June 2026

Implementing Ai Governance: Contrasting The Academic And Grey Literature

Fuad Khan
University of Turku, fuhakh@utu.fi

Matti Minkkinen
University of Turku, matti.minkkinen@utu.fi

Henrietta Jylhä
University of Turku, henrietta.jylha@utu.fi

Teemu Birkstedt
University of Turku, teemu.birkstedt@utu.fi

Matti Mäntymäki
University of Turku, matti.mantymaki@utu.fi

IMPLEMENTING AI GOVERNANCE: CONTRASTING THE ACADEMIC AND GREY LITERATURE

Completed Research Paper

Fuad Khan, University of Turku, Turku, fuhakh@utu.fi

Matti Minkkinen, University of Turku, Turku, matmin@utu.fi

Henrietta Jylhä, University of Turku, Turku, henrietta.jylha@utu.fi

Teemu Birkstedt, University of Turku, Turku, teemu.birkstedt@utu.fi

Matti Mäntymäki, University of Turku, Turku, matti.mantymaki@utu.fi

Abstract

The rapid development and increasing organizational deployment of AI, especially in high-risk application areas and under evolving regulation, create a growing need for effective AI governance. Hence, AI governance has attracted increasing attention in both academia and practice. While prior reviews have examined academic research on AI governance, practice-oriented grey literature has received less scholarly attention. This study examines how academic and grey literature differ in their treatment of AI governance. To this end, we review academic and grey literature on three focal objects of AI governance: AI data, AI algorithms, and AI models. Our findings suggest that academic literature tends to emphasize principles and pre-deployment governance, whereas grey literature more often foregrounds policies, implementation-oriented controls, and post-deployment oversight. Taken together, the two literatures appear to offer partly complementary perspectives on organizational AI governance. The study clarifies how AI governance is framed differently in scholarly and practice-oriented discourse and thereby contributes to Information Systems research and practice on organizational AI governance

Keywords: AI Governance, Artificial Intelligence Governance, AI, Artificial Intelligence, Literature Review, Grey Literature.

1 Introduction

The demand for Artificial Intelligence (AI) governance in organizations is rapidly increasing due to the growing use of AI in critical sectors such as healthcare, finance, transportation, and rising concerns about its risks (Jafari et al., 2026; Khoo et al., 2025; X. Wang & Xie, 2025; Papagiannidis et al., 2025). In response to the societal concerns regarding AI's risk, several regulatory governance initiatives, such as the EU AI Act, and evolving AI regulation in the United States (Novelli, 2025; Mäntymäki, 2023), and international voluntary ethical principles, such as the OECD principles and UNESCO's ethics of AI (OECD, 2024; UNESCO, 2023), emerged. These global developments indicate that AI governance, understood as "a system of rules, practices, processes, and technological tools that are employed to ensure an organization's use of AI technologies aligns with the organization's strategies, objectives, and values; fulfils legal requirements; and meets ethical AI principles followed by the organization" (Mäntymäki et al., 2022), is no longer optional but essential for organizations to ensure responsible AI use (Mäntymäki et al., 2022; Schneider et al., 2022; Papagiannidis et al., 2025). However, translating ethical principles and regulations into concrete AI governance practices remains challenging due to their abstract nature and lack of practical specificity. (Papagiannidis et al., 2025; Gengler & Schmalenbach, 2024; Birkstedt et al., 2023; Mittelstadt, 2019).

The volume of AI governance research, also in Information Systems (IS), has grown rapidly in recent years (Birkstedt et al., 2023). Literature reviews have synthesized the findings from prior studies and contributed to conceptualizing of AI governance (Batool et al., 2025; Papagiannidis et al., 2025; Birkstedt et al., 2023). However, developments in AI governance are not limited to the scholarly realm; for example, AI vendors, consultancies, and various non-governmental organizations (NGOs) such as UNESCO and OECD have been active in developing AI governance policies and practices targeted at organizations and guiding their implementation.

This study expands the scope of literature reviews on AI governance from peer-reviewed sources (Papagiannidis et al., 2025; Birkstedt et al., 2023) to cover grey literature, defined as publicly available, non-commercial documents from governments, academia, business, and industry (Paez, 2017; Benzies et al., 2006). Compared to peer-reviewed research, grey literature lacks well-established databases and guidelines for conducting reviews (Mahood et al., 2013). Hence, the potential to derive scientific insights from grey literature remains underutilized (Benzies et al., 2006). Thus, identifying potential complementary insights and points of divergence between the scholarly and practice-focused discourses can benefit both IS research and practice in AI governance implementation.

Against this backdrop, this study contrasts the academic literature on AI governance with the grey literature. We focus particularly on three focal objects of AI governance: AI data, AI algorithms, and AI models. In doing so, we address the following research question: "How do the academic and grey literature differ in conceptualizing and operationalizing AI Governance with respect to three dimensions: AI data, AI algorithms, and AI models? We conducted a systematic review of 48 empirical academic studies and 27 practice-oriented documents. Our findings contribute to the IS scholarship on AI governance (Birkstedt et al., 2023; Papagiannidis et al., 2025; Van Giffen & Ludwig, 2023; Batool et al., 2025, 2025; Mäntymäki et al., 2022; Schneider et al., 2022) by identifying gaps in the current body of knowledge and deciphering contrasts between the academic and grey literature with respect to implementing AI governance.

2 Background

2.1 Implementing AI Governance

The rapid technical development and increasing use of artificial intelligence (AI) in high-risk application areas, such as traffic, healthcare, and finance, demand AI systems to operate in alignment with societal values and organizations' objectives (Berente et al., 2021; Birkstedt et al., 2023; Zimmer et al., 2022). In response, organizations and governmental institutions have published ethical AI principles (Jobin et al. 2019; Morley et al., 2019; OECD, 2024; The Global Partnership on Artificial Intelligence, 2023; G7, 2023; UNESCO, 2023).

However, AI ethics principles have been criticized for being too abstract to translate into organizational practices (Mittelstadt, 2019; Schneider et al., 2022; Papagiannidis et al., 2025; Batool et al., 2025; Gengler & Schmalenbach, 2024). This creates a strong need for organizational AI governance, understood as "a system of rules, practices, processes, and technological tools that are employed to ensure an organization's use of AI technologies aligns with the organization's strategies, objectives, and values; fulfils legal requirements; and meets ethical AI principles followed by the organization" (Mäntymäki et al., 2022). The recently enacted AI-specific regulations, such as the EU's AI Act, further exacerbate this need.

While ethical AI principles and responsible AI have received considerable scholarly attention, the practical implementation of AI governance remains a challenge (e.g., Birkstedt et al., 2023) due to the lack of a principles-to-practice framework (Mittelstadt, 2019). Nevertheless, research in IS and other fields, such as computer science, has identified implementation gaps and proposed tools, frameworks, and methods for AI governance implementation. Notable examples include the hourglass model

(Mäntymäki et al., 2022), the ECCOLA method to implement AI ethics (Vakkuri et al., 2021), the AI governance implementation dimensions (Papagiannidis et al., 2022), the decomposed data, model, and system-level AI governance framework (Schneider et al., 2022), the impact assessment framework (Schiff et al., 2020), and the accountability framework (Batool et al., 2025).

In addition to academic research, several international bodies, including the OECD and UNESCO, as well as national bodies such as the National Institute of Standards and Technology (NIST), and AI developers such as OpenAI, Anthropic, Perplexity, and Google, have published AI governance implementation guidelines and frameworks. These include the OECD AI principles (AI Principles Overview - OECD.AI), UNESCO's proposal for an AI ethics approach (Ethics of Artificial Intelligence - UNESCO, 2023), and the NIST AI Risk Management Framework (U.S. National Institute of Standards and Technology, 2023).

Exploring both academic and grey literature in parallel can enrich the understanding of the state of the art in AI governance. Thus far, prior reviews on AI governance (e.g., Batool et al., 2025; Birkstedt et al., 2023; Papagiannidis et al., 2025) have considered only academic literature, while ethics guidelines in the grey literature have also been reviewed (Hagendorff, 2020; Jobin et al., 2019). Knowledge of the contrasts between academic and grey literature provides the basis for eventual integrative approaches to implementing AI governance. We present our structuring of the governance objects in the next section.

2.2 AI Data, AI Algorithms, and AI Models

In this research, we focus on three key elements of AI: AI data, AI algorithms, and AI models, since they are foundational to how AI systems operate and prevalent across diverse applications (Batool et al., 2025; Birkstedt et al., 2023). We therefore classify our results around these key elements. Data provides the input to train AI systems (Janssen et al., 2020; Schneider et al., 2022), AI algorithms define the logic of data processing and decision-making (Nishant et al., 2024; Samtani et al., 2023), and AI models are the reusable components that constitute AI systems (Ferrari et al, 2023; Schneider et al., 2022).

The difference between an AI model and an algorithm lies in their functionality, complexity, and application context. An algorithm's function is to provide logic (with mathematical and computational formulas) for data processing and analysis, whereas a model's function is to serve the end goal, e.g., classification or clustering of the data. In other words, an algorithm answers how the processing is executed, and a model answers what is being executed. (Achmadiah et al., 2025; Li & Bi, 2024; Hartog et al., 2024; Song, 2021; Phan et al., 2023; Li et al., 2019) Despite these distinctions, distinguishing these terminologies is seemingly difficult.

As data is the raw material for an AI model, the model's output quality depends significantly on the quality of the training data (Nishant et al., 2023; Schneider et al., 2022). Data governance refers to the systematic management of data usage in the AI development lifecycle (Janssen et al., 2020). There are several issues related to AI data, e.g., data privacy, biased datasets that lead to discrimination, and data and security breaches. Therefore, data governance includes activities to ensure trustworthy and responsible AI development. (Mäntymäki et al., 2022; Bu, 2021).

AI algorithm governance includes technical (along with legal and regulatory) activities to ensure accountability, transparency, and reliability in connection to algorithm type, application context, and risks (Doneda & Almeida, 2016). A biased coding of procedures may lead to an unfavourable decision by the AI. Ethical and legally compliant algorithmic design is ensured by evaluating the technical functionalities, e.g., logic, optimization functions, feature or variable management, and non-technical mechanisms, e.g., algorithmic transparency, documentation of logic and reasoning, design reviews, logic audits, and their validation. (Buhmann & Fieseler, 2021; Barn, 2019).

Model governance spans the entire lifecycle: design, development, deployment, and monitoring of the trained AI to prevent unintended consequences (Batool et al., 2025). In this case, model governance means ensuring accurate, robust, explainable, trustworthy, and ethically aligned outputs. In practice, the model's learned parameters, performance indicators, input-output behavioural patterns, and deployment settings are monitored continuously to prevent model drift, inaccuracy, overfitting, and hallucinations. Moreover, tools such as model cards, explainability tools, model risk assessment, impact assessment, incident tracking, and auditing are used for transparency and accountability. (Papagiannidis et al., 2022; United Nations System White Paper on AI Governance, 2024; Buhmann & Fieseler, 2021).

3 Methodology

3.1 Review of the academic literature

To answer our research question: "How do the academic and grey literature differ in conceptualizing and operationalizing AI Governance implementation?", we focused on contrasting academic and grey literature, specifically with respect to AI data, AI algorithms, and AI models. Based on Snyder (2019) literature review process, we utilized the following steps in reviewing academic literature: 1. Research questions, 2. Search terms, 3. Search process, 4. Inclusion/exclusion criteria. Our search process included the key search term *AI governance* (without quotation marks) to capture expansive results. The search timeline was set from January 2022 to March 2025, reflecting the rapid development of AI technologies and AI-specific regulatory developments (e.g., the EU AI Act and the evolving AI regulation in the United States).

With respect to the academic literature, we focused on empirical research published in the Senior Scholars' Premier Set of 11 IS journals and in the proceedings of ICIS and ECIS. We only included empirical studies. The search with the defined keyword and timeline was applied directly to publishers' websites. The search process yielded 1501 studies. Since our focus was on AI governance implementation, we screened the papers' titles, keywords, and abstracts to identify those that aligned with our focus. The screening yielded a pool of 117 articles discussing AI governance for further evaluation. In the evaluation, we employed three heuristics to determine the inclusion of the paper in the final pool of papers to be reviewed: 1) Does the paper inform about AI governance implementation?, 2) Does the paper inform about AI governance best practices?, and 3) Does the paper inform about AI governance mechanisms? 38 papers met at least one of these criteria and were thus included. Thereafter, we conducted a backward and forward search based on the papers' references. This led to the identification of 10 additional papers to be included. Hence, we identified 48 papers for inclusion in the review of the academic literature on AI governance (see Figure 1). The academic literature included in the review is available in Online Appendix B[1].

3.2 Review of the grey literature

Grey literature, understood here as publicly available, non-commercial documents from governments, academia, business, and industry (Paez, 2017; Benzies et al., 2006)**,** can provide meaningful insights that augment reviews of peer-reviewed literature (Pappas, C. & Williams, I., 2011). However, due to the volume of information, the lack of standard indexing and controlled vocabulary, and the lack of archiving, grey literature searches are often less systematic than those of scholarly literature (Turner et al., 2005). Therefore, developing a systematic search strategy is important to reinforce rigor in the review of grey literature (Higgins, J., & Green, S., 2011).

Hence, we developed the process for executing the grey literature search based on Godin et al. (2015). We began by consulting a group of AI governance experts to identify organizations that publish material potentially relevant to our study's objective. As a result of the expert consultation, we established a list of 26 organizations, including standard-setting bodies, audit firms, management and IT consulting firms, think tanks, and advocacy groups. Thereafter, we conducted a series of internet searches to obtain AI governance-focused documents published by these organizations. Where applicable, we also used search tools on the organizations' websites to locate grey literature documents. For example, research organizations such as the Centre for the Governance of AI, the Ada Lovelace Institute, and the Rand

[1] Appendix B

https://docs.google.com/document/d/1F_NZzJ2Yb_AvuCTRorzzRgG3LmqEMKDTCBr7PgI5QYc/edit?tab=t.0

Corporation publish their research on their websites. Overall, the grey literature search yielded 292 documents published between January 2022 and March 2025.

To ensure the quality of the grey literature included in the review, we established inclusion and exclusion criteria. Building on Tyndall's (2008) suggestions on evaluating grey literature, we established three inclusion criteria: 1) the document needs to include some description of its underlying methodology or data, 2) the document needs to explicate its scope or application area, and 3) the document's publishing data needs to be available. We excluded documents we evaluated as primarily marketing-oriented, i.e., not providing actionable guidance for organizations. We also excluded documents with a sole technical and legal focus. As a result of the quality evaluation, we included 27 grey literature documents in the review. Online Appendix C[2] provides details about the grey literature. Figure 1 presents the literature review process.

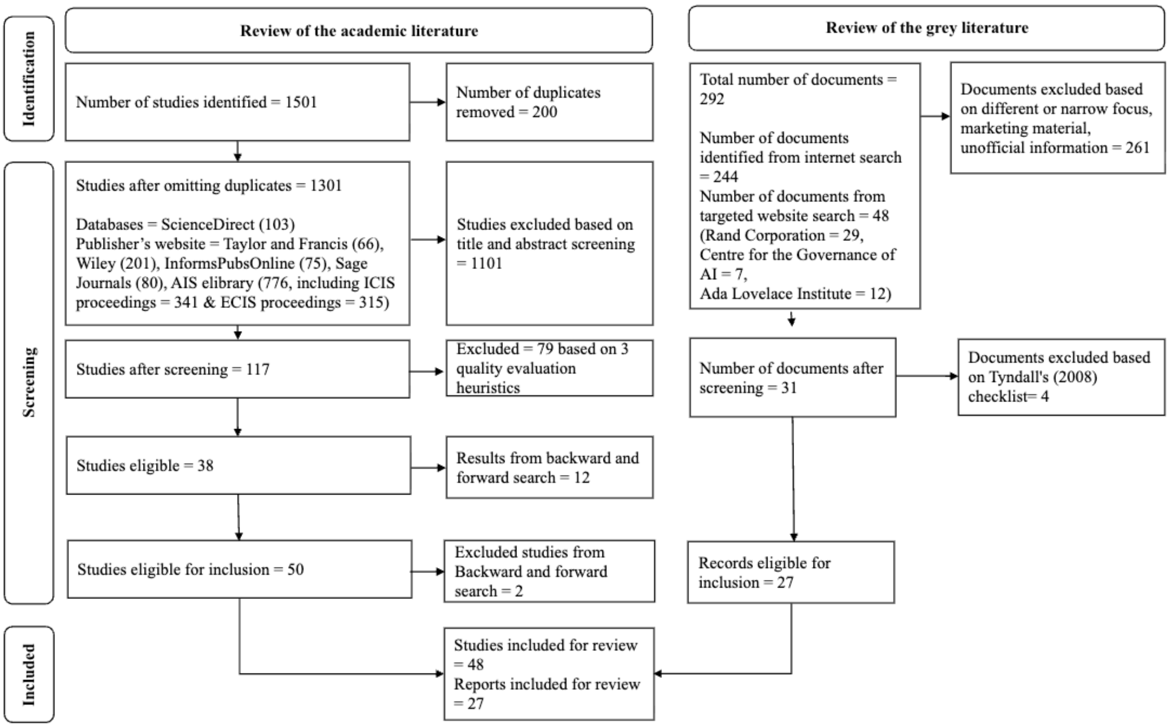

Figure 1. The literature review process.

3.3 Data analysis

Data analysis began with an initial coding of the collected material, guided by the research question. We then grouped similar codes into broader subcategories. In parallel, we examined how these subcategories and their constituent codes were discussed in the academic and grey literature. Through iterative comparison, we synthesized these observations into higher-order themes. In the final stage, we mapped the themes and subcategories onto our three focal objects of AI governance: AI data, AI algorithms, and AI models. The first author conducted the analysis, and the authors discussed interim results until agreement was reached.

4 Results

This section contrasts academic and grey literature on AI governance across three recurring differences in emphasis: first, how governance is structured in terms of principles and policies; second, how

[2] Appendix C

https://docs.google.com/document/d/1F_NZzJ2Yb_AvuCTRorzzRgG3LmqEMKDTCBr7PgI5QYc/edit?tab=t.0

governance is articulated at the level of implementation; and third, where governance attention is placed across the AI lifecycle. The comparison highlights dominant emphases rather than strict differences, because both corpora exhibit overlap and exceptions. Table 1 summarizes governance practices in the academic and grey literature across AI systems, AI algorithms, and AI models. The complete result is published in the online Appendix A[3]

4.1 Academic and grey literature differ in how explicitly they articulate principles and policies

In this study, we define principles as organizations' foundational values and commitments that are often normative and broad (Jobin et al., 2019). In the AI governance context, principles pertain to values and commitments to ensure transparency and explainability, justice, fairness and non-discrimination, non-maleficence, responsibility and accountability, privacy (Seppälä et al., 2021) in the AI development and use process. Policies, in turn, translate these values and commitments into prescriptive rules that guide the operational practices (Foote et al., 2005).

Both academic and grey literature discuss principles and policies relevant to AI governance, particularly regarding AI data. In the academic literature, these are often articulated as normative commitments such as explainability through AI output/decision and process-to-output / decision explanation (Papagiannidis et al., 2025; Tian et al., 2023), privacy through minimum data sharing (Janssen et al., 2020), accountability through data ownership responsibilities (Papagiannidis et al., 2022), fairness and transparency through data-driven ethical solutions (Rakova et al., 2021). These normative principles are connected to organizational policies concerning data access, use, and management. Grey literature also addresses such issues, particularly through commitments to data quality, privacy, security, protection, and retention deletion upon request, data ethical principles (KPMG, 2026; Ernst & Young, 2024; Perplexity, 2024; Boston Consulting Group, 2022; UNESCO, 2023; Bain & Company, n.d.).

The contrast becomes clearer in the AI algorithm and AI model governance. Academic literature often makes the normative basis of governance more explicit by framing accountability, and transparency (Shneiderman, 2020) as principles that inform algorithm-related governance choices. Grey literature, by contrast, more often foregrounds policy-oriented and organizationally salient concerns such as model quality by defining model quality characteristics, components, and metrics (European Telecommunications Standards Institute, 2022), security (Ernst & Young, 2024), audit and compliance such as access (Ada Lovelace Institute, 2023), without always making the underlying principles equally explicit. Thus, the two literatures overlap in their treatment of governance concerns but differ in how directly they articulate the normative and policy logic behind them.

4.2 Academic and grey literature differ in their articulation at the level of implementation

Both literatures discuss concrete means for implementing AI governance. However, they tend to do so in different ways. Academic literature often articulates governance through methods and mechanisms such as differential privacy and data anonymization (De Laat, 2021), algorithmic bias detection (Shneiderman, 2020), applying LIME algorithm (De Laat, 2021), building contestable AI (Lier et al., 2023), benchmarking (Shneiderman, 2020), and feature and weight transparency (Tutun et al., 2024). These are typically presented as governance-enabling mechanisms that can support goals such as privacy, fairness, interpretability, explainability, or performance. Hence, these governance methods and mechanisms are often grounded on normative principles in academic literature. For example, differential privacy and data anonymization (De Laat, 2021) derives from privacy-related principles such as separation of personal and sensitive data (Janssen et al., 2020). Similarly, the practice of sharing the

[3] Appendix A

https://docs.google.com/document/d/1F_NZzJ2Yb_AvuCTRorzzRgG3LmqEMKDTCBr7PgI5QYc/edit?tab=t.0

underlying algorithmic calculation (Loefflad et al., 2023) originates from the principle of algorithmic accountability and transparency (Shneiderman, 2020).

Grey literature more often articulates governance through implementation-oriented technical controls, safety and security procedures, and socio-technical organizational arrangements. For example, security is ensured through control mechanisms such as AI access control (Perplexity, 2024), privacy protection measures such as confidential computing (Microsoft, 2022), and federated learning (International Telecommunication Union, 2024). Similarly, safety is implemented through prompt engineering (International Association of Privacy Professionals, 2024), safety tuning (Google, 2025), and putting guardrails for redline crossing (OpenAI, 2023). Beyond these procedural activities, documentation artifacts such as model and system cards, lineage records (Google, 2025; Boston Consulting Group, 2022; International Association of Privacy Professionals, 2024) and data provenance (International Association of Privacy Professionals, 2024; Boston Consulting Group, 2022; Ada Lovelace Institute, 2023) for auditability and accountability demonstrates the socio-technical aspect of the implementation approach. However, unlike academic literature, grey literature implementation is not formulated on normative grounds, as data reveals AI algorithms and AI models objects lack principles and policy discussions.

However, the contrast is not simply that one literature is conceptual and the other practical. Rather, academic literature tends to foreground mechanisms for achieving governance objectives, whereas grey literature more often foregrounds the organizational embedding of governance in operational practice.

4.3 Academic and grey literature differ in temporal orientation across the AI lifecycle

A third difference concerns the lifecycle stages that receive most attention. Academic literature tends to emphasize pre-deployment governance activities including data preparation, model training, and approval for deployment. For example, data is prepared with activities such as bias testing (Shneiderman, 2020; De Laat, 2021), analyzing features (De Laat, 2021), data curation and reweighing (Shneiderman, 2020; De Laat, 2021), feature extraction and data partitioning (Zhdanov et al., 2021). Then, methods such as substantive rationality (Nishant et al., 2023), supervised learning method (Tanriverdi et al., 2023), and context-adaptive models (Lier et al., 2023) are applied to train the model with the prepared data. Finally, the model is validated through benchmarking (Shneiderman, 2020) and intervening mechanisms, such as quality gates (Mayer et al., 2024), ensuring that developed AI models meet quality requirements. These discussions are often tied to design-time choices intended to shape system behavior before deployment.

Grey literature emphasizes post-deployment governance, particularly monitoring, maintenance, incident response, and performance oversight related activities. For example, a model can be maintained through daily database backup (Perplexity, 2024) and regular model updates (Schuett et al., 2023). As model behavior is continuously monitored through model weight tracking (Schuett et al., 2023) and performance forecasting (Anthropic, 2024), corrections are performed, when necessary, through prompt engineering (International Association of Privacy Professionals, 2024) and safety tuning (Google, 2025).

Thus, in this corpus, governance appears more often as an ongoing operational task that continues after deployment. At the same time, this is a difference in emphasis rather than a strict divide, since academic literature also includes some post-deployment monitoring such as monitoring model stability (Schaschek & Engel, 2023), overoptimism (Liu & Kirshner, 2024), and changes in outcomes (Janssen et al., 2020), and grey literature includes pre-deployment concerns such as data quality investigations (G7, 2021; International Telecommunication Union, 2024), and data poisoning prevention (UNESCO, 2023).

Governance objects	Governance practices in the academic literature	Governance practices in the grey literature
AI Data	Minimum data sharing Janssen et al. (2020)	Commitment to data quality, privacy, security, protection, retention, deletion upon request, data ethical principles KPMG (2026); Ernst & Young (2024); Perplexity (2024); UNESCO (2023)
	Data ownership responsibilities Papagiannidis et al. (2022)	Confidential computing Microsoft (2022)
	Data driven approach to ethical challenges Rakova et al. (2021)	Federated learning International Telecommunication Union (2024)
	Data encryption and anonymization; Differential privacy De Laat (2021); Strobel & Banh (2024)	Daily database backup Perplexity (2024)
	Bias testing; Data curation and reweighing De Laat (2021)	Regular model updates; Model weights tracking Schuett et al. (2023)
AI Algorithms	AI output and process to output explanation; AI decision explanation, Process-to-decision explanation Papagiannidis et al. (2025); Tian et al. (2023)	Providing reasoning rationale European Telecommunications Standards Institute (2022)
	Algorithmic accountability and transparency; Algorithmic bias detection; Benchmarking Shneiderman (2020)	
AI Models	Contestable AI with counterfactual data Lier et al. (2023)	Defining quality characteristics of model components, Quality metrics for models European Telecommunications Standards Institute (2022)
	Placing quality gates before deployment Mayer et al. (2024)	Warranting model security Ernst & Young (2024)
		AI access control Perplexity (2024)
		Prompt engineering International Association of Privacy Professionals (2024)
	LIME algorithm De Laat (2021)	Safety tuning Google (2025)
		Guardrails for redline crossing OpenAI (2023)
		Model cards Google (2025); Boston Consulting Group (2022); International Association of Privacy Professionals (2024)

Table 1. Excerpt of AI Governance approaches in Academic and Grey Literature.

5 Discussion

5.1 Key findings

This study contrasted academic literature on AI governance with grey literature, with a particular focus on AI data, AI algorithms, and AI models. Drawing on a systematic literature review of academic literature with deliberate incorporation of grey literature, our findings show three differences in how these two bodies of literature approach AI governance implementation. First, the academic literature tends to ground governance in normative principles such as privacy, transparency, and accountability, whereas the grey literature tends to ground it in policies and operational concerns without articulating the underlying normative logic. Second, the academic literature articulates governance implementation through principled mechanisms, while the grey literature articulates it through organizational embeddings that are operationally important but normatively overlooked. Third, academic literature focuses on pre-deployment governance activities, including data preparation, model training, and pre-deployment approval, whereas grey literature focuses on post-deployment governance activities, such as monitoring and maintenance. In reflection of these findings, the differences do not merely signify the contrast in the framing or tonality, rather, such differences structurally complement each other - one corpus addressing dimensions of AI governance that the other leaves out. This complementarity leads towards a complete AI governance implementation in organizations.

5.2 Theoretical contributions

This study advances IS scholarship on AI governance (Batool et al., 2025; Birkstedt et al., 2023; Mäntymäki et al., 2022; Papagiannidis et al., 2022) in three ways. First, we have elaborated on how the academic literature foregrounds normative principles and grey literature foregrounds operational policies – sheds light on the work of Mäntymäki et al. (2022) and Birkstedt et al. (2023). Both studies highlighted the research gap: translation from (AI ethics) principles to practice and governance operationalization, respectively. Our research expands on this by framing the challenge not merely as a gap in the academic literature, but as a structural separation between the two epistemic communities. Academic literature produces principle-driven normative knowledge without operational foresight, while grey literature produces operational knowledge that is practical without any normative anchoring. Neither of these literatures addresses this gap independently.

Furthermore, our study directly responds to Batool et al. (2025), who proposed extending their governance analysis to grey literature, and Mäntymäki et al. (2022), who suggested a synthesis of both academic and grey literature. Thus, we contribute to the current IS scholarship by incorporating and synthesizing governance from both academic and grey literature. We argue that the differences in governance emphases between the two AI governance literatures are structural features of the knowledge production of the two corpora. Without incorporating grey literature, such observations would not have surfaced.

Second, we describe how the academic literature foregrounds mechanisms as an implementation approach to achieving governance objectives, whereas the grey literature more often foregrounds the organizational embedding of governance in operational practice. We argue that there is a logical difference between the literatures: the academic literature embodies a design logic that foregrounds governance mechanisms, whereas the grey literature employs an operational logic. Although prior literature identified the need for tools and frameworks (Mäntymäki et al., 2022) and an agenda for integrating AI into software development (Birkstedt et al., 2023), reviews did not include practitioner settings and thus excluded grey literature. Our study invokes this gap by including grey literature in the discussion. We support existing literature by identifying governance-enabling mechanisms (tools in prior literature) and extend them by highlighting a gap: principled governance mechanisms are not operationalized institutionally in the academic literature as they are in the grey literature. We view this gap as a situational or organizational challenge: underutilization of latent assets (governance mechanisms) across organizations to operationalize governance.

Third, by discussing how academic literature focuses on pre-deployment governance, whereas grey literature focuses on post-deployment governance, we address the current literature gap related to the coverage of the AI system. The academic literature emphasizes preventive actions before release, including data preparation, model training, and deployment approval. In contrast, grey literature discusses corrective actions through monitoring and maintenance activities in operations. Thus, together, the academic and grey literature cover a complete AI system lifecycle. With this finding, our study directly responds to the prior calls for complete lifecycle coverage in AI governance (Mäntymäki et al., 2022; Batool et al., 2025). Our contribution to the IS literature in this regard is that the different lifecycle focus between the two literatures is a systematic property of how each attends to governance. Based on this insight, we conclude that academic and grey literature are complementary in their temporal orientation.

Synthesizing our findings, we advance IS literature by proposing the structural complementarity thesis: academic and grey literature are not competing bodies; they are structurally interdependent. Academic literature provides normative foundations and principled mechanisms at the pre-deployment stage that grey literature often lacks. Grey literature provides operational policies and mechanisms at the post-deployment stage that governance mechanisms in the academic literature require to function. Thus, individually, they are insufficient in producing AI governance knowledge. We show, with evidence, that it is therefore value-adding to integrate grey literature into the AI governance literature review. We further show at which points (principles, policies, implementation approaches, and lifecycle stages) integration is required.

5.3 Practical implications

For AI governance practitioners, this study provides three critical implications. First, organizations formulating their AI governance framework should not exclusively rely on academic or grey literature. As we demonstrated, each literature is structurally different and complementary to the others. Therefore, a governance framework that starts with normative principles from the academic literature and maps them onto operational practices in the grey literature will likely be more robust. Second, the underdeveloped normative principles in grey literature may pose an organizational risk. Without principled foundations, organizations may struggle to justify their operational practices, leading to compliance issues. Hence, organizations should explicitly develop normative principles and connect them to concrete operational practices. Finally, our observation of differences in the AI lifecycle may suggest that many organizations are focusing on either pre-deployment or post-deployment governance practices. Such a design may create a design risk: emerging risks appearing at ungoverned lifecycle stages. A complete lifecycle coverage of governance would likely resolve this issue.

5.4 Limitations and future research

This study has several limitations. First, the broad keyword “AI governance” may have excluded other, more specific thematic sources. Second, the timeframe (January 2022 - March 2025) limits the identification of longitudinal insights and the evolution of practices. Third, grey literature selection was more subjective and selective than that of academic literature, excluding preprints and organizations with a niche focus. Finally, the focus on AI data, algorithms, and models leaves out several other potentially relevant areas of AI governance for future studies to address.

References

Achmadiah, M. N., Ahamad, A., Sun, C., & Kuo, W. (2025). Energy-Efficient fast object detection on edge devices for IoT systems. IEEE Internet of Things Journal, 1. https://doi.org/10.1109/jiot.2025.3536526

Accenture. (2021). Responsible AI from principles to practice. https://www.accenture.com/content/dam/accenture/final/a-com-migration/pdf/pdf-149/accenture-responsible-ai-final.pdf

Ada Lovelace Institute. (2023). Safe before sale: Learnings from the FDA's model of Life Sciences Oversight for foundation models. https://www.adalovelaceinstitute.org/wp-content/uploads/2023/12/2023_12_ALI_Safe-before-sale_Discussion_paper.pdf

Anthropic. (2024). Responsible scaling policy. https://assets.anthropic.com/m/24a47b00f10301cd/original/Anthropic-Responsible-Scaling-Policy-2024-10-15.pdf

Bain & Company. (n.d.). Responsible AI. Retrieved March 26, 2025. https://www.bain.com/about/responsible-ai/

Barn, B. S. (2019). Mapping the public debate on ethical concerns: algorithms in mainstream media. Journal of Information Communication and Ethics in Society, 18(1), 124–139. https://doi.org/10.1108/jices-04-2019-0039

Batool, A., Zowghi, D., & Bano, M. (2025). AI governance: a systematic literature review. AI And Ethics. https://doi.org/10.1007/s43681-024-00653-w

Benzies, K. M., Premji, S., Hayden, K. A., & Serrett, K. (2006). State-of-the-Evidence Reviews: Advantages and Challenges of including grey literature. Worldviews on Evidence-Based Nursing, 3(2), 55–61. https://doi.org/10.1111/j.1741-6787.2006.00051.x

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. MIS Quarterly, 45(3), 1433-1450. https://doi.org/10.25300/MISQ/2021/16274

Birkstedt, T., Minkkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: themes, knowledge gaps and future agendas. Internet Research, 33(7), 133–167. https://doi.org/10.1108/intr-01-2022-0042

Boston Consulting Group. (2022). AI Code of Conduct. https://media-publications.bcg.com/AI-Code-of-Conduct.pdf

Bu, Q. (2021). The global governance on automated facial recognition (AFR): ethical and legal opportunities and privacy challenges. International Cybersecurity Law Review, 2(1), 113–145. https://doi.org/10.1365/s43439-021-00022-x

Buhmann, A., & Fieseler, C. (2021). Towards a deliberative framework for responsible innovation in artificial intelligence. Technology in Society, 64, 101475. https://doi.org/10.1016/j.techsoc.2020.101475

De Laat, P. B. (2021). Companies Committed to Responsible AI: From Principles towards Implementation and Regulation? Philosophy & Technology, 34(4), 1135–1193. https://doi.org/10.1007/s13347-021-00474-3

Doneda, D., & Almeida, V. A. (2016). What is algorithm governance? IEEE Internet Computing, 20(4), 60–63. https://doi.org/10.1109/mic.2016.79

Ernst & Young. (2024). Responsible AI. https://www.ey.com/en_us/services/ai/responsible-ai

European Telecommunications Standards Institute. (2022). ETSI activities in the field of artificial intelligence preparing the implementation of the European AI Act (No. 979108262073). https://www.etsi.org/images/files/ETSIWhitePapers/ETSI-WP52-ETSI-activities-in-the-field-of-AI.pdf

Ferrari, F., van Dijck, J., & van den Bosch, A. (2023). Observe, inspect, modify: Three conditions for generative AI governance. New Media & Society. https://doi.org/10.1177/14614448231214811

Foote, D. A., Seipel, S. J., Johnson, N. B., & Duffy, M. K. (2005). Employee commitment and organizational policies. Management Decision, 43(2), 203–219. https://doi.org/10.1108/00251740510581920

G7. (2021). AI principles and code of conduct. https://www.ey.com/en_gl/insights/ai/g7-ai-principles-and-code-of-conduct

Gengler, E. & Schmalenbach, K. (2024) "Exploring Organizational AI Governance Maturity: A Case Study with Six Multinational Companies". ICIS 2024 Proceedings. 2. https://aisel.aisnet.org/icis2024/gov_strategy/gov_strategy/2

Google. (2025). Responsible AI Progress Report. https://ai.google/static/documents/ai-responsibility-update-published-february-2025.pdf

Godin, K., Stapleton, J., Kirkpatrick, S. I., Hanning, R. M., & Leatherdale, S. T. (2015). Applying systematic review search methods to the grey literature: a case study examining guidelines for school-based breakfast programs in Canada. Systematic Reviews, 4(1). https://doi.org/10.1186/s13643-015-0125-0

Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. Minds and Machines, 30(1), 99-120. https://doi.org/10.1007/s11023-020-09517-8

Hartog, P. B. R., Krüger, F., Genheden, S., & Tetko, I. V. (2024). Using test-time augmentation to investigate explainable AI: inconsistencies between method, model and human intuition. Journal of Cheminformatics, 16(1). https://doi.org/10.1186/s13321-024-00824-1

Higgins, J.P.T. & Green, S. (2011) Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. The Cochrane Collaboration. http://handbook-5-1.cochrane.org

International Association of Privacy Professionals. (2024). AI Governance in Practice report. https://iapp.org/media/pdf/resource_center/ai_governance_in_practice_report_2024.pdf

International Telecommunication Union. (2024). AI Governance Day – From Principles to Implementation 2024 report (No. 978-92-61-39291–8). https://s41721.pcdn.co/wp-content/uploads/2021/06/2401225_AI_Governance_Day_2024_Report-E.pdf

Jafari, A. A., Ozcinar, C., & Anbarjafari, G. (2026). Responsible AI: The good, the bad, the AI. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2601.21095

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), 101493. https://doi.org/10.1016/j.giq.2020.101493

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399. https://doi.org/10.1038/s42256-019-0088-2

Khoo, S., Foo, J., & Lee, R. K. (2025). With Great Capabilities Come Great Responsibilities: Introducing the Agentic Risk & Capability Framework for Governing Agentic AI Systems. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2512.22211

KPMG. (2026). Trusted AI governance: Elevating business and leadership outcomes through the responsible governance of AI. https://assets.kpmg.com/content/dam/kpmg/au/pdf/2024/kpmg-trusted-ai-governance-2024.pdf

Li, X., & Bi, S. (2024). Optimal AI model splitting and resource allocation for Device-Edge Co-Inference in Multi-User wireless sensing systems. IEEE Transactions on Wireless Communications, 23(9), 11094–11108. https://doi.org/10.1109/twc.2024.3378418

Li, X., Lv, Z., Wang, S., Wei, Z., & Wu, L. (2019). A reinforcement learning model based on temporal difference algorithm. IEEE Access, 7, 121922–121930. https://doi.org/10.1109/access.2019.2938240

Lier, S. K., Gerlach, J., and Breitner, M. H. (2023) "Who needs XAI in the Energy Sector? A Framework to Upgrade Black Box Explainability". ICIS 2023 Proceedings. 19. https://aisel.aisnet.org/icis2023/aiinbus/aiinbus/19

Liu, N. and Kirshner, S. (2024) "The Futures Too Bright: ChatGPT's Optimism Forecasting Bias".

Loefflad, C., Chen, M., & Grossklags, J. (2023). Factors Influencing Perceived Legitimacy of Social Scoring Systems: Subjective Privacy Harms and the Moderating Role of Transparency [Review of Factors Influencing Perceived Legitimacy of Social Scoring Systems: Subjective Privacy Harms and the Moderating Role of Transparency]. In ICIS 2023 Proceedings. 13.

Mahood, Q., Van Eerd, D., & Irvin, E. (2013). Searching for grey literature for systematic reviews: challenges and benefits. Research Synthesis Methods, 5(3), 221–234. https://doi.org/10.1002/jrsm.1106

Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Putting AI ethics into practice: the hourglass model of organizational AI governance. http://arxiv.org/abs/2206.00335

Mäntymäki, M., Minkkinen, M., Zimmer, M. P., Birkstedt, T., and Viljanen, M. (2023) "Designing an AI governance framework: From research-based premises to meta-requirements". ECIS 2023 Research Papers. 295. https://aisel.aisnet.org/ecis2023_rp/295

Mayer, V.; Schüll, M.; Aktürk, O.; and Guggenberger, T. (2024) "Designing Human-AI Hybrids: Challenges and Good Practices from a Multiple Case Study". ICIS 2024 Proceedings. 3. https://aisel.aisnet.org/icis2024/aiinbus/aiinbus/3

McKinsey & Company. (n.d.). Responsible AI principles. https://www.mckinsey.com/capabilities/quantumblack/how-we-help-clients/generative-ai/responsible-ai-principles

Microsoft. (2022). Microsoft Responsible AI Standard: General requirements. https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-brand/documents/Microsoft-Responsible-AI-Standard-General-Requirements.pdf?culture=en-us&country=us

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. Nature machine intelligence, 1(11), 501-507.

Mora-Cantallops, M., Sánchez-Alonso, S., García-Barriocanal, E., & Sicilia, M. (2021). Traceability for Trustworthy AI: A review of Models and tools. Big Data and Cognitive Computing, 5(2), 20. https://doi.org/10.3390/bdcc5020020

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2019). 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(4), 2141–2168. https://doi.org/10.1007/s11948-019-00165-5

Nishant, R., Schneckenberg, D., & Ravishankar, M. (2023). The formal rationality of artificial intelligence-based algorithms and the problem of bias. Journal of Information Technology, 39(1), 19–40. https://doi.org/10.1177/02683962231176842

OECD. (2024). AI principles. https://www.oecd.org/en/topics/sub-issues/ai-principles.html

OpenAI. (2023). Moving AI governance forward. https://openai.com/index/moving-ai-governance-forward/

Paez, A. (2017). Gray literature: An important resource in systematic reviews. Journal of Evidence-Based Medicine, 10(3), 233–240. https://doi.org/10.1111/jebm.12266

Papagiannidis, E., Enholm, I. M., Dremel, C., Mikalef, P., & Krogstie, J. (2022). Toward AI governance: identifying best practices and potential barriers and outcomes. Information Systems Frontiers, 25(1), 123–141. https://doi.org/10.1007/s10796-022-10251-y

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. The Journal of Strategic Information Systems, 34(2), 101885. https://doi.org/10.1016/j.jsis.2024.101885

Pappas, C., & Williams, I. (2011). Grey Literature: Its emerging Importance. Journal of Hospital Librarianship, 11(3), 228–234. https://doi.org/10.1080/15323269.2011.587100

Perplexity. (2024, April 23). Trust perplexity AI. https://trust.perplexity.ai/

Phan, Q., Wu, Y., & Phan, Q. (2023). Enhancing One-Day-Ahead probabilistic solar power forecast with a hybrid Transformer-LUBE model and missing data imputation. IEEE Transactions on Industry Applications, 60(1), 1396–1408. https://doi.org/10.1109/tia.2023.3325798

Rakova, B., Yang, J., Cramer, H., & Chowdhury, R. (2021). Where Responsible AI meets Reality. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW1), 1–23. https://doi.org/10.1145/3449081

Reddy, S., Allan, S., Coghlan, S., & Cooper, P. (2019). A governance model for the application of AI in health care. Journal of the American Medical Informatics Association, 27(3), 491–497. https://doi.org/10.1093/jamia/ocz192

Samtani, S., Zhu, H., Padmanabhan, B., Chai, Y., Chen, H., & Nunamaker, J. F. (2023). Deep Learning for Information Systems Research. Journal of Management Information Systems, 40(1), 271-301. https://doi.org/10.1080/07421222.2023.2172772

Seppälä, A., Birkstedt, T., and Mäntymäki, M. (2021). "From Ethical AI Principles to Governed AI". ICIS 2021 Proceedings. 10. https://aisel.aisnet.org/icis2021/ai_business/ai_business/10

Schaschek, M., and Engel, S., "Measuring Trustworthiness of AI Systems: A Holistic Maturity

Schiff, D. (2020). Out of the laboratory and into the classroom: the future of artificial intelligence in education. AI & Society, 36(1), 331–348. https://doi.org/10.1007/s00146-020-01033-8

Shneiderman, B. (2020). Bridging the gap between ethics and practice. ACM Transactions on Interactive Intelligent Systems, 10(4), 1–31. https://doi.org/10.1145/3419764

Schneider, J., Abraham, R., Meske, C., & Brocke, J. V. (2022). Artificial Intelligence governance for businesses. Information Systems Management, 40(3), 229–249. https://doi.org/10.1080/10580530.2022.2085825

Schuett, J., Dreksler, N., Anderljung, M., McCaffary, D., Heim, L., Bluemke, E., & Garfinkel, B. (2023). Towards best practices in AGI safety and governance. In Centre for the Governance of AI. https://cdn.governance.ai/AGI_Safety_Governance_Practices_GovAIReport.pdf

Song, B. (2021). Social-cultural perspective of fine arts policies using AI model. International Journal of Technology Management, 86(2/3/4), 167. https://doi.org/10.1504/ijtm.2021.118312

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Strobel, G. & Banh, L. (2024) "What Did the Doctor Say? Empowering Patient Comprehension with Generative AI". ECIS 2024 Proceedings. 15. https://aisel.aisnet.org/ecis2024/track18_healthit/track18_healthit/15

Tyndall, J. (2008). How low can you go? Towards a hierarchy of grey literature.

Tanriverdi, H., Akinyemi, J., and Terrence, N. (2023) "Mitigating Bias in Organizational Development and Use of Artificial Intelligence". ICIS 2023 Proceedings. 19. https://aisel.aisnet.org/icis2023/hti/hti/19

The Global Partnership on Artificial Intelligence. (2023). Scaling responsible AI solutions. In Challenges and Opportunities. https://gpai.ai/projects/responsible-ai/RAI05%20-

%20Scaling%20Responsible%20AI%20Solutions%20-%20Challenges%20and%20Opportunities.pdf

Tian, Y., Claudy, M., Xu, D. (J.), and LIAO, S. S. (2023) "Exploring the Role of AI Explanations in Delivering Rejection Messages: A Comparative Analysis of Organizational Justice Perceptions between HR and AI". ICIS 2023 Proceedings. 21. https://aisel.aisnet.org/icis2023/aiinbus/aiinbus/21

Turner, A. M., Liddy, E. D., Bradley, J., & Wheatley, J. A. (2005). Modeling public health interventions for improved access to the gray literature. Journal of the Medical Library Association, 93(4), 487.

Tutun, S., Topuz, K., Tosyali, A., Bhattacherjee, A., & Li, G. (2024). Explainable Artificial Intelligence for Mental Disorder Screening: A Computational Design Science Approach. Journal of Management Information Systems, 41(4), 958–981. https://doi.org/10.1080/07421222.2024.2415771

U.S. National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). https://doi.org/10.6028/nist.ai.100-1

UNESCO. (2023). Ethical Impact assessment: A tool of the recommendation on the ethics of artificial intelligence.https://unesdoc.unesco.org/in/documentViewer.xhtmlv=2.1.196&id=p::usmarcdef_0000386276&file=/in/rest/annotationSVC/DownloadWatermarkedAttachment/attach_import_353afbf2-65da-4cee-a99f-5347218f574d%3F_%3D386276eng.pdf&locale=en&multi=true&ark=/ark:/48223/pf0000386276/PDF/386276eng.pdf#917_23%20E%20Ethical%20Impact%20Assessment.indd%3A.58819%3A183

UN Systems White Paper on AI Governance (CEB/2024/1/Add.1). (2020). United Nations. https://unsceb.org/sites/default/files/2024-11/UNSystemWhitePaperAIGovernance.pdf

Vakkuri, V., Kemell, K., Jantunen, M., Halme, E., & Abrahamsson, P. (2021). ECCOLA — A method for implementing ethically aligned AI systems. Journal of Systems and Software, 182, 111067. https://doi.org/10.1016/j.jss.2021.111067

Van Giffen, B., & Ludwig, H. (2023). How Boards of Directors Govern Artificial Intelligence. MISQ Executive, 22(4), 7.

Vassilakopoulou, P., Parmiggiani, E., Shollo, A., and Grisot, M. (2022) "Responsible AI: Concepts, critical perspectives and an Information Systems research agenda," Scandinavian Journal of Information Systems: Vol. 34: Iss. 2, Article 3. Available at: https://aisel.aisnet.org/sjis/vol34/iss2/3 *and tools* (2nd ed.). McGraw-Hill Higher Education.

Vial, G., Cameron, A., Giannelia, T., & Jiang, J. (2022). Managing artificial intelligence projects: Key insights from an AI consulting firm. Information Systems Journal, 33(3), 669–691. https://doi.org/10.1111/isj.12420

Wang, X., & Xie, F. (2025). Global artificial intelligence governance research in the digital and intelligent era: advances, trends and countermeasures. *Journal of Knowledge Management*, *30*(1), 30–68. https://doi.org/10.1108/jkm-01-2025-0006

Zhang, F., Pan, Z., & Lu, Y. (2022). AIoT-enabled smart surveillance for personal data digitalization: Contextual personalization-privacy paradox in smart home. Information & Management, 60(2), 103736. https://doi.org/10.1016/j.im.2022.103736

Zhdanov, D., Bhattacharjee, S., & Bragin, M. A. (2021). Incorporating FAT and privacy aware AI modeling approaches into business decision making frameworks. Decision Support Systems, 155, 113715. https://doi.org/10.1016/j.dss.2021.113715

Zimmer, M. P., Minkkinen, M., & Mäntymäki, M. (2022). Responsible Artificial Intelligence Systems: Critical considerations for business model design. Scandinavian Journal of Information Systems, 34(2). https://aisel.aisnet.org/sjis/vol34/iss2/4

Novelli, C., Hacker, P., Morley, J., Trondal, J., & Floridi, L. (2025). A Robust Governance for the AI Act: AI Office, AI Board, Scientific Panel, and National Authorities. European Journal of Risk Regulation, 16(2), 566–590. doi:10.1017/err.2024.57