CAMBRIDGE
UNIVERSITY PRESS

DATA FOR POLICY PROCEEDINGS PAPER

Identifying stakeholder motivations in normative AI governance: a systematic literature review for research guidance

Frederic Heymans [ID] and Rob Heyman

imec-SMIT, VUB, Brussels, Belgium
Corresponding author: Frederic Heymans; Email: frederic.heymans@vub.be

Abstract

Ethical guidelines and policy documents destined to guide AI innovations have been heralded as the solution to guard us against harmful effects or to increase public value. However, these guidelines and policy documents face persistent challenges. While these documents are often criticized for their abstraction and disconnection from real-world contexts, it also occurs that stakeholders may influence them for political or strategic reasons. While this last issue is frequently acknowledged, there is seldom a means or a method provided to explore it. To address this gap, the paper employs a combination of social constructivism and science & technology studies perspectives, along with desk research, to investigate whether prior research has examined the influence of stakeholder interests, strategies, or agendas on guidelines and policy documents. The study contributes to the discourse on AI governance by proposing a theoretical framework and methodologies to better analyze this underexplored area, aiming to enhance comprehension of the policymaking process within the rapidly evolving AI landscape. The findings underscore the need for a critical evaluation of the methodologies found and a further exploration of their utility. In addition, the results aim to stimulate ongoing critical debates on this subject.

Policy Significance Statement

This study seeks ways to identify stakeholder motives such as interests, intentions, agendas, and strategies in relation to normative AI policy processes. By combining social constructivism and science & technology studies perspectives with desk research, the study proposes a theoretical framework and methodologies to deepen the understanding of policymaking in the evolving AI landscape. The findings not only offer insights into research disciplines, theoretical frameworks, and analytical methods but also emphasize the need for critical evaluation and further exploration. This contribution aims to stimulate ongoing research and critical debates, providing a starting point for a more profound exploration of policy processes. The study's overview of disciplines, frameworks, and methods invites researchers and policymakers to engage with this under-explored field, which can lead to a more comprehensive understanding of stakeholder motivations in AI governance.

1. Introduction

Artificial intelligence (AI)[1] is experiencing a resurgence, thanks to recent breakthroughs in machine learning and generative AI. This is leading to new applications and services across various fields but also new policy initiatives that seek to manage these advancements. Normative AI governance measures in both soft law, such as ethical guidelines, and hard law (legislation), are generally perceived as essential tools in guiding the development and implementation of AI innovations (Kim et al. 2023; Veale et al., 2023). These policy documents aim to address potential negative consequences and promote societal benefits.[2] Despite their good intentions, their operationalization faces numerous challenges.

The effectiveness of AI governance frameworks is challenged by multifaceted issues, revealing complex stakeholder dynamics. Deficiencies impede their effectiveness, which is beneficial to some industry stakeholders, and prompt them to strategize around delaying the development of clear and practical frameworks (Floridi et al., 2018). Furthermore, the lack of clarity in the inclusion of certain values, coupled with phenomena such as "ethics washing," due to hidden agendas, raises concerns about the legitimacy of ethical considerations of businesses and governments. Claims by Conti and Seele (2023), Munn (2022), and Hagendorf (2020) about potential political motives in policy documents and guidelines have surfaced. However, these claims lack a robust research method to substantiate beyond the guidelines themselves the presence of political motives. This triggers questions of transparency and authenticity in policymaking among various stakeholders. The insufficiently explored nature of the policymaking process, particularly within the context of innovation and technology, further compounds these challenges mentioned above.

A theoretical framework and methodological outlines are needed, aiming to uncover stakeholders' motivations[3] to influence policy. This is undertaken in this article by assembling an eclectic overview, bringing together research disciplines, theoretical frameworks and concepts that can help analyze stakeholder motivations, and analytical methods. The two research questions guiding this study are: (1) "How can we identify social constructivist factors in the shaping of AI governance documents?" and (2) "What research has been conducted to explore and scrutinize stakeholder motivations encompassing interests, strategies, intentions, and agendas?"

Through this inquiry, this research seeks to contribute valuable insights to the ongoing discourse on AI governance, shedding light on stakeholder motivations and enhancing our understanding of the policymaking process in this rapidly evolving field.

This article is structured as follows: Section 2 summarizes the current AI governance efforts and their limitations, demonstrates that the field of AI governance is dominated by and moving further toward self-regulation, and highlights the role that stakeholder dynamics plays in this context. Section 3 describes the bridge between policymaking and social constructivism theory. That account is further complemented by a contextualization of the link between the theory of the social construction of technology (SCOT) and stakeholder motivations. With this theoretical background, we outline how the formulation of AI governance policies is shaped by social constructs. Section 4 explains the research methodology. Section 5 discusses the findings, divided into disciplines, theoretical frameworks, and analytical models. Section 6 presents the findings, the research gaps, and future research directions. The final section presents the concluding remarks.

2. AI governance, limitations, and stakeholder dynamics

The rapid evolution and inherent complexity of AI pose challenges for traditional regulatory approaches (Hadfield and Clark, 2023; Smuha, 2021). This causes the field of AI governance to be currently

---

[1] In this article, we use the OECD's AI definition: "An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment" (OECD, n.d.).

[2] For an overview and discussion of the potential opportunities and negative consequences of AI, see Floridi et al. (2018) and Acemoğlu (2021).

[3] By stakeholder motivations we refer to interests, strategies, and agendas.

dominated by self-regulation (Schiff et al., 2020; De Almeida et al., 2021). Gutiérrez and Marchant (2021) see it as the leading form of AI governance in the near future. These frameworks offer advantages such as a lack of jurisdictional constraints, low barriers to entry, and a propensity for experimentation. Soft law can be easily modified and updated to keep pace with the continuously evolving nature of AI. This flexibility and adaptability allow stakeholders like governments, companies, and civil society organizations to participate in their development (Hadfield and Clark, 2023).

However, limitations often hinder their effectiveness. Documents tend to be overly abstract, theoretical, or inconsistent, making them difficult to operationalize. In addition, they frequently contain an imbalance of conflicting values, such as accuracy and privacy or transparency and nondiscrimination (Petersen, 2021), and disconnect from real-world contexts. For instance, many AI ethics guidelines and policies predominantly address concerns related to algorithmic decision-making, the integrity of fairness, accountability, sustainability, and transparency within corporate decision-making spheres where AI systems are embedded. These spheres are frequently compromised by competitive and speculative norms, ethics washing, corporate secrecy, and other detrimental business practices (Attard-Frost et al., 2022), which receive far less attention. The nonbinding nature of the documents also leads to disregard in some cases, further reducing their efficacy on AI innovations that enter society (Munn, 2022; Ulnicane et al., 2020).

In addition, stakeholder motivations can significantly influence policy outcomes (Valle-Cruz et al., 2020). In traditional policy processes, it is common that there is pressure on policymakers to limit or dilute rules to avoid overly strict legislative frameworks. AI governance differs from traditional approaches due to stakeholder dynamics. Here, actors leverage the unique characteristics of AI, such as its rapid evolution and inherent complexity, to advocate for self-regulation as the preferred approach and delay legislative intervention as much as possible (Butcher and Beridze, 2019; Hagendorff, 2020). In the meantime, companies can benefit from the absence of enforceable mechanisms for compliance in self-regulation frameworks (Papyshev and Yarime, 2022).

Within the context of stakeholder motivations, a distinction can be made between declared or stated interests and strategies, and hidden intentions or agendas. Both aspects have been understudied so far. Concerning public interests and strategies, Ayling and Chapman (2021) argue that meta-analyses of AI ethics papers narrowly focus on the comparison of ethical principles, neglecting stakeholder representation and ownership of the process. In terms of hidden interests and agendas, McConnell (2017) indicates that there is a dearth of analysis on the phenomenon of hidden agendas. These limitations highlight a broader gap in research on stakeholder motivations. While a wealth of analysis exists on specific AI governance documents (Christoforaki and Beyan, 2022), the role of stakeholder motivations within the complex socioeconomic landscape of policymaking remains largely unexplored.

Another reason for the field's importance is that policymakers, in practice, are not solely driven by motives to serve the public interest. Policy design is often viewed from an overly optimistic perspective (Arestis and Kitromilides, 2010). They can also be driven by malicious or venal motivations such as corruption or clientelism, rather than socially beneficial ones (Howlett, 2021). Oehlenschlager and Tranberg (2023) illustrate this with a case study on the role Big Tech has played in influencing policy in Denmark. Valle-Cruz et al. (2023) highlight that AI can have negative impacts on government, such as a lack of understanding of AI outcomes, biases, and errors. This highlights the need for research into political motivations for AI deployment in government and the ethical guidelines to guide implementation.

When translating principles into practice, stakeholder motivations can undermine good intentions, leading to issues such as ethics washing, where organizations pay lip service to ethical principles without actively implementing them in their AI systems and practices (Bietti, 2020; Floridi, 2019). Potential risks also include ethics bashing, where stakeholders intentionally undermine or criticize ethical principles to advance their interests (Bietti, 2020), or ethics shopping (Floridi, 2019), where stakeholders selectively adopt and promote ethical principles that align with their interests.

3. Theoretical background

3.1. Social constructivism

In this section, we define a framework to consider stakeholders in the social construction of AI governance. We first define social constructivism and then visit its application in policy design. In analyzing stakeholder motivations, it is imperative to examine the intricate interplay between social interactions, prevailing norms and values, and the nuances of language, as these elements shape and influence society (Elder-Vass, 2012). The theory is, therefore, a fundamental framework for understanding the mechanisms of policy processes and the role of stakeholders.

Flockhart (2016) identifies four key constructivist propositions:

1. A belief in the social construction of reality and the importance of social facts.
2. A focus on ideational as well as material structures and the importance of norms and rules.
3. A focus on how people's identities influence their political actions and the significance of understanding the reasons behind people's actions.
4. A belief that people and their surroundings shape each other, and a focus on practice (real-life situations) and action.

The propositions of social constructivism hold significant explanatory power within the context of the article's topic. Firstly, the emergence of societal norms, values, and shared beliefs through interactions leads to the formation of a collective reality, with social facts being produced by social practices (Yalçin, 2019). This social construction of reality and social facts influences policy processes, determining priorities, viable solutions, and resource allocation. Secondly, constructivist thought emphasizes the influence of meanings attached to objects and individuals on people's actions and perceptions, highlighting the intertwining of material and ideational structures in shaping societal organization (Adler, 1997). In addition, social constructivists emphasize the role of identity in influencing interests, choices in actions, and engagement in political actions, recognizing identity as essential for interpreting the underlying reasoning behind certain actions (Zehfuss, 2001; Wendt, 1992). Lastly, constructivism underscores the symbiotic relationship between individuals and societal structures, emphasizing the reciprocal influence of individuals and their practices on these structures, thus highlighting the fluid nature of policymaking within specific social and cultural contexts (Zhu et al., 2018).

3.1.1 Social constructivism in policy design

Pierce et al. (2014) refer to *Theories of the Policy Process* (Sabatier, 2019) as a canonical volume defining policy studies at that time. The edited volume excluded the work of constructivists, who were a minority of policy process scholars at that time. Constructivists emphasized the socially constructed nature of policy and reality, highlighting the importance of perceptions and intersubjective meaning-making processes in understanding the policy process. This theoretical framework only found a foothold in the next edition dating from 2007 (Pierce et al., 2014, p. 2). The approach consists of eight assumptions that influence policy design (Figure 1).

Schneider and Ingram's (1993) approach focuses on the analysis of who is allowed to be present as a target group within the design phase of policy documents. This choice is made by the *statutory designers* who are influenced by the factors outlined in Figure 1.

3.1.2 Statutory regulations and statutory designers

Schneider and Ingram (1993) define policy design as the content of public policy as found in the text of policies, the practices through which policies are conveyed, and the subsequent consequences associated with those practices (Pierce et al., 2014). In addition, statutory regulations are laws, rules, procedures, or voluntary guidelines initiated, recommended, mandated, implemented, and enforced by national governments to promote a certain goal (Patiño et al., 2020). Therefore, statutory designers can be understood as individuals or entities responsible for creating and implementing laws, rules, and procedures to shape

---

Model of the Individual

1. Actors cannot process all of the information relevant to make a decision, and therefore rely on mental heuristics to decide what information to retain.
2. Mental heuristics filter information in a biased manner, thereby resulting in a tendency for individuals to confirm new information that is consistent with preexisting beliefs and reject information that is not.
3. People use social constructions in a subjective manner that is evaluative.
4. Social reality is boundedly relative where individuals perceive generalizable patterns of social constructions within objective conditions.

Power

5. Power is not equally distributed among individuals within a political environment.

Political Environment

6. Policy creates future politics that feeds forward to create new policy and politics.
7. Policies send messages to citizens that affect their orientations and participation patterns.
8. Policies are created in an environment of political uncertainty.

Figure 1. *Assumptions of the theory of social construction and policy design.*
Source: *Pierce et al. (2014, p. 5).*

public policy and promote specific societal outcomes. These designers play a crucial role in shaping the content and practices of policies, ultimately influencing the impact of these policies on the target populations (Pierce et al., 2014; Patiño et al., 2020).

Schneider and Ingram's proposition introduces a classification of target populations based on social construction and power. They depict individuals on a gradient from undeserving to deserving on the social construction dimension and from powerful to lacking power on the power dimension. This is visually represented in a 2 × 2 matrix, creating four categories of target populations: advantaged, contenders, dependents, and deviants.

The advantaged are positively constructed and have high power, expected to receive a disproportionate share of benefits and few burdens, while the contenders, despite having high power, are negatively constructed and expected to receive subtle benefits and few burdens. The dependents, with low power but positive construction, are expected to receive rhetorical and underfunded benefits and hidden burdens, and the deviants, with low power and negative construction, are expected to receive limited to no benefits and a disproportionate share of burdens.

3.1.3 Social constructivism and AI governance

Building on the concept of social constructivism influencing policy design, we can identify similar dynamics in shaping AI governance documents. Schneider and Ingram's classification of target populations, introduced in Section 3.1.2, can help shed light on how power and social perception can influence who benefits from and who bears the burdens of AI governance.

Power dynamics and representation. Large tech companies and established research institutions, with their positive image as innovation drivers, hold significant sway in policy formation (Ulnicane et al., 2020). They can be seen as advantaged groups, which have "the resources and capacity to shape their constructions and to combat attempts that would portray them negatively" (Schneider and Ingram, 1993). Conversely, emerging AI start-ups (contenders) occupy a complex space in AI governance. While their potential for innovation and economic growth grants them power, they also face public concerns, such as fear of job displacement due to automation. This creates a challenge for policymakers. The societal benefits of AI advancements developed by start-ups might not be readily apparent to the public. This lack of public recognition, coupled with public concern, makes it difficult to design clear and effective policies that govern AI development by emerging start-ups (Sloane and Zakrzewski, 2022; Winecoff and Watkins, 2022).

Small and medium-sized enterprises (SMEs), looking to integrate AI in their operations, can be considered as dependents. While SMEs are positively constructed as potential drivers of innovation and economic growth, they often have limited power compared to larger corporations because of a lack of resources and expertise. As a result, they may receive rhetorical support and insufficient resources from AI governance efforts (Watney et al., 2021; Kergroach, 2021). For instance, policymakers may express encouragement for SME participation in AI initiatives, but these businesses may struggle to access adequate funding and infrastructure needed to effectively utilize AI technologies. In addition, they may face hidden burdens such as compliance costs and regulatory complexities. Meanwhile, marginalized communities and citizens, despite being potentially harmed by AI, have minimal influence on the design of policy (Donahoe and Metzger, 2019). They are often overlooked by policymakers, and their perspectives may be underrepresented in the final documents. Schneider and Ingram's classification system also shows that unequal distribution of power, one of the core assumptions of social constructivism, is evident within AI governance.

Bounded relativity and interpretation. Statutory designers, as discussed in Section 3.1.2, face a complex and evolving technology in AI governance. This is exemplified by the core challenge of defining AI itself in policy documents. Policymakers struggle with defining AI in a way that is both technically accurate and broad enough to encompass different applications of AI, from simple algorithms to complex deep learning systems (O'Shaughnessy, 2022). Reaching agreement on a definition is a major challenge, but so is finding cross-border consensus (Schmitt, 2021) or aligning different views on liability and responsibility (Zech, 2021), which are other examples of how statutory designers can rely on their ideologies in policy design. This reflects the concept of bounded relativity – stakeholders perceive and interpret the impacts of AI technologies based on their perspectives and experiences. For instance, tech companies may view AI algorithms in hiring processes as efficient tools, while jobseekers and civil rights advocates may see them as potential sources of discrimination and bias. These differing perspectives highlight how social constructions of AI are shaped by objective conditions (the capabilities and limitations of the technology), leading to divergent interpretations of its implications for society.

Dynamic policy environment. Just like any policy, AI policy shapes the social reality and how people understand and interact with the technology (Eynon and Young, 2020). By designing policy, AI governance documents can influence policy elements such as resource allocation and public discourse (Cheng et al., 2021). For example, AI legislation can strictly regulate the use of the technology, prompting policymakers to take additional measures to stimulate innovation. If policymakers focus only on restriction, this could potentially stifle innovation and drive entrepreneurs away to other jurisdictions (De Cooman and Petit, 2020; Scherer, 2015). Policymakers can also send messages to organizations or citizens through AI-related initiatives, such as ensuring trust or security, which can then adjust their orientation and participation accordingly.

Navigating uncertainty. AI governance operates in a dynamic and rapidly evolving environment, characterized by technological advancements, emerging risks, and evolving societal norms. Policymakers must navigate this uncertainty when crafting AI governance documents, balancing the need for regulatory flexibility with the imperative to address potential risks and societal concerns associated with AI technologies (Thierer et al., 2017).

3.2. Stakeholder motivations and the social construction of technology

Building upon the foundations of social constructivism theory, the SCOT theory takes this dynamic interplay a step further by examining how technology itself is a product of social factors, cultural contexts, and the perspectives and actions of different stakeholders (Bijker et al., 1987). SCOT asserts that technology is not merely a neutral tool but a complex entity whose form, function, and impact are linked to the perspectives and interests of various stakeholders. Within this theoretical field, Brück (2006)

conceptualizes technology as the embodiment of individuals' perceptions and conceptions of the world. This theory highlights the dynamic and reciprocal relationship between society and technology.

In light of this theoretical perspective, it is imperative to underscore the profound significance of stakeholders in the shaping of technological constructs and the formulation of policies aimed at integrating these technologies within the fabric of society. Stakeholders, whether they are industry leaders, policymakers, advocacy groups, or the broader public, bring their interests, values, beliefs, and power dynamics into the innovation and policy processes.

This can shape the direction and outcomes of technological innovations. For example, in the development of new medical technology, stakeholders such as medical professionals, pharmaceutical companies, regulatory bodies, patients, and advocacy groups all have a vested interest in the technology's success or failure. Their input and influence can determine factors such as the prioritization of research and development, the allocation of resources, the ethical considerations and guidelines, and the degree of accessibility and affordability of the technology.

AI governance documents are developed to guide the responsible implementation of AI in society. This is a task that involves defining which problems need to be avoided (values) and what methods (means) to use to reach this goal. This process often occurs behind closed doors,[4] leading to a need for more transparency about how problems were prioritized or how certain methods were chosen over others (Aaronson and Zable, 2023; Edgerton et al., 2023; Perry and Uuk, 2019). SCOT suggests that different social groups can interpret and use the same technology in different ways, based on their interests. In the case of AI governance, various stakeholders—such as policymakers, AI developers, end users, and the public—may have different views on what constitutes a "problem" and how it should be addressed. For example, policymakers might prioritize issues related to privacy and security, while AI developers might be more concerned with technical challenges like improving algorithmic fairness. End users, on the other hand, might focus on usability and the impact of AI on their daily lives. The main question thus becomes how ideational factors (worldviews, ideas, collective understandings, norms, values, etc.) impact political action (Saurugger, 2013) and whose perspectives are taken into account or worse, whose perspectives are excluded.

The theory serves as a valuable tool for comprehending the sociology of technology and its derivatives —exploring who communicates what, when, and why—and the underlying dynamics contributing to the pluralism of technology (Ehsan and Riedl, 2022). Metcalfe (1995), for example, indicates that similar dynamics as the ones that were mentioned in the context of social constructivism such as lobbying and hidden agendas, but also imperfect information, bureaucratic capture, and shortsighted politics, may lead to mistaken government interventions. To formulate effective technology policies, policymakers must have access to detailed microeconomic and social information. In addition, the insights from the SCOT theory can be applied to examine the diverse perceptions and practices surrounding AI. Eynon and Young (2020), for example, explored this aspect in the context of lifelong learning policy and AI, examining the perspectives of stakeholders in government, industry, and academia.

3.3. Stakeholder motivations

The significance of motivations to fulfil policy desires in the context of AI and ethics has been acknowledged in the scholarly literature (Cihon et al., 2021; Jobin et al., 2021; Krzywdzinski et al., 2022; Ulnicane et al., 2020). However, existing discussions on the role of motivations, and the power dynamics associated with safeguarding these, often oversimplify the complexity of this topic. Present depictions often simplify the situation into a division between the vested interests of established actors and the emerging concerns of newcomers (Bakker et al., 2014). These distinctions do not encompass the complex intricacies of interests and agency as AI and ethics continue to evolve.

The presence and absence of stakeholders in policymaking is such an aspect, critical for democratic governance. It is commonly agreed that the inclusion of diverse stakeholder groups in policy formulation

---

[4] It should be noted that the EU and other government initiatives do provide transparency with regards to who participated in expert groups and hearings leading to AI policy documents.

can lead to more informed and effective policies. Their participation can enhance the legitimacy of the policymaking process (Garber et al., 2017). However, full consultation of all stakeholders with an interest in a given policy issue is rarely achieved. This may be because of practical reasons such as resource constraints, but exclusion can occur for various other reasons, such as power dynamics or differing policy agendas (Headey and Muller, 1996; Balane et al., 2020).

Understanding why certain stakeholder groups are included or excluded is important for several reasons. It can reveal power imbalances and potential biases in the policymaking process (Balane et al., 2020; Jaques, 2006). But it can also inform efforts to improve stakeholder engagement strategies, which will enhance the quality and legitimacy of policy outcomes (Pauwelyn, 2023).

Dynamics may emerge that actors set in motion to achieve a particular policy desire. Stakeholders may secretly lobby or use political influence to help shape regulation in their favor (Stefaniak, 2022). The wielding of hidden agendas (Duke, 2022) is another dynamic. Stakeholders with varying agendas also often compete for priorities during policymaking on ethical grounds. These competing interests can dilute essential considerations such as fairness, enabling biases, and prejudice within decision-making processes related to AI system design (Mittelstadt, 2019; Stahl, 2021). The strategic manoeuvering by different stakeholders may impede progress toward developing effective governance frameworks for ensuring trustworthiness across various dimensions of AI technology.

Understanding these stakeholder motivations is essential for analyzing the dynamics around AI policies and ethics. By delving into the complexities of stakeholder interactions and engagement, researchers can uncover hidden agendas, potential conflicts, or collaborative opportunities. This nuanced understanding forms the foundation for the subsequent exploration of theoretical frameworks, research disciplines, and analytical methods.

4. Methodology

This article addresses the following research questions: (1) "How can we identify social constructivist factors in the shaping of AI governance policy documents?" and (2) "What research has been conducted to explore and scrutinize stakeholder motivations encompassing interests, strategies, intentions, and agendas?" The systematic literature review approach is best suited to answer this question. For this systematic review of the literature, the scoping study methodology was selected. This methodology offers a structured and comprehensive approach to mapping and synthesizing existing literature on a specific research topic. Scoping studies encompass a wide array of studies within the review process to chart the fundamental concepts that form the basis of a research domain, as well as the primary sources of evidence (Arksey and O'Malley, 2005). This methodology was suitable for this study because it enables a holistic exploration of the breadth and depth of the topic. It can indicate the boundaries of a field, the extent of research already completed, and any research gaps (Yu and Watson, 2017).

The scoping study methodology involves a systematic process of gathering, analyzing, and synthesizing a wide range of literature sources. This study uses a five-step framework (Arksey and O'Malley, 2005) that is considered the basis for a scoping study:

Stage 1: identifying the research question.
Stage 2: identifying relevant studies.
Stage 3: study selection.
Stage 4: charting the data.
Stage 5: collating, summarizing, and reporting the results.

In Stage 2 (identifying relevant studies), a search strategy was initially developed that included the formulation of keywords and identification of research tools. Therefore, "identifying," "stakeholders," "stakeholders interests," "stakeholder strategies," "stakeholder motivations," "hidden agendas," and "hidden motives" were selected as keywords. To enable a systematic retrieval of relevant scholarly literature, specialized academic databases and digital repositories were explored. In addition, AI-powered tools were used to perform an extensive search. The initial search was performed through Scopus, ACM
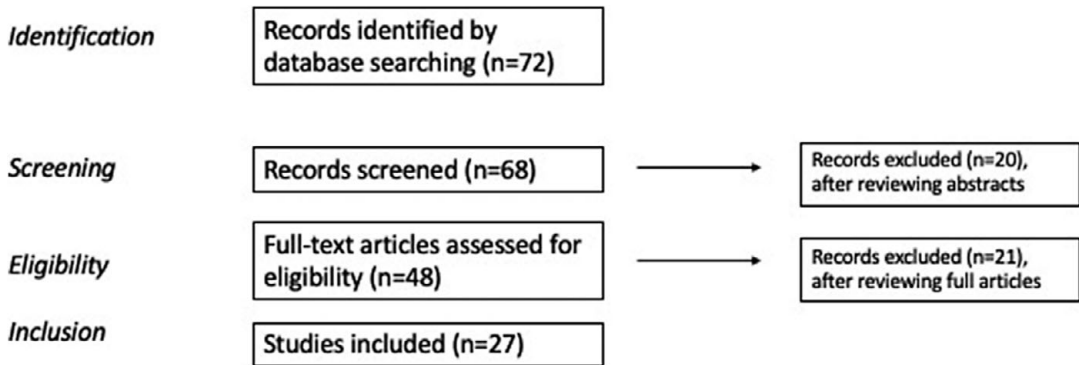
Figure 2. *Literature search and evaluation for inclusion.*

Digital Library, JSTOR, and Google Scholar. The AI tools for the literature review used were Elicit, Connected Papers, and SciSpace.

In Stage 3, the inclusion criteria were defined as academic journal articles and relevant reports by organizations published in English and available online in full text, deemed pertinent to the research aim. This entails selecting articles that are relevant to and contribute to addressing the research aim and question. Conversely, the exclusion criteria encompassed publications falling outside the scope of the aforementioned inclusion criteria.

During Stage 4, the initial step involved evaluating the abstracts of the chosen studies. If there was alignment with the research aim, a more in-depth examination of the complete studies ensued. This procedural sequence is depicted in Figure 2. The narrative of the findings assumes descriptive exploration rather than a static analysis of the results. The findings were divided into three subgroups: disciplines, theoretical frameworks, and analytical methods, to get an overview of both theoretical and practical approaches. Organizing the findings in this structure enhances the clarity, focus, and comprehensiveness and provides a nuanced exploration of stakeholder motivations.

Stage 5 entailed documenting and presenting our findings in the format of a literature review paper. In total, 27 studies were reviewed.

5. Findings

5.1. Disciplines

The first research discipline that can be used to dig into the main topic of this article is moral philosophy. This discipline provides a robust foundation for examining stakeholder motivations. According to Bietti (2020), moral philosophy provides a meta-level perspective for considering disagreements in technology governance, situating problems within broader contexts, and understanding them in relation to other debates. It broadens perspectives, overcomes confusion, and draws clarifying distinctions.

In political sciences and social movement studies, dynamic representation can help with understanding and analyzing complex systems, phenomena, or data over time. This discipline can be used to study how stakeholders' interests evolve or to measure the impact of a stakeholder group on a policy agenda. Bernardi et al. (2021) studied, for example, how legislative agendas respond to signals such as protests by public opinion.

Critical policy studies offer a multifaceted framework for evaluating stakeholders in various contexts. One such application involves analyzing stakeholders' argumentative turns, focusing on language, context, argumentation, and communicative practices within policy processes. In the context of public policy, communication serves as a pivotal conduit connecting the state with society or bridging the public and private sectors. This interaction can take the form of unilateral information dissemination or reciprocal dialogues. At its core, the policymaking process is entwined with ongoing discursive conflicts and communicative practices (Durnová et al., 2016).

Legrand (2022) identifies in this field of study political strategies applied by policymakers to disconnect individuals or groups from political participation. The author applies political exclusion as a methodological means to examine the malevolent aspects within the policymaking process, specifically applied to the context of Australia's asylum seeker policy.

Park and Lee (2020) employed a stakeholder-oriented approach to assess policy legitimacy. Their research delved into the communicative dynamics between state elites and societal stakeholders in South Korea, concentrating on anti-smoking policies under two distinct government administrations. In addition, frame analysis emerged as a valuable tool for comprehending discrepancies between policy intent and implementation.

Anthropologists may study stakeholders within specific cultural contexts, exploring how cultural norms and values influence stakeholders' motivations. Hoholm and Araujo (2011) see ethnography therein as a promising method to track an innovation process in real time. Real-time ethnography enhances our understanding of innovation processes by revealing the uncertainties, choices, and contextual interpretations faced by actors. It provides insights into agential moments, the construction of action contexts, and the complexities of selecting options, emphasizing the messy and multifaceted nature of innovation processes influenced by various conflicting factors. We explore ethnography further in the methodological section of the findings.

5.2. Theoretical frameworks

A self-evident theoretical framework for the examination of stakeholders is stakeholder theory, which originates from the field of business ethics. According to the definition of Mahajan et al. (2023), the theory promotes "… understanding and managing stakeholder needs, wants, and demands." Ayling and Chapman (2021) recognize that stakeholder theory provides a sound framework to, among other things, identify and describe all interested and affected parties in the application of technology and confirm that stakeholders have a legitimate interest in technology. Stakeholder theory was introduced in the 1980s as a management theory to counter the dominance of shareholder theory. But the theory can also be applied in a government/policy context (Flak and Rose, 2005; Harrison et al., 2015).

A concrete application can be seen in Neville and Mengucs' work (2006), which introduces an advanced framework within stakeholder theory. It emphasizes complex interactions among stakeholders, incorporating multiple forms of fit (matching, moderation, and gestalts), integrating stakeholder identification and salience theory, and proposing a hierarchy of influence among stakeholder groups (governments, customers, and employees) to understand stakeholder multiplicity and interactions in organizational contexts. Furthermore, the work of Miller (2022) provides a good example of the application of stakeholder theory, utilizing the stakeholder salience model to manage stakeholders in AI projects based on power, legitimacy, and urgency, and adding a harm attribute to identify passive stakeholders. Passive stakeholders are individuals affected by AI systems but lacking the power to influence the project. Miller's main theoretical contribution is the use of the different stages of an AI life cycle (planning and design, data collection, etc.) to show at which stage certain stakeholders can have an impact on an AI project or system.

McConnell (2017) introduces a specific framework that serves as a heuristic for assessing hidden policy agendas. This novel approach includes criteria such as who hides, what is hidden, who stakeholders conceal it from, and the tools/techniques used for concealment. In addition, it explores the consequences of hidden agendas that remain undisclosed and the outcomes once exposed. While scientifically proving the existence of a hidden agenda is challenging, McConnell proposes a four-step framework to provide a more foundational basis for claims about hidden agendas, offering a common analytical filter.

In subsequent work, McConnell (2020) introduces the concept of "placebo policies," policies implemented to address the symptoms of a problem without addressing the deeper causal factors. This implies that stakeholders advocating for placebo policies may strategically seek to appear proactive without taking substantive action. Such stakeholders, including politicians and government officials, aim to avoid political risks or criticism. The article provides a road map for researchers to study this

phenomenon, guiding them to examine policy intentions, focus, and effectiveness to identify whether a policy is a placebo.

The "Success for Whom"-framework can be applied to a particular policy to address and understand the success (or lack thereof) of that policy for stakeholders such as government officials or lobbyists. From that success or lack of success, stakeholder intentions can be derived. McConnel et al. (2020) developed a three-step road map and applied this in their study in a case study of housing policy in Australia.

The theoretical construct of coalition magnets encapsulates attractive policy ideas, for example, sustainability or social inclusion, strategically employed by policy entrepreneurs to foster coalitional alliances. With this, the concept can be a good tool to assess stakeholder intentions or strategies. Béland and Cox (2015) offer illustrative examples where these ideas have functioned as pivotal coalition magnets across diverse policy domains and periods. Coalition magnets also serve as instruments that are manipulated by policy entrepreneurs, enabling them to redefine prevailing policy problems. These magnets possess the unique ability to harmonize actors with previously conflicting interests or awaken new policy preferences in actors who were not previously engaged with the issue. Within their discourse, the authors provide a set of defining characteristics facilitating the identification of coalition magnets. Ideas characterized by ambiguity or infused with profound emotional resonance emerge as potent foundational elements for the creation of a coalition magnet.

5.3. *Analytical methods*

Content analysis is a valuable analytical instrument for identifying stakeholder motivations, due to its ability to provide both quantitative and qualitative insights, its systematic and replicable nature, the ability it gives to identify patterns and trends, and its adaptability to different types and scales of data (Wilson, 2016; Prasad, 2008). In the context of policy analysis, Hopman et al. (2012) conducted a qualitative content analysis of policy reports and transcribed interviews to identify and analyze the values and beliefs underlying Dutch youth policy. The study is innovative because it uses a theory on the content and structure of values. Text fragments expressing values were labeled with corresponding values of the theory, and the percentages of text fragments per value domain were calculated to rank them in order of importance. This resulted in insights into the latent value perspective in Dutch family policy and how these values shape policy strategies.

A powerful tool for uncovering implicit aspects of communication, making it well-suited for identifying stakeholder motivations, is discourse analysis. Its focus on language, framing, power dynamics, and contextual understanding could provide a nuanced perspective on the underlying forces at play in stakeholder discourse. Lynggaard and Triantafillou (2023) use discourse analysis to categorize three types of discursive agency that empower policy actors to shape and influence policies: maneuvering within established communication frameworks, navigating between conflicting discourses, and transforming existing discourses. This approach provides a nuanced understanding of the dynamics at play.

In the same academic methodological domain, Germundsson (2022) employs Bacchi's (2009) "What's the problem represented to be?" framework in the analysis of policy discourse. This application aims to deconstruct policy discourse and systematically examine the underlying presuppositions inherent in the delineation of issues. Specifically, Germundsson applies this framework in the context of scrutinizing the utilization of automated decision-support systems by public institutions in Sweden. Leifeld (2017) utilizes discourse network analysis within the scope of the research to uncover various elements, including intrinsic endogenous processes such as popularity, reciprocity, and social balance, evident in the exchanges between policy actors.

Social semiotics also offers interesting tools. Inwood and Zappavinga (2021) combine the Corpus Linguistics Sampling Method with a specific theoretical framework and coupling analysis to uncover values such as political ideology discursively contained in white papers of four blockchain start-ups. The Corpus Linguistics Sampling Method involves analyzing concordance lines—listing instances of specifically selected words—in white papers using AntConc software, focusing on the most frequent "3-grams" (sets of three co-occurring words) to establish prominent concerns and themes in the data set.

Klüver (2009) demonstrates the efficacy of quantitative text analysis, specifically with the Wordfish scaling method, in measuring policy positions of interest groups. This computerized content analysis tool proves valuable for analyzing the influence of interest groups, as exemplified in the study's application to a European Commission policy proposal regarding CO_2 emission reductions in the automotive sector.

Engaging with stakeholders through interviews allows researchers to gather firsthand insights, perspectives, and other information, directly from those who are affected by or involved in a certain action. Hansen et al. (2011) demonstrate this in the ICT4D study field. The interviews revealed a diverse range of agendas and interests in projects involving multiple organizations.

Studies employing a mix of methodologies demonstrate significant potential. Tanner and Bryden (2023) executed a two-phase research project on AI public discourse, employing AI-assisted quantitative analysis of videos and media articles, followed by qualitative interviews with policy and communication experts. The study illuminates a deficiency within the public discourse surrounding AI, wherein discussions concerning its broader societal ramifications receive limited attention and are predominantly shaped by technology corporations. Moreover, the research suggests that Civil Society Organizations are inadequately effective in contributing to and influencing this discourse. Ciepielewska-Kowalik (2020) combined surveys with in-depth interviews to uncover (un)intended consequences and hidden agendas related to education reform in Poland. Wen (2018) conducted a laboratory experiment using human-assisted simulation to investigate decision-making logic among Chinese public officials and citizens, aiming to explore hidden motivations within structured social systems in China. Results were supplemented by post-simulation and longitudinal interviews.

Ethnography proves to be a compelling method for examining stakeholders' motivations in a policy process, given its immersive and holistic nature. By embedding researchers within the social settings where policies are formulated, debated, and implemented, ethnography facilitates an in-depth understanding of stakeholders' lived experiences and perspectives. Utilizing participant observation, interviews, and the analysis of everyday interactions, ethnography unveils tacit knowledge, implicit norms, and nuanced social dynamics not readily apparent through conventional research methods. Emphasizing context and offering a rich narrative of the social milieu, ethnography contributes to a comprehensive comprehension of stakeholders' behaviors and motivations (Cappellaro, 2016).

In this methodological domain, various subcategories exist. Hoholm and Araújo (2011), studying innovation processes in real time, contend that real-time ethnography enhances theorizing processes by providing insights into uncertainties, contingencies, and choices faced by actors. They emphasize its capacity to contextualize interpretations of past and future projects. In addition, real-time ethnography elucidates how actors interpret and construct contexts of action, offering a better analytical understanding of controversies, tensions, and fissures arising from alternative choice paths. The authors dispel the simplistic view of ethnography, stressing the necessity to trace elements challenging or impossible to observe through other methods.

Critical ethnography is an approach that goes beyond traditional ethnography by incorporating a critical perspective on social issues and power structures. Thomas (1993) describes this as follows: "Critical ethnographers describe, analyze, and open to scrutiny otherwise hidden agendas, power centres, and assumptions that inhibit, repress, and constrain." We find an application of this in Myers and Young (1997), who use critical ethnography to uncover hidden agendas, power, and issues such as managerial assumptions in the context of an information system development project in mental health.

Within the scope of this paper, policy ethnography and political ethnography are particularly pertinent. While policy ethnography focuses on the study of policies and their implementation, political ethnography encompasses a broader spectrum, including diverse political activities and structures such as political institutions and movements. Dubois (2009) uses critical policy ethnography to examine welfare state policies in France. The study mainly provides a realistic understanding of the implementation of the policies and the impact on recipients of welfare, but the author indicates that the method also provided insight into the control practices of officials and their intentions. Namian (2021) shows through policy ethnography the values associated with essential structures in a policy field, specifically examining homeless shelters in the context of the Housing First policy.

Published online by Cambridge University Press

While explicit instances showcasing the utilization of political ethnography in investigating stakeholder motivations were not identified, the method remains an intriguing instrument for such research. For instance, Baiocchi and Connor (2008) offer a comprehensive yet not fully exhaustive survey of the method's application in examining politics, interactions between individuals and political actors, as well as the intersection of politics and anthropology.

6. Discussion

This study delves into existing research on identifying stakeholder motivations, revealing ample opportunities across various academic disciplines, theoretical frameworks, and analytical methods. The results show that a lot of opportunities exist for different academic disciplines. In addition, different theoretical frameworks and analytical methods can be employed. The results show that stakeholders are already approached in different ways in studies, but a limited focus on intentions, agendas, or strategies is observed.

In critical policy studies, there is much potential to deepen the understanding of stakeholder motivations, particularly within the dynamic contexts of innovation and AI. The policymaking process, especially in the context of innovation and AI, requires more in-depth research, especially against the current topical background. One recent example is the EU AI Act, where there have been widespread claims that the final text was diluted due to lobbying efforts by major technology companies (Vranken, 2023; Floridi and Baracchi Bonvicini, 2023). Critical policy researchers can, for example, explore power dynamics, language, and communicative practices within policy processes. The emphasis on discourse within this discipline allows researchers to uncover hidden agendas, power struggles, and how stakeholders shape policy narratives to advance their interests (Fischer et al., 2015).

In the context of the selected theoretical frameworks, McConnell's work assessing hidden policy agendas serves as an important foundation for studying stakeholder motives in policymaking. It could be very interesting for researchers to link methods to this framework to explore this topic in depth. Leveraging McConnel's conceptualization of placebo policies holds considerable promise within the domain of AI policy. This framework facilitates the identification of policy measures that ostensibly address an issue but do so in a cursory or temporary manner, neglecting deeper underlying factors. This analytical lens proves particularly insightful when examining themes such as fairness or transparency in AI governance. While stakeholder theory offers an interesting lens to explore power dynamics and relationships, it has limited applicability for in-depth understanding, necessitating additional research tools.

Disciplines and theoretical frameworks can be interesting building blocks from which research on stakeholder motivations can start or be supported and strengthened. Analytical methods, predominant in our results, offer more practical insights that can be put to work. Content analysis, despite its reductive nature (Kolbe and Burnett, 1991), can yield valuable results with an appropriate theoretical framework (Hopman et al., 2012). Discourse analysis, while widely used, requires complementation with other methodologies for a comprehensive understanding.

The integration of diverse research methodologies, as evidenced by studies such as Tanner and Bryden's AI discourse analysis, Ciepielewska-Kowalik's combined surveys and interviews on education reform, and Wen's multifaceted approach to uncovering hidden motivations in Chinese social systems, highlights the substantial potential of employing a mixed-methods approach. This versatility allows researchers to gain comprehensive insights.

Ethnography and its subdomains is the research method with the most potential in the context of this study. The various studies indicate that the method can gain unique insights and challenge existing views. Although we observed a concrete application to stakeholders' hidden agendas in one case (Myers and Young, 1997), we can find much evidence from the other studies that the methodology fits well with the objective of observations in an innovation policy context. In light of our study, this approach enables researchers to grasp the intricacies of stakeholders' decision-making processes, the rationales behind their actions, and the complex interplay of factors influencing policy outcomes.

Listing the various subcategories within ethnography reveals the potential for cross-fertilization between them. For instance, the synergy between subcategories like real-time ethnography and policy ethnography holds promise. This entails examining policy stakeholders in real time at various stages of a policy process, expanding beyond the predominant focus on the implementation stage in traditional policy ethnography.

In general, this review provides a valuable resource for scholars and practitioners exploring stakeholder dynamics. Further, this article also contributes indirectly to the epistemological problem of whether there can be conscience and in what ways hidden stakeholder motivations exist.

In its initial phase, this study provides tools for researchers to utilize. Subsequently, it holds the potential to yield insights for policymakers. As posited in the article, AI ethics guidelines and policy documents fall short for the following reasons: they are too abstract and theoretical, disconnect from real-world contexts, do not deal with balancing conflicting values, overemphasize self-regulation, have limited stakeholder inclusion, and are vulnerable to hidden agendas. Further research on stakeholder motivations within a policy context offers valuable information for refining these documents and improving the overall production process. The chosen approach provides a rather limited description of disciplines, frameworks, and methods. A critical evaluation of whether methods are suitable to apply in the complex (innovation) policy context requires further research.

7. Conclusion

The present study sought to address the research questions (1) "How can we identify social constructivist factors in the shaping of AI governance documents?" and (2) "What research has been conducted to explore and scrutinize stakeholder motivations encompassing interests, strategies, intentions, and agendas?" through a comprehensive review of literature spanning research disciplines, theoretical frameworks, and analytical methods. The results reveal a noticeable dearth of research on stakeholder motivation in policy processes. While several studies touch upon unilateral examinations of stakeholders' expressions, such as text or statement analyses, our results highlight a scarcity of studies delving into the "hidden" domain, uncovering concealed agendas or intentions.

Consequently, this study serves as a foundational contribution, intended to stimulate continued research and foster further critical debates on this topic, aiming for a more profound exploration of policy processes. This extends beyond merely understanding how guidelines and policy documents are created but can also be employed from a social-critical point of view to unmask the true intentions of stakeholders. Future research will build upon the insights gained in this study by critically evaluating the existing knowledge and exploring applicable theoretical frameworks and methods within the context of innovation policy.

In sum, this study exposes a research area that has already been explored to a very limited extent. It provides an overview of disciplines, theoretical frameworks, and methods with which researchers can engage in this field. While providing an overview of relevant disciplines, theoretical frameworks, and methods, this study not only prompts additional research into their utility but also highlights the potential inherent in combining different tools. Over time, such pathways may offer more insight into policy processes for researchers but may also be valuable for stakeholders engaged in the policy process.

References

Aaronson SA and Zable A (2023) Missing persons: the case of national AI strategies. *Social Science Research Network.* (No. 283). CIGI Papers. https://doi.org/10.2139/ssrn.4554650

Adler E (1997) Seizing the middle ground: constructivism in world politics. *European Journal of International Relations 3*, 319–363.

Arestis P and Kitromilides Y (2010) What economists should know about public policymaking? *International Journal of Public Policy 6*(1/2), 136. https://doi.org/10.1504/ijpp.2010.031211

Arksey H and O'Malley L (2005) Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology 8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Attard-Frost B, De Los Ríos A and Walters DR (2022). The ethics of AI business practices: a review of 47 AI ethics guidelines. *AI and Ethics 3*(2), 389–406. https://doi.org/10.1007/s43681-022-00156-6

Ayling J and Chapman A (2021) Putting AI ethics to work: are the tools fit for purpose? *AI and Ethics 2*(3), 405–429. https://doi.org/10.1007/s43681-021-00084-x

Bacchi CL (2009) *Analysing Policy: What's the Problem Represented to Be?* Frenchs Forest: Pearson.

Baiocchi G and Connor BC (2008) The ethnos in the polis: political ethnography as a mode of inquiry. *Sociology Compass 2*(1), 139–155. https://doi.org/10.1111/j.1751-9020.2007.00053.x

Bakker S, Maat K and Van Wee B (2014) Stakeholders' interests, expectations, and strategies regarding the development and implementation of electric vehicles: the case of the Netherlands. *Transportation Research Part A: Policy and Practice 66*, 52–64. https://doi.org/10.1016/j.tra.2014.04.018

Balane MA, Palafox B, Palileo-Villanueva LM, McKee M and Balabanova D (2020) Enhancing the use of stakeholder analysis for policy implementation research: towards a novel framing and operationalised measures. *BMJ Global Health 5*(11), e002661. https://doi.org/10.1136/bmjgh-2020-002661

Béland D and Cox RH (2015) Ideas as coalition magnets: coalition building, policy entrepreneurs, and power relations. *Journal of European Public Policy 23*(3), 428–445. https://doi.org/10.1080/13501763.2015.1115533

Bernardi L, Bischof D and Wouters R (2021) The public, the protester, and the bill: do legislative agendas respond to public opinion signals? *Journal of European Public Policy 28*(2), 289–310. https://doi.org/10.1080/13501763.2020.1729226

Bietti E (2020) *From Ethics Washing to Ethics Bashing: A View on Tech Ethics from Within Moral Philosophy.* https://doi.org/10.1145/3351095.3372860

Bijker WE, Hughes TP and Trevor P (eds) (1987) *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology.* Cambridge, MA: MIT Press.

Brück J (2006) Fragmentation, personhood and the social construction of technology in middle and late Bronze Age Britain. *Cambridge Archaeological Journal*, *16*(3), 297–315. doi:10.1017/S0959774306000187

Butcher J, and 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

Cappellaro G (2016) Ethnography in public management research: a systematic review and future directions. *International Public Management Journal 20*(1), 14–48. https://doi.org/10.1080/10967494.2016.1143423

Cheng L, Varshney KR and Liu H (2021) Socially responsible AI algorithms: issues, purposes, and challenges. *Journal of Artificial Intelligence Research 71*, 1137–1181. https://doi.org/10.1613/jair.1.12814

Christoforaki M and Beyan O (2022) AI ethics—a bird's eye view. *Applied Sciences 12*(9), 4130. https://doi.org/10.3390/app12094130

Ciepielewska-Kowalik A (2020) (Un)intended consequences of the 2016 educational reform on early childhood education and care in Poland. The story of a few applications that led to significant disorder. *Policy Futures in Education*, *18*(6), 806–823 https://doi.org/10.1177/1478210320923158

Cihon P, Schuett J and Baum SD (2021) Corporate governance of artificial intelligence in the public interest. *Information 12*(7), 275.

Conti LG and Seele P (2023) The contested role of AI ethics boards in smart societies: a step towards improvement based on board composition by sortition. *Ethics and Information Technology 25*, 51. https://doi.org/10.1007/s10676-023-09724-8

De Almeida PGR, Santos CDD and Farias JS (2021) Artificial intelligence regulation: a framework for governance. *Ethics and Information Technology 23*(3), 505–525. https://doi.org/10.1007/s10676-021-09593-z

De Cooman J and Petit N (2020). Models of law and regulation for AI. *Social Science Research Network*. https://doi.org/10.2139/ssrn.3706771

Donahoe E and Metzger MM (2019) Artificial intelligence and human rights. *Journal of Democracy 30*(2), 115–126. https://doi.org/10.1353/jod.2019.0029

Dubois V (2009) Towards a critical policy ethnography: lessons from fieldwork on welfare control in France. *Critical Policy Studies 3*(2), 221–239.

Duke SA (2022) Deny, dismiss and downplay: developers' attitudes towards risk and their role in risk creation in the field of healthcare-AI. *Ethics and Information Technology 24*(1). https://doi.org/10.1007/s10676-022-09627-0

Durnová A, Fischer F and Zittoun P (2016) *Discursive Approaches to Public Policy: Politics, Argumentation, and Deliberation. Palgrave Macmillan UK eBooks*, pp. 35–56. https://doi.org/10.1057/978-1-137-50494-4_3

Edgerton A, Seddiq O and Leinz K (2023, September 13) Tech Leaders Discuss AI Policy in Closed-Door Senate Meeting. Bloomberg. Available at https://www.bloomberg.com/news/articles/2023-09-13/tech-leaders-head-to-capitol-hill-to-propose-their-own-ai-rules

Ehsan U and Riedl MO (2022) Social construction of XAI: do we need one definition to rule them all? arXiv preprint arXiv:2211.06499.

Elder-Vass D (2012) *The Reality of Social Construction*. Cambridge University Press.

Eynon R and Young E (2020) Methodology, legend, and rhetoric: the constructions of AI by academia, industry, and policy groups for lifelong learning. *Science, Technology, & Human Values 46*(1), 166–191. https://doi.org/10.1177/0162243920906475

Fischer F, **Torgerson D**, **Durnová A and Orsini M** (2015) *Handbook of Critical Policy Studies*. Edward Elgar Publishing eBooks. https://doi.org/10.4337/9781783472352

Flak LS and Rose J (2005) Stakeholder governance: adapting stakeholder theory to E-Government. *Communications of the Association for Information Systems 16*. https://doi.org/10.17705/1cais.01631

Flockhart T (2016) *Constructivism and Foreign Policy*. Oxford University Press eBooks. https://doi.org/10.1093/hepl/9780198708902.003.0004

Floridi L (2019) Translating principles into practices of digital ethics: five risks of being unethical. *Philosophy & Technology 32*(2), 185–193. https://doi.org/10.1007/s13347-019-00354-x

Floridi L and Baracchi Bonvicini M (2023, November 23). Open Letter: Urging the approval of the EU AI Act. *Atomium European Institute*. Available at: https://www.eismd.eu/letter_to_mr_macron_mrs_meloni_mr_scholz_26-11-2_231128_160302.pdf (accessed 22 February 2024).

Floridi L, **Cowls J**, **Beltrametti M**, **Chatila R**, **Chazerand P**, **Dignum V**, **Luetge C**, **Madelin R**, **Pagallo U**, **Rossi F**, **Schäfer B**, **Valcke P and Vayena E** (2018) AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. *Minds and Machines 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Garber M, **Sarkani S and Mazzuchi T** (2017) A framework for multiobjective decision management with diverse stakeholders. *Systems Engineering 20*(4), 335–356. https://doi.org/10.1002/sys.21398.

Germundsson N (2022) Promoting the digital future: the construction of digital automation in Swedish policy discourse on social assistance. *Critical Policy Studies 16*(4), 478–496. https://doi.org/10.1080/19460171.2021.2022507

Gutiérrez CI and Marchant GE (2021) A global perspective of soft law programs for the governance of artificial intelligence. *Social Science Research Network*. https://doi.org/10.2139/ssrn.3855171

Hadfield GK and Clark J (2023) Regulatory Markets: The Future of AI Governance. arXiv preprint arXiv:2304.04914.

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, **Robertson T**, **Wilson L**, **Thinyane H and Gumbo S** (2011) Identifying stakeholder perspectives in a large collaborative project. *Association for Computing Machinery*. https://doi.org/10.1145/2071536.2071558

Harrison JS, **Freeman RE and De Abreu MCS** (2015) Stakeholder theory as an ethical approach to effective management: applying the theory to multiple contexts. *Revista Brasileira De Gestão De Negócios*, 858–869. https://doi.org/10.7819/rbgn.v17i55.2647

Headey B and Muller D (1996) Policy agendas of the poor, the public and elites: a test of Bachrach and Baratz. *Australian Journal of Political Science 31*, 347–368. https://doi.org/10.1080/10361149651094

Hoholm T and Araújo L (2011) Studying innovation processes in real-time: the promises and challenges of ethnography. *Industrial Marketing Management 40*(6), 933–939. https://doi.org/10.1016/j.indmarman.2011.06.036

Hopman M, **De Winter M and Koops W** (2012) The hidden curriculum of youth policy. *Youth & Society 46*(3), 360–378. https://doi.org/10.1177/0044118x11436187

Howlett M (2021) Avoiding a Panglossian policy science: the need to deal with the dark side of policy-maker and policy-taker behaviour. *Public Integrity 24*(3), 306–318. https://doi.org/10.1080/10999922.2021.1935560

Inwood O and Zappavigna M (2021) Ideology, attitudinal positioning, and the blockchain: a social semiotic approach to understanding the values construed in the whitepapers of blockchain start-ups. *Social Semiotics 33*(3), 451–469. https://doi.org/10.1080/10350330.2021.1877995

Jaques T (2006) Activist "rules" and the convergence with issue management. *Journal of Communication Management 10*(4), 407–420. https://doi.org/10.1108/13632540610714836

Jobin A, **Guettel L**, **Liebig L and Katzenbach C** (2021) AI Federalism: Shaping AI Policy Within States in Germany. arXiv preprint arXiv:2111.04454.

Kergroach S (2021) SMEs going digital: policy challenges and recommendations. In *OECD Going Digital Toolkit Notes*, Vol. 15. Paris: OECD Publishing. https://doi.org/10.1787/c91088a4-en

Kim D, **Zhu Q and Eldardiry H** (2023) Exploring approaches to artificial intelligence governance: from ethics to policy. *IEEE International Symposium on Ethics in Engineering, Science, and Technology (ETHICS)*. https://doi.org/10.1109/ethics57328.2023.10155067

Klüver H (2009) Measuring interest group influence using quantitative text analysis. *European Union Politics 10*(4), 535–549. https://doi.org/10.1177/1465116509346782

Kolbe RH and Burnett MS (1991) Content-analysis research: an examination of applications with directives for improving research reliability and objectivity. *Journal of Consumer Research 18*(2), 243. https://doi.org/10.1086/209256

Krzywdzinski M, **Gerst D and Butollo F** (2022) Promoting human-centred AI in the workplace. Trade unions and their strategies for regulating the use of AI in Germany. *Transfer 29*(1), 53–70. https://doi.org/10.1177/10242589221142273

Legrand T (2022) The malign system in policy studies: strategies of structural and agential political exclusion. *International Journal of Public Policy 16*(2/3/4), 88. https://doi.org/10.1504/ijpp.2022.10049342

Leifeld P (2017) Discourse network analysis: policy debates as dynamic networks. In Victor JN, Lubell MN and Montgomery AH (eds.), *The Oxford Handbook of Political Networks*. Oxford University Press.

Lynggaard K and Triantafillou P (2023) Discourse analysis and strategic policy advice: manoeuvring, navigating, and transforming policy. *Journal of European Public Policy 30*(9), 1936–1959. https://doi.org/10.1080/13501763.2023.2217846

Mahajan R, **Lim WM**, **Sareen ML**, **Kumar S and Panwar R** (2023) Stakeholder theory. *Journal of Business Research 166*, 114104. https://doi.org/10.1016/j.jbusres.2023.114104

McConnell A (2017) Hidden agendas: shining a light on the dark side of public policy. *Journal of European Public Policy 25*(12), 1739–1758. https://doi.org/10.1080/13501763.2017.1382555

McConnell A (2020). The use of placebo policies to escape from policy traps. *Journal of European Public Policy 27*(7), 957–976. https://doi.org/10.1080/13501763.2019.1662827

McConnell A, **Grealy L and Lea T** (2020). Policy success for whom? A framework for analysis. *Policy Sciences 53*(4), 589–608. https://doi.org/10.1007/s11077-020-09406-y

Metcalfe J (1995) Technology systems and technology policy in an evolutionary framework. *Cambridge Journal of Economics*. https://doi.org/10.1093/oxfordjournals.cje.a035307

Miller GJ (2022). Stakeholder roles in artificial intelligence projects. *Project Leadership and Society 3*, 100068. https://doi.org/10.1016/j.plas.2022.100068

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

Munn L (2022). The uselessness of AI ethics. *AI and Ethics*. https://doi.org/10.1007/s43681-022-00209-w

Myers MD and Young LW (1997) Hidden agendas, power and managerial assumptions in information systems development: an ethnographic study. *Information Technology & People 10*(3), 224–240.

Namian D (2021) Homemaking among the 'chronically homeless': a critical policy ethnography of housing first. *Housing Studies 37*(2), 332–349. https://doi.org/10.1080/02673037.2021.2009777

Neville BA, and Menguc B (2006) Stakeholder multiplicity: Toward an understanding of the interactions between stakeholders. *Journal of business ethics 66*, 377–391. https://doi.org/10.1007/s10551-006-0015-4

O'Shaughnessy M (2022, October 6) *One of the Biggest Problems in Regulating AI is Agreeing on a Definition*. Carnegie Endowment for International Peace. Available at: https://carnegieendowment.org/2022/10/06/one-of-biggest-problems-in-regulating-ai-is-agreeing-on-definition-pub-88100

OECD. (n.d.). *OECD.AI Principles Overview*. OECD.AI. Available at: https://oecd.ai/en/ai-principles

Oehlenschlager M and Tranberg P (2023) Big tech soft power in Denmark. In *Dataethics*.

Papyshev G and Yarime M (2022) The limitation of ethics-based approaches to regulating artificial intelligence: regulatory gifting in the context of Russia. *AI & Society*. https://doi.org/10.1007/s00146-022-01611-y

Park C and Lee J (2020) Stakeholder framing, communicative interaction, and policy legitimacy: anti-smoking policy in South Korea. *Policy Sciences 53*(4), 637–665. https://doi.org/10.1007/s11077-020-09394-z

Patiño SR, **Rajamohan S**, **Meaney K**, **Coupey E**, **Serrano E**, **Hedrick VE**, **Da Silva Gomes F**, **Polys NF and Kraak V** (2020) Development of a Responsible Policy Index to improve statutory and Self-Regulatory policies that protect children's diet and health in the America's region. *International Journal of Environmental Research and Public Health 17*(2), 495. https://doi.org/10.3390/ijerph17020495

Pauwelyn J (2023) Taking stakeholder engagement in international policy-making seriously: Is the WTO finally opening up? *Journal of International Economic Law 26*(1), 51–65. https://doi.org/10.1093/jiel/jgac061

Perry B and Uuk R (2019) AI governance and the policymaking process: key considerations for reducing AI risk. *Big Data and Cognitive Computing 3*(2), 26. https://doi.org/10.3390/bdcc3020026

Petersen TS (2021). Ethical guidelines for the use of artificial intelligence and the challenges from value conflicts. *Etikk I Praksis: Nordic Journal of Applied Ethics*. https://doi.org/10.5324/eip.v15i1.3756

Pierce JJ, **Siddiki S**, **Jones MD**, **Schumacher K**, **Pattison A and Peterson H** (2014) social construction and policy design: a review of past applications: social construction and policy design. *Policy Studies Journal 42*(1), 1–29. https://doi.org/10.1111/psj.12040

Prasad BD (2008) Content analysis. *Research Methods for Social Work 5*, 1–20.

Sabatier P (2019) *Theories of the Policy Process*. Routledge eBooks. https://doi.org/10.4324/9780429494284

Saurugger S (2013) Constructivism and public policy approaches in the EU: from ideas to power games. *Journal of European Public Policy 20*(6), 888–906. https://doi.org/10.1080/13501763.2013.781826

Scherer M (2015) Regulating artificial intelligence systems: risks, challenges, competencies, and strategies. *Social Science Research Network*. https://doi.org/10.2139/ssrn.2609777

Schiff D, **Biddle J**, **Borenstein J and Laas K** (2020) *What's Next for AI Ethics, Policy, and Governance? A Global Overview*, pp. 153–158. https://doi.org/10.1145/3375627.3375804

Schmitt L (2021) Mapping global AI governance: a nascent regime in a fragmented landscape. *AI and Ethics 2*(2), 303–314. https://doi.org/10.1007/s43681-021-00083-y

Schneider AL and Ingram H (1993) Social construction of target populations: implications for politics and policy. *American Political Science Review 87*(2), 334–347. https://doi.org/10.2307/2939044

Sloane M and Zakrzewski J (2022) German AI start-ups and "AI ethics": using a social practice lens for assessing and implementing socio-technical innovation. *ACM Conference on Fairness, Accountability, and Transparency*. https://doi.org/10.1145/3531146.3533156

Smuha NA (2021) From a 'race to AI' to a 'race to AI regulation': regulatory competition for artificial intelligence. *Law, Innovation and Technology 13*(1), 57–84. https://doi.org/10.1080/17579961.2021.1898300

Stahl BC (2021) Addressing ethical issues in AI. In Stahl BC (ed.), *Artificial Intelligence for a Better Future: An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies*, Springer International Publishing, pp. 55–79.

Stefaniak S (2022) The analysis of the lobbying actors regarding the adoption and implementation of the AI policy in Poland. *AHFE International.* https://doi.org/10.54941/ahfe1002734

Tanner J and Bryden J (2023) Reframing AI in Civil Society: Beyond Risk & Regulation. Available at: https://rootcause.global/framing-ai/

Thierer A, **O'Sullivan A and Russell R** (2017) Artificial intelligence and public policy. *Innovation Law & Policy eJournal.* https://doi.org/10.2139/ssrn.3021135.

Thomas J (1993) *Doing Critical Ethnography.* Newbury Park, CA: Sage Publications.

Ulnicane I, **Knight W**, **Leach T**, **Stahl BC and Wanjiku W** (2020) Framing governance for a contested emerging technology: insights from AI policy. *Policy and Society 40*(2), 158–177. https://doi.org/10.1080/14494035.2020.1855800

Valle-Cruz D, **Criado JI**, **Sandoval-Almazan R and Ruvalcaba-Gomez EA** (2020) Assessing the public policy-cycle framework in the age of artificial intelligence: from agenda-setting to policy evaluation. *Government Information Quarterly 37*(4), 101509. https://doi.org/10.1016/j.giq.2020.101509

Valle-Cruz D, **García-Contreras R and Gil-Garcia JR** (2023) Exploring the negative impacts of artificial intelligence in government: the dark side of intelligent algorithms and cognitive machines. *International Review of Administrative Sciences.* https://doi.org/10.1177/00208523231187051

Veale M, **Matus K and Gorwa R** (2023, June 2). AI and Global Governance: Modalities, Rationales, Tensions. Available at: https://doi.org/10.1146/annurev-lawsocsci-020223-040749

Vranken B (2023) *Big Tech Lobbying is Derailing the AI Act.* Corporate Europe Observatory. Available at: https://corporateeurope.org/en/2023/11/big-tech-lobbying-derailing-ai-act (accessed 22 February 2024).

Watney C, **Auer D** and **Progressive Policy Institute** (2021) *Encouraging AI Adoption by EU SMEs.* Progressive Policy Institute. Available at: https://eadn-wc05-3904069.nxedge.io/cdn/wp-content/uploads/2021/03/PPI_Encouraging-AI-adoption-by-EU-SMEs-3.24.21-2.pdf

Wen MY (2018) Chinese rule-abiding decision-making and hidden motives: simulation findings and implications. *Advances in Applied Sociology 8*, 95–105. https://doi.org/10.4236/aasoci.2018.82006

Wendt A (1992) Anarchy is what states make of it: the social construction of power politics. *International Organization 46*(2), 391–425. https://doi.org/10.1017/s0020818300027764

Wilson V (2016) Research methods: content analysis. *Evidence Based Library and Information Practice 11*(1S), 41–43. https://doi.org/10.18438/b8cg9d

Winecoff AA and Watkins EA (2022) Artificial concepts of artificial intelligence: institutional compliance and resistance in AI startups. In *AIES '22: Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society.* https://doi.org/10.1145/3514094.3534138

Yalçin Ö (2019) Postfoundational critical constructivism: renewing social and political theory of international relations. *The Journal of International Social Research 12*(63), 399–404. https://doi.org/10.17719/jisr.2019.3237

Yu X and Watson M (2017) Guidance on conducting a systematic literature review. *Journal of Planning Education and Research 39*(1), 93–112. https://doi.org/10.1177/0739456x17723971

Zech H (2021) Liability for AI: public policy considerations. *ERA Forum, 22*(1), 147–158. https://doi.org/10.1007/s12027-020-00648-0

Zehfuss M (2001) Constructivism and identity: a dangerous liaison. *European Journal of International Relations, 7*(3), 315–348. https://doi.org/10.1177/1354066101007003002

Zhu B, **Habisch A and Thøgersen J** (2018) The importance of cultural values and trust for innovation—a European study. *International Journal of Innovation Management 22*(2), 1850017. https://doi.org/10.1142/s1363919618500172