Embedding ethical AI risks in public sector risk management practice An explorative analysis using the Three Lines of Defense risk management model

Sattlegger, Antonia; Bharosa, Nitesh

DOI
10.1145/3657054.3657063

Publication date
2024

Document Version
Final published version

Published in
Proceedings of the 25th Annual International Conference on Digital Government Research, DGO 2024

Embedding ethical AI risks in public sector risk management practice

An explorative analysis using the Three Lines of Defense risk management model

Antonia Sattlegger
Faculty of Technology, Policy, and Management, Delft University of Technology
a.s.sattlegger@tudelft.nl

Nitesh Bharosa
Faculty of Technology, Policy, and Management, Delft University of Technology
n.bharosa@tudelft.nl

ABSTRACT

Artificial intelligence (AI) adoption by public sector organizations (PSOs) introduces various ethical risks stemming from a lack of integrating human values into AI design. Addressing these ethical risks is a complex collective responsibility among designers, developers, risk experts, and public sector managers. Embedding these risks in existing risk management practices is crucial for responsible AI adoption, as emphasized by the legal requirements of the EU AI Act. However, the responsibility for managing these ethical risks is often unclear. Public sector organizations face unique challenges due to the complex, uncertain, and rapidly evolving nature of AI technologies, further complicating the management of ethical risks. This paper explores using the Three Lines of Defense (TLoD) risk management model to understand and address these ethical risks in public sector AI adoption. The TLoD model structures risk management across three lines: operational management, risk oversight and compliance, and internal audit. This framework helps to distribute and integrate the collective responsibility for ethical AI risk management within public sector organizations, emphasizing alignment and collaboration among different actors. Through an exploratory study involving a survey and semi-structured interviews with professionals responsible for AI-related risk management in Dutch public sector organizations, we assess the TLoD model's usefulness in addressing ethical AI risks. The study examines the challenges and opportunities in applying the TLoD model to manage ethical risks and identifies the potential gaps in responsibility and oversight. The findings suggest that while the TLoD model offers a valuable lens for distributing risk management responsibilities, there are limitations in addressing the emergent and complex nature of ethical risks in AI adoption.

CCS CONCEPTS

• **Social and professional topics** → Professional topics; Management of computing and information systems; • **General and reference** → Document types; General conference proceedings.

KEYWORDS

AI Governance, risk management, TLoD, responsibility, data ethics

dg.o 2024, June 11–14, 2024, Taipei, Taiwan

ACM ISBN 979-8-4007-0988-3/24/06
https://doi.org/10.1145/3657054.3657063

ACM Reference Format:
Antonia Sattlegger and Nitesh Bharosa. 2024. Embedding ethical AI risks in public sector risk management practice: An explorative analysis using the Three Lines of Defense risk management model. In *25th Annual International Conference on Digital Government Research (dg.o 2024), June 11–14, 2024, Taipei, Taiwan.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3657054.3657063

1 INTRODUCTION

Emerging technologies, such as data-driven innovations, process automation and algorithmic systems, or, in short, artificial intelligence (AI), provide the public sector with new capabilities to enhance public value by increasing efficiency, service quality and boosting government responsiveness. However, these new capabilities give rise to ethical dilemmas and risks. The overuse or misuse of AI can devalue human skills, remove human responsibility, reduce human control, or erode human self-determination [15:691]. Public organizations may use predictive analytics which can enable discriminatory decision-making, such as predicting fraud in welfare service distribution. Human discretion and means of redress may be limited. The inexplicability of AI black boxes may exacerbate administrative burdens and exclusions through information architecture [30], inflicting harm onto individuals and marginalized groups and lead to distrust in government. The materialization of these ethical AI risks is particularly evident in the context of social welfare service. For instance, in the Netherlands, the tax administration employed a self-learning algorithm for creating risk profiles in childcare benefit applications, utilizing factors such as 'foreign-sounding names' and 'dual nationality' as indicators, leading to racial profiling, false fraud accusations, and demands for repayment [4, 12, 31]. In Australia, the flawed data-matching system Robodebt was used for automated debt assessment and to issue debt notices to welfare recipients [23, 35]. Both cases of irresponsible design and use of AI led to hardship among affected vulnerable citizens and caused distrust in government decision-making in broader society.

Such ethical risks arising with the adoption of AI by public sector organizations can be defined as "the lack of or incorrect integration of human values to AI decision-making and action, and the negative consequences thereof" [52:7]. This definition implies a perspective on ethical risk management as dynamic collective process, which requires the constant reflection on human values in a situated design process. The integration of human values in the design of AI is a complex collective responsibility of organizations adopting AI in their decision-making and service provision. This collective responsibility for identifying and mitigating ethical risks

is distributed through risk management – the identification, assessment, allocation, and mitigation of (ethical) risks. The relevance of establishing effective risk management systems is further emphasized by the legal requirements of the EU AI act. To ensure the responsible adoption of AI in public sector organizations, ethical risks need to be integrated in existing risks management practices. Designers and developers of AI, as well as dedicated risk experts such as risk managers, chief risk officers or (internal) auditors share an active moral responsibility to promote and achieve societally shared goals and values [32, 37:5] through the design of AI. However, the distribution of this responsibility among those designing AI is often not clear. Even if this responsibility is seemingly clearly attributed, those responsible may lack the necessary conditions to discharge their responsibility [38]. They may lack the awareness, knowledge or means to control the design conducive to identifying and successfully mitigating ethical risks [37]. These responsibility gaps are conducive to the realization of ethical risks.

The adoption of AI by public sector organizations may be especially prone to these responsibility gaps. The management of ethical risks arising with the adoption of AI in public sector organizations may pose challenges to traditional definition of risk and the management thereof by public sector organizations. The development and adoption of AI by public sector organizations is novel and innovative, or merely experimental in many instances [18]. It is characterized by a high degree of complexity, opacity, and interdependence [17]. Public sector organizations may struggle to manage these rapidly evolving and disruptive technologies [41, 42]. The transformative affordances of these technologies give rise to ambiguity and uncertainty about the norms governing the use of emerging technologies, as well as limited or inadequate regulation, rules, or precedent to define, identify and manage its ethical implications [6, 20, 21, 25]. The complexity, uncertainty and ambiguity which characterizes the adoption of AI in the public sector poses a challenge to the definition and respective management of ethical risks. First, classical notions of risk refer to the likelihood or probability of an event coupled with an assessment of the potential consequences or impact associated with that event. In short, "risk is the combination of probability of an event and its consequences" (ISO, 2002). Likewise, the EU AI act emphasizes the management of "known and reasonably foreseeable" risks (Article 9: risk management systems). However, this may not be the case for the ethical implications of emerging technologies as "unknown unknowns" [9:49]. Ethical risks are defined as "the lack of or incorrect integration of human values to AI decision-making and action, and the negative consequences thereof" [52:7]. The consequences of AI adoption on individuals and society are difficult to predict. Particularly, in the early stages of AI development, the full extent of potential risks and consequences may not be fully understood or appreciated. Ethical risks arising with the adoption of AI are emergent. Second, the management of emerging ethical risks arising with the adoption of AI may be challenging. Traditionally, PSOs are rather risk averse. Political and public accountability demands discourage risk taking and the accommodation of uncertainty by public managers [3]. Risk management practices, such as managing political and strategic risks are often primarily a responsibility of high-level management, while operational actors are tasked with documenting compliance and maintaining audit trails [26]. Dedicated risk experts such as internal auditors, chief risk officers and/or risk managers approach risk in a technocratic or bureaucratic manner [5, 29], typically in a "diagnostic" fashion [36], through use of largely quantifiable risk management tools, such as heat maps [7:8]. This segregation of formalized risk management contributes to a focus on compliance rather than strategy to mitigate emergent ethical risks [34]. The collective responsibility for the management of ethical risks is distributed through top-down delegation and bottom-up compliance and reporting.

A risk management model that provides a structured approach to distributing collective responsibilities may be found in the risk management model Three Lines of Defense (TLoD). The TLoD model offers a structured organizational paradigm that delineates risk management responsibilities across three lines – operations, risk oversight and compliance, and internal audit [22]. The TLoD model may be a useful lens to manage ethical AI risks for several reasons: First, the TLoD model is an actor-centric approach that emphasizes the distribution of actor's responsibilities throughout the organization. The model emphasizes the alignment, coordination, and collaboration between the lines, rather than merely establishing hierchical reporting and oversight responsibilities. This may enable the coordination and integration of the responsibility distribution. Second, we are interested in the embedding of ethical AI risks in existing risk management practices in public sector organizations. The TLoD model is a widely adopted model, particularly in the IT audit and security domain. Rather than a theoretical framework, we apply the TLoD as empirical lens to contribute to the understanding of addressing ethical AI risks in practice. As to whether it can address ethical AI risks in public sector organizations has not been researched. There has been limited empirical research on the effectiveness of the model in the corporate context [2, 10, 39, 40, 44]. No research has been conducted to address the model's usefulness in embedding ethical AI risks in PSOs. One notable exception in the context of AI can be found in Schuett's (2023) application of the model to private tech companies developing AI. He finds that the TLoD model for private tech companies developing AI "can plausibly contribute to a reduction of risk from AI", by addressing the diffusion of responsibility between researchers, engineers, legal and compliance departments, or leadership (p. 15). However, the current literature neither analyses the model in a public sector context, nor its usefulness to address ethical AI risks. We address this gap through the primary objective of this research is to understand the usefulness of the TLoD in embedding ethical AI risks in existing risk management practices. Therefore, this research explores the following research question: *to what extent is the TLoD a useful lens to manage ethical risks in the adoption of AI by public sector organizations?*

This paper proceeds as follows: First, we outline the explorative research design of this paper. Second, we define ethical risk and conceptualize risk management in the context of designing AI by PSOs. We further introduce the TLoD risk management model used in this research. Third, we will provide our empirical results. Lastly, we discuss the strengths and weaknesses of the TLoD model in the context of managing ethical AI risks.

2 CONCEPTUAL LENS – ETHICAL RISKS ARISING WITH THE ADOPTION OF AI

In this section we elaborate on the definition of ethical AI risks and the TLoD risk management as practical lens to address these risks.

2.1 Ethical risks arising with the adoption of AI in the public sector

Ethical risks arising with the adoption of AI by public sector organizations can be defined as "the lack of or incorrect integration of human values to AI decision-making and action, and the negative consequences thereof" [52:7]. This definition implies that ethical AI risks are not a finite list of possible adversaries. Rather it is a dynamic collective process, which requires the constant reflection on human values in a situated design process. This definition to ethical risks captures multiple perspectives on ethical risks arising with the adoption of AI in the public sector. First, ethical risks arise with the processing of data, such as unfair statistical discrimination [51]. Relating to this perspective are also epistemic concerns relating to inconclusive outputs which have ethical implications [28]. Non-deterministic machine learning algorithms generate probabilistic outputs, typically identifying associations and correlations among variables in the data without establishing causal connections [47]. This, Tsamados et al. [47:217] argue, encourages "the practice of apophenia: "seeing patterns where none actually exist, simply because massive quantities of data can offer connections that radiate in all directions". Second, Floridi [14:188] argues that the risk of lack of explicability is unique to AI systems. Ethical issues arise when AI systems are designed as black boxes. Their inner workings are complex, and the outputs are not easily interpretable or explainable. Users, including developers, regulators, and end-users, may find it challenging or impossible to understand how the system arrives at its decisions or predictions. A lack of transparency and epistemic understanding of opaque or invisible AI systems hinders individuals' ability to redress the outcomes [30, 31]. This opacity may undermine trust in and accountability of government decision-making, which we will discuss later. Third, ethicists of technology emphasize the responsibility gaps which can be created when human agency and responsibility are crowded out by artificial agents. Ethicists of technology raise a shift in moral responsibility with the introduction of AI in decision-making processes [1, 8, 19]. This leads to a reduction of human autonomy resulting in over-reliance on AI recommendations or decisions, potentially diminishing individual agency and the ability to make informed choices [24, 37]. AI systems may escape human oversight and control due to their complex, obliquitous, and autonomous nature [37]. In the absence of meaningful human control over such systems, responsibility is diffused in multiple ways. Santoni de Sio et al. [37] conceptualize various responsibility gaps: The active responsibility of those designing and using AI systems to promote their moral obligations may be undermined due to a lack of awareness of their responsibilities. The difficulty in attributing responsibility limits both moral accountabilities to explain one's reasoning, as well as political accountability to explain one's action to a broader public forum. Fourth, a socio-technical perspective conceptualizes algorithms as tools to simplify social complexities, encode and amplify systematic inequalities. Scholars in Science and Technology Studies (STS) and human-computer interaction (HCI) emphasizes the harms that can be inflicted upon individuals, and groups. Digital technologies are inseparable from the social dynamics in which such systems are developed and experienced [43]. In the public sector, these violations are sometimes conceptualized as human-rights violations [27]. Algorithms are essentially means to reduce the complexity of the social world. As van Es at al. (2022) note in [43], "algorithms and code reduce the complexity of the social world into a set of abstract instructions on how to deal with data and inputs coming from a messier reality" (p. 3). This "selection, reduction, and categorization" of social realities encodes, reinforces, and amplifies systematic inequalities in AI systems [43]. Shelby et al. [42] provide an comprehensive taxonomy of five major types of computational and contextual harms at the micro-, meso-, and macro-level of algorithmic systems, namely: "(1) how socially constructed beliefs and unjust hierarchies about social groups are reflected in model inputs and outputs (representational harms); (2) how these representations shape model decisions and their distribution of resources (allocative harms); (3) how choices made to optimize models for particular imagined users result in performance disparities (quality-of-service harms); (4) how technological affordances adversely shape relationships between people and communities (interpersonal harms); and (5) how algorithmic systems adversely impact the emergent properties of social systems, leading to increased inequity and destabilization (social system/societal harms)." (p. 6-7). Lastly, the widespread use of AI in society and particularly by government organizations can have transformative effects [48]. Yeung [52] eloquently argues that "the take-up of digital automation, algorithmic decision-making and data-driven technologies in public administration and public service delivery" (p. 3) or, what she refers to as new public administration paradigm of New Public Analytics, has "significant and troubling implications for practice of statecraft and the delivery of public services, for the relationship between states and individuals, including the nature of citizenship, and for the relationship between public and private power" (p. 3). Particularly, the lack of trust in AI based decision-making and public services may contribute towards an erosion of public trust into government institutions in general [45], especially, if such technologies are use in inappropriate and sensitive domains [33].

2.2 TLoD risk management model

The Three Lines of Defense (TLoD) risk management model, as proposed by the Institute of Internal Auditors [22], provides a structured approach to managing risks within an organization. The model emphasizes the distribution and integration of responsibilities across different actors, ensuring clear roles and relationships in overseeing risk management.

In this model, four key actors play distinct roles in risk management:

- The governing body is responsible for organizational oversight and is responsible for setting structures, processes, and objectives for effective governance. It demonstrates integrity, leadership, and transparency, and delegates responsibilities to management. Additionally, it establishes an independent internal audit function to provide assurance.

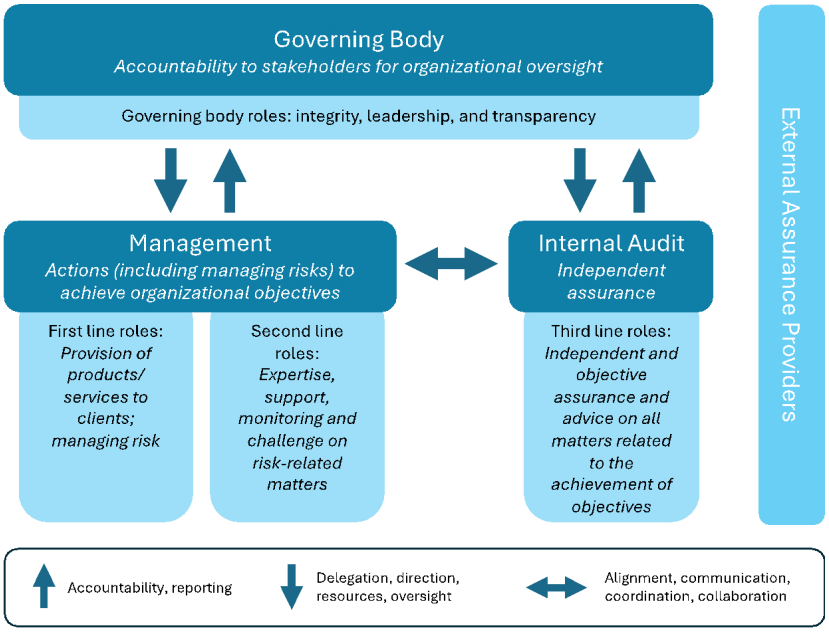

Figure 1: TLoD according to IAA (2022).

- The first line of defense is represented by the management, handling day-to-day operations and owning the associated risks. They identify, assess, and mitigate risks in their daily activities.
- The second line of defense is responsible for risk oversight and compliance. This line develops policies, ensures regulatory compliance, and guides the first line in managing risks.
- The third line of defense is the internal audit function, providing independent assurance to the governing body by assessing the effectiveness of risk management processes. Internal audit operates independently from management and focuses on auditing and evaluating the risk management systems.
- External assurance providers, such as consulting firms or certification bodies, offer additional independent oversight at the request of management or the governing body. They also provide assurance to external stakeholders, like government auditors or regulatory bodies.

The TLoD model delineates roles and responsibilities between the different actors. The integration of these responsibilities is fostered through communication, cooperation and collaboration between the governing body and the lines of defense is vital to ensure the alignment of activities and thereby the creation and protection of value (Principle 6 in IAA, 2020). The governing body defines the organization's core values and risk appetite while overseeing the implementation of risk management processes. Management, accountable to the governing body, must proactively communicate about risks and compliance. Internal audit ensures independent oversight and maintains a collaborative relationship with management to verify effective risk management. This alignment across the lines of defense is crucial for creating and protecting organizational value.

3 METHODOLOGY

This explorative study examines the application of the Three Lines of Defense (TLoD) risk management model as a framework for managing ethical risks associated with the use of AI in public sector organizations (PSOs). The study aims to assess the TLoD model is a useful lens for managing ethical risks in AI adoption by public sector organizations. We are particularly interested whether the model can provide a structured approach to effectively distributing responsibilities for the identification and mitigation of ethical AI risks. To address the research question, we employ an exploratory research approach [46]. We conducted a survey targeting key professionals responsible for managing ethical risks in AI development within Dutch public sector organizations. To gain further in-depth understanding of the management of ethical AI risks by PSOs in general and the usefulness of the TLoD in particular, we conducted additional background interviews.

3.1 Survey Design

The survey employed a cross-sectional design and was administered online using Qualtrics. It consisted of three sections: (1) expected (ethical) AI risks, (2) the maturity of AI (ethics) risk management in the respective organizations, and (3) the actual or ideal implementation of the TLoD model and reflection thereon. The survey includes both closed as well as open questions.

3.1.1 Sampling and Participants. Our target population included individuals responsible for managing (ethical) risks associated with

AI in public sector organizations, such as ethics advisors, risk managers, auditors, Chief Data Officers (CDO), Chief Information Officers (CIO), or professionals working with AI and algorithms. The unit of analysis was the public sector organizations these professionals represented. We targeted municipalities (9.7%), executive agencies (41.7%), and ministries (31.9%). In a few cases participants, such as strategic advisors at ministries or assurance organizations as well as consultants, work and reflect on multiple organizations (16.7%).

We utilized a combination of non-probability sampling techniques to reach participants:

- Purposive Sampling: We specifically targeted professionals responsible for AI-related risk management in the public sector.
- Convenience Sampling: The survey was shared through professional networks to reach a broader audience.
- Snowball Sampling: Through a snowballing process we reached out to further professionals beyond our own network.

Informed consent was obtained from participants. The survey was entirely anonymous, with no personal information collected.

3.1.2 Data collection and analysis. Data collection took place using the Qualtrics online survey platform. The survey was conducted over a two-month period, from February 26 to May 2, 2024. A total of 72 respondents started the survey, with 54 completing at least 70% of the survey. For data analysis, we used descriptive statistics to summarize survey responses and qualitative coding for open-text fields. Given the exploratory nature of the research, this approach provided a preliminary understanding of the data.

3.2 Semi-structured interview and document analysis

To gain further in-depth understanding of the management of ethical AI risks and the useability of the TLoD model, we conducted explorative interviews with practitioners and experts on the management of ethical risks. We conducted 11 interviews, with the interviews lasting between 40 and 60 minutes. The interviews were recorded for transcription. Additionally, we analyzed secondary data, including organizational reports, internal policy documents, and other relevant sources to contextualize the survey and interview findings.

3.3 Limitations

This study has several methodological limitations related to its exploratory nature. The sample size is too small for advanced statistical analysis or generalization. Additionally, the reliance on non-probability sampling techniques could introduce bias, further limiting the generalizability of the findings. As outlined, we are interested in the extent to which the TLoD can be a useful lens for both practice and academia. However, given the explorative approach and empirical data, we cannot generalize our findings. They provide a foundation for further (confirmative) research.

4 RESULTS – ETHICAL AI RISK MANAGEMENT PRACTICES BY PSOS

In our survey we find that most participants (94%) agree that embedding ethical AI risks in risk management practices is relevant (see Table 1). Despite this general agreement, some participants mention that ethical risks cannot be differentiated from other types of risks, rather ethical dimensions are underlying all other types of risks. Another respondent cautions to differentiate ethical risks from other types of risk: "We need to move past the ethical framing; every director you meet considers it important in peacetime, but when interests come into play, it's the first thing to fall apart.".

Generally ethical risks are not, yet, part of or explicitly addressed in risk management practices (see Table 2).

The TLoD model is broadly implemented by the surveyed organizations. We find that 47.06% of the respondent's state that their respective organizations have already implemented the model or are planning to do so (19.6%, see Table 3). The model is particularly adopted for the privacy and data security domain, it does not find broad application in the domain of (ethical) AI. However, when asked to plot the actors according to the current implementation of the model, we find much diversity and discrepancy from the TLoD model as presented by the IAA [22]. Moreover, multiple respondents add that while the model is "exists on paper, but implementation in practice is lacking". When asked to plot the ideal implementation, there is much diversity to, both in relation to the current implementation and to other participants ideal implementations. There is no generally accepted best practice.

4.1 First Line of Defense

The first line of defense takes primary responsibility for identifying, assessing, and mitigating ethical AI risks in their daily operations. This may be either explicitly or implicitly, as one respondent remarks: "Because there is no formal work process set up for ethical risk management, the responsibility for it falls on the individual developer of the information product." The distribution of this responsibility between the actors of the first line is not clearly distributed (see Table 4). Actors of the first line, see Table 5, are the owner and developer of the algorithm, responsible for the functional and technical aspects of the algorithm, the user of the algorithmic application, responsible for the use process of the algorithmic application, and the owner or supplier of the processed data. Generally, we find that, no new roles or relationships are established to identify, assess, or mitigate ethical risks. However, new tasks, such as risk or impact assessments are being developed, for which responsibilities are likewise unclear (see Table 4).

Risk identification – With most PSOs, the initial risk assessment or intake, including differentiation between high- and low(er) risk algorithmic applications, is often not formally established. Whether or not specific algorithmic applications pose a heightened level of ethical risk is most often a decision by the owner or user of the algorithmic application. While some PSOs are developing assessment tools in the form of four to seven screening questions, they are not always formally established.

Algorithmic impact assessment – Most PSOs state that the first formal identification and assessment of ethical risks is usually conducted through an algorithmic impact assessment. While

Table 1: Desirability of embedding ethical risks in risk management.

Desirability – In your opinion, is it desirable to embed ethical risks into risk management?		
Yes (94.0%)	No (2.0%)	I do not know (4.0%)

Table 2: Ethical AI risk management and mitigation.

Mean	Median	Standard Deviation	*Strongly Disagree (-2)*	*Disagree (-1)*	*Neutral (0)*	*Agree (1)*	*Strongly Agree (2)*
Ethical AI risk management – Ethical AI risks are part of our risk management.							
-0.2	0.0	1.1	10.0%	34.0%	36.0%	8.0%	12.0%
Ethical AI risk mitigation – Ethical AI risks are adequately addressed in our risk management.							
-0.5	-1.0	1.1	12.0%	48.0%	22.0%	12.0%	6.0%

Table 3: Application of the TLoD risk management model.

TLoD – Is your organization using the TLoD risk management model?	
Yes	47.1%
No, but I expect we'll use it in the future.	19.6%
No, but we've used it in the past.	2.0%
No, we don't use it, but I'm aware of it.	9.8%
No, I've never heard of it before.	21.6%

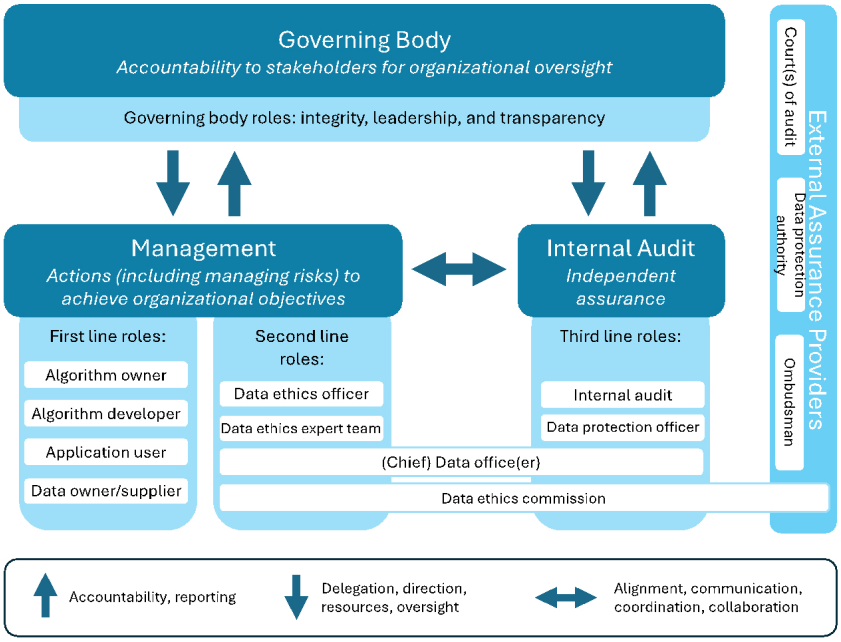

Figure 2: Adapted version of the TLoD according to IAA (2022). Respective public sector actors added by authors.

some organizations develop their own assessments, most are either already using or planning to use the Impact Assessment Fundamental Rights and Algorithms endorsed by the Dutch government (Impact Assessment voor Mensenrechten bij de inzet van Algoritmes, IAMA). This collective and deliberative approach enables the early identification and assessment of ethical risks. Ethical risks are conceptualized regarding how the algorithmic application contributes to public value creation and assessing whether and how human rights may be negatively impacted. In practice, responsibilities for conducting such algorithmic impact assessments are

Table 4: Assessment of responsibility attribution for the management of ethical AI risks.

	Mean	Median	Standard Deviation	Strongly disagree (-2)	Disagree(-1)	Neutral (0)	Agree (1)	Strongly agree (2)
1st Line	**Risk Identification** – The responsibilities for identifying, assessing, inventorying, and mitigating ethical risks are clearly assigned to the owner, user, and developer of algorithms in their daily operations.							
	-0.7	-1.0	1.1	26.1%	37.0%	19.6%	13.0%	4.3%
	Risk Assessment – It is clear who is responsible for initiating and conducting algorithmic impact assessments, such as the IAMA.							
	-0.2	0.0	1.3	21.7%	21.7%	15.2%	32.6%	8.7%
2nd Line	**Control** – It is clear who is responsible for overseeing the implementation and quality evaluation of impact assessments, such as the IAMA.							
	-0.4	0.0	1.2	21.7%	26.1%	28.3%	17.4%	6.5%
	Advice – If the owner, user, or developer of algorithms has a question about assessing ethical risks, it is clear who they can turn to for guidance.							
	-0.4	-1.0	1.2	20.0%	33.3%	22.2%	15.6%	8.9%
3rd Line	**Audit** – We conduct independent audits on risk management, such as how impact assessments are carried out, how the ethics committee operates, and whether these practices contribute to reducing harm.							
	-0.6	-1.0	1.3	29.5%	34.1%	13.6%	15.9%	6.8%

Table 5: First line of defense: actors, responsibilities, and risk management tasks.

First line of defense	Algorithmic design responsibilities . . .	Formal risk management responsibilities . . .
Owner of the algorithm	- functional and technical aspects of the algorithm, and - approving the algorithm for specific use cases and processes.	Algorithm owner and user may both be responsible for initiating and conducting: - an initial risk identification - subsequent algorithmic impact assessment,
User of the algorithmic application	- defining the use process of the algorithmic application.	and reporting to and seeking advice from second line
Developer of the algorithm	- developing the algorithm according to the technical requirements.	The developer and data owner may both be consulted in the following:
Owner and/or supplier of the data	- the quality and completeness of the data processed in the algorithm.	- an initial risk identification - subsequent algorithmic impact assessment Reporting to and seeking advice from second line

unclear (see Table 4). There is also confusion over who is responsible for implementing, following up and monitoring the assessment outcomes. Likewise, the responsibility for controlling the quality of the assessment is often not attributed. Research on the predecessor supports this diffusion of responsibilities [16].

Bias mitigation – Respondents generally recognize discrimination as most significant ethical risk, followed by explainability and privacy concerns. The responsibilities and activities related to bias identification and mitigation are frequently relatively well-established within the first line. Typically, the responsibility for bias mitigation is attributed to developers, who may delegate or collaborate with in-house or external data scientists in the research department. Despite the prevalence of efforts, there is no universal standard or procedure for bias mitigation. Bias mitigation assessments are also found to be disconnected from other ethical impact assessments (D13).

Table 5 provides an overview of an ideal attribution of responsibility in the first line of defense deduced from the TLoD model and respondents.

4.2 Second Line of Defense

The second line of defense is responsible for risk oversight and compliance. It oversees and supports the first line by providing independent risk oversight, developing policies, and ensuring compliance with regulations and internal guidelines. Many respondents have either already, or find it desirable to, establish a new actor, such as an ethics officer, or ethics expert team, or a combination thereof. In some organizations, particularly executive agencies, with a more mature formalized risk management, the preexisting data or information officer is attributed to second-line responsibilities. Table 6 provides an overview of an ideal attribution of responsibility in the second line of defense deduced from the TLoD model and respondents.

Ethics officer – The ethics officer – multiple other terms exist, such as algorithm expert, project, or program manager data ethics – develops the data ethics risk management, advises the first line in the implementation, and raises awareness for data ethics throughout the organization. This actor plays a facilitating role in conducting the AIA. Where an ethics expert team exists, the

Table 6: Second line of defense: actors, responsibilities, and risk management tasks.

Second line of defense	Algorithmic design responsibilities . . .	Formal risk management responsibilities . . .
Ethics officer, expert team or data officer	- development, monitoring and evaluation of ethical guidelines and policies	- advising and supporting the first line, e.g. in conducting AIAs - providing training to raise awareness about ethical risks - reporting and communicating ethical risks to the governing body

data ethics expert team officer coordinates the team. Where an ethics commission exists, this person takes on a facilitating role. Particularly in stages of low organizational AI ethics maturity, this actor is perceived to be a crucial forerunner and linchpin in creating awareness and establishing data ethics in the organization. As such, this person requires a unique combination of capabilities and knowledge to combine technical, ethical, and organizational aspects of this role. These individuals emphasize their advisory role and refrain from taking supervision responsibilities.

Data ethics expert team – The ethics officer is often, though not always, supported by a team of data ethics experts. The members of the ethics experts team often have a primary task elsewhere in the organization. They are usually an informal group of individuals with an interest and responsibility in data ethics. The need for formal structures and additional resources becomes more pressing as this group becomes more formalized. This group takes on similar tasks to the ethics officer.

Data Officer– Some PSOs attribute second-line responsibilities to the existing actor of the (chief) data officer or information or privacy officer. This attribution of responsibility to existing actors, rather than developing new functions, seems particularly prevalent in executive agencies. In these cases, the oversight responsibilities of implementing risk management in the first line seemed more formalized. Yet, such actors, particularly privacy officers, tended to give prevalence to the value of privacy above other values in discussing ethical risks. It seems essential that this position requires a unique set of competencies. Without these, the data officer may not provide substantial advice to the second line in dealing with ethical risks.

Algorithm register – The algorithm register is often mentioned as a central second-line responsibility. It can have multiple functions, one of which is the reporting of risks and mitigation strategies. There is much uncertainty about its use, such as: Which algorithms should be added at which time? Who is responsible for reporting an algorithm? Who is responsible for assessing the quality of reporting? Who is responsible for controlling the quality of reporting? The quality, completeness, and timeliness of the information varies. Its primary objective is often perceived as transparency and as a public accountability tool.

4.3 Third Line of Defense

The third line of defense conducts independent audits to assess the effectiveness of risk management practices through audits, evaluating the efficacy of risk management and control processes. Operating independently from management, specifically internal audit, it is accountable to the governing body while fostering collaboration with management for effective risk management. Most respondents appoint these responsibilities to three actors: the internal audit, the data protection officer, an ethics commission, or similar advisory body. In practice, however, audits are rarely conducted (see Table 4). Table 7 provides an overview of an ideal attribution of responsibility in the second line of defense deduced from the TLoD model and respondents.

Internal Audit – The internal audit is responsible for evaluating and advising on implementing risk management practices by the first and second line and reporting its findings to the government body. In practice, internal audits have thus far focused mainly on algorithmic applications rather than risk management (add source).

Data Protection Officer – The Data Protection Officer (DPO) is a new mandatory actor in PSOs. Their primary responsibility involves overseeing and providing advice on compliance with privacy legislation. Some PSOs task the DPO with overseeing and advising ethical risk management. The primary responsibility necessitates that the DPO possesses both knowledge of privacy laws and legal capabilities, adding responsibilities in the domain of data ethics would require uniquely different skills and knowledge, as we have emphasized with the EO.

Data Ethics Commission – Data ethics commission (DEC) have been established to address data ethics in PSOs. DECs are independent advisory bodies established by PSOs to advise on ethical concerns arising with data-driven innovations. DECs are established for multiple reasons, which are not always clear or univocally shared. In the context of ethical risk management, they are established as advisory bodies consisting of external experts who provide knowledge and enable reflections on ethical risks and mitigation strategies. Second, they are emphasized to be an external body composed of independent experts. As such, they are attributed a supervisory responsibility over managing ethical risks. Generally, PSOs place ethics commissions in the third line to emphasize their independence from management functions. Yet, in practice there is confusion about the role within and between the organization and the DEC members.

Generally, the primary emphasis of PSOs is on establishing risk management practices for ethical risks in the first line. Relative to the other two lines, responsibilities in the first are more clearly attributed here. The bias mitigation practices exemplify this. The second-line responsibilities for advising and supervision are less clearly attributed. This can be problematic, as the first line may not successfully discharge their responsibilities without sufficient

Table 7: Third line of defense: actors, responsibilities, and risk management tasks.

Third line of defense	Algorithmic design responsibilities ...	Formal risk management responsibilities ...
Ethics commission	- advising on data ethics	- independent advice for the first and second line - reporting to a government body - evaluating risk management practices in the first and second line, such as the quality of the AIA
Internal audit	- providing independent assurance and evaluating the effectiveness of risk management and control processes	- Independent evaluation of risk management practices
Data protection officer	- overseeing and providing advice on compliance with privacy legislation	- Independent supervision of implementation of ethical risk management governance

advice and oversight, for instance, because of the diffusion of responsibilities for the follow-up of risks identified through various risk assessments. The lack of integration through the second line may overburden the first line with possible conflicting outcomes of multiple risk assessments, such as bias, privacy, and algorithmic impact assessments. The third line is likewise rather insufficiently addressed. This is exemplified by the ambiguous responsibility attribution to the DECs. Consequently, there seems to be a responsibility gap in the oversight and support of ethical risk management. Generally, there is a lack of alignment and integration between the lines, respective actors, and risk management practices. The newly established actors, such as the EO and the DEC, have yet to be embedded in the existing risk management processes. Likewise, the responsibilities for newly created formal practices, such as AIAs and ARs is not clearly attributed.

5 DISCUSSION

Our findings suggest that the TLoD is a useful model for identifying actors and attribute responsibilities for ethical AI risk management in PSOs. Mapping out the various lines of defense and scrutinizing the relationships illuminates responsibility gaps in the management of ethical AI risks. The first line is attributed the responsibility for identifying, assessing, and mitigating ethical AI risks. However, our survey results indicate that the distribution within the first line is unclear. This emphasizes the importance of second line actors, such as the newly established ethics officer, in establishing awareness and a structure for ethical risks, as well as fostering shared (organizational) learning – or developing "moresprudence", as one interviewee calls it. The responsibilities of the third line, particularly the independent audit, are generally not established. This may not be surprising given the low maturity of (ethical AI) risk management. However, it also reflects the transition and learning period public sector organizations find themselves in dealing with the emerging risks of AI adoption. To address the needs of the second and first line, a third line should contribute to independent ethical reflection and shared (organizational) learning. These responsibilities are different from the independent oversight role and goes beyond the collaboration and collaboration that is emphasized by the IAA TLoD model. The digital ethics commissions are an example of the third line actors, who support the first and second line of defense through independent reflection, advice, and collective learning.

Without these adaptions to the TLoD model to manage ethical AI risks, this model in particular, and a risk management perspective in general, may lead to a reductionist approach to managing ethical risks. This risk is exemplified by embedding ethical AI risk management in privacy risk management, such as privacy officers, privacy assessments and privacy commissions. This may be problematic for two reasons: The actors responsible for data ethics require a unique set of knowledge and skills. In contrast, privacy officers need legal knowledge and skills to oversee advise on compliance with privacy legislation. In contrast, dealing with uncertainties and value dilemmas may embrace an interpretive epistemology, a constructivist ontology, and a broader range of ethical frameworks, such as consequentialism or virtue ethics. They grapple with uncertainty from the absence of precedent, regulations, and emergent ethical considerations. Recognizing these differences is essential to prevent a reductionist data ethics risk management approach. Second, existing risk management practices conceptualize risks as adverse outcomes. Integrating ethical risks thereby incentivizes a negative conception of ethical risk as harm. The TLoD model has been criticized for its narrow approach to risks and a lack of value orientation [2, 10, 40]. Our empirical analysis shows that PSOs prioritize ethical risks as harm instead of proactively integrating of human values into AI design. This is exemplified by the comparatively advanced formalization of bias assessment and mitigation. Relating to the negative perception of ethical risks, we find that risk is managed rather than prevented through pro-ethical design choices. The negative and ex-post perceptions of risk are not inherent features of the TLoD. Yet, our empirical insights show that in practice they tend to be interpreted and implemented in this manner.

6 CONCLUSION

In the TLoD we find an established (i.e. well documented and practiced) risk management model that distinguishes between actors, roles, and relationships to enable a structured distribution of responsibilities for managing and overseeing risks (IAA, 2020). The main question guiding this research is to what extent is the TLoD *a useful lens to manage ethical risks in the development of AI by public organizations?* In our explorative analysis, the TLoD provides a useful approach to distributing the responsibilities for ethical risk management according to the lines of defense. Actors of the first line, namely the algorithm owner, developer, and user,

are responsible for identifying, assessing, and mitigating ethical risk in their operational work. Actors of the second line, such as the ethics officer, the ethics expert team, or the data officer, are responsible for developing policies and ensuring compliance with ethical norms and standards. Actors of the third line, such as the ethics commission, internal audit, and data protection officer, evaluate the efficacy of ethical risk management and control processes. While our work does not provide any grounds for generalization, the interview respondents agree that the TLOD is a useful lens to attribute ethical risk management responsibilities to respective actors and, if necessary, establish new actors, such as the ethics officer or data ethics commission. In cases where such actors are missing or disagree with the allocation, responsibility gaps can be identified. The TLoD model's emphasis on accountability, oversight and coordination relationships between the lines provides a means to reintegrate and align the responsibilities. However, our analysis reveals limitations to embedding ethical risks in the TLoD model and existing risk management practices in general. Since the TLoD primarily focuses on intra-organizational information asymmetries, it does not fully capture the public accountability demands of PSOs, particularly in the realm of ethical risks playing out across multiple PSOs. Consequently, the TLoD needs to be adapted to include citizens' perspectives and engagement, considering them crucial indicators for proactively identifying, assessing, and mitigating ethical risks. Embedding ethical risks in risk management practices emphasizes a negative and ex-post perception of ethical risk. We argue that adopting a broader notion of moral responsibility in the TLoD risk management model can facilitate our comprehensive definition of ethical risk as an organizational assessment of the uncertainty of fulfilling a particular moral obligation. Moral concepts of responsibility emphasize the importance of respective responsibility conditions [11, 13, 19, 32, 49, 50]. To discharge their responsibility, actors require necessary conditions, such as knowledge, control, and human agency. In the context of ethical AI risk management, these conditions relate to, for example, the necessary information position, an individual's knowledge and skills, ability for moral reflection. Future research can further draw on philosophical theories of moral responsibility to conceptualize the actor's responsibilities and necessary conditions. Focusing on responsibilities and respective conditions in ethical AI risk management may contribute to a balance between a formalization of ethical AI risk management and a dynamic process enabling the reflection on and integration of human values in ethical AI design.

REFERENCES

[1] Angelika Adensamer, Rita Gsenger, and Lukas Daniel Klausner. 2021. "Computer says no": Algorithmic decision support and organisational responsibility. Journal of Responsible Technology 7, 8 (2021), 2666–6596. https://doi.org/10.1016/j.jrt.2021.100014

[2] Ulrich Bantleon, Anne d'Arcy, Marc Eulerich, Anja Hucke, Burkhard Pedell, and Nicole V.S. Ratzinger-Sakel. 2021. Coordination challenges in implementing the three lines of defense model. International Journal of Auditing 25, 1 (March 2021), 59–74. https://doi.org/10.1111/ijau.12201

[3] Barry Bozeman and Gordon Kingsley. 1998. Risk Culture in Public and Private Organizations. Public Adm Rev 58, 2 (1998), 109–118.

[4] Stefan Buijsman and Herman Veluwenkamp. 2022. Spotting When Algorithms Are Wrong. Minds Mach (Dordr) (2022). https://doi.org/10.1007/s11023-022-09591-0

[5] Justin B. Bullock. 2019. Artificial Intelligence, Discretion, and Bureaucracy. Am Rev Public Adm 49, 7 (October 2019), 751–761. https://doi.org/10.1177/0275074019856123

[6] Marianna Capasso. 2023. Responsible Social Robotics and the Dilemma of Control. Int J Soc Robot 15, 12 (December 2023), 1981–1991. https://doi.org/10.1007/S12369-023-01049-2/METRICS

[7] Martin Carlsson-Wall, Kalle Kraus, Anita Meidell, and Patrik Tran. 2019. Managing risk in the public sector – The interaction between vernacular and formal risk management systems. Financial Accountability and Management 35, 1 (February 2019), 3–19. https://doi.org/10.1111/faam.12179

[8] Neelke Doorn. 2012. Responsibility Ascriptions in Technology Development and Engineering: Three Perspectives. Sci Eng Ethics 18, 1 (March 2012), 69–90. https://doi.org/10.1007/s11948-009-9189-3

[9] Mary Douglas and Aaron Wildavsky. 1982. How Can We Know the Risks We Face? Why Risk Selection Is a Social Process. Risk Analysis 2, 2 (June 1982), 49–58. https://doi.org/10.1111/j.1539-6924.1982.tb01365.x

[10] Marc Eulerich. 2021. The new three lines model for structuring corporate governance – A critical discussion of similarities and differences. Corporate Ownership and Control 18, 2 (2021), 180–187. https://doi.org/10.22495/cocv18i2art15

[11] Nihlén Fahlquist, Neelke Doorn, and Ibo van de Poel. 2015. Design for the value of responsibility. In Handbook of ethics, values, and technological design. Springer, Dordrecht, 473–490.

[12] Menno Fenger and Robin Simonse. 2024. The implosion of the Dutch surveillance welfare state. Soc Policy Adm (2024). https://doi.org/10.1111/SPOL.12998

[13] John Martin Fischer and Mark Ravizza. 1993. Perspectives on moral responsibility. Cornell University Press.

[14] Luciano Floridi. 2019. Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical. Philos Technol 32, 2 (June 2019), 185–193. https://doi.org/10.1007/s13347-019-00354-x

[15] Luciano Floridi, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke, and Effy Vayena. 2018. AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. 28, (2018), 689–707. https://doi.org/10.1007/s11023-018-9482-5

[16] Aline Shakti Franzke, Iris Muis, · Mirko, and Tobias Schäfer. 2021. Data Ethics Decision Aid (DEDA): a dialogical framework for ethical inquiry of AI and data projects in the Netherlands. Ethics Inf Technol 23, (2021), 551–567. https://doi.org/10.1007/s10676-020-09577-5

[17] Stephan Grimmelikhuijsen and Albert Meijer. 2022. Legitimacy of Algorithmic Decision-Making: Six Threats and the Need for a Calibrated Institutional Response. Perspect Public Manag Gov 5, (2022), 232–242. https://doi.org/10.1093/ppmgov/gvac008

[18] Marissa Hoekstra, Anne Fleur Van Veenstra, and Cass Chideock. 2021. A Typology for Applications of Public Sector AI. In EGOV-CeDEM-ePart, 2021. 121–128.

[19] Jeroen Van Den Hoven. 1998. Moral Responsibility, Public Office and Information Technology. In Public administration in an information age: A handbook (6th ed.), I.Th.M. Snellen and W.B.H.J. van de Donk (eds.). IOS press, Amsterdam.

[20] Jeroen van den Hoven. 2022. Responsibility and innovation. J Respons Innov 9, 1 (2022), 133–137. https://doi.org/10.1080/23299460.2022.2050570

[21] Jeroen Van den Hoven, Gert Jan Lokhorst, and Ibo Van de Poel. 2012. Engineering and the Problem of Moral Overload. Sci Eng Ethics 18, 1 (March 2012), 143–155. https://doi.org/10.1007/s11948-011-9277-z

[22] IAA. 2020. The IIA's Three Lines Model - An update of the Three Lines of Defense.

[23] Alexandra James and Andrew Whelan. 2022. "Ethical" artificial intelligence in the welfare state: Discourse and discrepancy in Australian social services. Crit Soc Policy 42, 1 (2022), 22–42. https://doi.org/10.1177/0261018320985463

[24] Tommy Jensen, Johan Sandström, and Sven Helin. 2015. One Code to Rule Them All: Management Control and Individual Responsibility in Contexts. Retrieved from https://www.jstor.org/stable/44074854

[25] Olya Kudina and Peter Paul Verbeek. 2019. Ethics from within: google glass, the Collingridge dilemma, and the mediated value of privacy. Sci Technol Human Values 44, 2 (March 2019), 291–314. https://doi.org/10.1177/0162243918793711

[26] Habib Mahama, Mohamed Elbashir, Steve Sutton, and Vicky Arnold. 2020. Enabling enterprise risk management maturity in public sector organizations. (2020). https://doi.org/10.1080/09540962.2020.1769314

[27] Lorna McGregor, Daragh Murray, and Vivian Ng. 2019. International human rights law as a framework for algorithmic accountability. International and Comparative Law Quarterly 68, 2 (2019), 309–343. https://doi.org/10.1017/S0020589319000046

[28] Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. The ethics of algorithms: Mapping the debate. Big Data Soc 3, 2 (December 2016). https://doi.org/10.1177/2053951716679679

[29] Tommaso Palermo. 2014. Accountability and Expertise in Public Sector Risk Management: A Case Study. Financial Accountability & Management 30, 3 (August 2014), 322–341. https://doi.org/10.1111/FAAM.12039

[30] Rik Peeters and Arjan Widlak. 2018. The digital cage: Administrative exclusion through information architecture – The case of the Dutch civil registry's master data management system. Gov Inf Q 35, 2 (April 2018), 175–183. https://doi.org/10.1016/j.giq.2018.02.003

[31] Rik Peeters and Arjan C. Widlak. 2023. Administrative exclusion in the infrastructure-level bureaucracy: The case of the Dutch daycare benefit scandal.

Public Adm Rev 83, 4 (July 2023), 863–877. https://doi.org/10.1111/PUAR.13615
[32] Udo Pesch. 2015. Engineers and Active Responsibility. Sci Eng Ethics 21, (2015), 925–939. https://doi.org/10.1007/s11948-014-9571-7
[33] Iyad Rahwan. 2018. Society-in-the-loop: programming the algorithmic social contract. Ethics Inf Technol 20, (2018), 5–14. https://doi.org/10.1007/s10676-017-9430-8
[34] Tarek Rana and Lee Parker. 2024. The Routledge Handbook of Public Sector Accounting. Routledge, New York.
[35] Tapani Rinta-Kahila, Ida Someh, Nicole Gillespie, Marta Indulska, Shirley Gregor, Patrick Mikalef, Aleš Popovic, Jenny Eriksson Lundström, and Kieran Conboy. 2022. Algorithmic decision-making and system destructiveness: A case of automatic debt recovery. European Journal of Information Systems 31, 3 (2022), 313–338. https://doi.org/10.1080/0960085X.2021.1960905
[36] Andre Rodrigues, Adina Dudau, Georgios Kominis, and Alvise Favotto. 2023. Framing risk management within management control systems. The Routledge Handbook of Public Sector Accounting (November 2023), 189–201. https://doi.org/10.4324/9781003295945-18/FRAMING-RISK-MANAGEMENT-WITHIN-MANAGEMENT-CONTROL-SYSTEMS-ANDRE-RODRIGUES-ADINA-DUDAU-GEORGIOS-KOMINIS-ALVISE-FAVOTTO
[37] Filippo Santoni de Sio and Giulio Mecacci. 2021. Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them. Philos Technol 34, 4 (December 2021), 1057–1084. https://doi.org/10.1007/s13347-021-00450-x
[38] Antonia Sattlegger, Jeroen van den Hoven, and Nitesh Bharosa. 2022. Designing for Responsibility. In DG.O 2022: The 23rd Annual International Conference on Digital Government Research, June 15, 2022. ACM, New York, NY, USA, 214–225. https://doi.org/10.1145/3543434.3543581
[39] Jonas Schuett. 2023. Three lines of defense against risks from AI. AI Soc (November 2023). https://doi.org/10.1007/s00146-023-01811-0
[40] Kai-Uwe Seidenfuss, Angus Young, and · Mohan Datwani. 2023. Integrating governance, risk and compliance? A multi-method analysis of the new Three Lines Model. SN Business & Economics (2023), 1–28. https://doi.org/10.1007/s43546-023-00561-x
[41] Friso Selten and Bram Klievink. 2024. Organizing public sector AI adoption: Navigating between separation and integration. Gov Inf Q 41, (2024), 101885. https://doi.org/10.1016/j.giq.2023.101885
[42] Friso Selten and Albert Meijer. 2021. Managing Algorithms for Public Value. International Journal of Public Administration in the Digital Age 8, 1 (January 2021). https://doi.org/10.4018/IJPADA.20210101.oa9
[43] Renee Shelby, Shalaleh Rismani, Kathryn Henne, AJung Moon, Negar Rostamzadeh, Paul Nicholas, N'Mah Yilla, Jess Gallegos, Andrew Smart, Emilio Garcia, and Gurleen Virk. 2022. Sociotechnical Harms of Algorithmic Systems: Scoping a Taxonomy for Harm Reduction. (October 2022). Retrieved from http://arxiv.org/abs/2210.05791
[44] Sergeja Slapničar, Micheal Axelsen, Ivano Bongiovanni, and David Stockdale. 2023. A pathway model to five lines of accountability in cybersecurity governance. International Journal of Accounting Information Systems 51, (December 2023), 100642. https://doi.org/10.1016/J.ACCINF.2023.100642
[45] Tara Qian Sun and Rony Medaglia. 2019. Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. Gov Inf Q 36, 2 (April 2019), 368–383. https://doi.org/10.1016/j.giq.2018.09.008
[46] Robert Swedberg. 2020. Exploratory Research. In The Production of Knowledge: Enhancing Progress in Social Science (Strategies for Social Inquiry), Colin Elman, John Gerring and James Mahoney (eds.). Cambridge University Press, 17–41. Retrieved February 2, 2024 from https://books.google.com/books/about/The_Production_of_Knowledge.html?hl$=$nl&id$=$vlTMDwAAQBAJ
[47] Andreas Tsamados, Nikita Aggarwal, Josh Cowls, Jessica Morley, Huw Roberts, Mariarosaria Taddeo, · Luciano Floridi, and Luciano Floridi. 2022. The ethics of algorithms: key problems and solutions. 37, (2022), 215–230. https://doi.org/10.1007/s00146-021-01154-8
[48] Andreas Tsamados, Nikita Aggarwal, Josh Cowls, Jessica Morley, Huw Roberts, Mariarosaria Taddeo, and Luciano Floridi. 2022. The ethics of algorithms: key problems and solutions. AI Soc 37, 1 (March 2022), 215–230. https://doi.org/10.1007/s00146-021-01154-8
[49] Nicole A. Vincent. 2011. A Structured Taxonomy of Responsibility Concepts. In Library of Ethics and Applied Philosophy. Springer Science and Business Media B.V., 15–35. https://doi.org/10.1007/978-94-007-1878-4_2
[50] Nicole Vincent, Ibo van de Poel, and Jeroen van den Hoven. 2011. Moral responsibility - Beyond Free Will and Determinism. Springer. Retrieved from http://www.springer.com/series/6230
[51] Jan C. Weyerer and Paul F. Langer. 2020. Bias and Discrimination in Artificial Intelligence. . 256–283. https://doi.org/10.4018/978-1-7998-1879-3.ch011
[52] Bernd W Wirtz, Jan C Weyerer, and Ines Kehl. 2022. Governance of artificial intelligence: A risk and guideline-based integrative framework. (2022). https://doi.org/10.1016/j.giq.2022.101685
[53] Karen Yeung. 2022. The New Public Analytics as an Emerging Paradigm in Public Sector Administration. Tilburg Law Review 27, 2 (2022), 1–32. https://doi.org/10.5334/tilr.303