practice for identifying stakeholders (STAKEH), monitoring changes in the environment and social trends (ENVIRO), considering cases that were at a more advanced stage in the implementation of AI governance actions at a strategic level (SAIGOV(1)) (Appendix A - Annex 6). Of the three combinations that were calculated (Table 3), only combination 2 showed a proportion favorable to advanced-stage cases in AI governance actions at a strategic level, which presented 85.71 % of the cases (cases 3, 8, 9, 13, 22, 27) that have implemented, at any stage, a risk management process, practices for identifying stakeholders, and practices for monitoring environmental changes and social trends (consistency = 0.941645). Therefore, those processes and practices can be considered to have been associated with more advanced stages in implementing AI governance actions at a strategic level.

The positive impact of the four studied practices aligns with the researchers' arguments that a risk management approach for AI requires that traditional risk management processes (Chen & Deng, 2022; Duijm, 2015) be associated with stakeholders' identification (Schaefer et al., 2021; Wirtz et al., 2022; Wright & Schultz, 2018), with a risk-oriented audit process (De Oliveira, 2019; Erlina et al., 2020), and with a change management in the environment (González et al., 2020).

4.1.3. Analysis of proposition 1C

Considering only the cyber security management process, this analysis did not present combinations of processes or practices because this process was not decomposed into other practices. The fuzzy QCA analysis (Table 4) revealed that, of the cases that had implemented a process for managing security in AI systems (PSEC), 66.67 % (cases 3, 5, 6, 8, 9, 13, 19, 22, 25, 26) obtained a high score for AI governance actions at a strategic level (SAIGOV(1)) (Appendix A - Annex 6), which indicates an association between the existence of a security management process and advanced stages of AI governance actions at a strategic level. Such result confirms researchers' arguments that a) AI governance implies, among other principles, ensuring robust and safe AI systems (Dalrymple et al., 2024), b) since robust and safe AI systems require practices to deal with mechanisms to face cyberattacks created specifically to AI systems (Booth et al., 2023; Ee et al., 2024), AI governance also requires practices systematically organized to manage the cyber security of AI systems. Compared with the other propositions' analyses, this was the lower consistency (0.8312) in the considered associations, which can be further investigated in future research.

4.1.4. Analysis of proposition 1D

The analysis of system development practices required two levels of decomposition: firstly, the phases of the AI system development and support process – project (Proposition 1D1) and operation (1D2); and subsequently, decomposing each of those phases.

4.1.4.1. Analysis of Proposition 1D1. The fuzzy QCA applied to the project phase practices (Table 5) used the representation of rules and ethical dilemmas (RDILEM), practices to minimize biases (PBIAS), and practices to provide transparency (TRANSP) in the AI system development (Appendix A - Annex 6). The low average score for practices aimed at transparency (50.11, see Appendix A - Annex 4) confirms the challenge of obtaining an explanation for the algorithm's results (Buiten, 2019; Butterworth, 2018; Zuiderwijk et al., 2021). Among other factors, this scenario may have been amplified by the fact that 73.33 % of the sample outsourced at least part of their AI system development, and only 46.43 % of the sample had access to their AI system code. The low average score for the representation of principles and ethical dilemmas (51.11, see Appendix A - Annex 4) may be due to the lack of a clear definition of those principles or the lack of specialists to implement the practice, as Ahn and Chen (2022) have alerted. Among the cases that had deployed practices to represent the business rules, principles and ethical dilemmas, and practices for transparency in AI system development, 75 % (cases 3, 9, 13, 21, 26, 27) showed high scores for AI governance actions at a strategic level (SAIGOV(1)).

When going deeper into the fuzzy-value analysis, we notice that those cases also implemented practices to minimize biases. The "necessary condition" test showed greater consistency (Cs = 0.96263) for the combination of the three practices. Therefore, the analysis of Proposition 1D1 revealed that in the studied sample, organizations that implemented practices to represent rules related to principles and ethical dilemmas, practices to provide transparency, and practices to minimize biases in AI systems during their development are associated with a more advanced stage in the implementation of AI governance actions at a strategic level.

Such result aligns with the researchers' arguments that, to ensure trustworthy AI systems, it is necessary to implement practices for

Table 3
Fuzzy QCA applied to Proposition 1B.

Proposition	Combinations	Cases	Results	Coverage and Consistency
1B	1. ∼PAUDIT * STAKEH * ENVIRO	27,25	SAIGOV(1)	Cv:0.319031
		2,16	SAIGOV(0)	Cs:0.893411
	2. PRISM * STAKEH * ENVIRO	3,9,13,22,27,8	SAIGOV(1)	Cv: 0.49135
		2	SAIGOV(0)	Cs: 0.941645
	3.∼PRISM*PAUDIT*STAKEH* ∼ ENVIRO	5,21	SAIGOV(1)	Cv: 0.293426 Cs: 0.925764

Table 4
Fuzzy QCA applied to Proposition 1C.

Proposition	Combinations	Cases	Results	Coverage and Consistency
1C	PSEC	5,8,9,13,22,3,6,19,25,26 11,2,16,20,28	SAIGOV(1) SAIGOV(0)	Cv:0.719031 Cs: 0.8312

Table 5
Fuzzy QCA applied to Proposition 1D1.

Proposition	Combinations		Cases	Results	Coverage and Consistency
1D1	RDILEM * TRANSP		9,13,27,3,21,26 15,24	SAIGOV(1) SAIGOV(0)	Cv: 0.538408 Cs: 0.913146
	"Necessary Conditions" Test	RDILEM * TRANSP		SAIGOV(1)	Cv: 0.727125 Cs: 0.929412
		RDILEM * TRANSP * PBIAS		SAIGOV(1)	Cv: 0.679531 Cs: 0.96263

Table 6
Fuzzy QCA applied to Proposition 1D2.

Proposition	Combinations		Cases	Results	Coverage and Consistency
1D2	HOVER * FEEDB		3,6,9,13,22,25,5,14, 19,26,27 10,24,16,20,28	SAIGOV(1) SAIGOV(0)	Cv: 0.741869 Cs: 0.790561
	"Necessary Conditions" Test	HOVER* FEEDB			Cv: 0.627069 Cs: 0.891349
		AMONI*HOVER*FEEDB			Cv: 0.576529 Cs: 0.972318

improving transparency along with the AI system development (Adadi & Berrada, 2018; Arrieta et al., 2020; Das, 2020; Dazeley et al., 2021; Kale et al., 2022; Phillips et al., 2021; Schaefer et al., 2021), for identification of ethical principles and dilemmas to be applied in the business rules (González et al., 2020; Rajkomar et al., 2018), and for mitigating biases in data preparation and in modeling (Ashokan & Haas, 2021; Baeza-Yates, 2018; De Silva & Alahakoon, 2022; Leavy et al., 2020; Lin et al., 2021; Makhlouf et al., 2021; Ntoutsi et al., 2020; Oneto & Chiappa, 2020; Silberg & Manyika, 2019).

4.1.4.2. Analysis of proposition 1D2. The analysis of practices in the operation phase of the AI system development and maintenance process (Table 6) considered the automatic monitoring (AMONI), human oversight (HOVER), and user feedback (FEEDB) variables (Appendix A - Annex 6). The highest average score was identified among the practices of this phase in automatic monitoring (71.57, see Appendix A - Annex 4), revealing the greater ease of monitoring when one does not depend on human resources. The preliminary fuzzy QCA showed a combination composed of human oversight and user feedback in sixteen organizations, of which 68.75 % (cases 3, 5, 6, 9, 13, 14, 19, 22, 25, 26, 27) obtained high scores for the AI governance actions at a strategic level (SAIGOV(1))). It is important to note that fifteen cases implemented automatic monitoring, human oversight, and user feedback, confirming the perceptions of Rahwan et al. (2019), Wright and Schultz (2018), and De Silva and Alahakoon (2022). Additionally, the "Necessary Conditions" test indicates that a high score would unlikely be obtained for AI governance actions at a strategic level without implementing the three practices (consistency = 0.972318). Therefore, there is an association between organizations at a more advanced stage in AI governance actions at a strategic level and cases that had automatic monitoring, human oversight, and user feedback practices, as argued by De Silva and Alahakoon (2022), Laato et al. (2022), Strauß (2021), González et al. (2020), and Zicari et al. (2021), when they demanded monitoring AI in the real environment.

4.2. Analysis of proposition 2

The analysis of Proposition 2 (Table 7) included the fuzzy QCA variables (Appendix A - Annex 6) training in data, AI risks, and AI ethical principles directed at decision-makers (TDEMAK); training in data, AI system development, AI risks, and AI ethical principles targeting developers (TDEVEL); training in data for users (TUSER); and training in data, AI risks and AI ethical principles for auditors (TAUDIT). Focusing on figuring out whether the mentioned trainings are enablers of AI governance practices, the overall score of actions towards AI governance was considered as the outcome variable (actions at the strategic level, ancillary processes and practices, and practices that belong to the AI system development process) (GAIGOV(1)).

Delivering a consistency of 0.937173, the fuzzy QCA revealed that the combination characterized by training aimed at decision-makers, developers, and users showed 87.5 % (cases 8, 9, 13, 22, 24, 27, 28) of these cases with high scores for actions focused on AI governance. Such result indicates an association between training key stakeholders and higher stage implementation of AI governance, as argued by Calzada & Almirall, 2020; Micheli et al., 2020; Ruijer, 2021; Makarius et al., 2020; Herremans, 2021, Pinski et al., 2024; Schüller, 2022.

Table 7
Fuzzy QCA applied to Proposition 2.

Proposition	Combinations	Cases	Results	Coverage and Consistency
2	1. TDEVEL * ~ TAUDIT	27,19,20,28 14,25,26	GAIGOV (1) GAIGOV (0)	Cv: 0.452348 Cs: 0.927762
	2. TDEMAK * TDEVEL * TUSER	13,22,27,8,9,24,28	GAIGOV (1)	Cv: 0.494475 Cs: 0.937173

Table 8
Crisp-set Fuzzy QCA applied to Proposition 3.

Proposition	Combinations	Cases	Results
3	1. TUSEAI{1}*DTDEMAK{0}* DTDEVEL{0}	6,3	HAIGOV(1)
		1,4,2,10,15,18,21,23	HAIGOV(0)
	2. ACOD80{0}*TUSEAI{0}* DTDEMAK{1}	5,20,28	HAIGOV(1)
		22	HAIGOV(0)
	3. ACOD80{0}*TUSEAI{0}* DTDEVEL{1}	19,20,28	HAIGOV(1)
		22	HAIGOV(0)
	4. TUSEAI{1}*DTDEMAK{1}* DTDEVEL{1}	8,9,13,27,24,25	HAIGOV(1)
		26	HAIGOV(0)

The absence of auditor training revealed that focus on internal audits for AI systems has not been a priority for those organizations, although a small group has provided the four trainings.

4.3. Analysis of proposition 3

Proposition 3 was analyzed through a crisp-set QCA (Table 8) using the following dichotomous variables (Appendix A - Annex 6): training for decision-makers (DTDEMAK), training for developers (DTDEVEL), the organization's access to at least 80 % of their AI system code (ACOD80) (four of the five AI systems reported in the questionnaire), and more than three years of experience in AI system development (TUSEAI). For the outcome variable, a dichotomous variable was created to indicate the sum of the fuzzy scores of all practices for implementing AI governance (HAIGOV). HAIGOV = 1, if = GAIGOV$\geq$60, and HAIGOV = 0 if GAIGOV<60.

Combination 1 — consisting of cases with more than three years in AI system production, without training for decision-makers and for AI system developers — was associated with low (not high) overall scores for all actions towards AI governance (cases 1, 2, 4, 10, 15, 18, 21, 23). At the other end, Combination 4 — comprising cases with more than three years of AI system production that have offered training to decision-makers and AI system developers — was associated with more advanced stages of the implementation of AI governance (cases 8, 9, 13, 24, 25, 27). Combinations 2 and 3 were not assertive enough to any conclusion.

It is worth observing that having (or not) access to the AI system codes, which implies outsourcing AI systems, did not impact Combinations 1 and 4. According to Combination 4, training decision-makers and developers can be associated with an advanced stage in implementing AI governance practices. And according to Combination 1, not training decision-makers or developers can be associated with lower stages in implementing AI governance. Both situations align with Ahn and Chen (2022) and Benfeldt et al. (2020). Preposition 3 reinforces the need to train key stakeholders even when the public organization outsources the development of AI systems.

5. Analysis of the interview responses and documents

The analysis of the interview responses and documents provided by the organizations was carried out while attempting to understand how practices and processes categorized in the research model were applied and used in the analysis of propositions.

5.1. Processes and practices for AI governance

In the context of actions, processes, and practices, many approaches were given and challenges were found in the path towards AI governance in the studied public organizations (Appendix A - Annex 7)(CNJ, 2020; LIAA-3R, 2022; Nagbøl & Müller, 2020; Nagbøl, Müller, & Krancher, 2021; Vero, 2019). Since all efforts towards AI governance are distributed at many levels of the organization hierarchy, it reflects the organization's culture and its risk appetite.

5.2. Government standards and guidelines

Along with the interviews, one observes that organizations whose governments have already established AI ethical principles guidelines for all their agencies have adopted those principles completely and, in a few cases, they have added details to their policies or strategies to customize the guidelines to their singularities (See "Policies and Guidelines for AI Ethical Principles" in Appendix A - Annex 7 and Appendix A - Annex 9). Similarly, in some cases, governments create agencies specialized in developing standards for processes and practices related to AI governance (See Appendix A - Annex 8). In both cases, those organizations saved time, money, and human resources, as one can observe in Appendix A - Annex 10.

Regarding the construction of standards and transfer of knowledge, it is worth highlighting a partnership between the public and private sectors established by the Finnish Government (Aurora, 2019) and an agreement between Nordic countries to implement best practices for AI systems with a focus on ethical issues (Nordic Council of Ministers, 2018).

5.3. Outsourcing

In 73.33 % of the organizations, third parties were hired or partnered with to develop at least part of their AI systems. Considering that public organizations are not self-sufficient to produce AI systems on the scale and with the level of expertise they require, according to Hickok (2022), Zick et al. (2024) and Coglianese (2024), outsourcing is also a driver to implement practices for AI governance (Appendix A - Annex 10). A few organizations were inspired by the World Economic Forum's model (WEF, 2020, 2020b) for outsourcing AI system development compliant with ethical principles.

5.4. Partnership between the business unit and the IT unit

Among the interviewees, there is the perception of the "business + IT" joint action as a strategy to minimize biases, provide transparency, and implement data governance. A second group of reports was provided by professionals who implemented the risk management process and the AI system development process, with artifacts filled out by the IT and business staff (Nagbøl et al., 2021). In Annex 10, one can find some of the interviewees' responses and comments regarding the "business + IT" partnership.

6. Merging the analyses

6.1. Associations found in the QCA

The proposition results are summarized as follows:

Proposition 1 results: An association was found between AI governance actions at the strategic level ("create an AI governance structure," "elaborate an AI strategy," "establish an AI policy," "implement an AI governance process," "establish an AI ethical principles code") and the existence of data-related processes (data governance, data quality management, personal data protection management), risk-related processes and practices (risk management process, stakeholder definitions, monitoring changes in the environment), security management processes, AI system development practices (rules representing ethical principles and ethical dilemmas, biases minimization, transparency, automatic monitoring, human oversight, and feedback collection). No association was found between the audit process and the strategic actions for AI governance.

Proposition 2 results: An association was found between advanced stages of AI governance implementation and training targeting decision-makers, AI system developers, and digital services users, confirming that training such stakeholders is a driver for implementing AI governance.

Proposition 3 results: In the context of organizations with more than three years of experience in AI system production, regardless of the make-or-buy decision, training decision-makers and developers is associated with an advanced stage of implementing all AI governance practices. And not training decision-makers or developers is associated with lower stages in implementing AI governance. Thus, regardless of the make-or-buy decision, training is a driver of AI governance implementation.

6.2. Relevant contributions from the government to the whole public sector

Interviews and documents made available by governments showed the benefits of having a central government body that produces clear and accessible guidelines with recommendations for AI system development, which confirms Mikalef et al.'s (2022b) and Schaefer et al.'s (2021) perceptions. Some guides made up the governments' portfolio of standards for promoting AI governance in agencies and departments under their responsibility (Australian Government, 2019a, 2019b; Norwegian Data Protection Authority, 2018; Government of United Kingdom, 2017, 2019b, 2020a, 2020b, 2020c, 2020d, 2020e, 2021a, 2021b, 2021c, 2021d, 2021e, 2022a, 2022b, 2022c, 2022d, 2022e; Information Commissioner's Office, 2020, 2021; Ekspertgruppen om dataetik, 2018; Balahur et al., 2022; German Federal Ministry for Economic Affairs and Energy, 2020; AI HLEG, 2019; Government of Canada, 2019a, 2019b, 2020a, 2020b, 2021a, 2021b, 2022a, 2022b; Leslie, 2019; National Institute of Advanced Industrial Science and Technology, 2022; European Union Agency for Cyber Security, 2021; WEF, 2020a, 2020b; C4IR Brasil, 2022; Switzerland Federal Council, 2021, Council of Europe, 2021, European Data Protection Board, 2022) (Appendix A - Annex 8). In addition to supplying knowledge that is lacking in many public organizations, these specialized institutions speed up implementation and promote a standard for concepts that facilitates communication among government departments. In a strategic context, some governments (German Federal Government, 2020;, Presidencia de

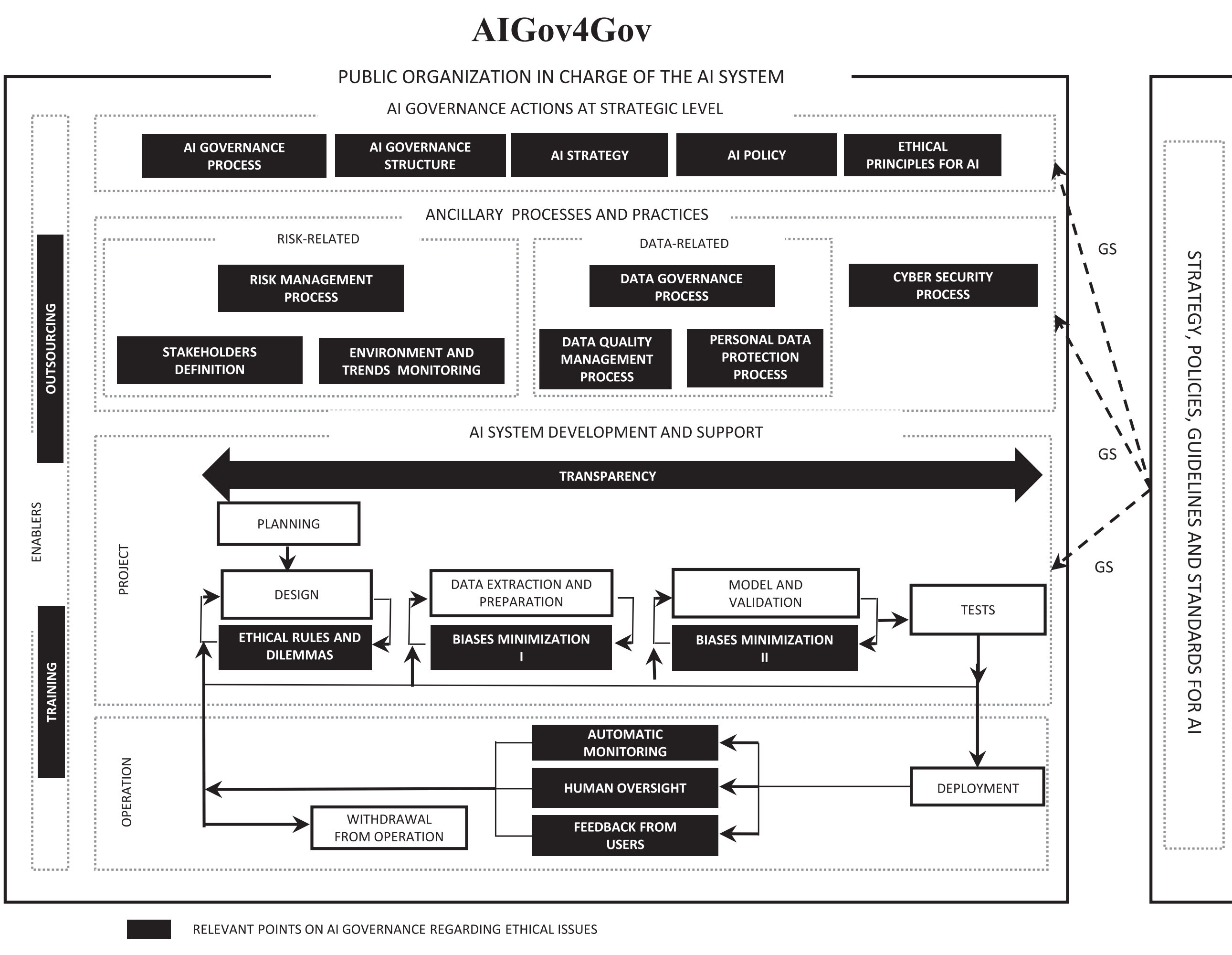

Fig. 6. AIGov4Gov – AI governance framework proposed for public organizations. (Source: Self-elaboration)

la Nación, 2019; Australian Government, 2021; Ministério da Ciência Tecnologia e Inovação do Brasil, 2021; Government of Canada, 2022b; Danish Government, 2019; Government of the Republic of Estonia, 2019; Ministry of Economic Affairs and Employment of Finland, 2017, 2019; Ministero dello sviluppo economico, 2019; Japanese Strategic Council for AI Technology, 2017; Ekspertgruppen om dataetik, 2018; Government of the Grand Duchy of Luxembourg, 2018; Norwegian Ministry of Local Government and Modernisation, 2020; Government of United Kingdom, 2021e; Government of Sweden, 2020; European Commission, 2018a) assign AI strategic guidelines to an agency that develops and monitors the national AI strategy and establishes policy and ethical principles for AI systems (Appendix A - Annex 9). In both cases, public organizations can not only adopt the government's central strategic guidelines and standards but can also build their own versions in line with the government's general definitions.

6.3. Drivers to AI governance implementation

Combining the findings of Propositions 2 and 3 with the interviews, training key stakeholders and outsourcing were considered drivers of AI governance implementation in public organizations.

6.4. AIGov4Gov framework

Consolidating the QCA results and the interviews' results, a conceptual view was built in a multilevel approach encompassing practices and processes found in the sample aimed at the production of AI systems that considered ethical principles, drivers to implement AI governance, and the Government Central Agency contributions to the public sector: the AIGov4Gov framework (Fig. 6).

Reaching the strategic, tactical, and operational levels, AIGov4Gov is an AI governance model for public organizations. At the strategic level, there are actions aimed at elaborating an AI strategy, establishing an AI policy, establishing an AI ethical principles code, implementing an AI governance process, and creating an AI governance structure for said governance. To support those actions, ancillary processes and practices are combined for a) data governance, supported by data quality management and personal data protection management; b) AI-related risk mitigating using a risk management process supported by a stakeholders' definition and by an environment and social trends monitoring in line with an audit process; c) Cyber security management. The expected increase of legislation to regulate AI worldwide was considered when deciding to maintain the audit process within the framework. At the tactical and operational levels, AIGov4Gov provides practices for AI system development and maintenance processes with a focus on ethical principles and aligned to researchers' argument for an AI system development process based on agile methods, which use continuous loops in each phase (Laato et al., 2022), and which consider ethical principles in the loops (Leijnen et al., 2020; Lu et al., 2024). Right after planning, development takes place in several interactions during problem specification and with the representation of ethical principles and dilemmas. Then, practices for minimizing biases are applied in successive iterations during the data extraction and preparation phase, as well as during the model construction and validation stage, followed by testing. Soon after deployment, practices to follow-up AI systems are implemented in the complexity of the real environment through automatic monitoring, human oversight, and collection of user feedback.

As an enabler of AI governance implementation, AIGov4Gov provides outsourcing in addition to training for key stakeholders, like decision-makers, developers, and users, customized for their role in the implemented processes and practices. Auditor training can also be considered when aiming for a scenario where AI legislation is a comprehensive reality.

Also included in AIGov4Gov is the interaction between the public organization in charge of the AI system and the agency in its sphere of government (if any) in charge of AI strategy, AI policy, and AI ethical principles for AI systems applicable to organizations under its responsibility. In such a situation, the organization in charge of the AI system can adopt its centralized government strategic guidelines or adapt them while complying with them. Similarly, guidelines are provided with standards for risk management, data governance, data quality management, personal data protection management, cyber security management, minimization of biases, and transparency throughout the AI system development process (GS Flow in Fig. 6). When the organization decides to outsource, the AI system development process still requires the "business + IT" partnership to enable the ancillary processes and practices for AI governance.

7. Conclusions

This study investigated how public organizations have incorporated the guidelines presented by academia, international standards, and legislation for AI system development, considering ethical principles in their governance and management processes. The results confirmed the perception that AI governance requires a multilayer model with strategic-level actions that guide processes and ancillary AI governance practices in a combined action. All processes and practices designed in the research model were observed in the sample. However, data-related, risk-related, and AI system development processes and practices were prioritized in the sample for AI governance implementation. The cyber security management process received lower adherence, and audit processes were still seldom adopted at the time of data gathering when very few countries had approved laws to regulate AI. Regarding development in the project phase, organizations at a more advanced stage in AI governance have prioritized practices representing ethical principles and ethical dilemmas, transparency practices, or practices to minimize biases. When those AI systems are in operation, organizations at a more advanced stage in AI governance have implemented practices for automatic monitoring, human oversight, and collection of user feedback.

Training key stakeholders and outsourcing are enablers of an AI governance implementation. It was observed that organizations that have outsourced their AI system development have also trained managers as well as AI developers, and the lack of training of those professionals is associated with less advanced stages in AI governance.

One could observe the huge opportunity that government agencies have to promote AI governance by defining guidelines for all organizations under their responsibility or recommending standards for AI governance despite the long time required to deploy them. In the context of public organizations, it is worth mentioning the need for policy-making that combines an internal multilayered approach with a continuous alignment with the government guidelines and standards. The findings also corroborated to a proposed framework — AIGov4Gov — encompassing combined processes and practices to establish AI governance in a public organization.

8. Limitations and agenda

Limitations found in this research:

a) Despite the broad and systematic process of obtaining the sample, this research carried out analyses in countries and organizations that published their AI systems, made contacts available, and agreed to participate in the research. Therefore, despite efforts to include representatives from all countries with a high level of AI system production, the sample does not follow the global AI ranking proportions, nor does it have a balanced representation of each continent.

b) The research focused on capturing the existence of practices and processes but did not delve into each process and its maturity model. Thus, each participant had his own perception regarding whether a practice/process was being implemented completely or partially. For the same reason, the research did not make a deeper analysis of the quality of the training offered to users, developers, decision-makers, and auditors.

c) Similar to the previous item, the government AI standards found in the sample were not classified considering their maturity.

d) The research encompassed only practices and processes representing all efforts to implement AI governance. Impacts on corporate governance and e-government were out of the scope.

e) The AI systems presented by the organizations in the sample did not consider generative AI, probably because the gathering criteria were systems that were in operation and were already part of the organization's official portfolio when the data were collected.

As an exploratory and descriptive study, this research paves the way for an agenda of new investigations that go deeper into the findings regarding each proposition, as well as an investigation of how practices and processes studied under the AI governance effort would impact corporate governance and e-government. Finally, a space is opened for deepening the AI system lifecycle through a maturity model for AI system development and support focusing on ethical principles.

9. Contribution

This research presents itself as innovative in terms of content, as it addressed the gap highlighted by Mäntymäki et al. (2022) and Mikalef, Conboy, et al. (2022) in the empirical knowledge of how organizations have interpreted and incorporated AI system development best practices into their processes and practices. The research has innovated by using crisp-set QCA, fuzzy QCA, and content analyses as it addressed the gaps presented by Zuiderwijk et al. (2021). Therefore, the following contributions to managers and researchers can be summarized: a) identification of how processes and practices aimed at applying ethical principles in AI system development have been combined and internalized in the governance and management models of public organizations; b) identification of how AI governance enablers have been used by public organizations; c) how a central government agency can booster AI governance in government agencies; d) a framework for AI governance in public organizations, in which processes and practices are articulated at the strategic, tactical, and operational levels in AI system production that consider ethical principles.

CRediT authorship contribution statement

Patricia Gomes Rêgo de Almeida: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Carlos Denner dos Santos Júnior:** Supervision, Writing- original draft, Investigation, Visualization, Validation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.giq.2024.102003.

References

Aasi, P., Rusu, L., & Han, S. (2014). The influence of culture on IT governance: A literature review. In *47th Hawaii international conference on system sciences* (pp. 4436–4445). https://doi.org/10.1109/HICSS.2014.546

Abdollahi, B., & Nasraoui, O. (2018). Transparency in Fair machine learning: The case of explainable recommender systems. In J. Zhou, & F. Chen (Eds.), *Human and Machine Learning. Human–Computer Interaction Series*. Cham: Springer. https://doi.org/10.1007/978-3-319-90403-0_2.

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management, 49*(July), 424–438. https://doi.org/10.1016/j.ijinfomgt.2019.07.008

Aburachid, L. M. C., & Greco, P. J. (2011). Validação de conteúdo de cenas do teste de conhecimento tático no tênis. *Estudos de Psicologia. Campinas, 28*(2), 261–267.

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access Review, 6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

Agarwal, L.. Defining organizational AI governance and ethics. Available at SSRN: https://ssrn.com/abstract=4553185.

Ahn, M. J., & Chen, Y. (2022). Digital transformation toward AI-augmented public administration: The perception of government employees and the willingness to use AI in government. *Government Information Quarterly, 39), Issue 2*. https://doi.org/10.1016/j.giq.2021.101664

AI HLEG. (2019). Ethics guidelines for trustworthy AI. European Commission. Retrieved from https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

Aiken, C. (2021). *Classifying AI systems. Center for Security and Emerging Technology*. Georgetown University. Retrieved from https://cset.georgetown.edu/publication/classifying-ai-systems/ Accessed September 23, 2024.

Alhosani, K., & Alhashmi, S. M. (2024). Opportunities, challenges, and benefits of AI innovation in government services: A review. *Discover Artificial Intelligence, 4*, 18. https://doi.org/10.1007/s44163-024-00111-w

Alshahrani, A., Dennehy, D., & Mäntymäki, M. (2021). An attention-based view of AI assimilation in public sector organizations: The case of Saudi Arabia. *Government Information Quarterly, 39), Issue 4*. https://doi.org/10.1016/j.giq.2021.101617

Anderson, M., & Anderson, S. L. (2018). Geneth: A general ethical dilemma analyzer. De Gruiter. Paladyn. *J. Behav. Robot., 9*, 337–357. https://doi.org/10.1515/pjbr-2018-0024

Andrews, L. (2018). Public administration, public leadership and the construction of public value in the age of the algorithm and 'big data'. *Public Administration, 97*(2), 296–310. https://doi.org/10.1111/padm.12534

Arrieta, A. B., Díaz-Rodríguez, N., Ser, J. D., Bennetot, A., Tabik, S., Barbado, A., … Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for artificial intelligence and digital technologies. *International Journal of Information Management, 62*. https://doi.org/10.1016/j.ijinfomgt.2021.102433

Ashokan, A., & Haas, C. (2021). Fairness metrics and bias mitigation strategies for rating predictions. *Information Processing & Management, 58*(5), Article 102646. ISSN 0306–4573 https://doi.org/10.1016/j.ipm.2021.102646.

Aurora, A. I. (2019). Aurora AI - Towards a human Centric Society. Retrieved from https://vm.fi/documents/10623/1464506/AuroraAI+development+and+implementation+plan+2019%E2%80%932023.pdf.

Australian Government. (2019). *Artificial intelligence – Australia's ethics framework*. Innovation and Science: Department of Industry. Retrieved from https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework Accessed September 23, 2024.

Australian Government. (2019b). Australian Privacy Principles Guidelines. Office of Australian Information Commissioner. Retrieved from https://www.oaic.gov.au/__data/assets/pdf_file/0009/1125/app-guidelines-july-2019.pdf Accessed September 23, 2024.

Australian Government. (2021). Australian's Artificial Intelligence Action Plan. Retrieved from https://www.industry.gov.au/publications/australias-artificial-intelligence-action-plan.

Awad, E., Dsouza, S., Bonnefon, J. F., Shariff, A., & Rahwan, I. (2020). Crowdsourcing moral machines. *Communications of the ACM, 63*(3), 48–55. https://doi.org/10.1145/3339904

Baeza-Yates, R. (2018). Bias on the web. *Communications of the ACM, 61*(6), 54–61. https://doi.org/10.1145/3209581

Balahur, A., Jenet, A., Torres, I., Charisi, V., Ganesh, A., Griesinger, C. B., Maurer, P., Mian, L., Salvi, M., Scalzo, S., Sol er Garrido, J., Taucer, F., & Tolan, S. (2022). *Data quality requirements for inclusive, non-biased and trustworthy AI. Putting-Science-Into-Standards*. Luxembourg: Publications Office of the European Union. https://doi.org/10.2760/365479

Bench-Capon, T., & Modgil, S. (2017). Norms and value based reasoning: Justifying compliance and violation. *Artificial Intelligence & Law Review, 25*, 29–64. https://doi.org/10.1007/s10506-017-9194-9

Benfeldt, O., Persson, J. S., & Madsen, S. (2020). Data governance as a collective action problem. *Information Systems Frontiers, 22*, 299–313. https://doi.org/10.1007/s10796-019-09923-z

Betarelli-Júnior, A. A., & Ferreira, S. F. (2018). *Introdução à Análise Qualitativa Comparativa e aos Conjuntos Fuzzi (FSQCA)*. Enap: Brasilia.

Bonnemains, V., Saurel, C., & Tessier, C. (2018). Embedded ethics: Some technical and ethical challenges. *Ethics and Information Technology, 20*, 41. https://doi.org/10.1007/s10676-018-9444-x

Bonsón, E., Lavorato, D., Lamboglia, R., & Mancini, D. (2021). Artificial intelligence activities and ethical approaches in leading listed companies in the European Union. *International Journal of Accounting Information Systems, 43*. https://doi.org/10.1016/j.accinf.2021.100535, 1467–0895.

Booth, J., Metz, D. W., Tarkhanyan, D. A., & Cheruvu, S. (2023). Machine learning security and trustworthiness. In *Demystifying intelligent multimode security systems*. Berkeley, CA: Apress. https://doi.org/10.1007/978-1-4842-8297-7_5.

Boyd, R., & Holton, R. J. (2018). Technology, innovation, employment and power: Does robotics and artificial intelligence really mean social transformation? *Journal of Sociology, 54*(3), 331–345.

Breier, J., Baldwain, A., Balinsky, & Liu, Y. (2020). Risk Management for Machine Learning Security. *arXiv*. https://doi.org/10.48550/arXiv.2012.04884, 2012.04884v1 [cs.CR].

Buiten, C. M. (2019). Towards intelligent regulation of artificial intelligence. *European Journal of Risk Regulation, 10*(1), 41–59. https://doi.org/10.1017/err.2019.8

Butterworth, M. (2018). The ICO and artificial intelligence: The role of fairness in the GDPR framework. *Computer Law and Security Review, 34*, 257–268. https://doi.org/10.1016/j.clsr.2018.01.004

C4IR Brasil. (2022). Guia de Contratações Públicas de Inteligência Artificial. Centro para a 4ª Revolução Industrial. Retrieved from https://c4ir.org.br/wp-content/uploads/2022/11/1648128585465GUIA-DE-CONTRATACOES-PUBLICAS-DE-AI_C4IR_v4.pdf Accessed September 23, 2024.

Calzada, I., & Almirall, E. (2020). Data ecosystems for protecting European citizens' digital rights. *Transforming Government: People, Process and Policy, 14*(2), 133–147. https://doi.org/10.1108/TG-03-2020-0047

Carretero, A., Gualo, F., Caballero, I., & Piattini, M. (2017). *MAMD 2.0: Environment for data quality processes implantation based on ISO 8000-6X and ISO/IEC 33000. 54* pp. 139–151). Elsevier BV. https://doi.org/10.1016/j.csi.2016.11.008

Carter, D. (2020). (2020). Regulation and ethics in artificial intelligence and machine learning technologies: Where are we now? Who is responsible? Can the information professional play a role? *Business Information Review, 37*(2), 60–68. https://doi.org/10.1177/0266382120923962

Cerka, P., Grigiene, J., & Sirbikyte, G. (2017). Is it possible to grant legal personality to artificial intelligence software systems? *Computer Law and Security Review, 33*(5), 685–699.

Chen, T., Liu, J., Xiang, Y., Niu, W., Tong, E., & Han, Z. (2019). Adversarial attack and defense in reinforcement learning-from AI security view. *Cybersecurity, 2*, 1–22. https://doi.org/10.1186/s42400-019-0027-x

Chen, X., & Deng, Y. (2022). An evidential software risk evaluation model. *Mathematics, 10*, 2325. https://doi.org/10.3390/math10132325

Cihon, P., Maas, M. M., & Kempo, L. (2020). Should artificial intelligence governance be centralized?: Design lessons from history. In *AAAI/ACM conference on AI, ethics, and society (AIES '20)* (pp. 228–234). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3375627.3375857.

CNJ. (2020). Resolução N° 332 de 21/08/2020. Retrieved from https://atos.cnj.jus.br/atos/detalhar/3429 Accessed September 23, 2024.

Codá, R. C., Farias, J. S., & Dias, C. (2022). Interactive value formation and lessons learned from Covid-19: The Brazilian case. *Journal of Quality Assurance in Hospitality and Tourism*. https://doi.org/10.1080/1528008X.2022.2135057

Coglianese, C. (2024). Procurement and artificial intelligence. In *Handbook on public policy and artificial intelligence* (pp. 235–248). Edward Elgar Publishing. https://doi.org/10.4337/9781803922171.00026.

Council of Europe. (2021). Guidelines on facial recognition. Consultative Committee of the Convention for the protection of individuals with regard to automatic processing of personal data. Convention 108. Retrieved from https://rm.coe.int/guidelines-facial-recognition-web-a5-2750-3427-6868-1/1680a31751 Accessed September 23, 2024.

Dalrymple, D., Skalse, J., Bengio, Y., Russell, S., Tegmark, M., Seshia, S., Omohundro, S., Szegedy, C., Goldhaber, B., Ammann, N., Abate, A., Halpern, J., Barrett, C., Zhao, D., Zhi Xuan, T., Wing, J., & Tenenbaum, J. (2024). *Towards Guaranteed Safe AI: A Framework for Ensuring Robust and Reliable AI Systems*. arXiv:2405.06624v3 [cs.AI]. https://doi.org/10.48550/arXiv.2405.06624.

Danish Government. (2019). *National Strategy for Artificial Intelligence*. Ministry of Finance and Ministry of Industry, Business and Financial Affairs. Retrieved from https://en.digst.dk/strategy/the-danish-national-strategy-for-artificial-intelligence/ Accessed September 23, 2024.

Danish Government. (2020). Lov om ændring af årsregnskabsloven. (Krav om rapportering af dataetik). Retrieved from https://www.retsinformation.dk/eli/lta/2020/741 Accessed September 23, 2024.

Das, A. (2020). *Opportunities and challenges in explainable artificial intelligence 9XAI: A survey*. arXiv:2006.11371 [cs.CV]. https://doi.org/10.48550/arXiv.2006.11371.

Dazeley, R., Vamplew, P., Foale, C., Young, C., Aryal, S., & Cruz, F. (2021). Levels of explainable artificial intelligence for human-aligned conversational explanations. *Artificial Intelligence, 299, 103525*. https://doi.org/10.1016/j.artint.2021.103525. ISSN 0004-3702.

De Almeida, P. G. R., dos Santos, C. D., & Farias, J. S. (2021). Artificial intelligence regulation: A framework for governance. *Ethics and Information Technology, 23*, 505–525. https://doi.org/10.1007/s10676-021-09593-z

De Oliveira, T. F. (2019). *Avaliação das Práticas de Auditoria Interna da Secretaria Federal de Controle Interno da CGU sob a Ótica da Auditoria Baseada em Riscos*. Brasil: Controladoria Geral da União. Retrieved from https://revista.cgu.gov.br/Revista_da_CGU/article/view/73/pdf_60 Accessed September 23, 2024.

De Silva, D., & Alahakoon, D. (2022). An artificial intelligence life cycle: From conception to production. *Patterns, 3*(6), Article 100489. https://doi.org/10.1016/j.patter.2022.100489

Dennis, L., Fisher, M., Slavkovik, M., & Webster, M. (2016). Formal verification of ethical choices in autonomous systems. *Robotics and Autonomous Systems, 77*, 1–14. ISSN 0921-8890, https://doi.org/10.1016/j.robot.2015.11.012.

Dias, O. C. (2011). *Análise Qualitativa Comparativa (QCA) Usando Conjuntos Fuzzy – Uma Abordagem Inovadora Para Estudos Organizacionais no Brasil*. Rio de Janeiro: XXXV Encontro da ANPAD.

Dignum, V. (2019). AI is multidisciplinar. *AI Matters, 5*(4), 19–21. https://doi.org/10.1145/3375637.3375644

Djeffal, C. (2018). *Sustainable AI Development (SAID): On the Road to More Access to Justice*. https://doi.org/10.2139/ssrn.3298980

Duijm, N. J. (2015). Recommendations on the use and design of risk matrices. *Safety Science, 76*, 21–31. ISSN 0925-7535. https://doi.org/10.1016/j.ssci.2015.02.014

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., … Williams, M. D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management, 57*, Article 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Ee, S., O'Brien, J., Williams, Z., El-Dakhakhni, A., Aird, M., & Lintz, A. (2024). Adapting cybersecurity frameworks to manage frontier AI risks: A defense-in-depth approach. *arXiv*. https://doi.org/10.48550/arXiv.2408.07933, 2408.07933v1 [cs.CY].

Eggers, S., & Sample, C. (2020). *Vulnerabilities in artificial intelligence and machine learning applications and data*. Idaho National Laboratory. US. Retrieved from https://inldigitallibrary.inl.gov/sites/sti/sti/Sort_57369.pdf Accessed September 23, 2024.

Eke, D. O., Chintu, S. S., & Wakunuma, K. (2023). Towards shaping the future of responsible AI in Africa. In D. O. Eke, K. Wakunuma, & S. Akintoye (Eds.), *Responsible AI in Africa. Social and cultural studies of robots and AI*. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-031-08215-3_8.

Ekspertgruppen om dataetik. (2018). Data i menneskets tjeneste Anbefalinger fra Ekspertgruppen om dataetik. Retrieved from https://www.em.dk/media/12013/ekspertgruppens-afrapportering-inkl-anbefalinger_final-a.pdf Accessed September 23, 2024.

Erlina, E., Nasution, A. A., Yahy, I., & Atmanegara, A. W. (2020). The role of risk based internal audit in improving audit quality. Erlina, Abdillah Arif Nasution, Idhar Yahya and Agung Wahyudhi Atmanegara, The Role of Risk Based Internal Audit in Improving Audit Quality. *International Journal of Management, 11*(12), 299–310.

European Commission. (2018). AI Watch. Retrieved from https://ai-watch.ec.europa.eu/index_en Accessed September 23, 2024.

European Commission. (2018a). Artificial Intelligence for Europe. Retrieved from https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence Accessed September 23, 2024.

European Commission. (2021a). Regulation of the European Parliament and of the Council Laying down harmonized rules on artificial intelligence. Artificial Intelligence Act and amending certain union legislative acts. Brussels. Retrieved from https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF Accessed September 23, 2024.

European Commission. (2021b). Selected AI cases in the public sector. Joint Research Centre [Dataset] PID. Retrieved from http://data.europa.eu/89h/7342ea15-fd4f-4184-9603-98bd87d8239a Accessed September 23, 2024.

European Data Protection Board. (2022). Guidelines 05/2022 on the use of facial recognition technology in the area of law enforcement. Version 1.0 Retrieved from https://edpb.europa.eu/our-work-tools/documents/public-consultations/2022/guidelines-052022-use-facial-recognition_en Accessed September 23, 2024.

European Parliament. (2022a). *Draft Report on the proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM2021/0206 – C9–0146/2021–2021/0106(COD))*.

European Parliament. (2022b). Regulatory divergences in the draft AI act: Differences in public and private sector obligations, Study, European Parliamentary Research Service (EPRS), Brussels. Retrieved from https://www.europarl.europa.eu/RegData/etudes/STUD/2022/729507/EPRS_STU(2022)729507_EN.pdf Accessed September 23, 2024.

European Union Agency for Cyber Security. (2021). *Securing Machine Learning Algorithms*. ISBN: 978–92–9204-543-2. https://doi.org/10.2824/874249.

European Union Agency for Cyber Security. (2022). Risk Management Standards – Analysis of standardisaton requeriments in supporting of cybersecurity policy. Retrieved from https://www.enisa.europa.eu/publications/risk-management-standards Accessed September 23, 2024.

FCT. (2021). Research in Data Science and Artificial Intelligence applied to Public Administration. Retrieved from https://www.fct.pt/wp-content/uploads/2022/06/Brochura_ResearchinDataScienceandAIappliedtoPA.pdf Accessed September 23, 2024.

Fernandes, G., Domingues, J., Tereso, A., & Pinto, E. (2021). A Stakeholders' perspective on risk Management for Collaborative. University-industry R&D programs. *Procedia Computer Science, 181*, 110–118. ISSN 1877-0509. https://doi.org/10.1016/j.procs.2021.01.110

Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). *Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI* (p. 1). Berkman Klein Center Research Publication.

Freitas, V. S., & Neto, F. B. (2016). Qualitative Comparative Analysis (QCA): usos e aplicações do método. *Revista Política Hoje, 2ª Edição, 24*, 103–117.

Gerkensmeier, B., & Ratter, B. M. W. (2018). Governing coastal risks as a social process—Facilitating integrative risk management by enhanced multi-stakeholder collaboration. *Environmental Science & Policy, 80*, 144–151. ISSN 1462-9011. https://doi.org/10.1016/j.envsci.2017.11.011

German Federal Government. (2020). Artificial intelligence strategy of the German Federal Government. Retrieved from https://www.ki-strategie-deutschland.de/home.html Accessed September 23, 2024.

German Federal Ministry for Economic Affairs and Energy. (2020). German Standardization Roadmap on Artificial Intelligence. Retrieved from https://www.din.de/resource/blob/772610/e96c34dd6b12900ea75b460538805349/normungsroadmap-en-data.pdf Accessed September 23, 2024.

Gianni, R., Lehtinen, S., & Nieminen, M. (2022). Governance of responsible AI: From ethical guidelines to cooperative policies. *Frontiers in Computer Science, 4*, Article 873437. https://doi.org/10.3389/fcomp.2022.873437

González, F., Ortiz, T., & Ávalos, R. S. (2020). Responsible use of AI for public policy: Data science toolkit. In *OECD e IDB*. Retrieved from https://publications.iadb.or

g/publications/english/document/Responsible-use-of-AI-for-public-policy-Data-science-toolkit.pdf Accessed September 23, 2024.
Government of Canada. (2019a). Directive on Identity Management. Retrieved from https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=16577 Accessed September 23, 2024.
Government of Canada. (2019b). Directive on Security Management. Retrieved from https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32611 Accessed September 23, 2024.
Government of Canada. (2020). Guidelines on Service and Digital. Retrieved from https://www.canada.ca/en/government/system/digital-government/guideline-service-digital.html#ToC4_5 Accessed September 23, 2024.
Government of Canada. (2020a). Algorithmic Impact Assessment. Retrieved from https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html Accessed September 23, 2024.
Government of Canada. (2021). Directive on Automated Decision-making. Retrieved from https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592 Accessed September 23, 2024.
Government of Canada. (2021d). Levels of security. National Security and defense. Retrieved from https://www.tpsgc-pwgsc.gc.ca/esc-src/protection-safeguarding/niveaux-levels-eng.html.
Government of Canada. (2022a). Responsible use of artificial intelligence – Our guiding principles. Retrieved from https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai.html#toc1 Accessed September 23, 2024.
Government of Canada. (2022). Pan-Canadian Artificial Intelligence Strategy. Retrieved from https://ised-isde.canada.ca/site/ai-strategy/en Accessed September 23, 2024.
Government of India. (2020). Artificial Intelligence – Use Case Compendium. Ministry of Housing and Urban Affairs. Retrieved from https://iccc.smartcities.gov.in/pdf/AI-Use-Case-Compendium.pdf Accessed September 23, 2024.
Government of Sweden. (2020). AI Sweden. Retrieved from https://www.ai.se/en/news/swedish-government-establishes-new-ai-commission Accessed September 23, 2024.
Government of the Grand Duchy of Luxembourg. (2018). Artificial Intelligence: a strategic view for Luxembourg. Retrieved from https://innovative-initiatives.public.lu/initiatives/artificial-intelligence-strategic-vision-luxembourg Accessed September 23, 2024.
Government of the Republic of Estonia. (2019). Kratt – Estonian Artificial Intelligence Deployment. Retrieved from https://f98cc689-5814-47ec-86b3-db505a7c3978.filesusr.com/ugd/7df26f_27a618cb80a648c38be427194affa2f3.pdf Accessed September 23, 2024.
Government of United Kingdom. (2017). Public sector use of the cloud. Retrieved from https://www.gov.uk/guidance/public-sector-use-of-the-public-cloud Accessed September 23, 2024.
Government of United Kingdom. (2019). Understanding artificial intelligence ethics and safety. Retrieved from https://www.gov.uk/guidance/understanding-artificial-intelligence-ethics-and-safety Accessed November 4, 2022.
Government of United Kingdom. (2020a). Data Ethics framework. Government Digital Service. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/923108/Data_Ethics_Framework_2020.pdf Accessed September 23, 2024.
Government of United Kingdom. (2020b). Guidelines for AI procurement. Retrieved from https://www.gov.uk/government/publications/guidelines-for-ai-procurement/guidelines-for-ai-procurement Accessed September 23, 2024.
Government of United Kingdom. (2020c). *Guidelines for AI procurement. A summary of best practices addressing specific challenges of acquiring Artificial Intelligence in the public sector.* UK Office for Artificial Intelligence. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/990469/Guidelines_for_AI_procurement.pdf Accessed September 23, 2024.
Government of United Kingdom. (2020d). The Government Data Quality Framework. Retrieved from https://www.gov.uk/government/publications/the-government-data-quality-framework/the-government-data-quality-framework-guidance Accessed February 26, 2023.
Government of United Kingdom. (2020e). A guide to using artificial intelligence in the public sector. https://www.gov.uk/government/publications/a-guide-to-using-artificial-intelligence-in-the-public-sector Accessed September 23, 2024.
Government of United Kingdom. (2021a). Ethics, Transparency and Accountability Framework for Automated Decision-Making. https://www.gov.uk/government/publications/ethics-transparency-and-accountability-framework-for-automated-decision-making/ethics-transparency-and-accountability-framework-for-automated-decision-making Accessed September 23, 2024.
Government of United Kingdom. (2021b). Algorithmic transparency template. Retrieved from https://www.gov.uk/government/publications/algorithmic-transparency-template/algorithmic-transparency-template Accessed September 23, 2024.
Government of United Kingdom. (2021c). Algorithmic Transparency Standard. https://www.gov.uk/government/collections/algorithmic-transparency-standard Accessed December 4, 2022.
Government of United Kingdom. (2021d). Using personal data in your business or other organization. Retrieved from https://www.gov.uk/guidance/using-personal-data-in-your-business-or-other-organisation#data-protection-and-gdpr Accessed September 23, 2024.
Government of United Kingdom. (2021e). Data Protection Impact Assessments. Retrieved from https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-assessments Accessed September 23, 2024.
Government of United Kingdom. (2022a). *Standard. Food Standards Agency*. Food Hygiene Rating Scheme – AI. Retrieved from https://www.gov.uk/government/publications/food-standards-agency-food-hygiene-rating-scheme-ai/food-standards-agency-food-hygiene-rating-scheme-ai Accessed September 23, 2024.
Government of United Kingdom. (2022b). Ethics self-assessment tool. UK statistics authority. Retrieved from https://uksa.statisticsauthority.gov.uk/the-authority-board/committees/national-statisticians-advisory-committees-and-panels/national-statisticians-data-ethics-advisory-committee/ethics-self-assessment-tool/.
Government of United Kingdom. (2022c). Equality Impact Assessment. Retrieved from https://www.gov.uk/government/consultations/emergency-evacuation-information-sharing/equality-impact-assessment Accessed September 23, 2024.
Government of United Kingdom. (2022d). Service Standard. Retrieved from https://www.gov.uk/service-manual/service-standard Accessed September 23, 2024.
Government of United Kingdom. (2022e). Data Sharing Governance Framework. Retrieved from https://www.gov.uk/government/publications/data-sharing-governance-framework/data-sharing-governance-framework Accessed February 26, 2023.
Gräf, M., Mehler, M., & Ellenrieder, S. (2024). AI strategy in action: A case study on make-or-buy for AI-based services. In *Publications of Darmstadt Technical University, Institute for Business Studies (BWL) 146709, Darmstadt Technical University, Department of Business Administration, Economics and Law*. Institute for Business Studies (BWL).
Gu, T., Dolan-Gavitt, B., & Garg, S. (2019). BadNets: Identifying vulnerabilities in machine learning model supply. *arXiv*. https://doi.org/10.48550/arXiv.1708.06733, 1708.06733v2 [cs.CR].
Gutierrez, C. I., & Marchant, G. (2021). *A global perspective of soft law programs for the governance of artificial intelligence*. Sandra Day O'Connor College of Law. Arizona State University.
Haneem, F., Kama, N., Taskin, N., Pauleen, D., & Abu Bakar, N. A. (2019). Determinants of master data management adoption by local government organizations: An empirical study. *International Journal of Information Management, 45*, 25–43. https://doi.org/10.1016/j.ijinfomgt.2018.10.007
Hernández-Nieto, R. (2002). *Contributions to statistical analysis*. Mérida: Universidad de Los Andes.
Herremans, D. (2021). aiSTROM–A roadmap for developing a successful AI strategy. IEEE. *Access, 9*, 155826–155838. https://doi.org/10.1109/ACCESS.2021.3127548
Hickman, E., & Petrin, M. (2021). Trustworthy AI and corporate governance: The EU's ethics guidelines for trustworthy artificial intelligence from a company law perspective. *European Business Organization Law Review, 22*, 593–625. https://doi.org/10.1007/s40804-021-00224-0
Hickok, M. (2022). Public procurement of artificial intelligence systems: New risks and future proofing. *AI & SOCIETY, 1-15*. https://doi.org/10.1007/s00146-022-01572-2
Holmström, J. (2022). From AI to digital transformation: The AI readiness framework. *Business Horizons, 65*(3), 329–339. https://doi.org/10.1016/j.bushor.2021.03.006
Hopster, J. (2021). What are socially disruptive technologies? *Technology in Society, 67*, Article 101750. https://doi.org/10.1016/j.techsoc.2021.101750
IEEE. (2019). Ethically aligned design. In *Committees of the IEEE global initiative on ethics of autonomous and intelligent systems. 2nd version*. Retrieved from https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead_v2.pdf Accessed September 23, 2024.
IEEE. (2020). P7010 - Wellbeing Metrics Standard for Ethical Artificial Intelligence and Autonomous Systems. Retrieved from https://standards.ieee.org/ieee/7010/7718/ Accessed September 23, 2024.
IEEE. (2021a). P7000 - Model Process for Addressing Ethical Concerns During System Design. Retrieved from https://standards.ieee.org/ieee/7000/6781/ Accessed September 23, 2024.
IEEE. (2021b). P7001 - Transparency of Autonomous Systems. Retrieved from https://standards.ieee.org/ieee/7001/6929/ Accessed September 23, 2024.
IEEE. (2021c). P7005 - Standard for Transparent Employer Data Governance Accessed September 23, 2024 https://standards.ieee.org/ieee/7005/7014/ Accessed September 23, 2024.
IEEE. (2021d). P7007 - Ontological Standard for Ethically Driven Robotics and Automation Systems. Retrieved from https://standards.ieee.org/ieee/7007/7070/ Accessed September 23, 2024.
IEEE. (2022). P7002 - Data Privacy Process. Retrieved from https://standards.ieee.org/ieee/7002/6898 Accessed September 23, 2024.
Information Commissioner''s Office. (2020). Big Data, artificial intelligence, machine learning and data protection. Version 2.2. Retrieved from https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf Accessed September 23, 2024.
Information Commissioner''s Office. (2021). International data transfers. Retrieved from https://ico.org.uk/for-organisations/dp-at-the-end-of-the-transition-period/data-protection-and-the-eu-in-detail/the-uk-gdpr/international-data-transfers/ Accessed September 23, 2024.
IPS-X. (2021). IPS-X survey. European cases. Retrieved from https://ipsoeu.github.io/ips-explorer/case/.
ISO. (2021a). ISO/IEC 24027 - Information technology — Artificial intelligence (AI) — Bias in AI systems and AI aided decision making. Retrieved from https://www.iso.org/standard/77607.html Accessed September 23, 2024.
ISO. (2021b). ISO/IEC 24372 - Information technology — Artificial intelligence (AI) — Overview of computational approaches for AI systems. Retrieved from https://www.iso.org/standard/78508.html Accessed September 23, 2024.
ISO. (2021c). ISO/IEC 24668 - Information technology — Artificial intelligence — Process management framework for big data analytics. Retrieved from https://www.iso.org/standard/78368.html Accessed September 23, 2024.

ISO. (2022a). ISO/IEC 38507 - Information Technology – Governance implications of the use of artificial intelligence by organizations. Retrieved from https://www.iso.org/standard/56641.html Accessed September 23, 2024.

ISO. (2022b). ISO/IEC 23894 – Information Technology – Risk management. Retrieved from https://www.iso.org/standard/77304.html Accessed September 23, 2024.

IT Governance Institute. (2003). Board Briefing on IT Governance. 2nd ed. Retrieved from http://www.gti4u.es/curso/material/complementario/itgi_2003.pdf Accessed September 23, 2024.

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy artificial intelligence. *Government Information Quarterly, 37*(3), Article 101493. https://doi.org/10.1016/j.giq.2020.101493

Japanese Strategic Council for AI Technology. (2017). Artificial Intelligence Technology Strategy. Retrieved from https://ai-japan.s3-ap-northeast-1.amazonaws.com/7116/0377/5269/Artificial_Intelligence_Technology_StrategyMarch2017.pdf Accessed September 23, 2024.

Jing, H., Wei, W., Zhou, C., & He, X. (2021). An Artificial Intelligence Security Framework. In *Journal of Physics: Conference Series, Volume 1948, The 2021 2nd International Conference on Internet of Things. Artificial Intelligence and Mechanical Automation (IoTAIMA 2021), 14–16, Hangzhou, China*.

Kale, A., Nguyen, T., Harris, F. C., Li, C., Zhang, J., & Ma, X. (2022). Provenance documentation to enable explainable and trustworthy AI: A literature review. *Data Intelligence, 1-41*. https://doi.org/10.1162/dint_a_00119

Kazim, E., & Koshiyama, A. S. (2021). A high-level overview of AI ethics, patterns, 2(9). *ISSN, 100314*, 2666–3899. https://doi.org/10.1016/j.patter.2021.100314

Khatri, V. (2016). Managerial work in the realm of the digital universe: The role of the data triad. *Business Horizons, 59*(6), 673–688. https://doi.org/10.1016/j.bushor.2016.06.001

Kitsios, F., & Kamariotou, M. (2021). Artificial intelligence and business strategy towards digital transformation: A research agenda. *Sustainability, 13*. https://doi.org/10.3390/su13042025

Korjani, M. M., & Mendel, J. M. (2012). Fuzzy set qualitative comparative analysis (fsQCA): Challenges and applications. In *Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS), 1-6*. https://doi.org/10.1109/NAFIPS.2012.6291026

Krippendorff, K. (2013). *Content analysis – An introduction to its methodology* (3td ed.). Sage.

Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy, 44*(6), Article 101976. https://doi.org/10.1016/j.telpol.2020.101976

Laato, S., Birkstedt, T., Mäantymäki, M., Minkkinen, M., & Mikkonen, T. (2022). AI governance in the system development life cycle: Insights on responsible machine learning engineering. In *In proceedings of the 1st international conference on AI engineering: Software engineering for AI* (pp. 113–123). https://doi.org/10.1145/3522664.3528598

Labadie, C., Legner, C., Eurich, M., & Fadler, M. (2020). FAIR enough? Enhancing the usage of Enterprise data with data catalogs. IEEE. In *22nd conference on business informatics (CBI)* (pp. 201–210). https://doi.org/10.1109/CBI49978.2020.00029

Leavy, S., O'Sullivan, B., & Siapera, E. (2020). Data, power and Bias in artificial intelligence. *ArXiv*. https://doi.org/10.48550/arXiv.2008.07341. abs/2008.07341.

Leijnen, S., Aldewereld, H., van Belkom, R., Bijvank, R., & Ossewaarde, R. (2020). *An agile framework for trustworthy AI. In NeHuAI@ ECAI* (pp. 75–78).

Leslie, D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. *The Alan Turing Institute.*. https://doi.org/10.5281/zenodo.3240529

LIAA-3R. (2022). *Diretrizes de auditabilidade e conformidade no desenvolvimento e testes de soluções de IA no âmbito do LIAA-3R / Grupo de Validação Ético-Jurídica (GVEJ) do LIAA-3R, iLabTRF3, iJuspLab* (2. ed., rev. e atual). São Paulo.

Ligot, D. V.. AI governance: A framework for responsible AI development. Available at SSRN: https://ssrn.com/abstract=4817726.

Lin, Y. T., Hung, T. W., & Huang, L. T. L. (2021). Engineering equity: How AI can help reduce the harm of implicit bias. *Philosophy and Technology, 34*(Suppl. 1), 65–90. https://doi.org/10.1007/s13347-020-00406-7

Locher, M. A., & Bolander, B. (2019). Ethics in pragmatics. *Journal of Pragmatics, 145*, 83–90. ISSN 0378–2166 https://doi.org/10.1016/j.pragma.2019.01.011.

Lu, Q., Zhu, L., Xu, X., Whittle, J., Zowghi, D., & Jacquet, A. (2024). *Responsible AI pattern catalogue: A collection of best practices for AI governance and engineering. Association for Computing Machinery. New York, NY, USA* (Vol. 56, p. 7). ISSN 0360-0300. https://doi.org/10.1145/3626234

Ma, L., Zhang, Z., & Zhang, N. (2018). Ethical dilemma of artificial intelligence and its research progress. *IOP Conference Series: Materials Science and Engineering, 392*, Article 062188. https://doi.org/10.1088/1757-899X/392/6/062188

Makarius, E. E., Mukherjee, D., Fox, J. D., & Fox, A. K. (2020). Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization. *Journal of Business Research, 120*, 262–273. https://doi.org/10.1016/j.jbusres.2020.07.045

Makhlouf, K., Zhioua, S., & Palamidessi, C. (2021). Machine learning fairness notions: Bridging the gap with real-world applications. *Information Processing & Management, 58*(5), Article 102642. https://doi.org/10.1016/j.ipm.2021.102642

Maluf, S. (1995). *Teoria Geral do Estado* (23ª ed., pp. 205–208). Editora Saraiva. São Paulo.

Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. *AI and Ethics, 2*(4), 603–609. https://doi.org/10.1007/s43681-022-00143-x

Manzini, E. J. (2004). *Entrevista semi-estruturada: análise de objetivos e de roteiros. Seminário Internacional sobre Pesquisa e Estudos Qualitativos, 2, 2004, Bauru. A pesquisa qualitativa em debate. Anais*. Bauru: USC. isbn:85-98623-01-6.

Martin, K., & Parmar, B. (). *AI and the Creation of Knowledge Gaps: The ethics of AI transparency*. Available at SSRN: https://ssrn.com/abstract=4207128 https://doi.org/10.2139/ssrn.4207128.

McGraw, G., Bonett, R., Shepardson, V., & Figueroa, H. (2020). The top 10 risks of machine learning security. *Computer, 53*(6), 57–61. https://doi.org/10.1109/MC.2020.2984868

Medaglia, R., Gil-Garcia, J. R., & Pardo, T. A. (2021). Artificial intelligence in government: Taking stock and moving forward. *Social Science Computer Review, 1–18*. https://doi.org/10.1177/08944393211034087

Meijerink, J., & Bondarouk, T. (2018). Uncovering configurations of HRM service provider intellectual capital and worker human capital for creating high HRM service value using fsQCA. *Journal of Business Research, 82*, 31–45. https://doi.org/10.1016/j.jbusres.2017.08.028

Mezgár, I., & Váncza, J. (2022). From ethics to standards – A path via responsible AI to cyber-physical production systems. *Annual Reviews in Control, 53*, 391–404. ISSN 1367-5788. https://doi.org/10.1016/j.arcontrol.2022.04.002

Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. *Big Data & Society, 7*(2). https://doi.org/10.1177/2053951720948087

Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A. (2022). Thinking responsibly about responsible AI and 'the dark side' of AI. *European Journal of Information Systems, 31*(3), 257–268. https://doi.org/10.1080/0960085X.2022.2026621

Mikalef, P., Lemmer, K., Schaefer, C., Ylinen, M., Fjørtoft, S. O., Torvatn, H. Y., … Niehaves, B. (2022). Enabling AI capabilities in government agencies: A study of determinants for European municipalities. *Government Information Quarterly, 39*(4), Article 101596.

Ministério da Ciência Tecnologia e Inovação do Brasil. (2021). Estratégia Brasileira de Inteligência Artificial. Governo do Brasil. Retrieved from https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ia_estrategia_documento_referencia_4-979_2021.pdf Accessed September 23, 2024.

Ministero dello sviluppo economico. (2019). Proposte per uma Strategia italiana per l'intelligenza artificiale. Retrieved from https://www.mise.gov.it/images/stories/documenti/Proposte_per_una_Strategia_italiana_AI.pdf.

Ministry of Economic Affairs and Employment of Finland. (2017). Suomen tekoälyaika. Retrieved from https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf?sequence=1&isAllowed=y Accessed September 23, 2024.

Ministry of Economic Affairs and Employment of Finland. (2019). *Leading the Way into the Era of Artificial Intelligence: Final Report of Finland's Artificial Intelligence Program 2019. Ministry of Economic Affairs and Employment of Finland* (p. 133). Retrieved from http://urn.fi/URN:ISBN:978-952-327-437-2 Accessed September 23, 2024.

Misuraca, G., & van Noordt, C. (2020). *Overview of the use and impact of AI in public services in the EU, EUR 30255 EN. 2020*. Luxembourg: Publications Office of the European Union. ISBN 978–92–76-19540-5, doi:10.2760/039619, JRC120399.

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From what to how: an initial review of publicly available ai ethics tools, methods and research to translate principles into practices. *Science and Engineering Ethics, 26*, 2141–2168. https://doi.org/10.1007/s11948–019-00165-5

MRE. (2022). De outros países no Brasil. Retrieved from https://www.gov.br/mre/pt-br/assuntos/Embaixadas-Consulados-Missoes/de-outros-paises-no-brasil Accessed September 23, 2024.

Nagbøl, P. R., & Müller, O. (2020). X-RAI: A framework for the transparent, responsible, and accurate use of machine learning in the public sector. In *IFIP EGOV-ePart-CeDEM conference. p. 259. CEUR workshop proceedings*.

Nagbøl, P. R., Müller, O., & Krancher, O. (2021). Designing a risk assessment tool for artificial intelligence systems. In the next wave of sociotechnical design. In *, 16. 16th international conference on design science research in information systems and technology, DESRIST 2021, Kristiansand, Norway, august 4–6, 2021, proceedings* (pp. 328–339). Springer International Publishing. https://doi.org/10.1007/978-3-030-82405-1_32.

National Institute of Advanced Industrial Science and Technology. (2022). Machine Learning Quality Management Guideline – 2nd English edition. Government of Japan – Digital Architecture Research Center. Retrieved from https://www.digiarc.aist.go.jp/en/publication/aiqm/aiqm-guideline-en-2.1.1.0057-e26-signed.pdf Accessed September 23, 2024.

Navarro, S., Llinares, C., & Garzon, D. (2016). Exploring the relationship between cocreation and satisfaction using QCA. *Journal of Business Research, 69*(4), 1336–1339.

NIST. (2022). AI Risk Management Framework: first draft. Retrieved from https://www.nist.gov/system/files/documents/2022/03/17/AI-RMF-1stdraft.pdf Accessed September 23, 2024.

Nordic Council of Ministers. (2018). AI in the Nordic-Baltic region. Retrieved from https://www.norden.org/en/declaration/ai-nordic-baltic-region Accessed September 23, 2024.

Norwegian Data Protection Authority. (2018). Datatilsynet. Retrieved from https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf Accessed September 23, 2024.

Norwegian Ministry of Local Government and Modernisation. (2020). National Strategy for Artificial Intelligence. Retrieved from https://www.regjeringen.no/contentassets/1febbbb2c4fd4b7d92c67ddd353b6ae8/en-gb/pdfs/ki-strategi_en.pdf Accessed September 23, 2024.

Ntoutsi, E., Fafaliosb, P., Gadirajua, U., Iosidisa, V., Nejdla, W., Vidalc, M., … Staab, S. (2020). Bias in Dat-driven AI systems – An introductory survey. *arXiv*. https://doi.org/10.48550/arXiv.2001.09762, 2001.09762v1 [cs.CY].

OECD. (2022a). Artificial Intelligence Observatory. Retrieved from https://oecd.ai/en/ Accessed September 23, 2024.

OECD. (2022b). United States – Local state and Federal Regulations on facial recognition technologies. Organization for Economic and co-operation Development Artificial Intelligence Observatory. Retrieved from retrieved from https://oecd.ai/en/dashboards/policy-initiatives/http:%2F%2Faipo.oecd.org%2F2021-data-policyInitiatives-26890.

OECD. (2022c). Policy Initiatives for Emerging AI-related regulation, Civil society. https://oecd.ai/en/dashboards/policy-initiatives?conceptUris=http:%2F%2Fai.oecd.org%2Fmodel%23Emerging_technology_regulation%7C%7Chttp:%2F%2Fai.oecd.org%2Ftaxonomy%2FtargetGroups%23TG16 Accessed September 23, 2024.

OECD/CAF. (2022). *The strategic and responsible use of artificial intelligence in the public sector of Latin America and the Caribbean*. Paris: OECD Public Governance Reviews, OECD Publishing. https://doi.org/10.1787/1f334543-en

Ojo, A., Mellouli, S., & Ahmadi Zeleti, F. (2019). A realist perspective on AI-era public management. In *In 20th annual international conference on digital government research* (pp. 159–170). ACM.

Oneto, L., & Chiappa, S. (2020). *Fairness in machine learning. arXiv:2012.15816 [cs.LG]*.

Özdemir, V., & Hekim, N. (2018). Birth of industry 5.0: Making sense of big data with artificial intelligence,"the internet of things" and next-generation technology policy. *OMICS: A Journal of Integrative Biology, 22*(1), 65–76. https://doi.org/10.1089/omi.2017.0194

Papagiannidis, E., Enholm, I. M., Dremel, C., Mikalef, P., & Krogstie, J. (2023). Toward AI governance: Identifying best practices and potential barriers and outcomes. *Information Systems Frontiers, 25*, 123–141. https://doi.org/10.1007/s10796-022-10251-y

Phillips, P. J., Hanan, C. A., Fontana, P. C., Broniatowski, D. A., & Przybocki, M. A. (2021). *Four principles of explainable artificial intelligence*. National Institute of Standards and Technology. U.S. Department of Commerce.

Pinski, M., Hofmann, T., & Benlian, A. (2024). AI literacy for the top management: An upper echelons perspective on corporate AI orientation and implementation ability. *Electronic Markets, 34*, 24. https://doi.org/10.1007/s12525-024-00707-1

Presidencia de la Nación. (2019). ARGENIA – Plan Nacional de Inteligencia Artificial. Argentina. Retrieved from https://oecd-opsi.org/wp-content/uploads/2021/02/Argentina-National-AI-Strategy.pdf Accessed September 23, 2024.

Ragin, C. C. (2008). *Redesigning Social Inquiry: Fuzzy Sets and Beyond* (pp. 85–97). Chicago: Univ. of Chicago Press.

Rahwan, I. (2017). Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology, Springer., 20*, 5–14. https://doi.org/10.1007/s10676-017-9430-8

Rahwan, I., Cebrian, M., Obradovich, N., et al. (2019). Machine behaviour. *Nature, 568*, 477–486. https://doi.org/10.1038/s41586-019-1138-y

Rajkomar, A., Hardt, M., Howell, M. D., Corrado, G., & Chin, M. H. (2018). Ensuring fairness in machine learning to advance health equity. *Annals of Internal Medicine, 169*(12), 866–872. https://doi.org/10.7326/M18-1990

Rhahla, M., Allegue, S., & Abdellatif, T. (2021). Guidelines for GDPR compliance in big data systems. *Journal of Information Security and Applications, 61*, Article 102896. https://doi.org/10.1016/j.jisa.2021.102896

Rihoux, B., & Ragin, C. (2008). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques*. London and Thousand Oaks, CA: Sage.

Roorda, S. A. H. (2021). *Facial recognition for public safety a supportive tool for the municipal decision-making process on using facial recognition for public safety, the FRPS risk governance method*. Master's thesis,. University of Twente.

Rose, J., Flack, L. S., & Sæbø, Ø. (2018). Stakeholder theory for the E-government context: Framing a value-oriented normative core. *Government Information Quarterly, 35*(3), 362–374. ISSN 0740-624X https://doi.org/10.1016/j.giq.2018.06.005.

Roselli, D., Matthews, J., & Talagala, N. (2019). *Managing Bias in AI. In Companion Proceedings of The 2019 World wide web conference (WWW '19)* (pp. 539–544). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3308560.3317590

Ruijer, E. (2021). Designing and implementing data collaboratives: A governance perspective. *Government Information Quarterly, 38*(4), Article 101612. https://doi.org/10.1016/j.giq.2021.101612

Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). Sage.

Schaefer, C., Lemmer, K., Samy Kret, K., Ylinen, M., Mikalef, P., & Niehaves, B. (2021). Truth or dare?–how can we influence the adoption of artificial intelligence in municipalities?. Retrieved from http://hdl.handle.net/10125/70899.

Schrader, D., & Ghosh, D. (2018). Proactively protecting against the singularity: Ethical decision making AI. *IEEE Computer and Reliability Societies Review, 16*(3), 56–63.

Schüller, K. (2022). Data and AI literacy for everyone. *Jan, 1*, 477–490. https://doi.org/10.3233/sji-220941

Shao, S., Zhao, R., Yuan, S., Dig, M., & Wang, Y. (2022). Tracing the evolution of AI in the past decade and forecasting the emerging trends. *Expert Systems with Applications., 209*, Article 118221. https://doi.org/10.1016/j.eswa.2022.118221

Sharma, G. D., Yadav, A., & Chopra, R. (2020). Artificial intelligence and effective governance: A review, critique and research agenda. *Sustainable Futures, 2*. https://doi.org/10.1016/j.sftr.2019.100004. ISSN 2666–1888.

Sigfrids, A., Leikas, J., Salo-Pöntinen, H., & Koskimies, E. (2023). Human-centricity in AI governance: A systemic approach. *Frontiers in Artificial Intelligence, 6*, Article 976887. https://doi.org/10.3389/frai.2023.976887. PMID: 36872934; PMCID: PMC9979257

Silberg, J., & Manyika, J. (2019). *Notes from the AI frontier: Tackling bias in AI (and in humans). 1(6)*. McKinsey Global Institute. Retrieved from https://www.mckinsey.com/~/media/mckinsey/featured%20insights/artificial%20intelligence/tackling%20bias%20in%20artificial%20intelligence%20and%20in%20humans/mgi-tackling-bias-in-ai-june-2019.pdf Accessed September 23, 2024.

Silveira, M. B., Saldanha, R. P., Leite, J. C. C., Silva, T. O. F. D., Silva, T., & Filippin, L. I. (2018). Construction and validation of content of one instrument to assess falls in the elderly. *einstein (Sao Paulo), 11;16*(2), Article eAO4154. Retrieved from https://doi.org/10.1590/S1679-45082018AO4154 accessed September 23, 2024.

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of big data challenges and analytical methods. *Journal of Business Research, 70*, 263–286.

Smuha, N. A. (2021). Beyond a Human Rights-Based Approach to AI Governance: Promise, Pitfalls, Plea. *Philos. Technol., 34*(Suppl 1), 91–104. https://doi.org/10.1007/s13347-020-00403-w

Stahl, B. C., Antoniou, J., Ryan, M., et al. (2022). Organisational responses to the ethical issues of artificial intelligence. *AI & SOCIETY, 37*, 23–37. https://doi.org/10.1007/s00146-021-01148-6

Stix, C. (2021). Foundations for the future: Institution building for the purpose of artificial intelligence governance. *AI and Ethics*. https://doi.org/10.1007/s43681-021-00093-w

Strauß, S. (2021). Deep automation Bias: How to tackle a wicked problem of AI? *Big Data Cogn. Comput., 5*, 18. https://doi.org/10.3390/bdcc5020018

Switzerland Federal Council. (2021). Guidelines on Artificial Intelligence for the Confederation. Retrieved from https://www.sbfi.admin.ch/dam/sbfi/en/dokumente/2021/05/leitlinien-ki.pdf.download.pdf/leitlinien-ki_e.pdf Accessed September 23, 2024.

Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and Society, 40*(2), 137–157. https://doi.org/10.1080/14494035.2021.1928377

Tangi, L., van Noordt, C., Combetto, M., Gattwinkel, D., & Pignatelli, F. (2022). *AI watch. European landscape on the use of artificial intelligence by the public sector, EUR 31088 EN*. Luxembourg: Publications Office of the European Union. ISBN 978-92-76-53058-9, doi:10.2760/39336, JRC129301.

Torlig, E. G. S., Resende-Júnior, P. C., & Fujihara, R. K. (2019). Proposição de uma Nova Orientação para Validação de Roteiros em Pesquisas Qualitativas. In *XLIII Encontro da ANPAD – EnANPAD 2019, São Paulo*.

Torlig, E. G. S., Resende-Júnior, P. C., Fujihara, R. K., Demo, G., & Montezano, L. (2022). Validation Proposal for Qualitative Research Scripts (Vali-Quali). *Administração: Ensino e Pesquisa, 23*(1). https://doi.org/10.13058/raep.2022.v23n1.2022, 4-29; Jan-Abr.

Vero. (2019). Finnish Tax Administration's ethical principles for AI. Retrieved from https://www.vero.fi/en/About-us/finnish-tax-administration/operations/responsibility/finnish-tax-administrations-ethical-principles-for-ai/#:~:text=Our%20AI%20follows%20laws%20and%20regulations&text=The%20use%20of%20AI%20does,our%20partners%20carefully%20and%20responsibly.

Vetrò, A., Torchiano, M., & Mecati, M. (2021). A data quality approach to the identification of discrimination risk in automated decision-making systems. *Government Information Quarterly, 38*(4). https://doi.org/10.1016/j.giq.2021.101619. ISSN 0740-624X.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems, 28*(2), 118–144. ISSN 0963-8687. https://doi.org/10.1016/j.jsis.2019.01.003

Vilminko-Heikkinen, R., & Pekkola, S. (2019). Changes in roles, responsibilities and ownership in organizing master data management. *International Journal of Information Management, 47*, 76–87. ISSN 0268-4012, https://doi.org/10.1016/j.ijinfomgt.2018.12.017.

Vining, R., McDonald, N., McKenna, L., Ward, M. E., Doyle, B., Liang, J., … Fogarty & Brennan, R. (2022). Developing a Framework for Trustworthy AI-Supported Knowledge Management in the Governance of Risk and Change. In *HCI International 2022-Late Breaking Papers. Design, User Experience and Interaction: 24th International Conference on Human-Computer Interaction, HCII 2022, Virtual Event, June 26–July 1, 2022, Proceedings* (pp. 318–333). Cham: Springer International Publishing.

WEF. (2020). Unblocking Public Sector AI – AI procurement in a Box: Project overview. Retrieved from https://mkto.deloitte.com/rs/712-CNF-326/images/Retningslinjer-for-offentlige-AI-anskaffelser.pdf Accessed September 23, 2024.

WEF. (2020b). Unblocking Public Sector AI. AI Procurement in a Box: Pilot case studies from the United Kingdom. Retrieved from https://www.weforum.org/reports/ai-procurement-in-a-box/#case-study-uk Accessed September 23, 2024.

Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2018). Artificial intelligence and the public sector—Applications and challenges. *International Journal of Public Administration, 42*(7), 596–615.

Wirtz, B. W., Weyerer, J. C., & Kehl, I. (2022). Governance of artificial intelligence: A risk and guideline-based integrative framework. *Government Information Quarterly*. https://doi.org/10.1016/j.giq.2022.101685. ISSN 0740-624X.

Wright, S. A., & Schultz, A. (2018). The rising tide of artificial intelligence and business automation: Developing an ethical framework. *Business Horizons, 61*(6), 823–832. https://doi.org/10.1016/j.bushor.2018.07.001

Xue, M., Yuan, C., Wu, H., Zhang, Y., & Liu, W. (2020). Machine learning security: Threats, countermeasures, and evaluations. *IEEE Access, 8*(74720–74742), 2020. https://doi.org/10.1109/ACCESS.2020.2987435

Zhou, J., & Chen, F. (2023). AI ethics: From principles to practice. *AI & SOCIETY, 38*, 2693–2703. https://doi.org/10.1007/s00146-022-01602-z

Zicari, R. V., Brodersen, J., Brusseau, J., Düdder, B., Eichhorn, T., Ivanov, T., … Westerlund, M. (2021). Z-inspection®: A process to assess trustworthy AI. *IEEE Transactions on Technology and Society, 2*(2), 83–97. https://doi.org/10.1109/TTS.2021.3066209

Zick, T., Kortz, M., Eaves, D., & Doschi-Velez, F. (2024). *AI Procurement Checklists: Revisiting Implementation in the Age of AI Governance*. arXiv:2404.14660v1 [cs.CY]. https://doi.org/10.48550/arXiv.2404.14660.

Zuiderwijk, A., Chen, Y., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly, 38*(3). https://doi.org/10.1016/j.giq.2021.101577. ISSN 0740-624X.

Patricia Gomes Rêgo de Almeida has Master degree at University of Rio Grande do Norte, Department of Electrical Engineering, and PhD degree at the University of Brasilia (UnB), Department of Business Administration, where she has developed research regarding artificial intelligence regulation (hard and soft law) and artificial intelligence governance. She is member of Linselab at University of Brasilia (UnB). She is the coordinator of Digital Innovation, Governance and Strategy at the Brazilian chamber of Deputies, where she is responsible for its Digital Transformation Strategy, Artificial Intelligence Governance and Data Fluency Program. She is coordinator of the Parliamentary Data Science Hub at the Inter-parliamentary Union, where she coordinated the writing of Guidelines for AI in Parliaments.

Carlos Denner dos Santos Júnior Junior is professor at University of Brasilia (UnB), Department of Business Administration, focusing on the intersection between administration and computing, developing cutting-edge scientific knowledge about the creation and management of new businesses and information technologies with applications in public administration and their impacts on the market in general. He has post-doctoral degrees at: Université du Quebec à Montréal (UQUAM), Universidade Federal de Pernambuco (UFPE), University of Nottingham, and Universidade de São Paulo (USP). He is the coordinator of the Research Group at CNPq Sociedados - on the Strategic and Competitive Use of Data (Open) and Software (Free), and the DINTER coordinator between UnB-UNIMONTES.