9-22-2023

THE INSTITUTIONAL LOGICS UNDERPINNING ORGANIZATIONAL AI GOVERNANCE PRACTICES

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

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

THE INSTITUTIONAL LOGICS UNDERPINNING ORGANIZATIONAL AI GOVERNANCE PRACTICES

Research paper

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

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

Abstract

Recent developments in artificial intelligence (AI) promise significant benefits but also invoke novel risks and harms to individuals, organizations, and societies. The rising role of AI necessitates effective AI governance. However, translating AI ethics principles into governance practices remains challenging. Our paper recasts the "AI ethics translation problem" from a unidirectional translation process to a bidirectional interaction between multiple institutional logics and organizational AI governance practices. We conduct a theory adaptation study using the AI governance translation problem as a domain theory and institutional logics and institutional pluralism as method theories. Using this framework, we synthesize key AI governance practices from the literature and outline four central institutional logics: AI ethics principlism, managerial rationalism, IT professionalism, and regulatory oversight. The institutional logics and AI governance practices reciprocally influence one another: logics justify practices, and practices enact logics. We provide an illustrative analysis of the ChatGPT chatbot to demonstrate the framework. For future research, our conceptual study lays a framework for studying how plural institutional logics drive AI governance practices and how practices can be used to negotiate conflicting and complementary institutional logics.

Keywords: AI, AI governance, Institutional logics, IT governance.

1 Introduction

Recent developments in artificial intelligence (AI), fueled by growing quantities of data and increasingly sophisticated processing algorithms, promise efficiency benefits across different sectors but also invoke novel risks and potential harms to individuals and societies (Butcher & Beridze, 2019; Jobin et al., 2019; Mäntymäki et al., 2022b). AI can be defined as the "frontier of computational advancements that references human intelligence in addressing ever more complex decision-making problems" (Berente et al., 2021). PricewaterhouseCoopers anticipates that AI could contribute 15.7 trillion dollars to the global economy by 2030 (cited in Strümke et al., 2022). At the same time, algorithmic systems may enable opaque discriminatory practices against minority groups (Keymolen, 2023), employees may face a loss of jobs due to increasing automation (Palladino, 2022), and there are concerns over safety incidents in AI systems (Falco et al., 2021). Due to the significant risks attached to AI technologies, AI regulation is presently drafted in the European Union and elsewhere (Stix, 2022), and AI has become a matter of organizational governance (Mäntymäki et al., 2022a; Schneider et al., 2022), distinct from previous IT governance approaches (Brown & Grant, 2005; Tiwana et al., 2013).

Governance of AI, defined as the rules, practices, processes, and tools to ensure alignment of AI systems and external requirements (Mäntymäki et al., 2022a), exhibits a paradox similar to the well-known privacy paradox, where individuals hold privacy in high esteem but do little to protect it in practice (Barth & de Jong, 2017). AI governance seems to be simultaneously a crucial endeavor and relatively unimportant in the day-to-day operations of organizations. While the necessity of AI governance is acknowledged to ensure the appropriate functioning and ethical and regulatory safeguards for AI systems,

empirical research on organizational AI governance indicates that organizations still devote little attention specifically to AI governance (Ibáñez & Olmeda, 2022; Papagiannidis et al., 2023; Stahl et al., 2022).

In current research, this problem area is conceptualized as the "translation problem" or "principles-to-practices gap" in AI ethics (Mittelstadt, 2019; Morley et al., 2020; Schiff et al., 2021). This refers to the phenomenon where there are numerous sets of AI ethics principles (Jobin et al., 2019), but deducing "concrete technological implementations from the very abstract ethical values and principles" remains a major challenge (Hagendorff, 2020). While there are tools for ethical AI, most require more work to be production-ready (Morley et al., 2020), and there are few guarantees that they cover the scope of the ethics principles such as fairness and explainability. Moreover, given the multi-actor nature of AI governance activities (Minkkinen et al., 2023), the translation problem of AI ethics indicates contradictions that reach beyond organizations into their surrounding institutional forces (Alford & Friedland, 1991).

From these starting points, this study investigates the following research question:

How can the relationship between organizations' AI governance requirements and AI governance practices be conceptualized?

We contribute to three scholarly domains. First, to the information systems (IS) literature (Brown & Grant, 2005; Berente et al., 2021; Ågerfalk, 2020), we highlight AI governance that addresses the characteristics of AI artifacts and their ethical risks. Second, we contribute to the literature on operationalizing AI ethics in practice (Eitel-Porter, 2021; Ibáñez & Olmeda, 2022; Mittelstadt, 2019; Morley et al., 2020; Seppälä et al., 2021; Stahl et al., 2022) by recasting the problem from a one-way "translation" issue to a bidirectional dynamic between multiple institutional logics and AI governance practices, where logics and practices reciprocally inform each other. Third, we contribute to the literature on multi-actor AI governance (Butcher & Beridze, 2019; Clarke, 2019b; Gasser & Almeida, 2017; Kaminski & Malgieri, 2021; Minkkinen et al., 2023; Shneiderman, 2020) by clarifying the institutional logics stemming from the institutional environment and faced by organizations.

The remainder of the paper proceeds as follows. First, we establish the methodological and theoretical background, positioning the study as a theory adaptation (Jaakkola, 2020) and discussing the domain theory (AI ethics translation) and method theory (institutional logics and institutional pluralism). Then, we present the practices and institutional logics in AI governance. We close with discussing how the institutional logics and AI governance practices inform each other, offering an illustrative analysis of the ChatGPT chatbot and articulating the implications of our study and future research directions.

2   Background

2.1   Theory adaptation

This article presents a *theory adaptation* (Jaakkola, 2020) study to revise the predominant understanding of translating AI ethics into AI governance, using AI ethics translation as the domain theory and institutional logics and institutional pluralism as the method theory (Lukka & Vinnari, 2014). We draw on two theoretical streams informing the theory adaptation: AI governance literature (e.g., Morley et al., 2020; Schiff et al., 2021) and the literature on institutional logics and institutional pluralism (e.g., Ajer et al., 2021; Alford & Friedland, 1991; Kraatz & Block, 2017). Following the principles of theory adaptation (Jaakkola, 2020), we seek to change how AI ethics translation (domain theory) is viewed by using institutional logics and institutional pluralism (method theory) (Lukka & Vinnari, 2014).

Within the institutional logic literature, we focus specifically on the IS literature on institutional logics (Ajer et al., 2021; Berente & Yoo, 2012; Bernardi & Exworthy, 2020; Boonstra et al., 2018; Hansen & Baroody, 2020), because institutional logics research has mushroomed in the management field and because the IS literature is pertinent to AI governance where the governed IT artifact is central.

2.2   Domain theory: Translating AI ethics into practice through AI governance practices

AI ethics is predominantly approached through establishing and discussing guideline documents that outline sets of principles (for overviews, see Thiebes et al., 2021; Hagendorff, 2020; Jobin et al., 2019). Commonly referred principles include fairness, transparency, accountability, non-maleficence, and privacy (Jobin et al., 2019; Dignum, 2020). These principles deal with requirements for AI systems, for example, that they should not discriminate against ethnic groups or genders (fairness) or that the operations of AI systems should be sufficiently visible to users and experts (transparency). Due to the abstract nature of the AI ethics principles, the *translation* of AI ethics principles into AI development and use practices has been the topic of recent scholarly attention (Mittelstadt, 2019; Morley et al., 2020; Li et al., 2023; Morley et al., 2023; Schiff et al., 2021). The AI ethics translation problem refers to how abstract ethical principles, such as fairness, could ensure responsible design, development, and use of AI systems. This entails translating AI ethics to AI governance, understood as the structures, processes, and tools that enable responsible design and use of AI (Mäntymäki et al., 2022a).

Researchers have noted that translating AI ethics principles into AI governance practices has been problematic. Even though software tools exist, few comprehensive solutions are available for organizational use (e.g., Morley et al., 2020; Schiff et al., 2021). Some scholars criticize the principle-based perspective for its high abstraction level and inability to account for the diverse application domains of AI systems (Mittelstadt, 2019; Morley et al., 2023). Others note that principles could nevertheless guide organizations' business processes (Clarke, 2019a) and the professional norms of AI developers (Seger, 2022).

IT governance research (Brown & Grant, 2005; Tiwana et al., 2013; Weill & Ross, 2004) provides one entry point into translating AI ethics into AI governance. IT governance research within the IS field has looked into, for example, IT governance's dimensions (Tiwana et al., 2013), antecedents and consequences (Bradley et al., 2012), business/IT alignment (De Haes & Van Grembergen, 2009), governance archetypes (Weill & Ross, 2004), and contingency factors (Brown & Grant, 2005). In a nutshell, IT governance is about specifying decision rights and accountabilities and ensuring desirable IT use and regulatory compliance (Brown & Grant, 2005; Weill & Ross, 2004). These aspects are also pertinent to AI governance because AI systems are IT systems with particular characteristics (Berente et al., 2021; Mäntymäki et al., 2022a).

However, current AI technologies exhibit features that challenge IT governance frameworks, necessitating AI governance as a separate concern. As incidents of biased and unsafe AI show (Wei & Zhou, 2023), ethical concerns and risks of misuse are more severe in the case of AI. For example, biases against ethnic minorities in facial recognition technologies have been widely critiqued (Raji et al., 2020). The amplified risks stem from the nature of AI systems as IT artifacts. AI systems act increasingly independently from human oversight, they improve by learning from data, and their workings are inscrutable to the wider public and often even to their developers (Berente et al., 2021). Hence, the increasingly agentic nature of AI systems as artifacts (Baird & Maruping, 2021; Ågerfalk, 2020) warrants research attention beyond adapting existing IT governance frameworks that govern less agentic IT artifacts. Because AI ethics principles (Jobin et al., 2019; Dignum, 2020) tackle AI-specific issues, their translation into governance remains a valid starting point for AI governance despite its challenges.

What is relevant to the present paper is that the translation perspective to AI ethics makes certain key assumptions about the domain of responsible AI. First, the principle-based approach requires a near-consensus on high-level principles. This is achieved relatively well at present, even though there are several sets of AI ethics guidelines (Jobin et al., 2019). Nonetheless, tensions between different principles or between principles and other requirements have received little attention (Whittlestone et al., 2019). Second, the translation perspective assumes that the role of ethical principles is to serve as a starting point for translation into practical governance, which is only one possible function of ethical principles that can also inform professional culture more indirectly (Seger, 2022). Third, the translation perspective provides little insight into the nature, foundations, and operating mechanisms of ethics

principles and possible competing considerations. In other words, AI ethics translation lacks a clear theoretical foundation as a phenomenon for IS research and other socio-technical scholarship.

While the debate on translating AI ethics is ongoing, it is clear that new approaches could enable researchers to elaborate particularly on tensions between ethical principles and other relevant considerations. Therefore, this paper presents institutional logics and institutional pluralism as method-theoretical lenses (Jaakkola, 2020; Lukka & Vinnari, 2014) through which the AI ethics translation problem can be revisited and made more theoretically understandable and practically tractable.

2.3   Method theory: Institutional logics and institutional pluralism

The institutional perspective on organizations and IS starts from the premise that organizations are constrained by social structures and forces in the form of several institutions rather than a monolithic "society" (Alford & Friedland, 1991). From an institutional theory perspective, institutional pressure drives organizations to implement AI governance practices. Institutional pressure (Figure 1) can be divided into *strong institutional pressure*, which refers to involuntary compliance with significant punitive consequences from nonconformity, and *weak institutional pressure*, where adoption is voluntary and nonconformity does not carry significant consequences (Berente et al., 2019).
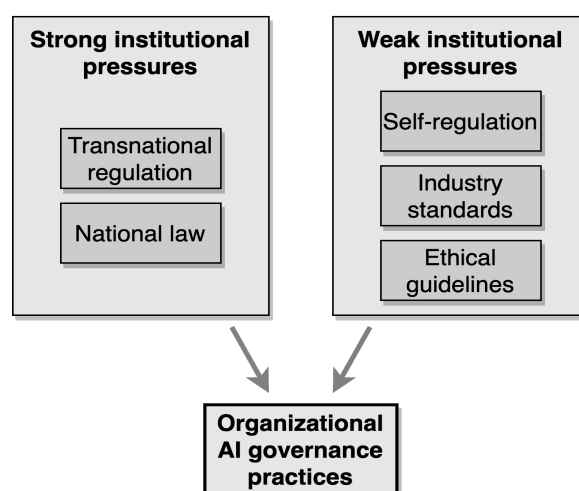
Figure 1. Institutional pressures for AI governance

Institutional logics are collective belief systems that shape actors' cognition and behavior. They are socially constructed sets of principles, practices, beliefs, rules, and systems through which organizations and individuals make sense of their social reality and of appropriate behaviors (Greenwood et al., 2011; Thornton et al., 2012). Institutional logics are described with varying terms as "broader cultural templates" (Pache & Santos, 2010), "socially constructed patterns" leveraged during action-taking (Ajer et al., 2021), and "socially constructed sets of practices, beliefs, rules and systems" (Thornton et al., 2012). The emphasis on individual and collective cognition and the normative binding nature of institutional logics are common features across these conceptualizations.

Institutional pluralism means a situation where organizations face institutional pressures from multiple logics that may complement each other or be in conflict, i.e., incongruent (Ocasio et al., 2017). Berente et al. (2019) conceptualize incongruent institutional logics as a set of logics that, utilized in a situation, cannot guide an actor's practices "without creating a dissonance that calls for fundamentally changing those practices." Incongruent institutional logics, thus, demand an active response from organizational actors to avoid dissonance leading to paralysis. Plural institutional logics help explain seemingly contradictory organizational behaviors because organizational actors "do not simply apply institutional rules, but they navigate and engage institutional orders in their everyday practices in ways that are consistent with particular logics" (Berente & Yoo, 2012). However, in recent work on institutional logics,

researchers emphasize that logics can also complement one another, creating alignment and organizational abilities to function across institutional domains (Hansen & Baroody, 2020).

The closely related concept of *institutional complexity* deals with organizational responses to conflicting demands posed by different institutional logics (Greenwood et al., 2011; Pache & Santos, 2010). A central divide in the literature is between focusing on organizational *strategies* to respond to institutional pluralism and organizational *structures and practices* in response to pluralism (Greenwood et al., 2011; Kraatz & Block, 2017). In this paper, we focus on the latter, practice-based approach to organizations' tackling of institutional pluralism and complexity.

Within the IS field, institutional logics have been used to study the implementation of enterprise systems (Berente et al., 2019), electronic health records (Hansen & Baroody, 2020), and enterprise architecture (Ajer et al., 2021). IS studies on institutional logics tend to fall under three basic types. First, studies have been conducted on the intra-organizational implementation of enterprise systems and enterprise architectures, where a new system is introduced into complex organizational systems with pre-existing institutional logics (Ajer et al., 2021; Berente & Yoo, 2012; Berente et al., 2019). In this case, the potential conflict between logics is between the new IS's logic and the context-specific existing logics, and the role of IT artifacts is left somewhat implicit. Second, IS researchers have used the institutional logics lens to study how logics, IT affordances, and organizational attention are intertwined in institutional change, particularly in times of crisis (Faik et al., 2020; Oborn et al., 2021). In this set of studies, the institutional logics are broader, referring to logics like the state logic or the family logic, and the role of the artifact is theorized more extensively than in the first set of studies. Third, IS research has examined multi-stakeholder situations in IT innovation and system adoption (Bernardi & Exworthy, 2020; Boonstra et al., 2018; Hansen & Baroody, 2020). In this case, different stakeholder groups are viewed as proponents of different logics, while the logics are context-specific, as in the first group of studies.

Our study aligns most closely with the third set of studies because AI governance is usually not a specific system that is implemented, and the logics involved are context-specific rather than broad societal logics, such as the family logic. However, because we discuss AI governance across different sectors, we bracket out any sector-specific institutional logics, such as healthcare professionalism, from this paper's investigation. IS studies on institutional logics have thus far focused on the healthcare sector (Ajer et al., 2021; Bernardi & Exworthy, 2020; Boonstra et al., 2018; Hansen & Baroody, 2020; Oborn et al., 2021), apart from a few notable exceptions (Berente & Yoo, 2012; Berente et al., 2019). Therefore, sector-specific institutional logics, such as healthcare professionalism, have been central in considering institutional pluralism, unlike our case, which discusses AI governance across sectors.

3   Results

3.1   AI governance practices

The operationalization of ethical AI occurs in organizations that deploy and develop AI systems. Therefore, it is important to focus on AI governance at the level of organizational practices (Mäntymäki et al., 2022a) which give meaning and thematic coherence to activities that may appear trivial by themselves (Smets et al., 2012). Table 1 outlines the organizational AI governance practices in the literature.

Table 1. Organizational AI governance practices

Practice	Sources
AI auditing	Mökander et al., 2021; Minkkinen et al., 2022a
Competence and knowledge development	Seppälä et al., 2021
Corporate sustainability reporting	Minkkinen et al., 2022b; Sætra, 2023

Data governance and data management	Schneider et al., 2022; Seppälä et al., 2021; Stahl et al., 2022
Explainability and transparency practices	Brundage et al., 2020; Laato et al., 2022b; Meske et al., 2022
Impact assessment	Kaminski & Malgieri, 2021; Metcalf et al., 2021
Organizational policies and ethics guidelines	Schneider et al., 2022; Seppälä et al., 2021
Regulatory compliance	Schneider et al., 2022
Risk management	Stahl et al., 2022; Tournas & Bowman, 2021
Software engineering workflows, AI design and development	Laato et al., 2022a; Shneiderman, 2020; Seppälä et al., 2021; Stahl et al., 2022
Stakeholder collaboration	Schneider et al., 2022; Seppälä et al., 2021; Zhu et al., 2021
Standardization	Cihon, 2019; Stahl et al., 2022
Validation, testing, and verification	Brundage et al., 2020; Yeung et al., 2020; Zhu et al., 2021

The literature shows that AI governance practices range from software engineering to regulatory expertise. Some practices, such as explainability and transparency, are related to one ethical principle, but generally, there is no clear relationship between AI ethics principles and governance practices.

3.2 Institutional logics in AI governance

In Table 2, we present four institutional logics in AI governance and their dimensions. Institutional logics are usually divided into dimensions. In the IS literature, often mentioned dimensions include principles (Ajer et al., 2021; Berente & Yoo, 2012; Berente et al., 2019), assumptions (Ajer et al., 2021; Berente & Yoo, 2012; Berente et al., 2019; Hansen & Baroody, 2020), identity (Berente & Yoo, 2012; Berente et al., 2019; Bernardi & Exworthy, 2020; Boonstra et al., 2018; Hansen & Baroody, 2020), and domain (Berente & Yoo, 2012; Berente et al., 2019; Hansen & Baroody, 2020), drawing primarily on Thornton and Ocasio (2008). Other dimensions include sources of legitimacy (Ajer et al., 2021; Bernardi & Exworthy, 2020; Boonstra et al., 2018; Hansen & Baroody, 2020), sources of authority (Bernardi & Exworthy, 2020; Boonstra et al., 2018; Hansen & Baroody, 2020), and basis of attention (Bernardi & Exworthy, 2020; Boonstra et al., 2018). In this paper, for the sake of parsimony, we adopt four commonly used dimensions: principles, assumptions, identity, and sources of legitimacy.

Table 2. Four institutional logics in AI governance

Dimensions	Logic of AI ethics principlism	Logic of managerial rationalism	Logic of IT professionalism	Logic of regulatory oversight
Principles	Adherence to high-level principles and guidelines (fairness, non-maleficence, accountability, and privacy) (Jobin et al., 2019; Seger, 2022). Embedding principles in practices of AI development and	Accountability and control (Berente & Yoo, 2012). Competition, efficiency, cost control, continuous improvement (Hansen & Baroody, 2020).	Data usability, data management, technical effectiveness, user participation (Hansen & Baroody, 2020).	Compliance, cost control, efficiency, standardization, continuous improvement (Hansen & Baroody, 2020)

	professional culture (Mittelstadt, 2019; Seger, 2022).			
Assumptions	Normative consensus provided by abstract principles (Gasser & Almeida, 2017; Resseguier & Rodrigues, 2021).	Accountability and control through standardization and visibility (Berente & Yoo, 2012). Financial incentives drive behavior (Hansen & Baroody, 2020). Benchmarking supports the improvement of outcomes (Hansen & Baroody, 2020).	Information integration improves outcomes, IT for IT's sake will fail (Hansen & Baroody, 2020).	Incentives and penalties guide behavior (Hansen & Baroody, 2020).
Identity	Ethics experts discern the relevant issues and provide a foundation for operationalization (cf. Morley et al., 2020).	Standardized structure implies rational bureaucracy and objective criteria for resource allocation (Berente & Yoo, 2012).	IT offers transparency and precision in a standardized way (Boonstra et al., 2018). Information brokers, vendors (Hansen & Baroody, 2020)	Rules empower regulatory actors such as regulators and administrators (Hansen & Baroody, 2020).
Sources of legitimacy	Expertise in AI technologies and AI ethics issues (Palladino, 2022), ethics guidelines (Jobin et al., 2019).	Managerial roles, performance management (Bernardi & Exworthy, 2020). Data evidence, objective data (Hansen & Baroody, 2020). Financial outcomes, profitability, and survival (Hansen & Baroody, 2020).	Education, rational standards based on a technical worldview, systems sciences (Boonstra et al., 2018).	Binding legislative documents. Standardization: establishing generally accepted standards (Hansen & Baroody, 2020)

The dimensions of the AI governance institutional logics are compiled from previous literature on comparable fields. While institutional logics have been discussed in IT governance (Boonstra et al., 2018; Offenbeek et al., 2013), work on institutional logics within AI governance is incipient. In a review of AI governance principles and their operationalization in practical tools, the institutional logic of private companies and the "tech community" were seen to inscribe a narrow technology- and solution-oriented institutional logic to the implementation of AI ethics (Palladino, 2022). Moreover, technical tools and social governance arrangements are deemed to be poorly integrated at present (Palladino, 2022).

AI ethics principlism. AI ethics principlism can be viewed as an institutional logic because one of its primary functions is to influence mindsets, professional culture, and professional norms in AI design and development (Seger, 2022). Principlism is closely related to the deontological approach to applied ethics, that is, adherence to rules and duties (Hagendorff, 2020). There is an abundance of research, particularly conceptual research, on the so-called translation problem or principles-to-practices gap in AI governance (Morley et al., 2020; Palladino, 2022; Schiff et al., 2021). Nevertheless, to the best of our knowledge, AI ethics principlism has not been previously conceptualized as an institutional logic.

Principlism refers to the practice of using ethical principles in tackling moral problems that arise in real-world situations, and it has been particularly discussed in medicine and bioethics (Clouser & Gert, 1990). Principlism in medical ethics embeds core norms into professional practice and helps to identify ethical challenges, guide health policy, and support clinical decision-making (Mittelstadt, 2019). AI ethics principlism refers to a focus on adherence to core principles and ethical guidelines as the preferred approach to dealing with AI ethics problems as well as the embedding of principles in professional practice in AI development (Seger, 2022; Mittelstadt, 2019; cf. Jobin et al., 2019; Schiff et al., 2021).

A principled approach to AI ethics operates on the level of abstract principles by design and neglects specific features of contexts and practices (Resseguier & Rodrigues, 2021). The list of core principles is relatively established, consisting of transparency, fairness, non-maleficence, accountability/responsibility, and privacy, sometimes coupled with less-often-mentioned principles such as human dignity (Jobin et al., 2019; Palladino, 2022). However, a feature of AI ethics principlism institutional logic is the abstract nature and lack of a clear definition of the principles (Palladino, 2022). This implies an assumption that the search for a normative consensus (Gasser & Almeida, 2017) on abstract principles is more important than clear definitions of the content and implications of principles.

The focus on ethical principles as guidance and legitimation means there is high functional indetermination, and roles and accountabilities in AI governance are typically unclear (Radu, 2021). In addition to guidelines, AI ethics principlism primarily uses AI experts and AI ethics experts as a source of legitimacy (Palladino, 2022). In particular, there is a tendency to form consultative bodies, such as AI ethics boards, with loosely defined mandates and opaque criteria for selecting relevant experts (Radu, 2021).

AI ethics principlism is motivated by promoting trust toward new technologies and, thereby, realizing their full potential (Palladino, 2022). The counterpart is the fear of stifling innovation with overly strict rules; hence, the focus is on guiding principles (Radu, 2021). While in the medical field, principlism has a strong institutional backing in curricula and medical practice, as well as accountability mechanisms such as ethics committees, AI ethics principlism is more loosely organized at present (Mittelstadt, 2019).

In the literature on the translation problem, AI ethics principlism is not discussed in "pure form" because it is already implicitly mixed with other logics, such as managerialism and IT development (Morley et al., 2020; Palladino, 2022). For example, the findings of Palladino (2022) can be interpreted as IT professionalism and managerialism subverting AI ethics principles and turning social issues into technologically solvable problems for legitimation (gaining trust) and self-interest (avoiding regulatory fines).

Managerial rationalism. The managerial rationalism logic is omnipresent in organizations, and it emphasizes accountability, control, standardization, and business-like efficiency (Berente & Yoo, 2012). Managerialism also prescribes self-preservation through effective business management (Hansen & Baroody, 2020). Logically, then, attention is turned to financial performance, cost control, revenue generation, and profitability (Hansen & Baroody, 2020). The legitimacy of this logic comes from objective measures of efficiency and profitability.

IT professionalism. The rationality is focused on the instrumentality of IT (Boonstra et al., 2018), where developers design systems that model reality and become useful tools for managers and others to achieve common ends (Boonstra et al., 2018). The logic places emphasis on the role of technical knowledge and on being in control, and it is concerned with using appropriate development methodologies, such as iterative and agile (Boonstra et al., 2018), as well as the quality of technical solutions and information system design (Hansen & Baroody, 2020). There is an overarching focus on achieving high-speed methods and a sufficient pace of development to meet increasing time pressures (O'Connor et al., 2023).

Regulatory oversight. The regulatory oversight logic is comparatively clear-cut because it refers to the effects of more binding rules than the principles discussed under AI ethics principlism. It is thus more coercive in nature than the other logics. Under the regulatory oversight logic, individuals or organizations identify compliance with existing and anticipated laws, governing bodies, and industry practices (Hansen & Baroody, 2020). Informed by this logic, organizations take actions because they believe that they have to and that they are mandated to act in a certain way (Hansen & Baroody, 2020). The regulatory oversight logic is evident in AI governance because new AI legislation is underway, most notably the upcoming AI Act in the European Union (Minkkinen et al., 2023; Stix, 2022).

4   Discussion

4.1   Institutional logics and AI governance practices

Drawing together the themes from the previous sections, Figure 2 outlines the bidirectional interaction between institutional logics and AI governance practices. Rather than a one-way translation from AI ethics principles to governance practices, multiple institutional logics (including AI ethics principlism) and governance practices continuously influence one another. Institutional logics *justify* and symbolically ground governance practices, i.e., logics provide the organizing principle and rationality as well as normative legitimation to the conduct of practices (Alford & Friedland, 1991; Berente & Yoo, 2012). The logics imbue the practices with meaning, making it sensible to devote time and attention to actions related to AI auditing, for example, even when not strictly required by legislation. In the other direction, practices materially *enact* the socio-cognitive institutional logics by deriving the practical implications from the logics, i.e., what activities should be conducted according to the logic (Boonstra et al., 2018; Hansen & Baroody, 2020; Smets et al., 2012). The notion of enactment comes close to the translation of AI ethics principles. However, in the case of institutional logics, AI ethics principlism is only one logic, offset by other logics. Therefore, AI ethics translation is not a one-directional issue but rather a bidirectional and dynamic interaction between institutional logics and AI governance practices.
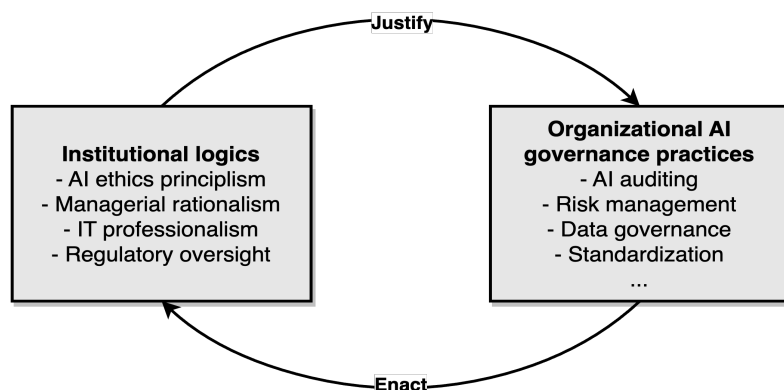
Figure 2. Dynamic interaction between institutional logics and AI governance practices

Figure 2 presents an abstract model where institutional logics, as a single unit, influences a single unit of AI governance practices. In reality, the relationship is not one-to-one but many-to-many. In other words, institutional logics and AI governance practices can be seen to form a matrix with AI governance practices as rows and institutional logics as columns. In an empirical study, the cells of this kind of matrix could be filled with indications of either support (e.g., with "+" symbols) or undermining (e.g., with "-" symbols) for each pair of logic and practice. Then, the rows would show which governance practices are strongly supported by logics (many "+" symbols), undermined by logics (many "-" symbols), or ambiguously supported (both symbols present).

Our conceptual study does not provide evidence of how individual logics support or undermine governance practices. However, some preliminary hypotheses could be made. For instance, AI auditing and impact assessment align with AI ethics principlism, and regulatory compliance aligns with regulatory oversight. Standardization is likely to be supported by managerial rationalism, IT professionalism, and regulatory oversight. On the other hand, auditing and impact assessment may conflict with managerial rationalism because they demand time and resources – ethics costs (Mittelstadt, 2019).

The degree of incongruence between institutional logics can be seen as the driving force of AI governance practices because it provides the central tension that requires organizations to act (to avoid conflicting with one or more logic) but makes action complex due to pluralism. Rather than a situation of "old" logics and a "new" logic introduced by an implemented IS (Berente et al., 2019; Berente &

Yoo, 2012; Ajer et al., 2021), the AI governance case exhibits a "force field" of multiple institutional logics that organizations negotiate by establishing and modifying their AI governance practices.

Practices, then, can mediate the contradictions and complementarities between institutional logics. One way of reconciling conflicting logics is loose coupling. Practices may be loosely coupled with institutional logics, which means that practices are not in conflict with logics, but they only loosely enact the logics, for example, by ceremonially enacting certain principles (Berente & Yoo, 2012). Another way of reconciling competing or complementary institutional logics is "reticulation," that is, the intertwining of practices (motivated by different institutional logics) via activities shared by both practices (Hansen & Baroody, 2020). Reticulation means that a higher-level practice, such as risk management, can have shared lower-level activities with another practice, such as AI auditing (Hansen & Baroody, 2020).

4.2   Illustrative analysis of ChatGPT

Even though our study is conceptual, we present an illustrative analysis of ChatGPT to test the applicability of our findings. ChatGPT is an AI chatbot developed by OpenAI based on a large language model (OpenAI, 2023a). It is able to ask questions and learn from human feedback, making its interaction with users seem more human-like than previous chatbots (Chatterjee & Dethlefs, 2023). ChatGPT and similar applications have been discussed under the term generative AI, meaning technologies that use deep learning to generate human-like content in response to complex and varied prompts (Lim et al., 2023).

ChatGPT is a fruitful case because the force field of congruences and incongruences between institutional logics is readily visible. From the AI ethics principlism perspective, already in 2018, OpenAI published a charter (OpenAI, 2018) that outlines four principles: broadly distributed benefits, long-term safety, technical leadership, and cooperative orientation. The company's product safety standards (OpenAI, 2023b), in turn, list the principles of minimizing harm, building trust, learning and iterating, and being a pioneer in trust and safety. The translation problem is evident here: How to turn these principles into AI governance practices?

For ChatGPT, the AI ethics principlism logic is supported by the regulatory oversight logic, where policy instruments such as the EU's AI Act (Stix, 2022) reinforce the principles of minimizing harm and dealing with safety. In addition, the IT professionalism logic supports AI ethics principlism because both logics espouse standardized practices and clear rules. Thus, AI ethics principlism, regulatory oversight, and IT professionalism all support standardization and sensible risk management practices.

However, the managerial rationalism logic strongly drives ChatGPT as a product. Although OpenAI was initially established as a non-profit, it quickly became a for-profit company. In January 2023, Microsoft invested $10 billion in OpenAI, and the company seeks to integrate ChatGPT into its products, such as Microsoft Office (Spataro, 2023). Managerial rationalism is allied with IT professionalism in seeking technical efficiency, which may facilitate effective AI governance practices or undermine them. AI ethics principlism and regulatory oversight, which espouse time-intensive governance practices, conflict with managerial rationalism and IT professionalism, which favor speed and agility. Hence, the temporal dimension of institutional logics is a central point of incongruence in the ChatGPT case. Indeed, generative AI looks to be a new frontier in the so-called AI race where fast movers reap significant benefits compared to companies with heavy governance practices.

In sum, ethical principles are in place for ChatGPT, but a crucial question is how regulatory oversight and IT professionalism could support governance practices given the intense time pressure expressed by managerial rationalism and IT professionalism. Reticulating activities across different practices (and logics) (Hansen & Baroody, 2020) could decrease the risk of mere ceremonial adherence to principles.

4.3   Implications for theory and practice

In relation to IS literature on IT governance (Brown & Grant, 2005; Tiwana et al., 2013) and AI (Berente et al., 2021; Ågerfalk, 2020), we call for an approach to governance that addresses the characteristics of AI artifacts and the concomitant ethical risks and tensions. To the literature on operationalizing AI ethics in practice (Eitel-Porter, 2021; Ibáñez & Olmeda, 2022; Mittelstadt, 2019; Morley et al., 2020; Seppälä

et al., 2021; Stahl et al., 2022), we recast the unidirectional translation problem into an issue of multiple institutional logics and AI governance practices reciprocally influencing each other. While this makes the issue theoretically complex, it is more true to the continuous justification of AI governance practices and enactment of institutional logics than the AI ethics translation perspective. To the literature on multi-actor AI governance (Butcher & Beridze, 2019; Clarke, 2019b; Gasser & Almeida, 2017; Kaminski & Malgieri, 2021; Minkkinen et al., 2023; Shneiderman, 2020), we enumerate the core institutional logics that drive AI governance. It is widely known that regulation and market pressures influence the governance of technologies. However, conceptualizing these pressures as institutional logics gives additional theoretical tools to discuss institutional pluralism and organizational responses to it.

In terms of practical implications, the results underscore the conflicting demands involved in AI governance from a management perspective. Instead of translating AI ethics principles as non-functional requirements into AI systems and their use, planning and conducting AI governance practices requires the continuous balancing of conflicting and complementary logics.

4.4   Limitations and future research directions

As a conceptual theory adaptation study (Jaakkola, 2020), this study can present preliminary findings to be tested in subsequent empirical studies and further theoretical elaboration. First, cross-sectoral, ethnographic, and longitudinal studies of organizational AI governance are warranted to validate the findings and dig deeper into the role of institutional logics in AI governance practices. Second, future studies could investigate the organizational responses and strategies available to organizations for tackling institutional pluralism (Greenwood et al., 2011; Pache & Santos, 2010). Third, this paper's lists of institutional logics and AI governance practices are not exhaustive. Other logics and practices could be discovered through literature reviews and empirical studies. However, in the interest of theoretical parsimony, as few significant logics as possible should be considered because the level of complexity rises as concepts are added. Thus, adding institutional logics may not yield new insights into how institutional logics and AI governance practices condition one another. Fourth and finally, the appropriate level of theoretical abstraction in studying AI governance remains for future studies to address.

References

Ågerfalk, P. J. (2020). Artificial intelligence as digital agency. *European Journal of Information Systems*, *29*(1), 1-8. https://doi.org/10.1080/0960085X.2020.1721947

Ajer, A. K. S., Hustad, E., & Vassilakopoulou, P. (2021). Enterprise architecture operationalization and institutional pluralism: The case of the Norwegian Hospital sector. *Information Systems Journal*, *31*(4), 610-645. https://doi.org/10.1111/isj.12324

Alford, R. R., & Friedland, R. (1991). Bringing society back in: Symbols, practices, and institutional contradictions. In W. W. Powell & P. J. DiMaggio (Eds.), *The new institutionalism in organisational analysis* (pp. 232-267). University of Chicago Press.

Baird, A., & Maruping, L. M. (2021). The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts. *MIS Quarterly*, *45*(1), 315-341. https://doi.org/10.25300/misq/2021/15882

Barth, S., & de Jong, M. D. T. (2017). The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review. *Telematics and Informatics*, *34*(7), 1038-1058. https://doi.org/10.1016/j.tele.2017.04.013

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

Berente, N., Lyytinen, K., Yoo, Y., & Maurer, C. (2019). Institutional logics and pluralistic responses to enterprise system implementation: a qualitative meta-analysis. *MIS Quarterly*, *43*(3), 873-902. https://doi.org/10.25300/MISQ/2019/14214

Berente, N., & Yoo, Y. (2012). Institutional Contradictions and Loose Coupling: Postimplementation of NASA's Enterprise Information System. *Information Systems Research*, *23*(2), 376-396. https://doi.org/10.1287/isre.1110.0373

Bernardi, R., & Exworthy, M. (2020). Clinical managers' identity at the crossroad of multiple institutional logics in it innovation: The case study of a health care organization in England. *Information Systems Journal*, *30*(3), 566-595. https://doi.org/10.1111/isj.12267

Boonstra, A., Yeliz Eseryel, U., & van Offenbeek, M. A. G. (2018). Stakeholders' enactment of competing logics in IT governance: polarization, compromise or synthesis. *European Journal of Information Systems*, *27*(4), 415-433. https://doi.org/10.1057/s41303-017-0055-0

Bradley, R. V., Byrd, T. A., Pridmore, J. L., Thrasher, E., Pratt, R. M. E., & Mbarika, V. W. A. (2012). An Empirical Examination of Antecedents and Consequences of IT Governance in US Hospitals. *Journal of Information Technology*, *27*(2), 156-177. https://doi.org/10.1057/jit.2012.3

Brown, A. E., & Grant, G. G. (2005). Framing the Frameworks: A Review of IT Governance Research. *Communications of the Association for Information Systems*, *15*. https://doi.org/10.17705/1cais.01538

Brundage, M., Avin, S., Wang, J., Belfield, H., Krueger, G., Hadfield, G., Khlaaf, H., Yang, J., Toner, H., Fong, R., Maharaj, T., Koh, P. W., Hooker, S., Leung, J., Trask, A., Bluemke, E., Lebensold, J., O'Keefe, C., Koren, M., . . . Anderljung, M. (2020). Toward Trustworthy AI Development: Mechanisms for Supporting Verifiable Claims. *arXiv*. http://arxiv.org/abs/2004.07213

Butcher, J., & Beridze, I. (2019). What is the State of Artificial Intelligence Governance Globally. *The RUSI Journal*, *164*(5-6), 88-96. https://doi.org/10.1080/03071847.2019.1694260

Chatterjee, J., & Dethlefs, N. (2023). This new conversational AI model can be your friend, philosopher, and guide. and even your worst enemy. *Patterns (N Y)*, *4*(1), 100676. https://doi.org/10.1016/j.patter.2022.100676

Cihon, P. (2019). Standards for AI Governance: International Standards to Enable Global Coordination in AI Research & Development.

Clarke, R. (2019a). Principles and business processes for responsible AI. *Computer Law & Security Review*, *35*(4), 410-422. https://doi.org/10.1016/j.clsr.2019.04.007

Clarke, R. (2019b). Regulatory alternatives for AI. *Computer Law & Security Review*, *35*(4), 398-409. https://doi.org/10.1016/j.clsr.2019.04.008

Clouser, K. D., & Gert, B. (1990). A critique of principlism. *J Med Philos*, *15*(2), 219-236. https://doi.org/10.1093/jmp/15.2.219

De Haes, S., & Van Grembergen, W. (2009). An Exploratory Study into IT Governance Implementations and its Impact on Business/IT Alignment. *Information Systems Management*, *26*(2), 123-137. https://doi.org/10.1080/10580530902794786

Dignum, V. (2020). Responsibility and artificial intelligence. In M. D. Dubber, F. Pasquale, & S. Das (Eds.), *The Oxford handbook of ethics of AI* (pp. 213-231). Oxford University Press. https://oxfordhandbooks.com/view/10.1093/oxfordhb/9780190067397.001.0001/oxfordhb-9780190067397-e-12

Eitel-Porter, R. (2021). Beyond the promise: implementing ethical AI. *AI and Ethics*, *1*(1), 73-80. https://doi.org/10.1007/s43681-020-00011-6

Faik, I., Barrett, M., & Oborn, E. (2020). How Information Technology Matters in Societal Change: An Affordance-Based Institutional Logics Perspective. *MIS Quarterly*, *44*(3), 1359-1390. https://doi.org/10.25300/MISQ/2020/14193

Falco, G., Shneiderman, B., Badger, J., Carrier, R., Dahbura, A., Danks, D., Eling, M., Goodloe, A., Gupta, J., Hart, C., Jirotka, M., Johnson, H., LaPointe, C., Llorens, A. J., Mackworth, A. K., Maple, C., Pálsson, S. E., Pasquale, F., Winfield, A., . . . Yeong, Z. K. (2021). Governing AI safety through independent audits. *Nature Machine Intelligence*, *3*(7), 566-571. https://doi.org/10.1038/s42256-021-00370-7

Gasser, U., & Almeida, V. A. F. (2017). A Layered Model for AI Governance. *IEEE Internet Computing*, *21*(6), 58-62. https://doi.org/10.1109/mic.2017.4180835

Greenwood, R., Raynard, M., Kodeih, F., Micelotta, E. R., & Lounsbury, M. (2011). Institutional Complexity and Organizational Responses. *Academy of Management Annals*, *5*(1), 317-371. https://doi.org/10.5465/19416520.2011.590299

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

Hansen, S., & Baroody, A. J. (2020). Electronic health records and the logics of care: complementarity and conflict in the U.S. healthcare system. *Information Systems Research*, *31*(1), 57-75. https://doi.org/10.1287/isre.2019.0875

Ibáñez, J. C., & Olmeda, M. V. (2022). Operationalising AI ethics: how are companies bridging the gap between practice and principles? An exploratory study. *AI & SOCIETY*, *37*(4), 1663-1687. https://doi.org/10.1007/s00146-021-01267-0

Jaakkola, E. (2020). Designing conceptual articles: four approaches. *AMS Review*, *10*(1-2), 18-26. https://doi.org/10.1007/s13162-020-00161-0

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

Kaminski, M. E., & Malgieri, G. (2021). Algorithmic impact assessments under the GDPR: producing multi-layered explanations. *International Data Privacy Law*, *11*(2), 125-144. https://doi.org/10.1093/idpl/ipaa020

Keymolen, E. (2023). Trustworthy tech companies: talking the talk or walking the walk. *AI and Ethics*. https://doi.org/10.1007/s43681-022-00254-5

Kraatz, M. S., & Block, E. S. (2017). Institutional pluralism revisited. In R. Greenwood, C. Oliver, T. B. Lawrence, & R. E. Meyer (Eds.), *The Sage handbook of organizational institutionalism*. SAGE Inc.

Laato, S., Birkstedt, T., Mäantymäki, M., Minkkinen, M., & Mikkonen, T. (2022a). *AI governance in the system development life cycle*. In *Proceedings of the 1st International Conference on AI Engineering: Software Engineering for AI*. New York, NY, USA: ACM. http://dx.doi.org/10.1145/3522664.3528598

Laato, S., Tiainen, M., Najmul Islam, A. K. M., & Mäntymäki, M. (2022b). How to explain AI systems to end users: a systematic literature review and research agenda. *Internet Research*, *32*(7), 1-31. https://doi.org/10.1108/intr-08-2021-0600

Li, B., Qi, P., Liu, B., Di, S., Liu, J., Pei, J., Yi, J., & Zhou, B. (2023). Trustworthy AI: From Principles to Practices. *ACM Computing Surveys*, *55*(9), 1-46. https://doi.org/10.1145/3555803

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, *21*(2), 100790. https://doi.org/10.1016/j.ijme.2023.100790

Lukka, K., & Vinnari, E. (2014). Domain theory and method theory in management accounting research. *Accounting, Auditing & Accountability Journal*, *27*(8), 1308-1338. https://doi.org/10.1108/aaaj-03-2013-1265

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

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

Meske, C., Bunde, E., Schneider, J., & Gersch, M. (2022). Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities. *Information Systems Management*, *39*(1), 53-63. https://doi.org/10.1080/10580530.2020.1849465

Metcalf, J., Moss, E., Watkins, E. A., Singh, R., & Elish, M. C. (2021). *Algorithmic Impact Assessments and Accountability*. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 735-746). New York, NY, USA: ACM. http://dx.doi.org/10.1145/3442188.3445935

Minkkinen, M., Laine, J., & Mäntymäki, M. (2022a). Continuous Auditing of Artificial Intelligence: a Conceptualization and Assessment of Tools and Frameworks. *Digital Society*, *1*(3), 21. https://doi.org/10.1007/s44206-022-00022-2

Minkkinen, M., Niukkanen, A., & Mäntymäki, M. (2022b). What about investors? ESG analyses as tools for ethics-based AI auditing. *AI & SOCIETY*. https://doi.org/10.1007/s00146-022-01415-0

Minkkinen, M., Zimmer, M. P., & Mäntymäki, M. (2023). Co-Shaping an Ecosystem for Responsible AI: Five Types of Expectation Work in Response to a Technological Frame. *Information Systems Frontiers*, *25*(1), 103-121. https://doi.org/10.1007/s10796-022-10269-2

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, *1*(11), 501-507. https://doi.org/10.1038/s42256-019-0114-4

Mökander, J., Morley, J., Taddeo, M., & Floridi, L. (2021). Ethics-Based Auditing of Automated Decision-Making Systems: Nature, Scope, and Limitations. *Sci Eng Ethics*, *27*(4), 44. https://doi.org/10.1007/s11948-021-00319-4

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. *Science and Engineering Ethics*, *26*(4), 2141-2168. https://doi.org/10.1007/s11948-019-00165-5

Morley, J., Kinsey, L., Elhalal, A., Garcia, F., Ziosi, M., & Floridi, L. (2023). Operationalising AI ethics: barriers, enablers and next steps. *AI & SOCIETY*, *38*(1), 411-423. https://doi.org/10.1007/s00146-021-01308-8

O'Connor, M., Conboy, K., & Dennehy, D. (2023). Time is of the essence: a systematic literature review of temporality in information systems development research. *Information Technology & People*, *36*(3), 1200-1234. https://doi.org/10.1108/itp-11-2019-0597

Oborn, E., Pilosof, N. P., Hinings, B., & Zimlichman, E. (2021). Institutional logics and innovation in times of crisis: Telemedicine as digital 'PPE'. *Information and Organization*, *31*(1), 100340. https://doi.org/10.1016/j.infoandorg.2021.100340

Ocasio, W., Thornton, P. H., & Lounsbury, M. (2017). Advances to the institutional logics perspective. In R. Greenwood, C. Oliver, T. B. Lawrence, & R. E. Meyer (Eds.), *The Sage handbook of organizational institutionalism* (pp. 509-531). SAGE Inc.

OpenAI. (2023a). *Introducing ChatGPT*. https://openai.com/blog/chatgpt

OpenAI. (2023b). *Product safety standards*. https://openai.com/safety-standards

OpenAI. (2018). *OpenAI Charter*. https://openai.com/charter

Pache, A.-C., & Santos, F. (2010). When Worlds Collide: The Internal Dynamics of Organizational Responses to Conflicting Institutional Demands. *Academy of Management Review*, *35*(3), 455-476. https://doi.org/10.5465/amr.35.3.zok455

Palladino, N. (2022). A 'biased' emerging governance regime for artificial intelligence? How AI ethics get skewed moving from principles to practices. *Telecommunications Policy*, 102479. https://doi.org/10.1016/j.telpol.2022.102479

Papagiannidis, E., Enholm, I. M., Dremel, C., Mikalef, P., & Krogstie, J. (2023). Toward AI Governance: Identifying Best Practices and Potential Barriers and Outcomes. *Information Systems Frontiers*, *25*(1), 123-141. https://doi.org/10.1007/s10796-022-10251-y

Radu, R. (2021). Steering the governance of artificial intelligence: national strategies in perspective. *Policy and Society*, *40*(2), 178-193. https://doi.org/10.1080/14494035.2021.1929728

Raji, I. D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J., & Denton, E. (2020). *Saving Face: Investigating the Ethical Concerns of Facial Recognition Auditing*. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. New York, NY, USA: ACM. http://dx.doi.org/10.1145/3375627.3375820

Resseguier, A., & Rodrigues, R. (2021). Ethics as attention to context: recommendations for the ethics of artificial intelligence. https://open-research-europe.ec.europa.eu/articles/1-27/v1

Sætra, H. S. (2023). The AI ESG protocol: Evaluating and disclosing the environment, social, and governance implications of artificial intelligence capabilities, assets, and activities. *Sustainable Development*, *31*(2), 1027-1037. https://doi.org/10.1002/sd.2438

Schiff, D., Rakova, B., Ayesh, A., Fanti, A., & Lennon, M. (2021). Explaining the Principles to Practices Gap in AI. *IEEE Technology and Society Magazine*, *40*(2), 81-94. https://doi.org/10.1109/mts.2021.3056286

Schneider, J., Abraham, R., Meske, C., & Vom Brocke, J. (2022). Artificial Intelligence Governance For Businesses. *Information Systems Management*, 1-21. https://doi.org/10.1080/10580530.2022.2085825

Seger, E. (2022). In Defence of Principlism in AI Ethics and Governance. *Philosophy & Technology*, *35*(2), 45. https://doi.org/10.1007/s13347-022-00538-y

Seppälä, A., Birkstedt, T., & Mäntymäki, M. (2021). *From ethical AI principles to governed AI*. In *ICIS*. https://aisel.aisnet.org/icis2021/ai_business/ai_business/10

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

Smets, M., Morris, T., & Greenwood, R. (2012). From Practice to Field: A Multilevel Model of Practice-Driven Institutional Change. *Academy of Management Journal*, *55*(4), 877-904. https://doi.org/10.5465/amj.2010.0013

Spataro, J. (2023). *Introducing Microsoft 365 Copilot – your copilot for work*. https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/

Stahl, B. C., Antoniou, J., Ryan, M., Macnish, K., & Jiya, T. (2022). Organisational responses to the ethical issues of artificial intelligence. *AI & SOCIETY*, *37*(1), 23-37. https://doi.org/10.1007/s00146-021-01148-6

Stix, C. (2022). The Ghost of AI Governance Past, Present, and Future. In J. Bullock & V. Hudson (Eds.), *The Oxford Handbook of AI Governance*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780197579329.013.56

Strümke, I., Slavkovik, M., & Madai, V. I. (2022). The social dilemma in artificial intelligence development and why we have to solve it. *AI and Ethics*, *2*(4), 655-665. https://doi.org/10.1007/s43681-021-00120-w

Thiebes, S., Lins, S., & Sunyaev, A. (2021). Trustworthy artificial intelligence. *Electronic Markets*, *31*(2), 447-464. https://doi.org/10.1007/s12525-020-00441-4

Thornton, P. H., & Ocasio, W. (2008). Institutional logics. In R. Greenwood, C. Oliver, R. Suddaby, & K. Sahlin-Andersson (Eds.), *The Sage handbook of organizational institutionalism* (pp. 99-128). Sage.

Thornton, P. H., Ocasio, W., & Lounsbury, M. (2012). *The institutional logics perspective: a new approach to culture, structure, and process*. Oxford University Press.

Tiwana, A., Konsynski, B., & Venkatraman, N. (2013). Special Issue: Information Technology and Organizational Governance: The IT Governance Cube. *Journal of Management Information Systems*, *30*(3), 7-12. https://doi.org/10.2753/MIS0742-1222300301

Tournas, L. N., & Bowman, D. M. (2021). AI Insurance: Risk Management 2.0. *IEEE Technology and Society Magazine*, *40*(4), 52-56. https://doi.org/10.1109/mts.2021.3123750

Wei, M., & Zhou, Z. (2023). *AI Ethics Issues in Real World: Evidence from AI Incident Database*. In *56th Hawaii International Conference on System Sciences*. Maui, Hawaii. https://hdl.handle.net/10125/103236

Weill, P., & Ross, J. W. (2004). *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results*. Harvard Business Press. http://books.google.fi/books?id=xI5KdR21QTAC&hl=&source=gbs_api

Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). *The Role and Limits of Principles in AI Ethics*. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 195-200). New York, NY, USA: ACM. http://dx.doi.org/10.1145/3306618.3314289

Yeung, K., Howes, A., & Pogrebna, G. (2020). AI Governance by Human Rights–Centered Design, Deliberation, and Oversight: An End to Ethics Washing. In M. D. Dubber, F. Pasquale, & S. Das (Eds.), *The Oxford Handbook of Ethics of AI* (pp. 75-106). Oxford University Press. https://oxfordhandbooks.com/view/10.1093/oxfordhb/9780190067397.001.0001/oxfordhb-9780190067397-e-5

Zhu, L., Xu, X., Lu, Q., Governatori, G., & Whittle, J. (2021). AI and Ethics – Operationalising Responsible AI. In F. Chen & J. Zhou (Eds.), *Humanity Driven AI*. https://doi.org/10.1007/978-3-030-72188-6_2