Geography of Innovation and Sustainability Transitions

GEIST Working Paper series

The social embeddedness of trust in AI: How existing trust relations to decision-makers and institutions influence trust in AI decision aids for public administration

GEIST Working Paper No. 2025(02)

Tamara Schnell

Carl von Ossietzky Universität Oldenburg, Institute for Social Sciences, Working group "Organization and Innovation", Ammerländer Heerstr. 114-118, D - 26129 Oldenburg. Corresponding author: tamara.schnell@uni-oldenburg.de

Ricarda Schmidt-Scheele

Carl von Ossietzky Universität Oldenburg, Institute for Social Sciences, Working group "Organization and Innovation", Ammerländer Heerstr. 114-118, D - 26129 Oldenburg.

Hanse-Wissenschaftskolleg, Institute for Advanced Study, Delmenhorst, Germany

Cite as: Schnell and Schmidt-Scheele (2025) The social embeddedness of trust in AI: How existing trust relations to decision-makers and institutions influence trust in AI decision aids for public administration. *GEIST – Geography of Innovation and Sustainability Transitions, 2025*(02), GEIST Working Paper series.

The social embeddedness of trust in AI: How existing trust relations to decision-makers and institutions influence trust in AI decision aids for public administration

Authors: Tamara Schnell [a *], Ricarda Schmidt-Scheele [a b]

[a] *Carl von Ossietzky Universität Oldenburg, Institute for Social Sciences, Working group "Organization and Innovation", Ammerländer Heerstr. 114-118, D - 26129 Oldenburg*

[b] *Hanse-Wissenschaftskolleg, Institute for Advanced Study, Delmenhorst, Germany*

* *Corresponding author*

Contact: tamara.schnell@uni-oldenburg.de; ricarda.schmidt-scheele@uni-oldenburg.de

Abstract

AI decision aids are increasingly adopted in public administration to support complex decisions traditionally carried out by employees of public authorities. While prior research has emphasized decision-makers' trust in AI, less attention has been paid to stakeholders who are exposed to and affected by these emerging AI-supported decision-making processes and outcomes. In such contexts, it is not only the AI itself – its process, performance, and purpose – that is assessed for trustworthiness, but also existing constellations of decision-makers and institutions that govern decision-making. We argue that trust in AI is socially embedded. Drawing on sociological theories of trust, we propose a framework that conceptualizes trust in AI decision aids as shaped by existing trust relations with decision-makers and institutions involved in decision-making – the 'shadow of the past'.

To explore this, we examine a case study of an AI-augmented geographic information system (AI-GIS) developed to support spatial planning for onshore wind energy in the course of sustainably energy transition dynamics in Germany. Based on 38 interviews with stakeholders from seven groups involved in spatial planning and wind energy development, we analyze initial (mis)trust in the AI-GIS. Using a combination of qualitative comparative analysis (QCA) and qualitative content analysis, we identify four distinct configurations that condition stakeholders' (mis)trust. Each reflects a unique interplay of interpersonal and institutional trust relations.

The study offers a more nuanced understanding of trust in AI as a relational, context-dependent phenomenon, highlighting the relevance of institutions and existing trust relations for understanding and guiding AI adoption. It therefore directly contributes to the literature on sustainability transitions and their place-specific dynamics. AI systems are considered viable technical solutions for the transformation of energy, water, or food systems. Accordingly, trust in these AI systems needs to be understood as highly context-dependent: Trust is developed and experienced within specific institutional settings, regulatory cultures, and histories of technology adoption. Hence, our paper directs attention to *who* trusts *what* AI, *where*, and *under what* institutional arrangements, and urges this to be a central question in the sustainability transitions literature.

Keywords
Trust in AI; Social embeddedness of trust; Trust in institutions; Qualitative Comparative Analysis (QCA); Spatial planning; Onshore wind energy

Highlights

- Trust in AI is an individualized yet socially embedded concept
- Trust in AI is shaped by prior trust relations to decision-makers and institutions
- Our framework embeds trust in AI within existing actor and institutional relations
- Mixed-method approach combines QCA and qualitative content analysis to study trust
- Identification of four configurations that condition (mis)trust in AI

Abbreviations

AI	Artificial Intelligence
GIS	Geographic Information System
AI-GIS	AI-augmented GIS
QCA	Qualitative comparative analysis

1. Introduction

Artificial intelligence (AI) encompasses a variety of machine learning techniques that enable technologies to perform tasks traditionally associated with human intelligence, such as reasoning, learning, problem-solving, and decision-making (McCarthy 2007; Omrani et al. 2022; Grilli et al. 2024). This is made possible by AI algorithms' capacity to process vast amounts of data in real time (Bailey et al. 2019) and "to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation" (Kaplan und Haenlein 2019, S. 15). Given their potential to optimize tasks and decision-making, AI systems are considered viable technical solutions for the grand societal challenges of our time and thus play a role in sustainability transitions, i.e., the transformation of energy, water, or food systems (Wang et al. 2025; Camaréna 2020). Therefore, innovative AI systems are continuously being developed, such as AI-enabled smart irrigation systems that continuously monitor soil moisture, plant stress, and weather data to optimize water usage in agriculture (Pandey und Mishra 2024), AI-driven digital twins and cyber-physical platforms such as the ACWA testbed for intelligent water management (Batarseh et al. 2023), or multi-layered green-AI architectures engineered for circular economies – optimizing urban waste logistics and battery-recycling workflows with demonstrable energy and resource recovery gains (Ranpara 2025). AI applied as decision-support systems (Hoff und Bashir 2015; Glikson und Woolley 2020) – i.e., algorithms designed to generate decision alternatives or prescribe courses of action for specific objectives (Shrestha et al. 2021; Shrestha et al. 2019) – is hence adopted across heterogeneous organizations and sectors and has become a central topic of scholarly interest (Rijswijk et al. 2023; Sjödin et al. 2023; Nahar 2024). These AI decision aids intervene in tasks and responsibilities that have previously been performed by human actors (Shrestha et al. 2021).

An emerging body of literature provides nuanced, real-world insights into the challenges of integrating AI decision aids into organizational settings (Berente et al. 2021; Daly et al. 2025; van der Werff et al. 2021) as well as into public administration, where decision-making has traditionally relied on the discretion of public servants (Bullock 2019; Agbabiaka et al. 2025; Bao et al. 2025). Here, studies highlight the emergence of shared decision-making spaces between humans and AI, in which decision-makers maintain certain levels of control while choosing to rely on algorithmic recommendations (Shrestha et al. 2019; Murray et al. 2021). AI decision aids hence deeply interfere with, alter, and potentially replace existing interpersonal relationships and need to integrate into – and often compete within – established institutional structures (Shrestha et al. 2019; Bao et al. 2025; Trunk et al. 2020).

Trust is a critical precondition for humans to adopt and rely on AI decision aids in decision-making. Accordingly, trust in AI has emerged as a central topic of interdisciplinary inquiry (Omrani et al. 2022; Hengstler et al. 2016; Capestro et al. 2024; Gan und Lau 2024; Daly et al. 2025; Agbabiaka et al. 2025). Existing theoretical and conceptual approaches have examined the meaning, antecedents, and consequences of trust in AI, and have identified a large range of human- and AI-related factors that shape trust, often moderated by contextual circumstances (Kaplan et al. 2023). Here, empirical studies suggest that trust in AI develops differently among actors depending on their roles, responsibilities, and levels of involvement in AI implementation (Daly et al. 2025; van der Werff et al. 2021; Lockey et al. 2021).

While most scholarly attention has focused on the trust of decision-makers – those actors who determine the extent to which AI is adopted and used – less attention has been paid to those stakeholders who are exposed to and affected by these emerging AI-supported decision-making processes. We argue that in such contexts, it is not only the AI decision aid itself – its process, performance, and purpose – that is assessed for trustworthiness, but also the existing constellation of decision-makers and institutions – rules, regulations, and practices – that govern decision-making. In other words, we argue that trust in AI is socially embedded and therefore highly context-dependent: Trust is developed and experienced within specific institutional settings, regulatory cultures, and histories of technology adoption. In other words, *who* trusts *what* AI, *where*, and *under what* institutional arrangements becomes the central analytical question. This becomes particularly relevant for AI systems that are to be adopted in transition processes with highly context-specific structural conditions and differing institutional dynamics. For the purpose of our research, we thereby follow the call of the EU High-Level Expert Group on Artificial Intelligence (AI HLEG) (2019, S. 5) that

> *"(…) it is not simply components of the AI system but the system in its overall context that may or may not engender trust. Striving towards Trustworthy AI hence concerns not only the trustworthiness of the AI system itself, but requires a holistic and systemic approach, encompassing the trustworthiness of all actors and processes that are part of the system's socio-technical context."*

Such a focus is crucial for understanding trust in AI decision aids for public administration, where decisions concern a variety of individual and organizational, private and public sector stakeholders, as well as an array of laws, standards, and best practices (Bullock 2019; Berman et al. 2024). However, current literature on trust in AI lacks a conceptual foundation for analyzing this social embeddedness of trust in AI. To address this gap, we draw on sociological theories that emphasize how trust is multifaceted and context-dependent (Rijswijk et al. 2023; Solberg et al. 2022) and how it unfolds in concrete social relationships (Latusek und Cook 2012; Schilke et al. 2021). This relational understanding of trust maintains that when actors cooperate in pursuit of a common task and – often inevitably – need to rely on one another, their relationships are embedded in broader social structures – that is, in a web of social relations and institutions (Granovetter 1985; Bachmann und Kroeger 2017; Bachmann und Inkpen 2011). Concretely, this means that when actors assess the trustworthiness of decision-makers and decision-guiding institutions, past experiences and prior interactions with this particular social environment play a crucial role, denoted as the 'shadow of the past' (Poppo et al. 2008).

The aim of this paper is to make this relational perspective productive for conceptualizing trust in AI as embedded in pre-existing relations with decision-making actors and decision-guiding institutions. We, therefore, address the following research question: *How do existing trust relations to decision-makers and institutions shape stakeholders' (mis)trust in AI decision aids for public administration?*

To address this question, we develop a conceptual framework that views trust in AI decision aids as embedded in a dynamic network of pre-existing trust relations – a 'shadow of the past'. We apply this framework to an empirical case on the early-stage implementation of an AI-augmented geographic information system (AI-GIS) designed to support the identification of areas for onshore wind energy development in Germany. This tool has been developed as a technological solution to accelerate energy transition processes by better attending to the place specificity of wind energy spatial planning. Decision-making in spatial planning for wind energy falls under the responsibility of public administration at the level of federal states and is guided by numerous – and often contested – regulations, marking this field as historically conflictual (Tucci 2022; Katzner et al. 2019; Kiunke et al. 2022; Nowak et al. 2023). In so-called potential area analyses, experts from public planning offices seek to identify low-conflict zones with high suitability, guided by both expert judgment and regulatory constraints. Final site designation decisions have implications for a range of stakeholders, including land owners, municipalities, project planners, plant operators, nature conservationists, and local residents.

For our analysis, we conducted 38 interviews across seven stakeholder groups, each with distinct trust relations and diverging degrees of involvement in the decision-making process. These interviews enable us to analyze how (mis)trust in the AI decision aid is shaped by stakeholders' prior trust relations with the institutions and decision-makers involved.

Our paper contributes to a deeper understanding of the contextual and relational nature of trust in AI. While pre-existing trust relations are often individual and situational, our dataset allows us to identify recurring patterns that can be conceptualized and empirically substantiated. Concretely, we apply qualitative comparative analysis (QCA) to identify configurations of existing trust relations with decision-makers and institutions that condition trust in the AI decision aid. Based on these configurations, we derive four types of distinct trust relations – each reflecting a unique pattern of how trust in AI is shaped by prior experiences in real-world decision-making contexts. Beyond contributing to debates on trust in AI, our argument also speaks directly to conceptual discussions in the sustainability transitions literature. By examining how pre-existing trust relations to decision-makers and institutions shape the reception of AI decision aids in contested energy planning processes, we contribute a relational and place-sensitive perspective to these debates. In particular, we highlight trust as a key social mechanism through which novel digital technologies become integrated into – or resisted within – established institutional arrangements, thereby extending the lens to the emerging role of AI in sustainability transitions.

In the remainder of the paper, we first review the state of research on trust in AI and introduce the sociological literature on relationality and social embeddedness of trust (Section 2). We then synthesize these insights and develop our conceptual framework (Section 3). We introduce the details of our empirical case and methodological approach (Section 4) and present the empirical findings (Section 5). We conclude by discussing the contributions and practical implications of our study as well as outlining limitations and directions for future research (Section 6).

2. State of research: Trust in AI and the social embeddedness of trust

In order to investigate how existing trust relations with decision-makers and institutions shape actors' (mis)trust in an AI decision aid, we begin by reviewing the relevant literature on trust in AI (Section 2.1). We examine three core dimensions conceptualized and studied concerning human-AI trust relationships: the AI (→ trustee), the human actor who places trust (→ trustor), and the context of their relationship[1]. In Section 2.2, we examine research on the social embeddedness of trust. This research serves as the theoretical foundation for our conceptualization of 'Trust in AI' as a relational, individualized, and context-dependent concept.

2.1. Factors influencing (mis)trust in AI

Trust is a central aspect of the relationship between humans and technologies (Omrani et al. 2022; Hoff und Bashir 2015; Schoorman et al. 2007). Because the processes and underlying logics of technologies – including AI algorithms – often remain incomprehensible and intransparent for human actors (Barredo Arrieta et al. 2020; Baker and Xiang 2023), their reliance on such technologies for decision-making largely depends on trust. As such, trust serves as a mechanism to cope with and reduce social and technical complexity (Daly et al. 2025; Schilke et al. 2021). Early concepts of trust in technology and automation, therefore, associate trust with a technology's perceived usefulness, particularly in situations involving uncertainty and vulnerability (Davis 1989; Lee und Moray 1992; Hoff und Bashir 2015; Lee und See 2004).

Building on this, studies on trust in AI draw on sociological and social psychological accounts that define trust as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the

[1] The literature on trust in AI spans a wide range of applications and does not always systematically differentiate between types of AI. In our review, we follow the classification proposed by Glikson and Woolley (2020), which distinguishes between robotic, virtual, and embedded AI. While our primary focus is on embedded AI – namely, AI systems used as decision-support, we also include relevant literature that extends beyond this specific focus.

ability to monitor or control that other party." (Mayer et al. 1995, S. 712). Applying these core principles of interpersonal trust to trust in AI highlights that trustors are – at least to some extent – willing to depend on AI technologies, because they expect some gains or positive outcomes, even though such dependence entails risks and a certain loss of agency and control (Glikson und Woolley 2020; Schoorman et al. 2007; Mayer et al. 1995). This means that while the technological potential of AI offers different gains in the form of decision-support – e.g., impact assessment, optimization, or prediction (Trunk et al. 2020) – human actors also take considerable risks due to limited control over algorithmic outcomes (Solberg et al. 2022) and ultimately remain responsible for the consequences of those decisions (Shrestha et al. 2019). Like in interpersonal trust relationships, these aspects of gains and risks in relationships between humans and AI technologies require the active decision to trust or mistrust of the trustors (those who lend trust). In the context of trust in AI, this decision is not only affected by humans' personality traits, characteristics, and attitudes, but also by the attributes of the AI system and the specific interaction context.

Accordingly, the existing literature has identified a variety of factors that contribute to humans' trust in AI. These factors are closely interlinked and can be categorized into (i) individual-level characteristics of the human trustor, (ii) specific attributes of an AI-based system as a potential trustee, and (iii) contextual factors that moderate human-AI relationships (Kaplan et al. 2023).

Studies on human-related antecedents of trust in AI have examined a variety of factors, including human trustors' perceived competence, expertise, or experience with an AI (Omrani et al. 2022; Kaplan et al. 2023; Hoff und Bashir 2015), individual attitudes towards AI (Daly et al. 2025), as well as relatively stable personality traits that shape an individual's propensity to trust (Küper und Krämer 2024). Given this variety, Glikson und Woolley (2020) propose a clustering that distinguishes between factors influencing cognitive dimensions of human trust in AI – i.e., rational evaluations of AI's trustworthiness (Schoorman et al. 2007) – and emotional dimensions of trust – i.e., emotion-driven and affect-driven perceptions of a technology (McAllister 1995). Accordingly, they derive how AI-related attributes like transparency, reliability, and accountability can foster cognitive trust in AI (Glikson und Woolley 2020; Ryan 2020; Floridi et al. 2018). In this context, prominent research agendas like 'Explainable AI' (Baker und Xiang 2023; Barredo Arrieta et al. 2020) have prompted empirical analyses (Sullivan et al. 2022) into "whether users are willing to take factual information or advice and act on it" (Glikson und Woolley 2020, S. 631). Factors relating to emotional dimensions of trust cover aspects like tangibility, anthropomorphism, and the interactive, human-like abilities of an AI (Glikson und Woolley 2020). This existing line of research assumes that adopting AI-based technology requires a 'leap of faith' as well as some level of affection or felt comfort on behalf of the human trustor (Hoff und Bashir 2015). Finally, existing studies also highlight the relevance of contextual factors for shaping the directional relation between human trustors and AI trustees. These factors constitute the 'context of AI use' (Daly et al 2025) and include task-related factors, such as whether humans perceive the tasks at hand to be risky and complex (Kaplan et al. 2023), important or trivial (Siau und Wang 2018), technically or socially oriented (Glikson and Woolley 2020), or whether the use of AI is voluntarily or mandatory (Lichtenthaler 2020) Some studies also conceptualize the context of human-AI relations more broadly, for instance, by examining individuals' perceptions of the innovating firms' reputation behind the AI technology (Hengstler et al. 2016), or institutional factors that structure and regulate AI usage (Siau und Wang 2018).

Taken together, human-related antecedents of trust as well as the specific attributes of AI-based systems offer profound insights into the complex trust relations between human trustors and AI trustees. However, two related gaps in the literature persist in relation to our research agenda. The first gap pertains to human-related antecedents of trust in AI. Although there is increasing evidence that trust develops differently among actors depending on their roles, responsibilities, and levels of involvement in AI implementation (Daly et al. 2025; van der Werff et al. 2021; Lockey et al. 2021), most scholarly attention continues to focus either on actors who have the agency to decide on the (non-)usage of AI – e.g., organizational employees in firms or public agencies (Lichtenthaler 2020; Berman et al. 2024) – or on experimental data featuring hypothetical evaluations of AI trustworthiness across domains and applications (Aoki 2021; Araujo et al. 2020; Hoff und Bashir 2015). Yet, to better understand the conditions under which trust or mistrust in AI systems used in public administration develops (Bao et al. 2025), we also need to consider other actors who are directly exposed to and affected by emerging AI-supported decision-making processes as trustors. This requires empirical research that investigates how trust in AI develops across individual and organizational actors in concrete decision-making contexts (Glikson & Woolley 2020) and how these actors' perceptions shape trust in practice (Berman et al. 2024; Christin 2017).

Building on this, the second research gap concerns how the context of human-AI interaction is currently conceptualized. As shown above, context has primarily been studied in terms of the immediate decision-making tasks that are to be delegated to AI systems. Yet, if we broaden the perspective to examine trustors affected by AI decision-making, we must also expand our understanding of the context in which human-AI trust relations are embedded. Following Kitchin, S. 1 (2017), in practice, the works of algorithms "unfold contextually and contingently" and are embedded in complex socio-technical assemblages – that is, in the actor dynamics, practices, regulations, and technological functionalities that shape concrete decision-making processes.

As such, AI decision aids are not implemented in a vacuum, and, therefore, need to be studied as embedded in concrete social contexts (Shrestha et al. 2019). In organizational as well as public administration settings, this means AI decision aids do merely suggest particular decision-making solutions – such as improved knowledge management or more systematic decisions – but also interact with and potentially disrupt existing practices, approaches, or even human decision-makers previously responsible for such tasks (Trunk et al. 2020). They essentially 'enter' established structures, actor dynamics, and institutional frameworks and must integrate into, and often compete within, pre-existing constellations of actors and institutions (Shrestha et al. 2019; Murray et al. 2021). Hence, in such contexts, it is not only the AI decision aid itself (e.g., its process, performance, and purpose) that is assessed for trustworthiness, but also the existing constellation of decision-makers and institutions – rules, regulations, and practices – that govern decision-making. In short, we need conceptual approaches that understand the social embeddedness of trust in AI.

To address these research gaps, we draw on sociological accounts of trust that emphasize its embeddedness in broader social relations (Schilke et al. 2021; Granovetter 1985). We elaborate on this approach in the next section and use it to inform our conceptual framework in Section 3.

2.2. The social embeddedness of trust of in AI

The previous section has established that AI decision aids are not implemented in a vacuum. Rather, they enter into and interact with established structures, practices, and institutional dynamics – often disrupting or competing with existing actors, approaches, and decision-making frameworks (Trunk et al. 2020; Shrestha et al. 2019; Murray et al. 2021). Consequently, we have emphasized that the development of trust in AI decision aids needs to be examined in relation to these existing constellations. For this purpose, we now turn to sociological theory that conceptualizes trust as an inherently relational phenomenon – emerging from the social interaction of individual actors within specific contexts and directed towards particular actions (Cook 2005; Schilke et al. 2021) – and derive conceptual implications for studying trust in AI.

The relational understanding that trust is "rooted in concrete social relationships" (Latusek und Cook 2012, S. 513) differs fundamentally from 'generalized' notions of trust, which view trust to be determined by actors' dispositions, personality traits, or cultural influences (Hardin 2002; Gheorghiu et al. 2009; Schilke et al. 2021). Instead, the notion of relational trust maintains that when individuals choose to interact or collaborate in pursuit of a common objective, the building of trust among them is shaped by their interconnected relationships as well as by the norms, practices, and regulations that guide their interactions. In short, trust is embedded in wider social structures – that is, in a web of interpersonal relations and institutional frameworks (Granovetter 1985; Bachmann und Kroeger 2017; Bachmann und Inkpen 2011). Following Granovetter (1985), the social embeddedness of trust in interpersonal relations means that actors' trust in each other builds over time and emerges from the stability of the underlying social structures, because when actors are embedded in a web of interpersonal relationships, there is a strong incentive to act reliably and cooperatively. Moreover, trust is also embedded in and shaped by the presence or absence of institutions, i.e., laws, rules, norms, shared values, and expectations. They are a decisive factor as they define the scope of acceptable behavior so that trustors can expect trustees to act in accordance with these 'rules of the game' (Scott 2014). These expectations of trustors are, in turn, based on their trust in the legitimacy of these institutional structures. Rijswijk et al. (2023, S. 2) note that trust in institutions differs from interpersonal trust „[…] as it refers to generalized trust placed in abstract systems, a feeling of taking for granted that they will function as they always do." Studies highlight that the stability and reliability of institutions may either reduce the need for interpersonal trust among actors or facilitate perceptions of reliability and trust among them (Bachmann und Inkpen 2011). In turn, the very absence of institutional stability has been proven to be detrimental for interpersonal trust (Latusek und Cook 2012). However, this instability may also have ramifications for the trust placed in institutional structures themselves. Accordingly, an increase in institutional instability leads to actors placing less reliance on the validity of such 'objectified' structures (Bitektine und Haack 2015). Instead, these actors increasingly rely on individual assessments of the appropriateness of institutional structures that guide a specific social context [Reference under review; Blind for Review].

Building on this line of research as well as on our argumentation in Section 2.1, AI decision aids, too, are socially embedded – and so is the trust in them. This is because AI systems do not constitute external 'things' or 'technologies', but their adoption is closely interwoven in practices and regulations as well as negotiated and differently understood by users and society (Bareis 2024; Suchman 2023). Human-AI interactions, thus, are embedded in a complex socio-technical system, involving not only the actual AI systems but also a variety of different stakeholders and institutional frameworks (Bareis 2024). For Eschenbach (2021, S. 1618), this perspective implies that the critical questions "then become whether one can trust the socio-technical system of AI (…), and whether the constituent members of this system are trustworthy". Similarly, the EU High-Level Expert Group on Artificial Intelligence (AI HLEG) (2019, S. 5) acknowledges that trust in AI "concerns not only the trustworthiness of the AI system itself, but requires a holistic and systemic approach, encompassing the trustworthiness of all actors and processes that are part of the system's socio-technical context". Taken together, as AI systems become socially embedded – that is, integrated into established socio-technical systems – this embeddedness must be a central consideration in analyses of trust in AI. First, this requires to conceptually grasp the socio-technical system as the 'context of human-AI interactions' more systematically (see Section 2.1). Second, it necessitates broadening the understanding of trust itself as an individualized, yet 'embedded' concept (Bachmann 2011). While trust depends on the individual experiences and assessments of trustors with a set of trustees in a concrete social context (Cook 2005; Schilke et al. 2021), trustors' behavior and willingness to trust others are influenced by broader institutional environments (Bachmann 2011).

The literature on trust offers various perspectives on how trust develops within these structures. One such perspective highlights the significance of the 'shadow of the past' – how trustors' experiences with a (socio-technical) context influence their current and future trust assessments (Poppo et al. 2008). Trustors use their accumulated experiences to extrapolate and estimate trustees' behaviors in the future: "Accumulated relationship experience from the past offers important cues to the kind of behavior expected from the trustee" (Schilke et al. 2021, S. 244). In other words, the "history of prior relations and interactions" (Poppo et al. 2008, S. 39) matter for trustors when assessing the trustworthiness of other actors. Given that the development of trust in other actors is shaped by institutional arrangements (Bachmann und Inkpen 2011), those arrangements, too, are subject to trustworthiness assessments in the 'shadow of the past'. However, with a few exceptions (Nourani 2023; Tsuchiya 2024), actors' past experiences with a given socio-technical system and the effect on their willingness to interact and trust AI have not been thoroughly investigated.

To sum up, a perspective on how trust in AI is embedded in a socio-technical system offers a relevant extension of the current literature, which often conceptualizes trust in AI as a characteristic of the trustor-trustee relationship with limited consideration of the interaction context. A broader view is particularly important for understanding the complex socio-technical environments in which AI decision aids are deployed within public administration. Understanding how affected stakeholders assess the trustworthiness of AI that is integrated into these systems is crucial. Yet, the literature lacks clear operationalizations and conceptual frameworks that allow for empirical exploration of this embedded trust. Section 3, therefore, proposes such a framework.

3. Conceptual framework: Impact of existing trust relations on trust in AI decision aids

Our framework operationalizes trust in AI as an individualized yet embedded concept: While we understand trust in AI as individualized assessments made by human trustors, we emphasize how their individualized experiences with the socio-technical system in which an AI is to be embedded – and thus their trust in this pre-existing system – in turn shapes their trust in AI. More concretely, this means that their existing trust relations with decision-making actors and decision-guiding institutions in concrete decision-making contexts ('shadow of the past') become a central reference for trustworthiness assessments of AI decision aids (see Figure 1). To this end, we integrate insights from research on trust in AI (Section 2.1) and relational perspectives on the embeddedness of interpersonal trust (Section 2.2).

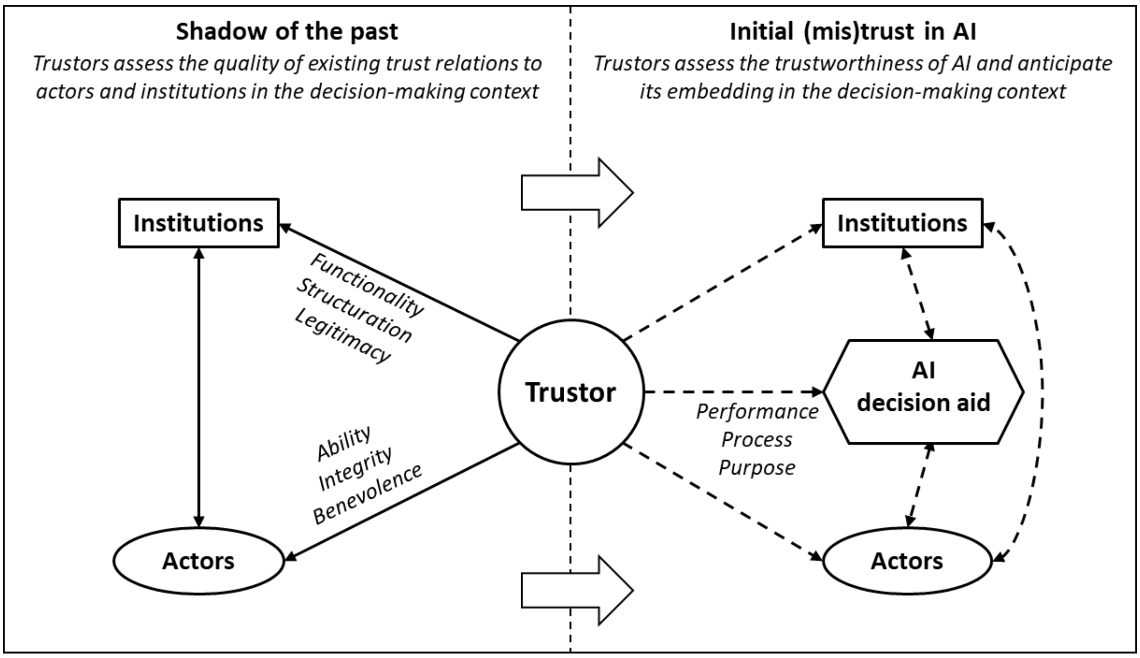


Figure 1: *Framework for analyzing the Impact of existing trust relations on trust in AI decision aids (own figure)*

In our framework, trustors, i.e., those stakeholders who are exposed to and affected by emerging AI-supported decision-making processes, are at the center of analysis. When trustors assess AI decision aids, the 'shadow of the past' (left side of the framework in Figure 1) reflects their existing relationships within the concrete decision-making context before the AI decision aid had been introduced; it involves existing trust relations with decision-making actors and decision-guiding institutions. The quality of these existing relations depends on the trustors' experience with the trustworthiness of these actors and institutions, as well as their interplay (Bachmann und Inkpen 2011; Cook 2005). Building on Mayer et al. (1995), we conceptualize trust in actors as being based on the judgment of their *ability*, *benevolence*, and *integrity*. To conceptualize the past trust relation to institutions, we transfer this tripartite division and combine it with insights from institutional theory to conceptualize the trustworthiness of institutions as based on assessments of their *functionality*, *structuration*, and *legitimacy* (Deephouse und Suchman 2008; Bachmann und Inkpen 2011)

For the actual assessment of AI trustworthiness, our framework focuses on *initial* trust in AI. Initial trust denotes a level of trust that trustors place in AI systems before they have had any direct experience or interaction with it.[2] Especially in early phases of AI implementation, initial trust plays a decisive role (Cabiddu et al. 2022). Glikson and Woolley (2020) maintain that there remains a need for non-laboratory field research that investigates conditions for initial trust in AI within more complex (organizational) environments. Hence, following our argumentation, past experiences – in all forms – should therefore be systematically studied as relevant factors for determining initial trust in AI-based decision aids.

The initial (mis)trust in AI (right side of framework in Figure 1) then results from trustors' anticipations about how and to what extent the concrete decision-making context would change when the AI decision aid enters as an alternative or complementary trustee. For this purpose, on the one hand, the trustor assesses the trustworthiness of the AI. Following Solberg et al. (2022), the three dimensions to capture the trustworthiness of AI decision aids considered here are *performance*, *process*, and *purpose.* On the other hand, trustors embed this assessment about the trustworthiness of the AI decision aid in their existing web of existing trust relationships with actors and institutions.

In the following, we discuss how the individual components of the framework are conceptualized in more detail.

3.1. The 'shadow of the past' – Trustors' assessment of existing trust relations

The 'shadow of the past' reflects the trustors' web of existing trust relations to actors and institutions in the decision-making context in which the AI decision aid is to be introduced. Following relational perspectives on trust (see Section 2.2), this web depends on trustors' situated experiences, and so trust relations differ in their quality. Consequently, in order to capture the 'shadow of the past', trustors must have gained experience with the trustworthiness of decision-making actors and decision-guiding institutions in the concrete decision-making context, from which the quality of their trust relationships can be deduced. The framework acknowledges that trustors can be involved in the current decision-making processes to varying degrees – from direct involvement to merely being affected by

[2] We acknowledge that different conceptual understandings of initial trust in AI exist. While some studies frame it as 'first impressions' that are derived intuitively and without engaging much with the AI or the context of usage (Siau und Wang 2018), we follow studies that emphasize the development of initial trust as a critical phase during actors' early encounters and experiences with an AI (Cabiddu et al. 2022; Glikson und Woolley 2020). Here, research shows that also initial trust is based on a variety of factors, including AI-based and human-based factors.

the outcomes. Crucially, what matters is that they have gained direct experiences with the actors and institutions that shape these processes, which also informs their perceptions of AI decision aids introduced in such contexts.

To conceptualize the quality of trust relations to decision-making actors, we build on Mayer et al. (1995) and distinguish three dimensions that capture how trustors assess the trustworthiness of other actors: *ability*, *benevolence,* and *integrity*. *Ability* refers to trustors' judgments of whether trustees have the necessary skills to perform the decision-making. Ability consists, for example, of trustees' knowledge, personal qualities, skills, and social position. *Integrity* refers to the perceived reliability of trustees to behave and decide as they say they will. This is related to the extent to which trustees adhere to normative values and principles, such as fairness and honesty. *Benevolence* refers to trustees' motivation and willingness to help, i.e., whether trustees are perceived as loyal and helpful and, therefore, seen as willing to take on the task of decision-making. Depending on trustors' role and social position in the decision-making process – for instance, as advisors, decision-makers, or stakeholders affected by the outcome – these roles will be reflected in the assessment of the trustworthiness of the decision-making actors, either positively or negatively.

Our framework also builds on the recognition that trustors place diverging degrees of trust in institutions, based on their experience with their trustworthiness in guiding and regulating the decision-making process (see Section 2.2). To capture these diverging trust relations to institutions, we conceptualize trustors' relation to institutions as dependent on their individual assessments of institutions' trustworthiness. To conceptualize the quality of trust relations with decision-guiding institutions, we build on neo-institutional theory (Deephouse und Suchman 2008; Suchman 1995) as well as on the institutional trust research (Bachmann und Inkpen 2011; Bachmann 2011; Rijswijk et al. 2023) to distinguish three dimensions of the trustworthiness of institutions in analogy to the well-established dimensions of interpersonal trust described above (Mayer et al. 1995). Accordingly, trustors assess institutions based on their *functionality*, *structuration*, and *legitimacy*. Trust in the *functionality* refers to trustors' assessment that the institution is performing its designated tasks in guiding the decision-making, i.e., that it functions as it should. Trust in the *structuration* of institutions refers to the extent to which the institution is judged to be stable, reliable, and consistent enough to guide decision-making. Particularly as a result of institutional change and in connection with technological innovations, it is possible that for instance laws do not (yet) appear differentiated enough to guide decision-making properly. Trust in the *legitimacy* of institutions is based on the assessment that the institution is proper, or appropriate for guiding decision-making in the specific context and, therefore, worthy of maintaining.

We acknowledge that trust in actors and trust in institutions are conceptually and empirically intertwined. Assessments of actors' integrity, for instance, are influenced by whether actors adhere to normative values and principles, i.e., institutional frameworks. Likewise, trust in the functionality and structuration of institutions is not entirely detached from trust in the actors that implement and enforce those institutions. The framework recognizes this dynamic interplay but keeps both dimensions distinct to ensure conceptual clarity.

Taken together, the core assumption of our framework is that trustors' experiences with the trustworthiness of decision-making actors and decision-guiding institutions in a concrete context ('shadow of the past'), impact their initial trust in an AI decision aid. The next section explains how this influence is conceptually captured in our framework.

3.2. Initial (mis)trust in AI decision aids – Embedding in concrete decision-making contexts

The framework conceptualizes trustors' initial (mis)trust in AI decision aids as influenced by what we have previously described as the 'shadow of the past'. Accordingly, we propose that trustors' experiences with decision-making actors and decision-guiding institutions shape their initial (mis)trust in an AI decision aid. Yet, our framework posits that these pre-existing trust relations do not fully determine trust in AI decision aid. On the one hand, existing research on the 'shadow of the past' has revealed both direct, indirect, and non-linear influences on the formation of trust in other fields (Poppo et al. 2008; Gulati und Sytch 2008). On the other hand, existing research on trust in AI has established how trustors' assessments of AI (decision aids) depend on a range of AI-related factors (see Section 2.1). Our framework is therefore grounded in the understanding that such assessments of AI trustworthiness are informed both by trustors' existing relationships with decision-making actors and decision-guiding institutions, and by their anticipations regarding the ability and functions of the AI itself.

To capture the elements that come into play when assessing the trustworthiness of an AI decision aid, we employ three dimensions that categorize and differentiate these elements (Solberg et al. 2022): *performance*, *process*, and *purpose*. These dimensions are derived from the well-established model on interpersonal trust – ability, integrity, and benevolence (Mayer et al. 1995) – and are adapted to reflect the peculiarities of AI systems. Accordingly, *performance*-based trust refers to the assessment that the AI decision aid performs its task "capably and reliably" (Solberg et al. 2022, S. 196) and thus refers to the ability-dimension of interpersonal trust. *Process*-based trust in AI refers to trustors' understanding of the underlying algorithms, methods, and functions of the technology. Hence, process-based trust builds on trustors' knowledge about what the AI decision aid "[…] does, why it does it, and how it works." (Solberg et al. 2022, S. 196). Here, central aspects such as transparency and explainability of AI are subsumed. The third dimension, *purpose*-based trust, reflects what is referred to as "benevolence" in interpersonal trust (Mayer et al. 1995). It involves the assessment of the AI's intended purpose and the rationale behind its development (Solberg et al. 2022). When evaluated positively, trustors believe that the AI decision aid serves their interests and objectives – for instance by simplifying the decision-making or making it more efficient.

Based on these AI-related assessments and trustors' 'shadow of the past', trustors form expectations about how the AI decision aid will function within the existing decision-making context and how it will interact with the decision-makers and institutions guiding the process. In this way, trustors anticipate the embedding of the AI decision aid in the concrete decision-making context (illustrated by dashed lines in Figure 1), which results in either initial trust or mistrust in the AI.

By incorporating the 'shadow of the past', our framework enables the analysis of initial (mis)trust in AI decision aids in relation to trustors' individual experiences with decision-making actors and institutions in a specific context. Thus, while we continue to conceptualize trust as an individualized phenomenon, our framework highlights how it is embedded in and shaped by prior interactions with broader social structures.

4. Methodology

We apply our framework to analyze the initial (mis)trust in an AI-augmented geographic information system (AI-GIS), developed to assist in the decision-making process for spatial planning of wind energy development in Germany; more concretely, for identifying and designating potential areas for wind energy plants. In the following, we describe the characteristics of this case, outline the empirical data used, and introduce the two empirical methods applied for our analysis.

4.1. The case of an AI-GIS decision aid for wind energy spatial planning

As the empirical setting for our study, the initial (mis)trust in an AI decision aid is investigated, which is currently gaining traction in the wind energy sector: AI-augmented Geoinformation systems (AI-GIS) that can support decision-making on identifying and designating sites for onshore wind energy (see info box below). We chose this setting because decision-making in spatial planning for wind energy currently falls under the responsibility of public administration at the level of federal states and is guided by numerous – and often contested – regulations (Tucci 2022; Katzner et al. 2019; Kiunke et al. 2022; Nowak et al. 2023). The implementation of the AI-GIS as decision aid, thus, intervenes in the structure of existing decision-making actors and institutions and adds to a highly conflictual decision-making context:

With Germany being politically committed to restructuring its energy system towards more sustainable forms of production and consumption, wind energy (onshore and offshore) plays a key role in replacing fossil fuel-based energy production. While these objectives are generally accepted, their implementation involves a number of controversial decision-making by a multitude of actors with different interests, strategies, resources, and skills (Farla et al. 2012). Decision-making in such sustainability transitions is, therefore, inherently political, as it touches on the question of who is authorized to decide what sustainable change means. As Fuenfschilling (2019) notes in this regard: "The reconciliation of different actor perspectives and interests towards [sustainability as] a normative, uncertain and contested goal is a difficult and time-consuming process." (Fuenfschilling 2019, S. 220)

The identification and designation of sites for wind power plants in Germany constitutes a programmatic case that is characterized by difficult and conflictual decision-making processes. These touch upon competing interests of various stakeholders, are regulated by strict institutional frameworks, and involve significant technical, environmental, and spatial challenges.

The difficulties with these decision-making processes lie in the availability of open land required for wind power plants. As this is limited, the expansion of wind energy competes with other interests, such as species and nature conservation, the military, or tourism – each of these interests being driven by different stakeholder groups. On an operational level, it is the task of spatial planners to reconcile these interests and anticipate potential conflicts while considering established institutional framework conditions that guide this decision-making. Concerning our conceptual framework, these spatial planners are the decision-making actors in this empirical case. In Germany, this role is institutionalized, meaning that spatial planning is in the hands of respective state authorities. The level of authority responsible for the decision-making process depends on the respective federal state. Annex 1 contains an overview of the responsible authorities in the five federal states considered in our analysis. The authorities build on their experience and established practices in spatial planning by weighing up the various (potential) conflicts of interest. Stakeholders who want to plan a wind farm (e.g., project developers or municipalities) are bound by the decision of the official spatial planning authorities and can only plan and build wind farms within the specified landscape framework plan.

In addition to the determined role of decision-makers, various institutions guide the identification and designation of sites for wind energy in Germany. These institutions are primarily manifested as laws and regulations (for instance, specifications on the minimum distances of power plants to residential buildings, nature and species protection requirements, or water protection zones). Also, more unconscious institutions, i.e., normative or cultural-cognitive, guide the decision-making process of spatial planning. These include, for instance, role expectations towards spatial planners and project developers, expectations regarding participation in the planning process, the implementation and consideration of (ecological, economic, and social) sustainability, or procedural and distributive justice [Reference under review; Blind for Review].

A peculiarity of spatial planning in Germany is that all federal states are obliged to make 2% of their land available for onshore wind energy (EEG (2023) 2022). Consequently, despite spatial differences, the federal states have to achieve the same targets, which on the one hand gives the impression of justice, but on the other hand, results in the need to adopt differing regulations in order to achieve these targets. More concretely, this means that, for instance, particularly densely forested federal states have to include forests in their wind energy planning, whereas others can explicitly exclude these areas from their planning. This is often subject to controversy as it shows the situatedness and institutional differences that have concrete consequences for spatial planning. Nature conservation and species protection, in particular, are repeatedly brought forward in this context to illustrate the different positions on this decision (Katzner et al. 2019; Kiunke et al. 2022). As a result, despite (seemingly) clear regulations for the decision-making process as well as the responsibility of government-authorized decision-makers, complications and controversies repeatedly arise concerning the designation of sites for wind energy and often lead to delays or cancellations of wind energy projects (Tucci 2022).

INFO BOX: The AI-GIS under investigation was developed as part of the state-funded collaborative project [Blind for Review] and presented as a prototype in 2024. For the data model underlying the AI-GIS, Germany was divided into 50x50 meter tiles, and for each tile, criteria were captured (e.g., the distance to the nearest residential area or the presence of wind energy-sensitive bird species). Based on these data, the AI-GIS determines for each tile how suitable it is for the construction of wind turbines. In total, over 60 criteria from the categories of wind speed, meteorology, water, landscape protection, aviation, nature conservation, grid infrastructure, settlement structure, transport infrastructure, forests, avifauna, and topography were integrated – 33 criteria as exclusion criteria, 57 criteria as evaluation criteria. The AI was then trained using data on existing wind farms as training examples for suitable tiles. Old (commissioned before 2010) and small (< 150m total height) wind turbines were excluded. In addition, the AI was trained with random tiles without wind farms as training examples for unsuitable tiles. The set of tiles that were randomly selected was limited to tiles that are highly likely to be unsuitable based on expert-based area scoring. The AI-GIS prototype offers the option of evaluating tiles as a heat map, as well as a detailed view for entire areas. It shows which features are particularly relevant for the evaluation of an area, allows the comparison of areas and the selection of the best sub-areas in large areas. Finally, changes to institutional framework conditions, such as laws, can also be simulated.

We follow a qualitative comparative approach to empirically analyze the initial trust of 38 stakeholders in this AI-GIS by applying the framework presented above (see Section 3).

4.2. Data selection

In our empirical analysis, we examine the initial trust of 38 stakeholders from seven stakeholder groups. Stakeholders from each group have been involved in or affected by at least one concrete decision-making process on the identification and designation of sites for onshore wind energy before the interview and the introduction of the AI-GIS. They thus all have inherently distinct trust relations with the respectively other actors and institutions involved in the decision-making process: Project developers and initiators (1), (local) politicians (2), state authorities for spatial planning, approval, and forestry (3), environmental ministries (4), (local) associations and citizens' initiatives (5), science (6), and other (7). Annex 2 contains an overview of the interviewed stakeholders, including their backgrounds. These different groups reflect the heterogeneity of trustors in the concrete decision-making context. The groupings further indicate the extent to which the interviewees themselves were or are involved in the decision-making process, which is a key variable in the empirical analysis. In addition to the different stakeholder groups, the heterogeneity of decision-makers and institutions that prevail in Germany is considered. For this reason, the stakeholders were also selected depending on where they have gained their experiences with spatial planning. Hence, five different federal states are considered: Lower Saxony, North Rhine-Westphalia, Hesse, Rhineland-Palatinate, and Bavaria. Annex 1 contains a table with information on the decision-making authorities in the respective federal states.

We conducted the stakeholder interviews in a semi-structured manner using a guideline in which all trust dimensions of the framework – interpersonal trust, trust in institutions, and trust in the AI decision aid – were addressed. The interviews were structured as follows: First, the interviewees were asked to explain their experiences with onshore wind energy and concrete sites and projects. Next, they outlined their experiences with other stakeholder groups as well as potential conflicts. The main focus here was on decision-makers (spatial planning authorities) and their competencies. Then, they were asked to reflect on institutional framework conditions relevant to selecting potential areas and wind energy projects. Here, three categories of institutions were discussed: regulatory (e.g., laws and regulations), normative (e.g., expectations of decision-makers' actions, fundamental values, and norms), and cultural-cognitive institutions (e.g., self-image of the communities, expertise). In particular, interviewees were asked to reflect on their perception of the functionality, structuration, and legitimacy of these institutions. Finally, the interviewees were asked for their evaluation of the prospects of using the AI-GIS decision aid as an assistance tool for identifying potential areas for wind energy. To this end, interviewees were given a brief description of the features, underlying data, and aims of the AI-GIS to illustrate its performance, process, and purpose. The interviewees were given the opportunity to ask follow-up questions for their assessment (which was used to varying degrees). Even though the interviewees were provided with extensive information about the AI-GIS, they did not have the opportunity to try out the tool at the time of the interview, so that the evaluation reflects their initial trust.

The interviews lasted an average of one hour and were audio-recorded and transcribed. While a data set of n = 38 interviews enables profound insights into the trust dynamics surrounding the complex decision-making of wind energy site selection, the richness of our data set (stakeholder groups, various trust dynamics prior to the AI-GIS introduction) makes a systematic identification of trust patterns more difficult. Therefore, investigating the interplay of past trust relations and the embeddedness of trust in AI requires methods that can grasp this complexity. Hence, to answer our research question, we analyze our data in two steps. First, we apply a qualitative comparative analysis (QCA) to identify patterns in the conditions for trusting the AI-GIS (Sections 4.3 and 5.1). We understand this as an exploratory step to elaborate on possible parallels between interviewees. Second, based on the findings of the QCA, we further examine the interplay of different trust relations in an in-depth qualitative content analysis (Sections 4.4 and 5.2).

4.3. Qualitative comparative analysis (QCA)

Qualitative comparative analysis (QCA) is a systematic method for analyzing case-related data. As a configurational comparative approach, QCA identifies patterns of conditions leading to specific outcomes (Ragin 1987). In social science research, various forms of qualitative data – including interview or focus group data but also secondary documents – are transformed into categorical or dichotomous variables that reflect the presence, absence, or degree of specific conditions. Hence, by applying QCA, we aim to identify combinations of conditions that are consistently associated with the *presence* or *absence* of 'Trust in AI'. These configurations can be interpreted as sufficient conditions in the sense of QCA logic, without deriving a causal effect.

Following our framework, we analyze 'Trust in Actors' (TrustAct) and 'Trust in Institutions' (TrustInst) as conditions ('shadow of the past') and 'Trust in AI' (TrustAI) and 'Mistrust in AI' (NoTrustAI) as the outcomes. We also added the condition 'Inclusion in Decision-making' (DecIncl). Based on the stakeholder groups, this condition measures the degree of their involvement in the decision-making process on identifying and designating sites for onshore wind energy. Hence, each interviewee represents a trustor with a distinct web of existing trust relations to decision-making actors and decision-guiding institutions. We analyze their initial trust and mistrust in the AI-GIS decision aid concerning these variables.

A crucial step in QCA is the calibration of data. In our case, this first involved converting our qualitative interview data into set membership scores (Ragin 2008). Our study employs a fuzzy-set QCA (fsQCA) approach where values range from 0 (clear mistrust) to 1 (clear trust), with intermediate stages of 0.25 (tends to mistrust), 0.5 (ambivalent trust relation), and 0.75 (tends to trust).[3] To create a truth table, we then carried out a mean-based calibration. Table 1 contains the definitions of all conditions and outcomes, as well as further information on the scale and calibration types.

Variable (Label)	**Conceptual Definition**	**Type**	**Original Scale**	**Calibration Type**	**Thresholds**
TrustAct	Trust relation to decision-making actors, based on experience with their trustworthiness (ability, integrity, benevolence)	Condition	0/ 0.25/ 0.5/ 0.75/ 1	Fuzzy set	0/ 0.728/ 1
TrustInst	Trust relation to decision-guiding institutions, based on experience with their trustworthiness (functionality, structuration, legitimacy)	Condition	0/ 0.25/ 0.5/ 0.75/ 1	Fuzzy set	0/ 0.478/ 1
DecIncl	Degree of involvement in the decision-making process of identifying and designating potential areas for wind energy	Condition	0/ 0.25/ 0.5/ 0.75/ 1	Fuzzy set	0/ 0.478/ 1
TrustAI	Initial trust in AI-augmented GIS, based on its anticipated trustworthiness and embedding in assisting the decision-making (performance, process, purpose)	Outcome	0/ 0.25/ 0.5/ 0.75/ 1	Fuzzy set	0/ 0.588/ 1
NoTrustAI	Initial mistrust in AI-augmented GIS, based on its anticipated trustworthiness and embedding in assisting the decision-making (performance, process, purpose)	Outcome	0/ 0.25/ 0.5/ 0.75/ 1	Fuzzy set	0/ 0.588/ 1

Table 1: *Definitions of conditions and outcomes for the fsQCA.*

Our analysis followed the typical procedure of a fsQCA, beginning with the construction of a truth table in which our cases were arranged based on their condition-outcome configurations. For this purpose, we cumulated one outcome variable from the three trustworthiness dimensions. This allowed us to consider the multidimensionality of trustworthiness assessments to some extent while the individual

[3] Except from the condition of 'DecIncl', where the same levels refer to the degree of involvement in the decision-making process.

dimensions are examined again in more detail in the qualitative content analysis (Sections 4.4 and 5.2). After generating the truth table, we applied Quine-McCluskey minimization to reduce it and identify sufficient pathways. The final step involved assessing consistency and coverage to ensure that the identified pathways reliably explain the presence or absence of the outcome and have sufficient explanatory power (Ragin 2008). Following the argumentation of Legewie (2022) on the application of QCA in exploratory studies, the consistency threshold was set to 0.8. The analysis was conducted using the R package "QCA for R. A Comprehensive Resource." by Dusa (2019).

4.4. Qualitative content analysis

The findings of the QCA were complemented and triangulated with an in-depth qualitative content analysis (Kuckartz und Rädiker 2023). Accordingly, the interviews were transcribed and systematically analyzed using the software MAXQDA. Here, we further investigated the dynamic interplay of trust relations underlying trust in the AI-GIS. While QCA reveals patterns of necessary and sufficient conditions, this more in-depth approach allowed us to explore the underlying mechanisms that shape trust in AI. To this end, we systematically analyzed the interview data regarding our analytical framework (see Table 2). To systematically capture the trust relations from our framework, a hybrid deductive-inductive coding strategy was employed: Codes were developed deductively based on our conceptual framework and subsequently enriched and refined inductively to account for the specificity of the concrete decision-making context.

Code	**Conceptual Definition**	**Example from our data set**
Act_Abil	Act_Abil is coded when the interviewee expresses their experience-based evaluation of the ability of decision-makers.	*"There are amateurs sitting there. […] They don't know anything about developing a planning area. They are not able to read a building code properly." (Trustor09, Mayor)*
Act_Integ	Act_Integ is coded when the interviewee expresses their experience-based evaluation of the integrity of decision-makers.	*"And we've had a really, really good working relationship ever since. It's an exceptionally good cooperation, it's almost friendly. […] Because we were always honest with each other, we were always honest." (Trustor20, Project developer)*
Act_Benev	Act_Benev is coded when the interviewee expresses their experience-based evaluation of the benevolence of decision-makers.	*"I always appreciated the cooperation with the district. [...] The district itself also wants wind energy, so it supports the goal of climate change." (Trustor09, Mayor)*
Inst_Funct	Inst_Funct is coded when the interviewee expresses their experience-based evaluation of the functionality of institutional structures, guiding the decision-making process.	*"This is a rape of the law. It does not actually do justice to the intention of the legislator. […] These are all just things to deliberately destroy or at least obstruct a project." (Trustor12, Project developer)*

Inst_Struct	Inst_Struct is coded when the interviewee expresses their experience-based evaluation of the structuration of institutional structures, guiding the decision-making process.	*"I always wished for much better support from the legislature. Clearer regulations, better statements on the subject of accelerated procedures, and all these issues. So, there's a lot that could be improved in terms of legislation to implement the whole thing if you want to – after all, it's our declared goal that we want the energy transition." (Trustor09, Mayor)*
Inst_Legit	Inst_Legit is coded when the interviewee expresses their experience-based evaluation of the legitimacy of institutional structures, guiding the decision-making process.	*"When the federal government drew up this whole Wind Energy on Land Act, they put a lot of effort into the expertise. There were various meetings with relevant specialist institutions, associations, and practitioners." (Trustor36, Ministry, Spatial planning)*
AI_Perform	AI_Perform is coded when the interviewee expresses their expectations of the performance of the AI-GIS decision aid in assisting the decision-making process.	*"Yes, I could imagine that artificial intelligence could immediately show the planners whether their locations affect any biotopes or nature conservation areas, so that we could simply have this delivered from the outset." (Trustor39, Nature Conservation Ass.)*
AI_Process	AI_Process is coded when the interviewee expresses their expectations of the process underlying the AI-GIS decision aid in assisting the decision-making process.	*"It depends very much on how you feed the system with information. In my opinion, it's not enough to just enter partial pieces somehow, but it should actually be done on a very, very large scale." (Trustor35, Ministry)*
AI_Purpose	AI_Purpose is coded when the interviewee expresses their expectations of the purpose of the AI-GIS decision aid in assisting the decision-making process.	*"It might be quite helpful if we had a neutral scheme like artificial intelligence, where the emotions are outside so that we can work through something like this in a purely objective way." (Trustor26, Project developer)*

Table 2: *Code system for analyzing initial trust in the AI-GIS as shaped by existing trust relations.*

Based on this coding system, the results of the QCA were deepened by analyzing the qualities of trust relations. Particular emphasis was placed on how past trust relations ('shadow of the past') interacted with expectations of the AI-GIS' trustworthiness. In the following Section 5, the results of the two analytical approaches are presented.

5. AI-GIS decision aid for wind energy spatial planning: The embeddedness of initial trust in existing trust relations

As outlined in Section 3, we conceptualize trust in AI as shaped by trustors' past trust relations to other actors and institutions within a specific decision-making context. Accordingly, we now analyze if and how initial (mis)trust in an AI-GIS decision aid not only depends on trustors' expectations of its trustworthiness but also on trustors' existing relation to other actors and institutions, and their experience with their trustworthiness in executing and guiding decision-making on wind energy spatial planning. Our findings, thus, map the interplay between past trust relations ('shadow of the past') and initial (mis)trust in AI. We first derive four configurations of conditions that lead to (mis)trust in the AI-GIS (QCA findings presented in Section 5.1). Based on these configurations, we explain and elaborate on the four types of trust relations in more depth (qualitative content analysis presented in Section 5.2).

5.1. Configurations leading to initial (mis)trust in the AI-GIS decision aid (QCA)

In this section, we present the results of the fsQCA and interpret configurational patterns that lead to either initial trust (*TrustAI*) or mistrust (*NoTrustAI*) in the AI-GIS assisting the decision-making of identifying and designating wind energy potential areas. The analysis was based on three core conditions: i) trust in actors, i.e., current decision-makers (TrustAct); ii) trust in institutions that guide and regulate the current decision-making process (TrustInst); and iii) the degree of trustors' involvement in the current decision-making process of identifying and designating potential areas for wind energy (DecIncl). The analysis included 34 trustors/ cases[4], each representing a stakeholder from one of the seven aforementioned stakeholder groups (see Section 4.2).

Configurations leading to trust in the AI-GIS decision aid

The intermediate solution model (M1) for the outcome *TrustAI* consists of two sufficient pathways. While both pathways were derived from configurations meeting the initial consistency threshold of 0.8, the resulting minimized configurations exhibit slightly lower consistency values due to logical minimization. Nevertheless, both pathways still demonstrate reasonably strong empirical support (see Table 3).

		incls	PRI	covS	covU	Cases (Trustors)
Config 1	~DecIncl*TrustAct → TrustAI	0.791	0.672	0.557	0.356	*10,11,19,22,32; 3,13,14,25,26, 30,34*
Config 2	DecIncl*~TrustInst → TrustAI	0.830	0.711	0.435	0.235	*4,20; 1,7,12,16,17, 18,21,24,28*
	M1	0.774	0.683	0.791		

Table 3: *Intermediate solutions for initial trust in the AI-GIS decision aid (TrustAI)*

Note: ~ indicates the absence of the condition; no algebraic sign indicates the presence of the condition.

Configuration 1, with a consistency of 0.791 and a unique coverage of 0.356, indicates that when trustors are not directly involved in the decision-making process (~DecIncl) yet have trust in the current decision-makers (TrustAct), they are more likely to trust the AI-GIS with assisting the decision-making of identifying and designating potential sites. This suggests a delegation dynamic where trust in decision-making actors compensates for a lack of their own participation. At the same time, it indicates that trustors who have gained overall positive experiences with delegating critical tasks to other actors

[4] Of the n=38 interviews, four were excluded from the QCA because they did not allow the data to be clearly coded according to the truth values.

continue to be willing to take the risk of trusting others – even AI agents. *Configuration 2*, with a higher consistency of 0.830 and unique coverage of 0.235, implies that involvement in the decision-making process (DecIncl), when coupled with a lack of trust in institutions (~TrustInst), can also lead to trust in the AI-GIS decision aid. In such cases, the AI-GIS might be perceived as a corrective mechanism or neutral support that offsets institutional shortcomings.

The overall solution coverage of M1 was 0.791, and the overall consistency was 0.774, indicating that the identified configurations explain a substantial share of the outcome and are empirically consistent.

Configurations leading to mistrust in the AI-GIS decision aid

A separate fsQCA was conducted on the negated outcome (NoTrustAI) to explore causal asymmetry. The intermediate solution model (M2) also consists of two sufficient pathways for explaining initial mistrust in the AI-GIS (see Table 4).

		incls	**PRI**	**covS**	**covU**	**Cases (Trustors)**
Config 3	~DecIncl*~TrustAct → NoTrustAI	0.849	0.741	0.409	0.270	*8,9,23,27,33*
Config 4	~TrustAct*TrustInst → NoTrustAI	0.953	0.878	0.250	0.110	*2,5,6*
	M2	0.859	0.777	0.519		

Table 4: *Intermediate solutions for the absence of initial trust in the AI-GIS decision aid (NoTrustAI)*

With a consistency of 0.849 and unique coverage of 0.270, *Configuration 3* suggests that trustors who are excluded from the decision-making process (~DecIncl) and simultaneously lack trust in current decision-makers (~TrustAct) are more inclined to mistrust the AI-GIS. This indicates that if the current decision-makers are not trusted and trustors are not involved, they are so disconnected from the current decision-making process that they do not trust an AI assistant to close the trust gap. *Configuration 4* has the highest individual consistency (0.953) and a unique coverage of 0.110. The path implies that when institutional trust is present (TrustInst) but trust in decision-making actors is lacking (~TrustAct), initial mistrust in the AI-GIS is likely. This indicates that mistrust in current decision-makers is so dominant that neither the trust in institutional structures nor the AI-GIS decision aid is expected to close this trust gap.

The overall solution consistency is 0.859, and coverage is 0.519, again demonstrating distinct causal mechanisms for trust versus mistrust.

Table 5 summarizes the four configurations of conditions for the presence or absence of trust in an AI-GIS. Filled dots indicate the presence of a condition, while unfilled dots indicate its absence. The findings underscore that trust in AI is not driven by a single condition but by configurations of trust relations and participatory experience. Our results also demonstrate three core features of configurational causality: *equifinality*, *asymmetry*, and *context sensitivity*.

Configuration	**DecIncl**	**TrustAct**	**TrustInst**	**TrustAI (Outcome)**	**Consistency**	**Coverage**
Config 1	○	●		✓	0.791	0.557
Config 2	●		○	✓	0.830	0.435
Config 3	○	○		✕	0.849	0.409
Config 4		○	●	✕	0.953	0.250

Table 5: *Intermediate solution table of configurations serving as conditions for the presence or absence of trust in the AI-GIS decision aid*
[Legend: ● = condition present, ○ = condition absent, [ ] = irrelevant]

First, the presence of multiple sufficient configurations for one outcome – either 'TrustAI' or 'NoTrustAI' – demonstrates *equifinality*, meaning that different combinations of conditions can independently lead to the same outcome. Second, the comparison between configurations that lead to TrustAI versus NoTrustAI confirms *asymmetry*: the causal paths resulting in the presence of trust are not simply the inverse of those producing mistrust. This highlights the complex and distinct mechanisms underlying the formation of initial trust and mistrust in AI decision aids. Finally, our analysis reveals *context sensitivity* as the effect of any given condition, such as trust in actors, is contingent upon its combination with other factors. For instance, trust in actors supports trust in AI only when institutional trust is low or inclusion in the decision-making is absent. This reinforces the configurational nature of causal influence.

While the fuzzy-set QCA has revealed distinct and empirically consistent configurations that lead to trust and mistrust in the AI-GIS decision aid, the method remains inherently cross-case and set-theoretic: It identifies patterns of association but does not capture the rich, within-case dynamics through which trust is interpreted, negotiated, and expressed by individual trustors.

To address this, the next step of our analysis involves a qualitative re-examination of the interview data for each of the configurational pathways identified above. This in-depth exploration allows us to validate the configurations by examining whether the theorized combinations of conditions are reflected in interviewees' narratives and reasoning. Through this qualitative lens, we aim to further contextualize the configurational findings, deepen the interpretation of key conditions, and move from causal sufficiency to interpretive plausibility.

5.2. Embeddedness of (mis)trust in AI: Four types of trust relations (qualitative content analysis)

The qualitative analysis reveals clear patterns in the reasoning of stakeholders – i.e., trustors – within each configuration identified by the QCA[5]. These patterns reflect how different forms of trust interplay and how initial (mis)trust in the AI-GIS is shaped by pre-existing relational structures ('shadow of the past'). While we observe similarities among trustors within the same configuration identified in Section 5.1, important differences emerge in their underlying motivations and criteria for evaluating trustworthiness. This heterogeneity suggests that the 'shadow of the past' cannot simply be equated with the configurations from the QCA.

In the following, we re-examine the four configurations on the basis of the interview data with regard to the mechanisms that underlie the 'shadow of the past' and the anticipated embedding of AI in it, to build four types of trust relations.

Relation Type 1 – Trust in AI embedded in interpersonal trust as backing up trusted decision-makers

In Configuration 1, trust in the AI-GIS appears to be strongly mediated by the past trust relation of the trustors with the decision-makers responsible for spatial planning. While the trustors have not been directly involved in the planning process (as indicated by the absence of 'DecIncl'), they nevertheless express initial trust in the AI system. Crucially, this trust is less grounded in the technology itself than in the perceived *ability*, *integrity*, and *benevolence* of the decision-makers who are about to employ it (e.g., Trustor32: 63). Hence, the interview data suggest that trustors tend to interpret the AI-GIS as a supporting tool – an assistant to the actors they already deem trustworthy. Here, the AI- system is

[5] Note that the QCA reveals the individual cases/trustors that are covered by each configuration, see Table 1.

rarely seen as an autonomous decision-maker but rather as a component that operates under the guidance of human judgment. This framing allows trustors to maintain continuity in their trust orientation: they extend trust to the AI-GIS because it is embedded in a familiar and positively evaluated context of interpersonal trust relations:

> *"Well, I think there are different phases, I can imagine that in the first phase, it [the AI-GIS] can certainly be helpful to identify an area. [...] So if there are facts that are also well-founded by any tools, that is, of course, an even better basis. [...] But later on, I think it's easier to create trust from person to person than with technology. But I think it can certainly be a good basis." (Trustor22, Educational Designer for Wind Energy: 51)*

A recurring theme across interviews in this configuration (as well as others) is the notion of the 'human factor'. This term emerges as a kind of shorthand for the aspects of planning that are seen as inherently non-automatable: compromise, empathy, contextual judgment, and informal problem-solving. The 'human factor' is mentioned both as a source of unpredictability and as a necessary component of trustworthy decision-making processes:

> *"Another big point here is acceptance – whether from the authorities, from the community, or citizens – it can be one way today and another tomorrow. That's one of the things that AI will never achieve, because humans change their minds. [...] It's simply a matter of sympathy and convincing someone. And I don't think AI will be able to solve these factors. But it will be able to pre-analyze many areas for the decision-makers." (Trustor29, Project developer: 69)*

Interestingly, trustors' skepticism toward the AI-GIS tends to focus less on the system itself and more on the data that underpins it. While the AI decision aid is not expected to replicate human capabilities (which tends to evoke mistrust in the *performance* dimension), it is seen as potentially valuable if its outputs are based on a credible, well-curated database (e.g., Trustor20: 243; Trustor12: 426). In this way, trust is relocated from the AI's decision logic to its informational foundation. This indicates a trust profile in which the *purpose* dimension (i.e., the intention behind the deployment of AI) plays a dominant role. The trustors see the AI-GIS as an opportunity to facilitate or legitimize decisions made by trusted decision-making actors – not to fundamentally alter the decision-making process itself:

> *"We are talking about data that we can collate. And there's an urgent need for a platform that has all the data and is easy to work with. [...] So compiling such hard facts would be a step forward in spatial planning and project development." (Trustor29, Project developer: 69)*

At the same time, this focus on the data level opens up a subtle line of criticism against institutions, which also appears in other parts of the interviews. Although (mis)trust in institutions is not a decisive condition in this configuration, according to QCA, many interviewees explicitly refer to laws and regulations in their arguments. For instance, they point to the difficulties of creating a land use plan with legal certainty (Trustor04: 87), or how specific laws hinder the decision-making process (e.g., 10 ha regulation in Bavaria) (Trustor39: 25). Accordingly, although the perspective of an AI-GIS is supported, the underlying motivation is a desire for objective and unambiguous framework conditions in which already trusted actors can make their decisions:

> *"For now, I would like to see us get to the point where politicians set clear guidelines. That will work faster than the AI-GIS – to be honest, I don't know how far artificial intelligence is from being truly operational in this area. I expect – and this is one thing we can do now – our politicians to enact laws in such a way that a judge cannot play the kind of games we have to endure at the moment. And that can be done by enacting very clear laws." (Trustor26, Project developer: 96)*

This tendency to distrust the institutions in which trust in AI is embedded is central in Configuration 2, discussed below.

From the patterns characterizing Configuration 1, we derive a distinct type of trust relation ‘Relation Type 1’ (see Figure 2). In this trust relation, trustors place confidence in established decision-makers and see AI as a useful but bounded enhancement of existing planning routines. Their trust is not technological in nature – it is mediated through people, practices, and the organizational context in which the AI is to be deployed. This type does not expect AI to replace the human dimension of decision-making but to support it in objective and efficient ways.

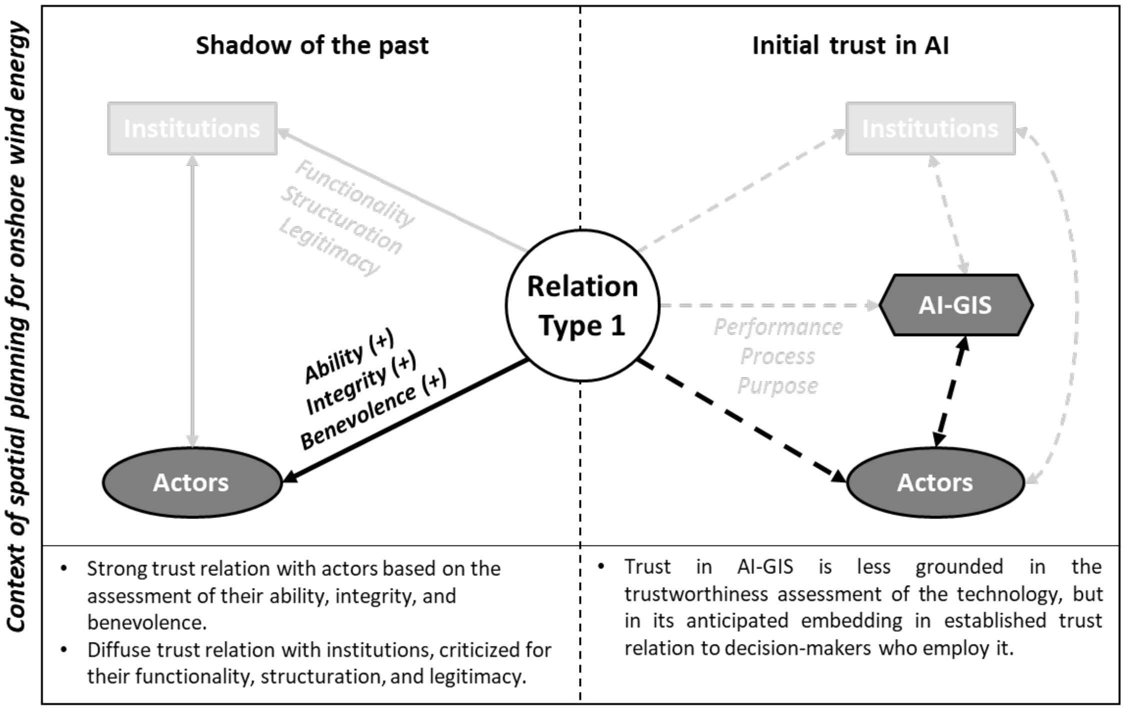


Figure 2: *Relation Type 1 – Trust in AI-GIS embedded in interpersonal trust*
Legend: Dark **= relationships relevant for the Relation type;** (+) **assessed as trustworthy = strengthens trust relation;** (-) **assessed as not trustworthy = weakens trust relation**

Relation Type 2 – Trust in AI embedded in institutional mistrust

Configuration 2 presents a distinct pattern in which trust in the AI-GIS does not arise because it is anticipated as an extension of existing structures, but rather as a response to a lack of trust in institutional frameworks. In this case, the trustors are actively involved in the decision-making process (DecIncl), yet they express skepticism or even disillusionment toward institutional norms, regulations, or procedures concerning spatial planning (TrustInst). Despite – or perhaps because of – this institutional distrust, they articulate a surprisingly high level of initial trust in the AI-GIS decision aid.

Interviewees frequently position the AI-GIS as a corrective mechanism, capable of counterbalancing what they perceive as inefficiencies, biases, or politicization within current decision-making. This orientation toward AI as an alternative source of *legitimacy* marks a notable shift in trust dynamics: the technology is not seen as enhancing the existing system but as compensating for its failures:

> *“The pressure came from outside, from above, from the Bavarian state parliament. The majority fraction here – at that time they changed the Bavarian building regulations in July 2020 and so the turbines under construction were then canceled. [...] And that was the death sentence for wind energy.” (Trustor27, Mayor)*

> *"And Bavaria is really well researched in terms of geography, so the data is all available. The problem is bringing the data together, and you could speed that up with artificial intelligence. So, I really see AI as pointing the way to the future for humanity, so the AI-GIS really is a good project." (Trustor27, Mayor)*

In contrast to Configuration 1, where trust in the AI-GIS is embedded in trust relations to decision-makers, the trustors in this configuration appear more ambivalent toward both institutions and actors – the latter to a lesser extent. While some degree of trust in current decision-makers persists, it is overshadowed by a broader disillusionment with systemic structures, i.e., mistrust in the *structuration* and *functionality* of institutions, which results in mistrusting their *legitimacy*. As a result, the AI-GIS is not framed merely as a tool used by trusted people, but as a potentially more neutral or objective alternative to traditional decision-making procedures (*process*):

> *"Sure, if you define criteria, then I think it works if you feed the AI with them and also have this background information. I think it can have a supportive effect. And it might also be more likely to be accepted than if people make decisions themselves. Because we're often accused of being played by politics, and if an AI were to do that now, it would perhaps be more likely to be accepted. At some point, if you familiarize yourself with it." (Trustor28, Approval authority)*

The interviews of this configuration reflect a subtle but important shift from past mistrust relations ('shadow of the past') to anticipated trust relations in the future, more concretely: from dysfunctional and illegitimate institutions (and actors) to objective procedures and data-driven outputs. Hence, the trustors anticipate the AI-GIS to feature a procedural legitimacy. For them, its appeal lies in the AI-GIS' perceived ability to generate consistent, transparent, and rule-based decisions that provide a clear contrast to the perceived arbitrariness of existing institutional practice. Accordingly, the *performance* and *process* dimensions are highly relevant for assessing trustworthiness. Nevertheless, this assessment of trustworthiness is also conditional. It hinges upon assumptions about how the AI is designed, implemented, and governed. Some interviewees voice concerns about transparency or the possibility of hidden biases. These concerns, though, are often expressed alongside a willingness to trust the AI-GIS, especially in light of their dissatisfaction with the status quo:

> *"I winced a little at the over 60 of restriction criteria that are supposed to be included in the AI. Because the question is what I want. If I want to have areas for wind energy, then I'm not going to build a tool out of many obstacles. [...] I don't want to challenge the technology; perhaps something good can actually come out of it in the end. All in all, that is my great wish, because I have experienced the decision-making as very unfortunate at times in recent years." (Trustor14, Spatial planning authority)*

From this configuration emerges the distinct Trust Relation Type 2 (see Figure 3): trustors who are to some extent involved in the planning process but question the *functionality*, *structuration*, or *legitimacy* of the institutional frameworks governing it. For this type, AI offers a potential avenue towards depersonalizing and depoliticizing decision-making. Trust in the AI-GIS is not inherited from past relations but emerges as a strategic hope for procedural reform or disruption.

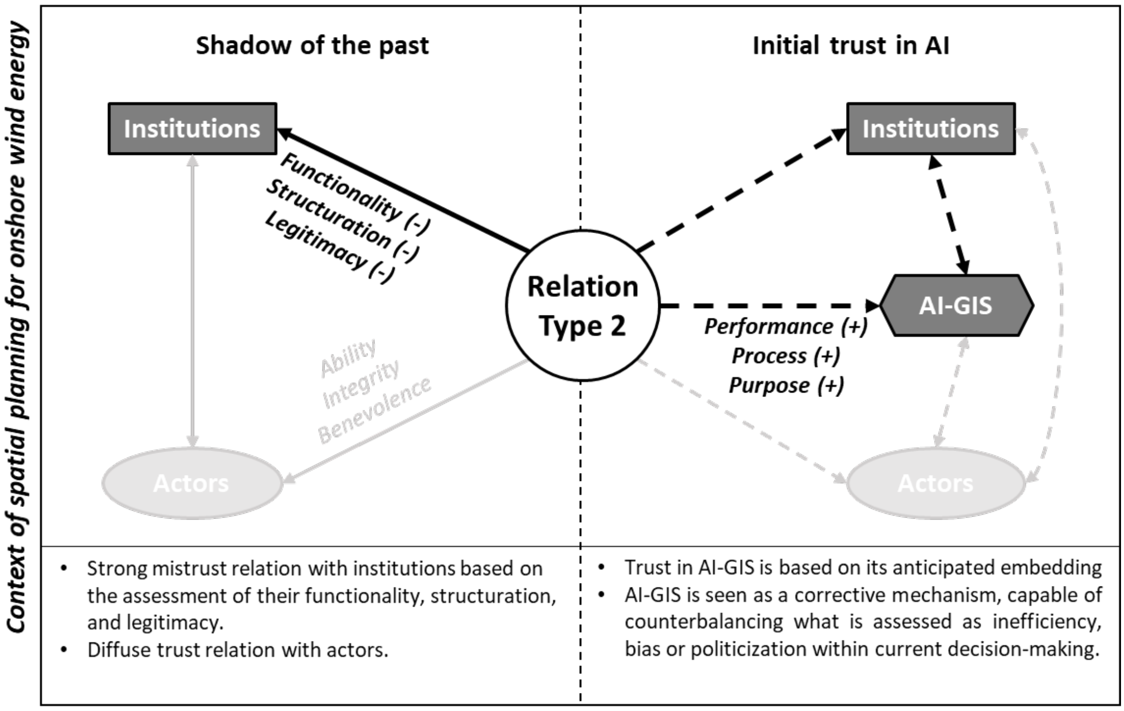


Figure 3: *Relation Type 2 – Trust in AI-GIS embedded in institutional mistrust*

Relation Type 3 – Mistrust in AI embedded in interpersonal and institutional mistrust

Configuration 3 reveals a trust dynamic that is characterized by across-the-board skepticism. Trustors in this configuration all belong to stakeholder group 5, i.e., citizens' initiatives and nature conservation initiatives. They have not been involved in planning processes (absence of DecIncl) and simultaneously lack trust in the actors responsible for decision-making. Against this background, they also express explicit distrust in the AI-GIS. Rather than being seen as a corrective or supportive tool, the AI-GIS is viewed as yet another manifestation of a flawed or illegitimate system.

In contrast to the previous configurations, trustors do not have differentiated trust relations to decision-makers, institutions, or the AI-GIS – they perceive all of them as part of the same problematic structure. Their mistrust is not narrowly focused but systemic, and often is accompanied by a broader critique of the political or normative direction of the planning process itself. In this sense, mistrust in all (potential) trustees is linked to trustors' opposition to the overarching objective of the German energy transition and the associated expansion of wind energy, in the form in which it is currently pursued. One example concerns the controversial decision of designating potential areas in forests:

> *"If this continues, we will have both: Climate change and destroyed forests. But if we just give away all the land because we think, 'Yes, we have to do this now to save the world.', then I say: No! What will happen at best is that we will achieve our climate targets on paper, yet we will still go down, but with the medal around our necks. We've done it, great. But at what cost? Environmental costs. But on paper, we are climate neutral. I believe that we are pouring a lot of money into these stories, which makes me wonder whether the euro wouldn't do much better in other areas if we really wanted to cut CO_2 emissions." (Trustor30, Citizens' Initiative)*

This 'shadow of the past' is so powerful that the AI-GIS is seen as distant, unaccountable, and potentially manipulative because it is interpreted as a tool of power – a way to formalize and manifest mistrusted structures and disguise political decisions as technical necessities (mistrust in institutional *legitimacy*). Hence, the *purpose* of the AI-GIS is assessed as not trustworthy. Rather than being perceived as a transparent or objective aid, the AI-GIS is treated with suspicion, especially concerning how data is selected, interpreted, and applied. The lack of trust in the *ability*, *integrity*, and *benevolence* of current decision-makers spills over into the technical system, especially given the assumption that it is those same decision-makers who are about to deploy and ultimately control the AI-GIS:

> *"It starts with the data finders, I mean, concerning the idea of artificial intelligence. These expert reports on wind speed are, of course, completely imprecise; important parameters, the roughness of the forest image, have not been assessed at all. Then the data density is also questionable, as they have measuring points 35 kilometers away, which are not at all comparable with the altitude here on site from a topographical point of view. [...] And GIS planning, that's certainly not a trick, there are thousands of images, and if you look at what data basis has been used, everything is a lie. [...] And the parties practically all caved in because they wanted to be involved. And we were accused of holding up progress." (Trustor34, Citizens initiative)*

As this quote highlights, the rejection of the AI-GIS is not just technological skepticism. It reflects a deeper political and procedural alienation. Trustors express their mistrust in actors and institutions by pointing out the logic by which decisions are made. In line with this, in the interviews, the role of power relations and market dynamics involved in the construction of legitimate sustainability is repeatedly discussed. For example, politicians are accused of subordinating species protection to the expansion of wind energy because the latter is a source of income (mistrust in actors' *integrity*). The same applies to the encroachment of forest areas. The AI-GIS, in this view, represents not a new innovation but an instrument that reinforces existing power asymmetries:

> *"Because this topic, how all species are connected, is simply very, very complex and, above all, abstract for people who don't deal with it. And then there's probably also the fact that technical climate protection – there's just an incredible amount of potential in terms of money. And in species conservation, it's not so obvious how you can earn money with it." (Trustor10, Nature conservation Ass.)*

These trustors should not be dismissed as disgruntled neighbors or angry citizens. Rather, they question power mechanisms that are involved in sustainability transitions which, from their perspective, also undergird decision-making on spatial planning for wind energy. Behind the seemingly unnuanced lack of trust lies a deeper understanding of the complexity of decision-making processes and the political dimensions influencing them.

> *"The legal situation for the expansion is not as ideal as one might expect. But to be honest, it [sustainability] is also more complicated than simply expanding a large number of wind turbines. [...] And I think a compromise has to be found. In my opinion, this cannot be at the expense of biodiversity and nature. Because I believe that, above all, people must reduce their impact. So, we have to downsize. Live more sustainably, consume less." (Trustor11: Nature Conservation Ass.)*

This configuration gives rise to the Trust Relation Type 3 (see Figure 4). The trustors express a fundamental lack of trust across all (potential) trustees – actors, institutions, and the AI-GIS. Their position is shaped by the perception of illegitimacy of institutional structures shaping the decision-making context, missing integrity of decision-makers, and a broader critique of the political direction of the energy transition. For this type, the AI-GIS is not a tool of improvement but an extension of a system they neither trust nor feel part of.

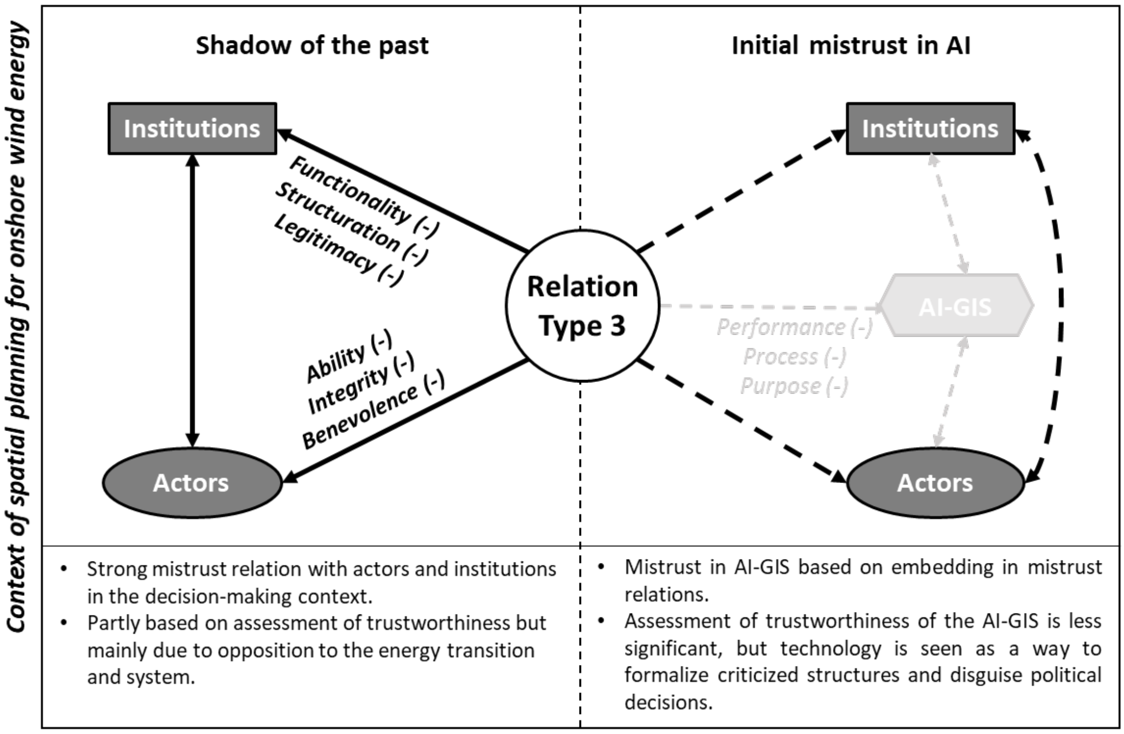


Figure 4: *Relation Type 3 – Mistrust in AI-GIS embedded in interpersonal and institutional mistrust*

Type 4 – Mistrust in AI with limited embedding in institutional trust and interpersonal mistrust

In Configuration 4, trustors express trust in institutions – laws, regulatory frameworks, and formal procedures – that guide and regulate spatial planning. At the same time, however, they convey clear skepticism toward the decision-makers responsible for enacting those rules in practice. This is particularly striking given that all trustors belonging to this configuration are involved in the decision-making process to some extent, even if this condition is not decisive for this configuration according to QCA. The duality of institutional trust and actor distrust here ultimately results in mistrust in the AI-GIS. This outcome highlights the complexity of the interplay between institutional and interpersonal trust when assessing the trustworthiness of AI systems.

A key characteristic of this configuration is the local specificity of mistrust. All trustors in this configuration are situated in the same regional context (Lower Saxony), and their doubts toward actors are grounded in past planning decisions that are perceived as lacking transparency or legitimacy. In the interviews, trustors repeatedly refer to specific sites that were included in spatial development plans against their judgment or contrary to their understanding of proper procedures. One interviewee elaborates on a controversial area:

> *"So, it's a very problematic area. It's very problematic in terms of nature conservation law, it's problematic in terms of species protection. So, there are many things that are very problematic. [...] That was also always the point in the negotiations, how to deal with it. And yes, it is problematic." (Trustor07, Species protection authority)*

As a result of the disagreement over the decision-making, the decision-makers are distrusted, in particular concerning their *ability* and *integrity*. According to the trustors, the area was not included in the spatial plan based on a *legitimate* institutional framework (trust in institutional *legitimacy*) but due to the influence of a single company (mistrust in actors' *integrity*):

> *"There is actually a very, very old story behind this area. [...] It is important to know that parallel to the district's potential study, two large areas in the municipality had already been*

> *presented as potential future wind farm sites by a company that is quite well-known – without the municipality even knowing that this was the case. The company later went bankrupt. At the time, however, the company had advertised its cause with mailshots and brochures. [...] And the district – and I have to be a little restrained in my choice of words – told the municipality that it should please push its planning in this direction. The municipality was very surprised and also carried out its own potential study and found something against both areas. [...] So the municipality would not have been able to implement either area. It remains open to question why the district gave these areas to the municipality. These are speculations as to why. But that's the way it was, and the municipality had to deal with it." (Trustor02, Freelancer for spatial planning (municipality))*

The strong trust in institutions raises a central paradox in this configuration: although an AI-GIS could, in theory, compensate for perceived shortcomings with respect to the trustworthiness of decision-makers, it is still not trusted. This is primarily due to two factors: First, trustors express mistrust in the AI-GIS's *performance* and *process*. They do not see the AI-GIS as capable of capturing the complex, dynamic, and often informal nature of spatial planning. As one interviewee succinctly puts it:

> *"At the moment, I'm not quite sure what the difference is between this and land use planning. Because I think it needs a certain depth to go into a potential analysis or land use analysis." (Trustor08, Species protection authority)*

Second, and perhaps more importantly, trustors do not perceive a deficit of objectivity or legitimacy in the current (institutional) system and, therefore, do not see the need for technological supplementation. Because they maintain trust in the overarching rules, there is no perceived need for an AI system to enhance transparency. Their distrust in actors does not appear to erode their institutional trust, nor does it translate into openness to technological solutions. Instead, the relationships between institutions and actors are treated as analytically and emotionally distinct. This means that no connection is made between the distrust in actors and the trust in institutions in the 'shadow of the past', resulting in the fact that there is also no connection in the anticipated trust relations. However, there is one anticipated connection between the AI-GIS and institutions: Trustors deny the *purpose* of an AI-GIS (objectivity) due to their strong trust in institutional *legitimacy*. Embedding, therefore, only takes place in parts and not for the mistrust relation to actors.

This decoupling has important implications. Trustees do not see AI-GIS as a solution to the problems they have identified, nor do they see it as an extension of the system. Instead, the AI-GIS is perceived as superfluous.

To summarize, in this Relation Type 4 (see Figure 5), trustors uphold the legitimacy of formal norms, procedures, and legal structures but have become skeptical of how these are applied in practice by individual decision-makers. Nevertheless, institutions are not criticized regarding their functionality, which indicates a decoupling of the relations to institutions and actors. For the trustors, the prospect of an AI-GIS neither addresses the core problem nor adds value. In terms of our analytical framework, trust in the AI-GIS is not shaped by the 'shadow of the past', as this is not perceived as an interconnected web, but as more isolated trust relations.

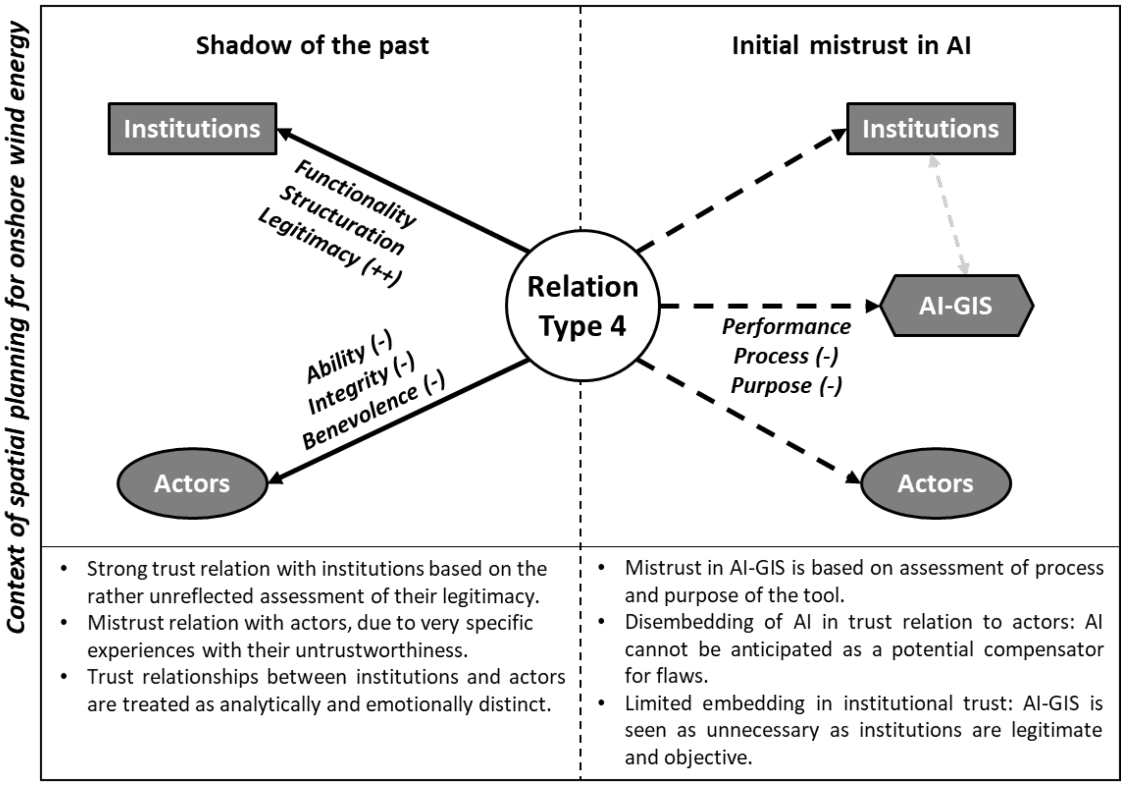


Figure 5: *Relation Type 4 – Mistrust in AI-GIS with limited embedding in institutional trust and interpersonal mistrust*

6. Discussion and Conclusion

The core objective of this paper has been to conceptualize the complex socio-technical environments in which AI decision aids are deployed within public administration. Understanding how affected stakeholders assess the trustworthiness of AI that is integrated into these systems is crucial. We argued that in such contexts, it is not only the AI decision aid itself (e.g., its process, performance, and purpose) that is assessed for trustworthiness, but also the existing constellation of decision-makers and institutions – rules, regulations, and practices – that govern decision-making. In other words, trust in AI is socially embedded. Yet, the literature currently lacks clear operationalizations and conceptual frameworks that allow for empirical exploration of this embedded trust in AI. Accordingly, our central contribution lies in the development and application of an analytical framework that conceptualizes trust in AI as socially embedded. The framework proposes that trust in AI decision aids is shaped by trustors' experiences with decision-makers and decision-guiding institutions in a concrete context ('shadow of the past'), and by trustors' expectations about how the AI decision aid will be embedded within these existing decision-making contexts. We applied this framework to the empirical case of an AI-augmented geographic information system (AI-GIS) developed to assist the decision-making of identifying and designating potential areas for onshore wind energy in Germany. The study focuses on how stakeholders' initial (mis)trust in this emerging AI decision aid is shaped by their trust relations to existing decision-making actors and institutions, as well as by how the AI-GIS is anticipated to be embedded within these relations in the future. We identified and examined four distinct trust configurations and derived associated types of trust relations. In this way, we offer a nuanced typology that reveals how different constellations of trust anchored in previous experiences with a specific decision-making context give rise to divergent degrees of trust toward AI systems. These four configurations illustrate that trust in AI is neither linear nor uniform, but conditional, context-sensitive, and relationally anchored.

In the remainder of this section, we first lay out the core conceptual, methodological, and empirical contributions of our paper. We then discuss practical implications, reflect on the limitations of our research, and point to directions for future research.

6.1. Core contributions

The conceptual framework proposed in this paper contributes to existing research on trust in AI by broadening the perspective from the dyadic relation between human trustors and AI trustees towards foregrounding the concrete socio-technical systems in which AI decision aids are employed. In this regard, our approach resonates with the growing assumption that trust in AI emerges differently in real-world situations than in experimental or hypothetical scenarios (Glikson und Woolley 2020; Shrestha et al. 2019; Lichtenthaler 2020). By grounding trust in socio-technical systems that embody concrete decision-making histories, we show how trust in AI is embedded in and shaped by these social structures. Our framework provides a theoretically grounded starting point to capture and interpret these social dynamics in which trust in AI is embedded, by drawing on the personal experience of human trustors with the decision-making actors and decision-guiding institutions. In this way, our framework synthesizes insights from existing research on trust in AI and relational perspectives on the embeddedness of interpersonal trust. By building on the tripartition of the assessment of trustworthiness of human actors (Mayer et al. 1995) – ability, integrity, and benevolence – and its application to AI systems (Solberg et al. 2022) – performance, process, and purpose –, we also develop a tripartition of the assessment of trustworthiness of institutions (functionality, structuration, and legitimacy). This enables a more differentiated approach to trustors' subjective experiences of (controversial) institutional structures, which are particularly prevalent in the institutionally complex context of sustainability transitions (Hacker und Binz 2021). We thereby conceptually add to the existing trust in AI literature (Kaplan et al. 2023) by proposing an operationalization of trust in AI as an individualized yet embedded concept: While trust in AI involves individualized assessments made by human trustors, we emphasize how their individualized experiences with the socio-technical system in which an AI is to be embedded – and thus their trust in this pre-existing system – in turn shapes their trust in AI.

Methodologically, our study contributes by combining qualitative comparative analysis (QCA) with in-depth qualitative content analysis. The QCA allowed us to identify relevant configurations of trust conditions that lead to (mis)trust in AI, while the qualitative material provided the depth needed to interpret the meaning and underlying patterns of these configurations. We believe that this combination offers a promising approach to studying complex and contingent trust dynamics – especially in socio-technical contexts where formal structures, lived experiences, and expectations intersect. The combination of both methods allows to navigate between abstraction and contextual depth, yielding a multi-layered and theoretically grounded understanding of trust in AI and its embeddedness in concrete social contexts. Rather than offering an exhaustive picture, the approach provides a focused yet rich account of how trust is configured in socio-technical settings.

Our empirical case of an AI-augmented geographic information system (AI-GIS) developed to assist the decision-making of identifying and designating potential areas for onshore wind energy allowed for a rich analysis of initial (mis)trust in the AI-GIS as being shaped by pre-existing trust relations with decision-makers and institutions in this concrete decision-making context. The case illustrates broader dynamics of AI technologies that are currently developed within the context of sustainability transitions, where digital tools are often portrayed as 'technological fixes' to complex socio-technical challenges (Katzner et al. 2019). Importantly, our study contributes to ongoing debates in the transitions literature, which highlight that transition pathways are influenced by spatial and institutional differences (Ribeiro et al. 2025). By demonstrating how trust functions as a relational mechanism, we position the social embeddedness of trust in AI as a crucial yet under-researched factor in the geography of socio-technical change. Based on our mixed-method approach, we were able to identify four configurations of trust relations. Each configuration highlights a distinct pattern of how trust or mistrust in AI systems is linked to existing trust relations and experiences ('shadow of the past') as well as the anticipated embedding of the AI-GIS in this context: Relation Type 1 exemplifies most strongly the influence of the shadow of the past: trust in the AI-GIS is not created in a vacuum, but emerges from pre-existing interpersonal and institutional trust. Familiar and credible decision-makers act as a bridge to trust the AI-GIS, even when stakeholders themselves are not directly involved in the decision-making. At the same time, cautious optimism is expressed concerning the anticipated embedding of the AI-GIS, which may enhance efficiency and objectivity – though never as a substitute for human judgment. Relation Type 2, by contrast, reflects a more aspirational orientation. Here, a lack of institutional trust – rooted in negative past experiences – coexists with forward-looking hopes that the AI-GIS can correct systemic deficiencies. In this configuration, trust is not inherited from familiar structures but formed in opposition to them. The AI-GIS is valued precisely because it is perceived as separate from the institutional status quo. This highlights the importance of understanding trust in AI not only as a continuation of established legitimacy but also as a potential alternative to failing systems. Relation Type 3 represents the reversed extreme of the shadow of the past. Here, cumulative experiences of exclusion, manipulation, or broken promises produce not just mistrust in decision-makers and institutions, but also a rejection of potential AI support. The AI-GIS is not seen as a solution but as an extension of systemic failure – a new layer of opacity rather than transparency. In such contexts, technological innovations are unlikely to generate trust unless the broader governance structures are also reconfigured. Lastly, Relation Type 4 presents a more ambivalent trust relation. Trust in institutions remains high – rules, norms, and regulatory principles are perceived as legitimate – yet, this is accompanied by persistent mistrust in the decision-makers tasked with its implementation. Notably, this distrust is not projected onto institutions themselves, nor is it transferred to the AI-GIS. Instead, the potential of AI to address shortcomings in decision-makers' performances is not recognized, because trust in institutional procedures is perceived as sufficient. Trustors in this configuration do not see the AI-GIS as a necessary or valuable addition. This configuration highlights how trust relations can remain compartmentalized, and how a lack of perceived need – rather than fear – can lead to the rejection of AI decision aids. Taken together, our empirical findings add to an emerging literature that highlights how trust in and attitudes towards AI systems vary depending on actors' roles (Daly et al. 2025; van der Werff et al. 2021) – leading to the realization that actors can have both positive and negative attitudes and willingness to trust AI, depending on the specific situation (Lichtenthaler 2020). Our findings dive deeper into how exactly these specific situations and 'contexts of interaction' (Kaplan et al. 2023) matter: Trustors form expectations about how AI decision aids will function *within* existing decision-making contexts and how it will interact with the decision-makers and institutions guiding the process.

6.2. Practical implications

The study offers practical insights into how trust in AI operates in the specific field of spatial planning, particularly concerning emerging digital tools. Our findings demonstrate that regional and institutional contexts – such as prior planning conflicts, the transparency and perceived legitimacy of decision-making, as well as the experienced trustworthiness of public authorities – significantly shape how AI systems are received. The study thus contributes to broader discussions on the governance of AI in the public sector by highlighting that trust in AI is shaped not only by its technological functionality and performance but also by the concrete decision-making contexts into which it is introduced. This leads to two practical implications of our research. First, it points to the relevance of those actors who are exposed to and affected by emerging AI-supported decision-making processes. Their trust in AI emerges differently and follows different dynamics than the trust of public authorities' employees – i.e., the users of AI systems. Therefore, isolated efforts towards enhancing the AI's transparency, explainability, or reliability may likely fall short in addressing those stakeholders' real concerns with the AI adoption. Second, our research raises awareness of the fact that trust in AI is embedded in existing institutional, historical, and political environments. As such, while AI and other digital tools may offer solutions to specific operational challenges, they also carry the risk of reinforcing or exacerbating existing tensions, inequalities, or pre-existing governance issues. Therefore, efforts to implement AI decision aids in public administration should not only address challenges to the concrete adoption of AI by public officials but need to attend to the broader socio-political environments in which they are deployed. Technology developers, managers, and government officials need to anticipate to what extent contested institutional structures, stakeholder power dynamics, or broad scepticism towards public administration may become a source for conflict when it comes to stakeholders' trust in AI.

6.3. Limitations and directions for future research

While the conceptual framework and empirical typology of trust configurations offer valuable conceptual and empirical insights, several limitations of this research should be acknowledged that might inform future research. First, this study provides an in-depth exploration of how the quality of existing trust relations within concrete decision-making processes shapes stakeholders' initial (mis)trust in AI-based decision aids. For this purpose, we framed the complex decision-making environments as socio-technical systems that not only involve the AI decision aid itself, but also trust relations to decision-making actors and decision-guiding institutions (see also right side of Figure 1). This interplay of technologies, actors, and institutions reflects the core analytical elements of socio-technical systems (Fuenfschilling und Truffer 2016) and, thus, allows our framework to be analytically stringent and potentially applicable to a range of contexts. At the same time, these broad categories could be fleshed out into more fine-grained systems 'components' in future research. This could include: other technologies, software, or protocols already in place before AI implementation; oversight mechanisms, auditing procedures, best practices that govern decision-making beyond regulatory institutions; or a finer differentiation between 'types' of decision-making actors, such as civil servants or politicians (EU High-Level Expert Group on Artificial Intelligence (AI HLEG) 2019). Second, our research integrates human-related antecedents to trust in AI by focusing on trustors' situated and individualized experiences within a given decision-making context. While such a social constructivist lens (Berger und Luckmann 2011) corresponds to existing research that has investigated human trustors' perceived competence, expertise, or experience with an AI (Omrani et al. 2022; Kaplan et al. 2023; Hoff und Bashir 2015), our framework does not account for more stable personality traits that shape an individual's propensity to trust (Küper und Krämer 2024) or sociodemographic factors. Future research could therefore investigate whether and how these variables can have a moderating or altering effect on the emergence of trust in AI in concrete situations.

Thirdly, our study combined qualitative comparative analysis (QCA) with in-depth qualitative content analysis. While this mixed-method approach allowed us to identify and interpret configurations of trust conditions, both methods have inherent limitations: QCA reduces the complexity of data by focusing on specific combinations of conditions that inevitably exclude alternative configurations. Our qualitative analysis – despite its depth – was shaped by a deductive coding process that steered interpretation along predefined theoretical categories. This may have limited the openness to entirely new or unexpected dimensions within the data. Future research could therefore adopt further empirical methods to scrutinize our findings and potentially alter our framework. On the quantitative side, conjoint analyses could be used to investigate what elements and conditions – i.e., trust relations from the past – are comparatively most important for humans' assessments of AI trustworthiness. On the qualitative side, content analysis following grounded theory and inductive coding could discover trust dynamics and patterns that are not yet theoretically framed.

Lastly, our empirical research is based on a single case study that is characterized as a conflictual decision-making context. Future research could further elaborate and adapt our conceptual framework to other empirical contexts, AI decision aids, or domains of application. It could be investigated whether and how similar trust dynamics emerge in other public sectors where the adoption of AI decision aids is not merely a technical matter but also a deeply contextual one – such as healthcare, education, or law enforcement. Additionally, longitudinal research could provide insights into how trust in AI decision aids evolves over time, particularly in response to institutional changes, policy reforms, or shifts in governance practices. Such studies could examine whether and how trust changes as digital technologies become more embedded in existing social structures.

Declaration of competing interest

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

Data availability

The data that has been used is confidential.

Declaration of generative AI and AI-assisted technologies in the writing process

During the final preparation of our manuscript, we used ChatGPT (version GPT-4o) for language editing to detect grammatical errors and enhance readability. The authors carefully reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgments

Tamara Schnell acknowledges funding in the context of the research project WindGISKI (funding code 67KI21003B), funded by the German Federal Ministry for the Environment, Nature Conservation, Nuclear Safety and Consumer Protection (BMUV). Ricarda Schmidt-Scheele acknowledges funding by the Ministry of Science and Culture (Lower Saxony) as part of the program “zukunft.niedersachsen”. We thank our colleagues at the working group *Organisation and Innovation* for their critical thoughts and helpful comments on earlier versions of the manuscript. We also thank all interviewees for their time and willingness to share their perspectives on the matter. Special thanks go to Lena Dahlhaus, who assisted with the QCA using R.

Literature

Agbabiaka, Olusegun; Ojo, Adegboyega; Connolly, Niall (2025): Requirements for trustworthy AI-enabled automated decision-making in the public sector: A systematic review. In: *Technological Forecasting and Social Change* 215, S. 124076. DOI: 10.1016/j.techfore.2025.124076.

Aoki, Naomi (2021): The importance of the assurance that “humans are still in the decision loop” for public trust in artificial intelligence: Evidence from an online experiment. In: *Computers in Human Behavior* 114, S. 106572. DOI: 10.1016/j.chb.2020.106572.

Araujo, Theo; Helberger, Natali; Kruikemeier, Sanne; Vreese, Claes H. de (2020): In AI we trust? Perceptions about automated decision-making by artificial intelligence. In: *AI & Soc* 35 (3), S. 611–623. DOI: 10.1007/s00146-019-00931-w.

Bachmann, Reinhard (2011): At the crossroads: Future directions in trust research. In: *Journal of Trust Research* 1 (2), S. 203–213. DOI: 10.1080/21515581.2011.603513.

Bachmann, Reinhard; Inkpen, Andrew C. (2011): Understanding Institutional-based Trust Building Processes in Inter-organizational Relationships. In: *Organization Studies* 32 (2), S. 281–301. DOI: 10.1177/0170840610397477.

Bachmann, Reinhard; Kroeger, Frens (2017): Trust, power or money: What governs business relationships? In: *International Sociology* 32 (1), S. 3–20. DOI: 10.1177/0268580916673747.

Bailey, Diane; Faraj, Samer; Hinds, Pamela; Krogh, Georg von; Leonardi, Paul (2019): Special Issue of Organization Science: Emerging Technologies and Organizing. In: *Organization Science* 30 (3), S. 642–646. DOI: 10.1287/orsc.2019.1299.

Baker, Stephanie; Xiang, Wei (2023): Explainable AI is Responsible AI: How Explainability Creates Trustworthy and Socially Responsible Artificial Intelligence.

Bao, Haixu; Liu, Wenfei; Dai, Zheng (2025): Artificial intelligence vs. public administrators: Public trust, efficiency, and tolerance for errors. In: *Technological Forecasting and Social Change* 215, S. 124102. DOI: 10.1016/j.techfore.2025.124102.

Bareis, Jascha (2024): The trustification of AI. Disclosing the bridging pillars that tie trust and AI together. In: *Big Data & Society* 11 (2), Artikel 20539517241249430. DOI: 10.1177/20539517241249430.

Barredo Arrieta, Alejandro; Díaz-Rodríguez, Natalia; Del Ser, Javier; Bennetot, Adrien; Tabik, Siham; Barbado, Alberto et al. (2020): Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. In: *Information Fusion* 58, S. 82–115. DOI: 10.1016/j.inffus.2019.12.012.

Batarseh, Feras A.; Kulkarni, Ajay; Sreng, Chhayly; Lin, Justice; Maksud, Siam (2023): ACWA: An AI-driven Cyber-Physical Testbed for Intelligent Water Systems.

Berente, Nicholas; Gu, Bin; Recker, Jan; Santhanam, Radhika (2021): Managing Artificial Intelligence. In: *MIS Quarterly* 45 (3), S. 1434–1450. DOI: 10.25300/MISQ/2021/16274.

Berger, Peter L.; Luckmann, Thomas (2011): The social construction of reality. A treatise in the sociology of knowledge. New York: Open Road Media Integrated Media. Online verfügbar unter https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=1804796.

Berman, Alexander; Fine Licht, Karl de; Carlsson, Vanja (2024): Trustworthy AI in the public sector: An empirical analysis of a Swedish labor market decision-support system. In: *Technology in Society* 76, S. 102471. DOI: 10.1016/j.techsoc.2024.102471.

Bitektine, Alex; Haack, Patrick (2015): The “Macro” and the “Micro” of Legitimacy: Toward a Multilevel Theory of the Legitimacy Process. In: *AMR* 40 (1), S. 49–75. DOI: 10.5465/amr.2013.0318.

Bullock, Justin B. (2019): Artificial Intelligence, Discretion, and Bureaucracy. In: *The American Review of Public Administration* 49 (7), S. 751–761. DOI: 10.1177/0275074019856123.

Bundesregierung (2022): Erneuerbare Energien Gesetz 2023. “We’re tripling the speed of the expansion of renewable energies”. Begründet von Bundesregierung.

Cabiddu, Francesca; Moi, Ludovica; Patriotta, Gerardo; Allen, David G. (2022): Why do users trust algorithms? A review and conceptualization of initial trust and trust over time. In: *European Management Journal* 40 (5), S. 685–706. DOI: 10.1016/j.emj.2022.06.001.

Camaréna, Stéphanie (2020): Artificial intelligence in the design of the transitions to sustainable food systems. In: *Journal of Cleaner Production* 271, S. 122574. DOI: 10.1016/j.jclepro.2020.122574.

Capestro, Mauro; Rizzo, Cristian; Kliestik, Tomas; Peluso, Alessandro M.; Pino, Giovanni (2024): Enabling digital technologies adoption in industrial districts: The key role of trust and knowledge sharing. In: *Technological Forecasting and Social Change* 198, S. 123003. DOI: 10.1016/j.techfore.2023.123003.

Christin, Angèle (2017): Algorithms in practice: Comparing web journalism and criminal justice. In: *Big Data & Society* 4 (2), 205395171771885. DOI: 10.1177/2053951717718855.

Cook, Karen S. (2005): Networks, Norms, and Trust: The Social Psychology of Social Capital∗ 2004 Cooley Mead Award Address. In: *Soc Psychol Q* 68 (1), S. 4–14. DOI: 10.1177/019027250506800102.

Daly, Sarah J.; Wiewiora, Anna; Hearn, Greg (2025): Shifting attitudes and trust in AI: Influences on organizational AI adoption. In: *Technological Forecasting and Social Change* 215, S. 124108. DOI: 10.1016/j.techfore.2025.124108.

Davis, Fred D. (1989): Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. In: *MIS Quarterly* 13 (3), S. 319. DOI: 10.2307/249008.

Deephouse, David L.; Suchman, Mark (2008): Legitimacy in Organizational Institutionalism. In: Royston Greenwood, Christine Oliver, Roy Suddaby und Kerstin Sahlin (Hg.): The SAGE Handbook of Organizational Institutionalism. 1 Oliver's Yard, 55 City Road, London EC1Y 1SP United Kingdom: SAGE Publications Ltd, S. 49–77.

Dusa, Adrian (2019): QCA with R. A Comprehensive Resource.: Springer International Publishing.

Eschenbach, Warren J. von (2021): Transparency and the Black Box Problem: Why We Do Not Trust AI. In: *Philos. Technol.* 34 (4), S. 1607–1622. DOI: 10.1007/s13347-021-00477-0.

EU High-Level Expert Group on Artificial Intelligence (AI HLEG) (2019): Ethics Guidelines for Trustworthy AI. Hg. v. European Commission. Brussels.

Farla, Jacco; Markard, Jochen; Raven, Rob; Coenen, Lars (2012): Sustainability transitions in the making: A closer look at actors, strategies and resources. In: *Technological Forecasting and Social Change* 79 (6), S. 991–998. DOI: 10.1016/j.techfore.2012.02.001.

Floridi, Luciano; Cowls, Josh; Beltrametti, Monica; Chatila, Raja; Chazerand, Patrice; Dignum, Virginia et al. (2018): AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. In: *Minds and machines* 28 (4), S. 689–707. DOI: 10.1007/s11023-018-9482-5.

Fuenfschilling, Lea (2019): An institutional perspective on sustainability transitions. In: *Chapters*, S. 219–236. Online verfügbar unter https://ideas.repec.org/h/elg/eechap/17966_12.html.

Fuenfschilling, Lea; Truffer, Bernhard (2016): The interplay of institutions, actors and technologies in socio-technical systems — An analysis of transformations in the Australian urban water sector. In: *Technological Forecasting and Social Change* 103, S. 298–312. DOI: 10.1016/j.techfore.2015.11.023.

Gan, QingQiu; Lau, Raymond Yiu Keung (2024): Trust in a ‘trust-free’ system: Blockchain acceptance in the banking and finance sector. In: *Technological Forecasting and Social Change* 199, S. 123050. DOI: 10.1016/j.techfore.2023.123050.

Gheorghiu, Mirona A.; Vignoles, Vivian L.; Smith, Peter B. (2009): Beyond the United States and Japan: Testing Yamagishi's Emancipation Theory of Trust across 31 Nations. In: *Soc Psychol Q* 72 (4), S. 365–383. DOI: 10.1177/019027250907200408.

Glikson, Ella; Woolley, Anita W. (2020): Human trust in artificial intelligence: Review of empirical research. In: *Academy of Management Annals* 14 (2), S. 627–660.

Granovetter, Mark (1985): Economic Action and Social Structure: The Problem of Embeddedness. In: *AJS* 91 (3), S. 481–510.

Grilli, Luca; Mariotti, Sergio; Marzano, Riccardo (2024): Artificial intelligence and shapeshifting capitalism. In: *J Evol Econ* 34 (2), S. 303–318. DOI: 10.1007/s00191-024-00865-7.

Gulati, Ranjay; Sytch, Maxim (2008): Does familiarity breed trust? Revisiting the antecedents of trust. In: *Manage Decis Econ* 29 (2-3), S. 165–190. DOI: 10.1002/mde.1396.

Hacker, Miriam E.; Binz, Christian (2021): Navigating institutional complexity in socio-technical transitions. In: *Environmental Innovation and Societal Transitions* 40, S. 367–381. DOI: 10.1016/j.eist.2021.09.003.

Hardin, R. (2002): Trust and Trustworthiness. New York: Russell Sage.

Hengstler, Monika; Enkel, Ellen; Duelli, Selina (2016): Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices. In: *Technological Forecasting and Social Change* 105, S. 105–120. DOI: 10.1016/j.techfore.2015.12.014.

Hoff, Kevin Anthony; Bashir, Masooda (2015): Trust in automation: integrating empirical evidence on factors that influence trust. In: *Human factors* 57 (3), S. 407–434. DOI: 10.1177/0018720814547570.

Kaplan, Alexandra D.; Kessler, Theresa T.; Brill, J. Christopher; Hancock, P. A. (2023): Trust in Artificial Intelligence: Meta-Analytic Findings. In: *Human factors* 65 (2), S. 337–359. DOI: 10.1177/00187208211013988.

Kaplan, Andreas; Haenlein, Michael (2019): Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. In: *Business Horizons* 62 (1), S. 15–25. DOI: 10.1016/j.bushor.2018.08.004.

Katzner, Todd E.; Nelson, David M.; Diffendorfer, Jay E.; Duerr, Adam E.; Campbell, Caitlin J.; Leslie, Douglas et al. (2019): Wind energy: An ecological challenge. In: *Science (New York, N.Y.)* 366 (6470), S. 1206–1207. DOI: 10.1126/science.aaz9989.

Kitchin, Rob (2017): Thinking critically about and researching algorithms. In: *Information, Communication & Society* 20 (1), S. 14–29. DOI: 10.1080/1369118X.2016.1154087.

Kiunke, Theresa; Gemignani, Natalia; Malheiro, Pedro; Brudermann, Thomas (2022): Key factors influencing onshore wind energy development: A case study from the German North Sea region. In: *Energy Policy* 165, S. 112962. DOI: 10.1016/j.enpol.2022.112962.

Kuckartz, Udo; Rädiker, Stefan (2023): Qualitative content analysis. Methods, practice and software. Second. Thousand Oaks: SAGE Publications.

Küper, Alisa; Krämer, Nicole (2024): Psychological Traits and Appropriate Reliance: Factors Shaping Trust in AI. In: *International Journal of Human–Computer Interaction*, S. 1–17. DOI: 10.1080/10447318.2024.2348216.

Latusek, Dominika; Cook, Karen S. (2012): Trust in Transitions. In: *Kyklos* 65 (4), S. 512–525. DOI: 10.1111/kykl.12004.

Lee, J.; Moray, N. (1992): Trust, control strategies and allocation of function in human-machine systems. In: *Ergonomics* 35 (10), S. 1243–1270. DOI: 10.1080/00140139208967392.

Lee, J. D.; See, K. A. (2004): Trust in automation: Designing for appropriate reliance. In: *Human factors* 46 (1), S. 50–80.

Legewie, Nicolas (2022): Qualitative Comparative Analysis. In: Nina Baur und Jörg Blasius (Hg.): Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer Fachmedien Wiesbaden, S. 707–724.

Lichtenthaler, Ulrich (2020): Extremes of acceptance: employee attitudes toward artificial intelligence. In: *JBS* 41 (5), S. 39–45. DOI: 10.1108/JBS-12-2018-0204.

Lockey, Steven; Gillespie, Nicole; Holm, Daniel; Someh, Ida Asadi (2021): A Review of Trust in Artificial Intelligence: Challenges, Vulnerabilities and Future Directions. In: Tung Bui (Hg.): Proceedings of the 54th Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences: Hawaii International Conference on System Sciences (Proceedings of the Annual Hawaii International Conference on System Sciences).

Mayer, Roger C.; Davis, James H.; Schoorman, David F. (1995): An Integrative Model of Organizational Trust. In: *The Academy of Management Review* 20 (3), S. 709–734.

McAllister, D. J. (1995): Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. In: *Academy of Management Journal* 38 (1), S. 24–59.

McCarthy, John (2007): What is artificial intelligence? Stanford University. Stanford, CA. Online verfügbar unter http://www-formal.stanford.edu/jmc/whatisai.pdf, zuletzt geprüft am 25.02.2025.

Murray, Alex; Rhymer, Jen; Sirmon, David G. (2021): Humans and Technology: Forms of Conjoined Agency in Organizations. In: *AMR* 46 (3), S. 552–571. DOI: 10.5465/amr.2019.0186.

Nahar, Sharmin (2024): Modeling the effects of artificial intelligence (AI)-based innovation on sustainable development goals (SDGs): Applying a system dynamics perspective in a cross-country setting. In: *Technological Forecasting and Social Change* 201, S. 123203. DOI: 10.1016/j.techfore.2023.123203.

Nourani, Mahsan (2023): SHADOWS OF THE PAST: THE EFFECTS OF USER'S DIFFERENCES AND PAST EXPERIENCES ON HUMAN-AI PARTNERSHIP. University of Florida (Dissertation). Online verfügbar unter https://www.proquest.com/openview/28124d79a129317bbb97b37f2c829cbb/1?cbl=18750&diss=y&pq-origsite=gscholar, zuletzt geprüft am 30.05.2025.

Nowak, Maciej J.; Brelik, Agnieszka; Oleńczuk-Paszel, Anna; Śpiewak-Szyjka, Monika; Przedańska, Justyna (2023): Spatial Conflicts concerning Wind Power Plants—A Case Study of Spatial Plans in Poland. In: *Energies* 16 (2), S. 941. DOI: 10.3390/en16020941.

Omrani, Nessrine; Rivieccio, Giorgia; Fiore, Ugo; Schiavone, Francesco; Agreda, Sergio Garcia (2022): To trust or not to trust? An assessment of trust in AI-based systems: Concerns, ethics and contexts. In: *Technological Forecasting and Social Change* 181, S. 121763. DOI: 10.1016/j.techfore.2022.121763.

Pandey, Dhananjay K.; Mishra, Richa (2024): Towards sustainable agriculture: Harnessing AI for global food security. In: *Artificial Intelligence in Agriculture* 12, S. 72–84. DOI: 10.1016/j.aiia.2024.04.003.

Poppo, Laura; Zhou, Kevin Zheng; Ryu, Sungmin (2008): Alternative Origins to Interorganizational Trust: An Interdependence Perspective on the Shadow of the Past and the Shadow of the Future. In: *Organization Science* 19 (1), S. 39–55. DOI: 10.1287/orsc.1070.0281.

Ragin, Charles C. (1987): The comparative method. Moving beyond qualitative and quantitative strategies. Berkeley: University of California Press.

Ragin, Charles C. (2008): Redesigning Social Inquiry: University of Chicago Press.

Ranpara, Ripal (2025): Energy-Efficient Green AI Architectures for Circular Economies Through Multi-Layered Sustainable Resource Optimization Framework.

Ribeiro, Inácio Pinto; Lopes, Hélder Silva; Dinis, Maria Alzira Pimenta; Remoaldo, Paula C. (2025): Geography of Sustainability Transitions: Mapping Spatial Dynamics and Research Trends Between 1995 and 2024. In: *Environments* 12 (5), S. 148. DOI: 10.3390/environments12050148.

Rijswijk, Kelly; Vries, Jasper R. de; Klerkx, Laurens; Turner, James A. (2023): The enabling and constraining connections between trust and digitalisation in incumbent value chains. In: *Technological Forecasting and Social Change* 186, S. 122175. DOI: 10.1016/j.techfore.2022.122175.

Ryan, Mark (2020): In AI We Trust: Ethics, Artificial Intelligence, and Reliability. In: *Science and engineering ethics* 26 (5), S. 2749–2767. DOI: 10.1007/s11948-020-00228-y.

Schilke, Oliver; Reimann, Martin; Cook, Karen S. (2021): Trust in Social Relations. In: *Annual Review of Sociology* 47, S. 239–259.

Schoorman, F. D.; Mayer, R. C.; Davis, J. H. (2007): An integrative model of organisational trust: Past, present, and future. In: *AMR* 32 (2), S. 344–354.

Scott, W. Richard (2014): Institutions and organizations. Ideas, interests, and identities. Fourth edition. Los Angeles: SAGE.

Shrestha, Yash Raj; Ben-Menahem, Shiko M.; Krogh, Georg von (2019): Organizational Decision-Making Structures in the Age of Artificial Intelligence. In: *California Management Review* 61 (4), S. 66–83. DOI: 10.1177/0008125619862257.

Shrestha, Yash Raj; Krishna, Vaibhav; Krogh, Georg von (2021): Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges. In: *Journal of Business Research* 123, S. 588–603. DOI: 10.1016/j.jbusres.2020.09.068.

Siau, Keng; Wang, Weiyu (2018): Building Trust in Artificial Intelligence, Machine Learning, and Robotics. In: *Cutter Business Technology Journal* 31 (2), S. 47–53.

Sjödin, David; Parida, Vinit; Kohtamäki, Marko (2023): Artificial intelligence enabling circular business model innovation in digital servitization: Conceptualizing dynamic capabilities, AI capacities, business models and effects. In: *Technological Forecasting and Social Change* 197, S. 122903. DOI: 10.1016/j.techfore.2023.122903.

Solberg, Elizabeth; Kaarstad, Magnhild; Eitrheim, Maren H. Rø; Bisio, Rossella; Reegård, Kine; Bloch, Marten (2022): A Conceptual Model of Trust, Perceived Risk, and Reliance on AI Decision Aids. In: *Group & Organization Management* 47 (2), S. 187–222. DOI: 10.1177/10596011221081238.

Suchman, Lucy (2023): The uncontroversial 'thingness' of AI. In: *Big Data & Society* 10 (2), Artikel 20539517231206794. DOI: 10.1177/20539517231206794.

Suchman, Mark C. (1995): Managing Legitimacy: Strategic and Institutional Approaches. In: *AMR* 20 (3), S. 571. DOI: 10.2307/258788.

Sullivan, Yulia; Bourmont, Marc de; Dunaway, Mary (2022): Appraisals of harms and injustice trigger an eerie feeling that decreases trust in artificial intelligence systems. In: *Ann Oper Res* 308 (1-2), S. 525–548. DOI: 10.1007/s10479-020-03702-9.

Trunk, Anna; Birkel, Hendrik; Hartmann, Evi (2020): On the current state of combining human and artificial intelligence for strategic organizational decision making. In: *Bus Res* 13 (3), S. 875–919. DOI: 10.1007/s40685-020-00133-x.

Tsuchiya, Motohiro (2024): Overcoming the Long Shadow of the Past: Defense AI in Japan. In: Heiko Borchert, Torben Schütz und Joseph Verbovszky (Hg.): The Very Long Game. Cham: Springer Nature Switzerland (Contributions to Security and Defence Studies), S. 487–504.

Tucci, F. (2022): Hemmnisse beim Ausbau der Windenergie an Land. Ergebnisse einer Branchenbefragung. Hg. v. Fachagentur Windenergie an Land. Berlin.

van der Werff, Lisa; Blomqvist, Kirsimarja; Koskinen, Sirpa (2021): Trust Cues in Artificial Intelligence: A multilevel Case Study in a Service Organization. In: Nicole Gillespie, C. Ashley Fulmer und Roy J. Lewicki (Hg.): Understanding Trust in Organizations. A Multilevel Perspective. 1st ed. Oxford: Taylor & Francis Group (SIOP Organizational Frontiers Series).

Wang, Qiang; Li, Yuanfan; Li, Rongrong (2025): Integrating artificial intelligence in energy transition: A comprehensive review. In: *Energy Strategy Reviews* 57, S. 101600. DOI: 10.1016/j.esr.2024.101600.

Annex 1 – Responsibilities for land use planning according to the federal states covered by our case studies

In Germany, the federal states are responsible for the identification and designation of sites for wind energy use, whereby the specific regulations and responsible authorities vary depending on the federal state. Below is an overview of the different responsibilities.

Lower Saxony Spatial planning is managed by the Lower Saxony Ministry of Food, Agriculture, and Consumer Protection, which is the highest state planning authority. Implementation is carried out by the offices for regional state development, which act as the higher state planning authorities. These offices are responsible for the creation and updating of regional planning in which, among other things, priority and suitability of areas for wind energy use are defined. https://www.ml.niedersachsen.de/startseite/themen/raumordnung_landesplanung/zukunftsfaehiges-niedersachsen-durch-raumordnung-und-landesplanung-4856.html (last accessed: 08.05.2025)
North Rhine-Westphalia The Ministry of Economic Affairs, Innovation, Digitization, and Energy is responsible for regional planning. The district governments and the “Regionalverband Ruhr” act as regional planning authorities and are responsible for designating areas for wind energy use. They draw up regional plans that define suitable areas for wind energy use. https://landesplanung.nrw.de/aktuelle-fachthemen/ausbau-der-windenergie-nordrhein-westfalen (last accessed: 08.05.2025)
Rhineland-Palatinate The Ministry of the Interior and Sport is the highest state planning authority. The structural and licensing authorities act as higher state planning authorities and work together with the districts to identify and designate suitable areas for wind energy use. https://mdi.rlp.de/themen/raumentwicklung-in-rheinland-pfalz/landesentwicklungsprogramm (last accessed: 08.05.2025)
Hesse The Ministry of Economic Affairs, Energy, Transport and Regional Development is the highest state planning authority. The state councils (“Regierungspräsidien”) serve as the higher state planning authorities and are responsible for regional planning, including the designation of areas for wind energy use. https://www.lea-hessen.de/energiewende-in-hessen/windenergie/ (last accessed: 08.05.2025)
Bavaria The State Ministry of Economic Affairs, Regional Development, and Energy acts as the highest state planning authority. The governments of the districts assume the role of supreme state planning authorities and work together with the 18 regional planning associations to determine suitable areas for wind energy use. It is important to note that the 10H regulation applies in Bavaria, which states that the distance between a wind turbine and residential buildings must be at least ten times the height of the turbine, which influences the designation of areas for wind energy use. https://www.energieatlas.bayern.de/thema_wind/themenplattform_windenergie/raumordnung_regionalplanung (last accessed: 08.05.2025)

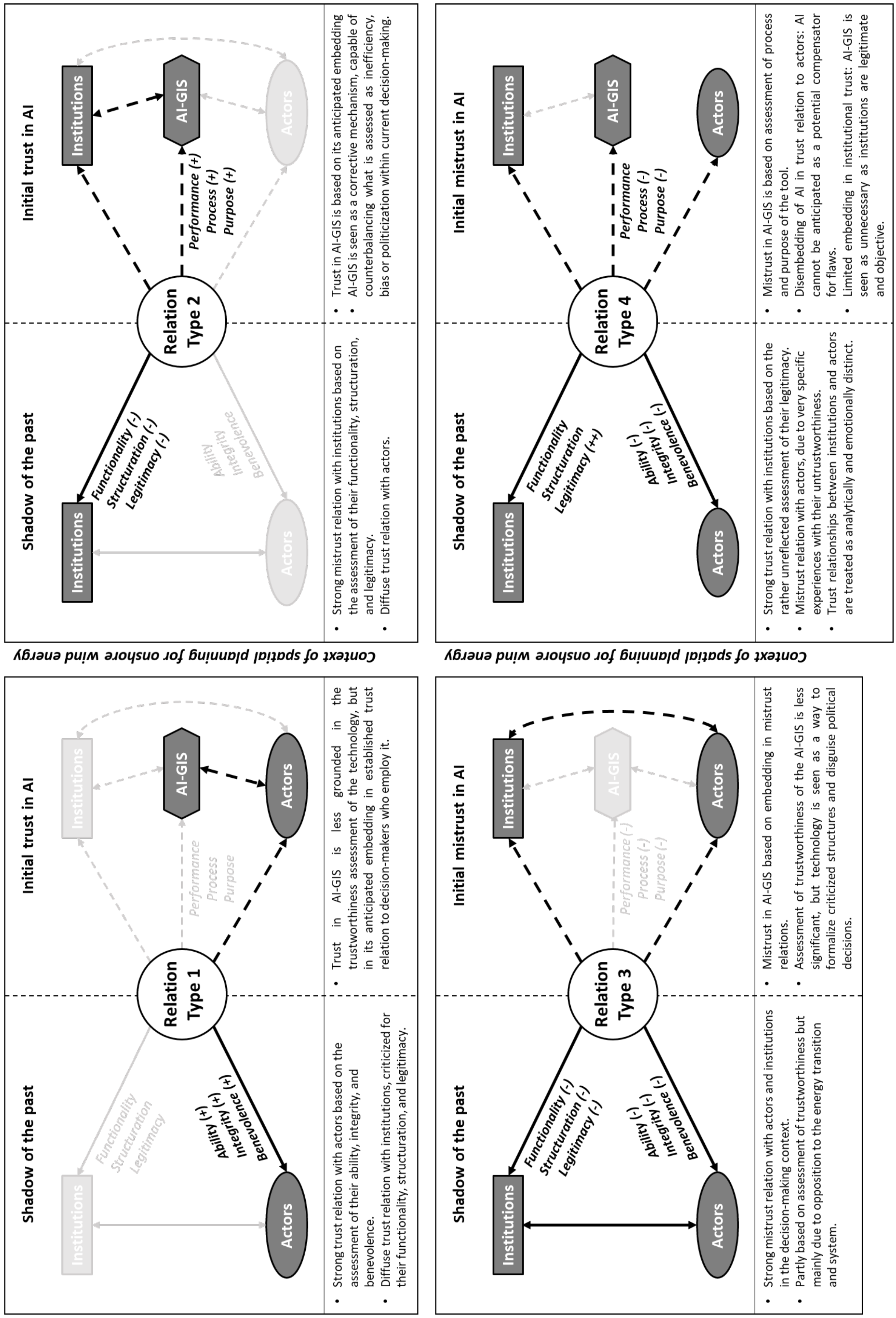
Annex 2 – Overview of Interviews (Cases/ Trustors)

Case ID	Interviewee (Trustor)	Stakeholder group
Case1	Trustor01: Regional building authority	3
Case2	Trustor02: Regional spatial planning, Freelancer	3
-	Trustor03: Windfarm Management	1
Case3	Trustor04: Windfarm Management	1
Case4	Trustor05: Community Councilor	2
-	Trustor06: Approval authority	3
Case5	Trustor07: Immission Control	3
Case6	Trustor08: Species Protection	3
Case7	Trustor09: Mayor	2
Case8	Trustor10: Nature Conservation	5
Case9	Trustor11: Nature Conservation	5
Case10	Trustor12: Project Developer	1
Case11	Trustor13: Project Developer	1
-	Trustor16: Project Developer	1
Case12	Trustor17: Mayor	2
Case13	Trustor20: Project Developer	1
-	Trustor21: Project Developer, Communication	1
Case14	Trustor22: Educational Designer (Wind Energy)	7
Case15	Trustor23: Mayor	2
Case16	Trustor24: Mayor	2
Case17	Trustor25: Community, Tourism department	2
Case18	Trustor14: Regional spatial planning, Authority	3
Case19	Trustor26: Project Developer	1
Case20	Trustor27: Mayor	2
Case21	Trustor28: Approval authority	3
Case22	Trustor29: Project Developer	1
Case23	Trustor30: Citizens' Initiative	5
Case24	Trustor31: Mayor	2
Case25	Trustor32: Science	6
Case26	Trustor33: Science	6
Case27	Trustor34: Citizens' Initiative	5
Case28	Trustor35: Ministry	4
Case29	Trustor36: Ministry, Spatial planning	4
Case30	Trustor37: Science, Seismology	6
Case31	Trustor38: Forestry, Authority	3
Case32	Trustor39: Nature Conservation	5
Case33	Trustor40: Regional Club	5
Case34	Trustor41: Public Utilities	1

Legend: IP03, IP06, IP16 and IP21 were excluded from the QCA analysis for methodological reasons (see Section 4.3).

Stakeholder groups: Project developers and initiators (1), (local) politicians (2), state authorities for spatial planning, approval and forestry (3), environmental ministries (4), (local) associations and citizens' initiatives (5), science (6), and other (7).

Annex 3 – Overview: Types of (Mis)Trust Relations



For further information about the GEIST Working Paper series, see:
www.geist-wp.com

Department of Human Geography
Lund University
Contact person: Johan Miorner
E-Mail: johan.miorner@keg.lu.se
https://portal.research.lu.se/sv/persons/johan-mi%C3%B6rner

Environmental Social Sciences
Eawag - Swiss Federal Institute of Aquatic Science and Technology
Contact person: Christian Binz
Überlandstrasse 133, 8600 Dübendorf, Switzerland
Tel.: +41-587-656493
E-Mail: christian.binz@eawag.ch
https://www.eawag.ch/en/department/ess/

Department of Geography and Regional Research
University of Vienna
Contact person: Michaela Trippl
Universitätsstraße 7/5/A0528, 1010 Vienna, Austria
Tel.: +43-1-4277-48720
E-Mail: Michaela.trippl@univie.ac.at
https://humangeo.univie.ac.at/