



Delft University of Technology

Responsible governance of generative AI Conceptualizing GenAI as complex adaptive systems

Janssen, Marijn

DOI

[10.1093/polsoc/puae040](https://doi.org/10.1093/polsoc/puae040)

Publication date

2025

Document Version

Final published version

Published in

Policy and Society

Citation (APA)

Janssen, M. (2025). Responsible governance of generative AI: Conceptualizing GenAI as complex adaptive systems. *Policy and Society*, 44(1), 38-51. <https://doi.org/10.1093/polsoc/puae040>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.


Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Responsible governance of generative AI: conceptualizing GenAI as complex adaptive systems

Marijn Janssen 

Faculty of Technology, Policy & Management, Delft University of Technology, Delft, the Netherlands

Corresponding author: M. Janssen, Faculty of Technology, Policy & Management, Delft University of Technology, Delft, the Netherlands. Email: m.f.w.h.a.janssen@tudelft.nl

Abstract

Organizations increasingly use Generative Artificial Intelligence (AI) to create strategic documents, legislation, and recommendations to support decision-making. Many current AI initiatives are technology-deterministic, whereas technology co-evolves with the social environment, resulting in new applications and situations. This paper presents a novel view of AI governance by organizations from the perspective of complex adaptive systems (CASs). AI is conceptualized as a socio-technological and adaptive system in which people, policies, systems, data, AI, processes, and other elements co-evolve. The CAS lens draws attention to focusing AI governance on the entire organization, taking an outward perspective and considering public values and societal concerns. Although there is no shortage of AI governance instruments, they differ in their effectiveness, and combinations of appropriate mechanisms should be selected to deal with AI's evolving nature and complexity. A major challenge is that no responsibility, and therefore accountability, is taken due to the lack of understanding of the full socio-technological CAS. As such, joint accountability is needed in which involved parties work together.

Keywords: AI governance; generative AI; responsible governance; IT governance; complex adaptive systems

Highlight

- Generative AI has become omnipresent
- Many AI governance instruments exist
- Responsible governance starts with societal expectations and public values
- AI governance should cover the entire organization instead of one part
- Governance should co-evolve with AI applications and developments

Generative Artificial intelligence (GenAI) has been heralded for its transformative power to change organizations and society with its content-generation capabilities. GenAI applications are often based on *Large Language Models* (LLMs) that use large datasets to understand, summarize, generate, and predict new content. GenAI can generate new content, including audio, code, images, text, simulations, and videos. Beyond these applications, it can also be used to create designs and prototypes, summarize text, answer questions, create classifications, translate, generate, debug, and verify software code, analyze data, reason, and more. While this provides tremendous opportunities for organizations in different areas, oversight is needed to create value and minimize the risks for organizations embracing Artificial Intelligence (AI) (Taeihagh, 2021). Blackberry's 2023 survey (Yu, 2023a) found that 75% of the companies surveyed considered banning GenAI, whereas, 1 year later, a *CapGemini* (2024) survey showed that 97% of the organizations allowed their employees to use GenAI. These changes and findings exemplify the struggle of companies and the need for governance. Yet, the complex and changing nature of GenAI may produce undesirable results, resulting in a *responsibility gap*. The monetary returns of this novel technology might be overemphasized, but societal implications are given less attention.

Many organizations are exploring ways to discover and harness the power of ChatGPT and similar tools (Dwivedi et al., 2023). Even in public decision-making, governments increasingly regard the introduction of AI as inevitable (Cave & ÓhÉigeartaigh, 2018), and GenAI is already being integrated into administrative and decision-making processes. AI will likely influence and even transform the policy cycle (Janssen & Helbig, 2018). GenAI can help generate all kinds of government policy documents, software code, and even legislation, areas that used to be the domain of experts but now require little more than feeding a prompt into a GenAI tool. However, the ease of the process says nothing about the suitability of the tasks, the quality of the outcomes, and values. The easier it becomes to use AI, the more oversight is needed, but the more complicated governance will prove to be. To illustrate, it serves to compare the introduction of GenAI to the first car. If there is only one car, the need for regulation is limited, but if society as a whole embraces cars, regulation and governance becomes critical in order to avoid accidents and nuisance. In a similar vein, by scaling up the use of GenAI in organizations there is a need for governance of GenAI by organizations that pay due attention to societal values.

GenAI is a general-purpose technology that might have adverse effects. Even if organizations adhere to legislation, customers, citizens, and society at large might not accept the outcome or decision made by AI. For example, while the Dutch agency responsible for disbursing benefits followed the letter of the law by not giving certain groups of people benefits—which saw them fall into poverty or take out predatory loans—society spoke up against what it believed to be unfair treatment. Similar situations affected a Dutch bank and the Dutch Tax and Customs Administration, and there has been a widespread backlash against companies using GenAI to recruit new employees, as AI automatically excluded certain demographics.

The variety of AI use cases creates a highly complex and fragmented field of study in which scholars face the daunting challenge of disentangling and clarifying the various domains and fields of analysis (Maragno et al., 2021). In view of this great variety of use cases, AI governance is highly dependent on the situation and context. Recent AI frameworks and discussions focus heavily on AI law and regulations (Almeida et al., 2023; Wirtz et al., 2020) and less on *responsible AI governance*. Governance lays down clear accountabilities, processes, and procedures, as well as relational mechanisms (Schneider et al., 2024), but this is insufficient to deal with the problematic ramifications of GenAI and guide future developments and usage in the desired direction. Responsible governance means identifying and addressing societal concerns, identifying problematic consequences, clearly defining the accountabilities of software vendors, providers, users, and other stakeholders, and empowering people and society. Responsible governance is a move from an inward organizational perspective to an outward organizational perspective on governance.

Responsible governance goes beyond AI, data, technology, or other forms of IT governance. AI governance has already been studied by many researchers (Janssen et al., 2020; Mäntymäki et al., 2022; Taeihagh, 2021; Wirtz et al., 2020), who generally advocate a layered approach (Gasser & Almeida, 2017), balancing benefits and risks (Taeihagh, 2021), and structural, procedural, and relational mechanisms (Peterson, 2004; Schneider et al., 2024). Yet, governance is not static and is subject to changes in the technological and organizational landscape. Responsible governance should deal with the idiosyncratic, ever-changing nature of GenAI. Bottom-up processes can result in unintended or unforeseeable consequences on individual-level behaviors (Nan, 2011). *Responsible governance, however, does not put the*

data, AI, IT, or other technology central, but is developed from an ethical and societal values point of view with due consideration of the complex interactions between the myriad actors involved. Crucial for the development of responsible governance are the interactions with society to avoid alienation and respect for values such as inclusiveness, transparency, and non-discrimination. Responsible governance is about establishing oversight mechanisms based on societal values and norms.

This paper presents a novel view of AI governance from the complex adaptive systems (CASs) perspective. We view organizations from a CAS lens, showing the large number of complex nonlinear interactions among agents. From a CAS point of view, accountabilities are more difficult to define due to the many different cause-and-effect loops at play in the myriad relationships involved, setting AI governance apart from traditional, rigid IT governance frameworks. CAS theory suggests that influencing the behavior of individual agents can influence the system as a whole by creating new patterns. Our governance mechanisms focus on society, influencing individual agents and creating patterns of accountability.

The next section briefly reviews the literature on CAS and AI governance, after which the opacity of GenAI and its governance challenges are discussed. Next, governance mechanisms are analyzed, showing the wealth of instruments at our disposal. The governance mechanisms focus on influencing individual agents, resulting in overall system behavior. In the subsequent section, we identify how governance can evolve from the project to the organization level and how it should evolve into responsible governance. In doing so, we stress the adaptive and changing nature of GenAI and the fact that AI governance evolves with changes in GenAI, data, experience, and other elements. Finally, various principles for responsible governance are presented, and conclusions are drawn.

AI governance conceptualized as CASs

What are CASs?

Although AI can be programmed and controlled to some extent, its outcomes cannot be predicted due to the self-learning nature of AI. Furthermore, GenAI outcomes depend on the interrelationships between AI systems, the data provided, user feedback, operator interventions, and other stakeholders (Janssen & Kuk, 2016). A CASs view looks at agent–technology interactions and dynamic relationships over time. Agents influence each other, resulting in emergent, global system-level properties (Carmichael & Hadžikadić, 2019). Nonlinear interactions occur at different times and different levels (societal, organization, project) and timescales, which are influenced by previous interactions. In CAS, the term “agent” is used broadly to refer to individuals, algorithms, data, information systems, and so on, including agents residing in the environment. The environment plays an important role in any CAS.

GenAI can be characterized as a CAS due to the multiple and diverse players involved, their evolving relationships and interactions, the use of new data, the introduction of new types of technologies, the finding of new applications, and changing regulations. Nan (2011) argues that the CAS lens produces deeper and more holistic analyses and provides a natural view for researchers to conceptualize the dynamic nature of GenAI, influenced as it is by both technology and human actors. From a CAS lens, an organization does not exist in isolation but is part of society and, therefore, co-evolves with its environment and emerging trends (Lewin & Regine, 1999). No single entity has the power or position to govern the whole system, and interactions among the agents within the organization and with society create the dynamics. Each agent can make a unique contribution to the system, and the interactions among all agents make up the system as a whole. While agents work with local information and local interactions and do not see the larger picture, agent interactions result in emerging behavior that makes the total system more than the sum of its parts. Consequently, predicting how new technology will unfold is highly complicated.

Adopting a CAS-based approach effectively means that evolution and complexity are taken as key starting points (Sullivan, 2011). Hill (2011) found that global health systems can be studied as CAS, which opens up ways of influencing the system through local points of engagement. Lewin and Regine found that organizations function much like a flock of birds, with individuals following simple rules and interacting with others to form a cohesive and dynamic whole. Eisenhardt and Sull (2001) suggest creating a set of relatively simple results to guide decisions when dealing with complexity and guiding organizations. Governance mechanisms can guide how AI behaves and direct GenAI within an organization viewed as a CAS.

With GenAI in particular, the large number of agents and continuous changes involved are a significant barrier to effective governance. Various stakeholders are involved, all with different levels of experience and education; the technology itself changes, how the technology is used changes, the application landscape is heterogeneous, and the technology harnesses data of different types and qualities. What is more, the complexities can change over time and depend on how the technology is used. Many external organizations are involved, including AI software vendors, developers, internal stakeholders such as project managers, system administrators, and so on. As GenAI technology and its applications change, the role of stakeholders may co-evolve and subsequently change. This results in the need to capture the whole CAS. However, a CAS operates at different levels, all of which evolve at different rates. GenAI is implemented and realized at the project level, while the organization level may involve multiple projects, all of which are influenced by other organizations that provide the software and the technology. Finally, at the societal level, people start using GenAI in their daily lives and have expectations about how GenAI is developed and provided by tech companies.

From a CAS point of view, a couple of lessons can be drawn for the organization-led governance of GenAI. CAS suggests that behavioral systems cannot be engineered while these systems are influenced by multiple interactions among agents that influence each other, as overall performance depends on the interactions and dynamics among interacting agents. Therefore, governance should not try to control but rather guide and direct the behavior of agents, thus influencing system-level properties. The system cannot be understood by looking at its individual parts, but complexity should be taken as a starting point, and the environment (society) should be included. A change in one element can cause other aspects to co-evolve and change accordingly. Nevertheless, CAS suggests that it only takes relatively simple principles to govern a complex system. One such governance principle can be used to evaluate the societal impact of GenAI before its release, or another one could be that clear accountability should be defined upfront for the whole system. As such, responsible governance mechanisms need to be defined to guide co-evolution.

IT governance streams

Different strands of IT governance used by organizations have followed different development paths. [Brown and Grant \(2005\)](#) mention two main streams: centralization/decentralization and contingency. The first stream deals with forms of IT governance, while the second focuses on IT governance contingencies ([Brown & Grant, 2005](#)). The first is centered around organizational centralization and decentralization of decision-making ([Brown & Grant, 2005](#)). The degree to which decision-making authorities are allocated to central or decentral parts of an organization changes over time and can be viewed as a “pendulum swing” ([Peterson, 2004](#)). The other stream revolves around how organizational IT governance fits in with the environment, investigating how multiple, interacting contingency factors influence the modes of governance and identifying factors like economies of scope and absorptive capacity, as well as the IT knowledge of line managers ([Sambamurthy & Zmud](#)), firm size, industry, and organizational structure ([Brown & Grant, 2005](#)). In general, both streams focus on systematically determining who makes each type of decision (a decision right), who provides input to a decision (an input right), and how these people (or groups) are held accountable for their role ([Weill, 2004](#)).

A third stream in AI governance efforts focuses on the organizational governance processes of balancing risks and returns ([Abraham et al., 2019](#); [Dafoe, 2018](#); [Perks & Beveridge, 2003](#); [Schneider et al., 2024](#)). This means maximizing AI’s advantages to create organizational value and minimize the risks. AI capabilities are essential in creating organizational value ([van Noordt & Tangi, 2023](#)). This stream takes organizational objectives as a starting point and does not consider the broader societal objectives as responsible governance does. Often, this type of governance is part of the annual planning and control process of organizations.

A fourth stream takes an organizational control approach ([Wirtz et al., 2020](#)). For this, it is important to first distinguish between control and governance. Control stems from a machine-oriented view of organizations, in which a manager controls employees by providing commands, much as an operator would control a machine. Control has a deterministic connotation and often does not take into account the behavior, the interactions among elements, and the critical capabilities of humans into account. Some work in AI governance seems to be more about controlling AI rather than governing it. Governance refers to situations in which control is often not entirely possible, as humans and their behavior play a major role in the CAS. AI control is about ensuring that the AI software system works, such as somebody

who is deploying the system, whereas responsible governance is about directing and supervising the whole organizational system consisting of data, humans, algorithms, and software. Governance deals with human behavior that can always be unpredictable to some extent.

For AI, there is a fifth stream of governance that focuses on how organizations involve human decision-makers. A key dilemma for organizations is whether AI systems should be allowed to make decisions independently or whether they support human decision-makers, i.e., the question of who makes the final decision (Mittelstadt et al., 2016; Van Noordt & Misuraca, 2022). Often, discussions focus on human-in-the-loop decision-making, human decision-making discretion, and humans' ability to challenge the outcomes of AI-based recommendations and decisions. For organizations, efficiency and the ability to control the outcome of AI play a major role.

A final stream looks more at *regulation* to send organizations in the desired direction. Governance regulations are often found in national strategies, and organizations should comply with them (Radu, 2021). Often, the focus is on regulating AI without interfering with its advancement (Scherer, 2015). Principles of responsibility and explainability are proposed to secure fairness, privacy, non-discriminatory accountability, and so on (Gasser & Almeida, 2017). The European AI Act is one of the first comprehensive pieces of AI legislation to ensure proper conditions for developing and using this innovative technology. Regulations operate at the federal and national levels, but no specific organizational governance mechanisms are provided to realize them. Organizations need to be compliant with regulations. Some AI applications might not even be allowed, whereas higher levels of control are required for others. Regulations influence the need for responsible AI governance.

Our view on governance follows a CAS lens in which responsible AI governance is focused on the complete systems, including society, in which human behavior and various complexities play a role. We embrace a contingency approach in which governance is context-specific, dependent on the situation, and subject to changes. We look at both decision-making and processes.

GenAI in practice: opaqueness and smog

The capabilities of GenAI include content generation, the ability to generalize, and reinforcement learning based on human feedback (Nah et al., 2023). Organizations use GenAI to increase their efficiency and reduce the time their employees need to spend on their work by harnessing those capabilities. With GenAI, staff can even write a report about areas they do not necessarily know much about. GenAI effectively serves as *cognitive scaffolding*, aiding staff in performing a task at a level that it would otherwise be unable to attain. In that sense, it is tantamount to giving your car keys to somebody who cannot drive. Responsible governance tries to avoid this by imposing on organizations the responsibility to ensure that their staff have the needed driving skills.

The notion of AI as cognitive scaffolding also raises questions about the correctness of the outcomes, as non-experts might not be able to scrutinize the outcomes properly, and neither might experts, for that matter. Aaronson (2023) advances the concept of *Data Dysphoria*, describing factual, coherent answers into which GenAI has mixed fictional elements, such as a false name or a false academic citation. The issue with Data Dysphoria is that it makes it very complicated for organizations and their customers to detect what is correct and what is not. Hallucinations are when an LLM perceives nonexistent patterns or objects, creating nonsensical, misleading, or inaccurate outputs (Fayyad, 2023). Hallucinations are often sentences that sound plausible and realistic but are not, which might be hard to detect. GenAI is highly opaque, and its outcomes resemble smog in the sense that they may or may not be correct.

AI increases the complexity of governance and introduces new dependencies and unknowns. AI should not be looked at in isolation by organizations but rather as a technology that is highly dependent on input data, the selection of AI types and implementation, and human use. GenAI often relies on data and feedback loops for reinforcement learning. We classify the causes of the problems into the following categories.

- *Origin and quality of training data.* GenAI models are often trained on unknown collections of datasets, in which data quality is unknown, and unknown training and validation mechanisms are used. Even if the training data is listed, then the question is the quality of the vast amount of training data. Data are rarely perfect, can be outdated, incomplete, or simply incorrect. Humans might be able to interpret if some data are incorrect, whereas GenAI models might not detect this at all.

- *Opaque working of GenAI models.* LLMs consist of architectures made of many fine-tuned components. The models behind the chat functionality are rendered inaccessible to public scrutiny by their proprietors; however, even if they are opened, the questions remain if scrutiny can be provided, as even the developers might not understand the workings of the models.
- *Model collapse.* The use of GenAI data as input can result in *model collapse* problems, i.e., generator degeneration, continuous generation of the same sample points, and inability to continue learning (Gonog & Zhou, 2019). Models collapse is specific to GenAI as it generates new data and no longer knows the original and correct data. As a consequence, model outputs can become more wrong, and the stored synthetic data might not be suitable for training LLM anymore. When new models arise in the future, they might not be able to learn from the data. As such, data provenance of the original and nongenerated data is essential.
- *Wrongly generated predictions and outcomes.* GenAI aims to generate new data and does not provide factual outcomes. Hallucinations can be viewed as a feature of GenAI (Fayyad, 2023). The resulting output might not be accurate or could even be completely wrong, yet these errors are convincingly presented. Furthermore, the outcomes change over time. LLMs usually do not develop cause-effect models of the world but only look at what logically comes next.
- *Replicating and amplifying bias.* Motoki et al. (2023) found that ChatGPT presents a significant and systematic political stance and that LLMs can generally extend or amplify the existing bias.
- *Deteriorating of the quality of answers over time.* GenAI output can change over time based on user feedback in prompts and new data. For example, Chen et al. (2023) found a deterioration in the quality of the answers in some areas over time. The behavior of the “same” LLM service can change substantially in a relatively short time, highlighting the need for governance (Chen et al., 2023).
- *Security and privacy challenges.* The training data might contain private or other types of sensitive information. This information might be revealed by prompt engineering (Giray, 2023). Furthermore, others might try to manipulate the LLM to provide other outcomes that they are interested in for political or other purposes.

The CAS lens draws attention to the variety of interacting agents involved. For organizations, responsible governance is not only about internal governance within the organization itself but also about connecting with and even trying to influence the social environment. After all, the social environment has expectations, determining what is deemed acceptable and what the organization's responsibilities are. Overall, the social environment limits the use of GenAI and its implications on people. However, the concentration of power in big-tech companies means that they can serve their own interests instead of focusing on responsibility (Khanal et al., 2024). Organizations need a delicate balance with an external environment that poses internal governance requirements and limits what can be done.

The complexity of GenAI might increase further with new updates, training data, reinforcement learning, new employees, new data, and the system's and new users' patching. This results in smog and opaqueness of the systems and the need to look at the situation from a CAS point of view as the situation evolves continuously. All these challenges and the complex stakeholder network in play require organizations to be responsible for GenAI. Nevertheless, GenAI governance is not a panacea for all problems; the first question is whether the challenges are acceptable and can be dealt with by governance. If the foundations are not right, then governance might not be able to deal with the challenges and will only produce red tape and more opaqueness and smog. The next question is whether governance can actually help before adding any governance mechanisms. For example, can governance mechanisms ensure proper training of data and help detect Data Dysphoria, hallucinations, and incorrect GenAI outcomes?

GenAI governance mechanisms

Many stakeholders are usually involved in the implementation and use of GenAI in organizations. The people who buy the GenAI and train it with organization-specific training data are often not the people who use the results and are expected to manage the GenAI. Their knowledge of and experience with GenAI often differ and need to be considered. The large number of stakeholders involved complicates accountability, as boundaries between responsibilities cannot be defined clearly and might change due to evolution.

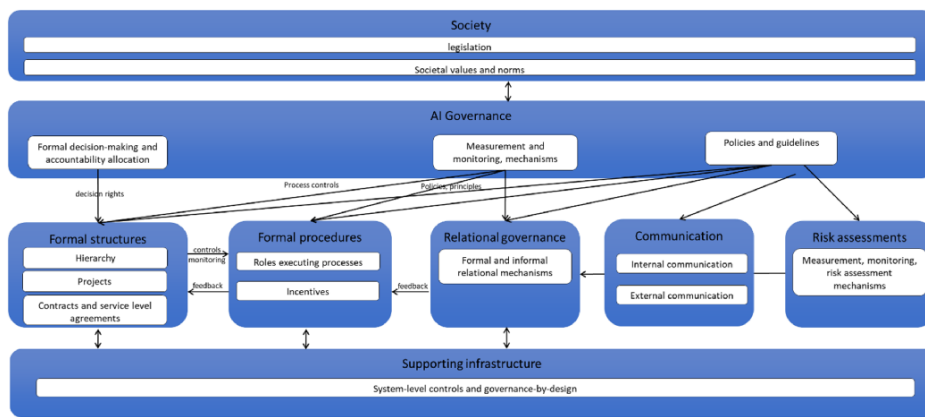


Figure 1. Overview of governance mechanisms.

Furthermore, for effective governance, a variety of questions need to be answered, such as how to detect the use of incorrect data/algorithms. Lack of overview poses a major challenge to effective governance and makes it unclear what actions are needed when something goes wrong. How can you govern what you do not know? Furthermore, there are many hidden complexities to consider in AI governance.

From a CAS view, AI can be conceptualized as a socio-technological adaptive system in which people, policies, systems, data, AI, processes, and other elements co-evolve. The basic inputs are various types of information, the initial GenAI systems, infrastructure, and humans. Once humans use GenAI, the system will learn from usage, resulting in feedback loops. Mistakes can also be corrected and fed back into the system using feedback loops. As the system evolves, governance should also evolve and consider the new situation.

The CAS views also show the model collapse challenge. GenAI created new policy knowledge, which can, but might not be correct. Policy documents created by humans can serve as input, but in the future, policy documents created by GenAI will also be used as input. This results in data provenance challenges, with organizations having to figure out which policy documents were generated by humans and are likely correct, and which were not due to Data Dysphoria, hallucinations, and problems with LLMs. The CAS views show that the evolving nature and complexity of the situation should be taken as a starting point, and societal expectations and requirements should be considered. Before initiating any governance, *reducing complexity is a good strategy. Complexity is hard to govern and almost impossible to control.* The main governance challenges include (1) understanding what needs to be governed and (2) determining who will take the lead in defining AI governance (e.g., accountabilities). In organizations, GenAI might be used without defined governance, in which case regulation can play a pivotal role.

Organizational governance is guided by legislation and societal values, expectations, and norms (Janssen et al., 2020; Van de Poel, 2020). This is shown at the top of Figure 1, representing the influence of an organization's environment. The arrow depicts a mutual influence as the organizations and environment co-evolve. For the sake of clarity, the Figure does not show all actors involved.

Organizational governance needs to allocate responsibilities, monitoring and intervention mechanisms, and policies and guidelines for using GenAI. Accountabilities are defined by allocating responsibilities, and monitoring mechanisms should ensure oversight and enable management to take corrective actions. Decision-making authorities raise the question of who is accountable. Often, authorities are distributed over organizational units, resulting in a chain of decisions among management layers, as shown on the left-hand side of Figure 1. Societal expectations of governance mechanisms can be translated into policies and guidelines. The organizational governance level allocates the decision-making authorities, defines the planning and control cycle and risk assessment (Janssen et al., 2020), and guides organizational efforts. The governance system should be continuously evaluated and should co-evolve with GenAI and society.

Overviews of governance mechanisms exist in the literature. Categorizations of governance mechanisms include procedural, structural, and relational mechanisms, as well as decision, communication,

and alignment processes (Weill & Ross, 2004). Another approach to categorizing governance mechanisms distinguishes between contractual and relational governance. The essence of IT governance is related to decision-making authorities (Sambamurthy & Zmud, 1999). There are five governance mechanisms: decision-making authorities, formal processes, relationships, communication, and risk assessment.

Management makes decisions through *formal procedures and processes* (formal procedures for short, see the middle of the Figure 1) in which decisions are made. Processes can deal with incidents and problems, but they also include major choices, such as the choice of LLMs software providers, training data, customer complaints, and so on. The processes contain the annual planning and control cycle but are also needed for evaluation and fine-tuning.

Humans have *relational* governance mechanisms with others inside and outside the organizations, as shown on the right side of Figure 1. Essential elements of relational governance are formal and informal communications that create awareness of processes, roles, and other governance elements. Relationships contribute to building awareness, influence the direction of decision-making within the governance framework, and are aimed at creating trust. For responsible governance, the relationship with society is key.

Communication is a well-known success factor in IT governance (Nfuka & Rusu, 2011). Responsibilities, procedures, and policies need to be communicated within the organization. Communication with the external environment is essential. Surveys might be used to collect insights or ethical boards could be introduced to ensure the involvement of the societal perspective.

Finally, *risk assessment* mechanisms are essential (Janssen et al., 2020; Taeihagh, 2021). When risks are identified, they should be mitigated or reported immediately. Mechanisms for annual risk assessment and auditing are also needed. Also, when changes or updates occur, reassessment will be needed.

Figure 1 provides a more fine-grained categorization of governance mechanisms for GenAI to emphasize the variety and differences. The areas can be used differently by organizations. In each area in Figure 1 and Table 1, the most appropriate governance mechanisms for the situation can be selected. For example, if there is a lack of AI control, then the formal structures need to be improved, or if there is much uncertainty, risk assessment should be strengthened. Organizations can use Figure 1 to ensure that appropriate governance mechanisms are selected in all areas to strengthen each other. In this way, the whole CAS can be addressed. The governance mechanisms focus on influencing individual agents and creating patterns of accountability. Governance mechanisms should facilitate learning and improvement to deal with the complexity and evolving nature of GenAI.

Accountability is a crucial element of any governance system, and GenAI poses new challenges to accountability. Accountability consists of answering questions about one's actions or inactions and being responsible for consequences (Roberts, 2002). In most common situations, an operator or manufacturer of a machine can be held accountable for the consequences of how the machine works. As such, manufacturers are required by law to ensure that their products are safe and secure, as detailed regulations have often been laid down. This is based on *predictive* accountability. Yet, with autonomously learning machines, AI developers and human users cannot be held accountable for the behavior of AI systems due to their lack of control and influence (Matthias, 2004). Furthermore, responsibilities might not be easily defined. For local projects, project managers can be held accountable, but once governance is integrated into the organization, clear responsibilities and roles need to be defined at the organizational level. Creating responsible governance can be challenging as managers might not understand GenAI, and being held accountable for something they do not understand might harm their career. Furthermore, it might be hard to create faultless systems, since GenAI usually makes mistakes, and creating systems without mistakes might be simply impossible. A car can be tested to guarantee that it will work safely in certain conditions, but with AI, it is hard to prove it is safe and secure due to changing environments, the co-evolution of interacting agent, its self-learning capabilities, and its overall complexity. The use of AI might be a gamble in which nobody wants to be held accountable when something goes wrong.

Instead of defining fine-grained responsibilities for AI based on predictive accountability, there should be shifts toward *joint accountability* in responsible AI governance. The latter refers to a situation in which the departments and organizations involved accept that they are accountable to each other for fulfilling the goal of GenAI by sharing the responsibility for the commitment they made, including the consequences. Joint accountability requires an emphasis on relational mechanisms such as feedback

mechanisms, sharing experiences and knowledge, shared risk analysis, and working together to ensure that GenAI works properly. Actors involved include the user organizations, the developer, and the software company providing GenAI. For this, the stakeholders should work together and empathize with each other. A focus on joint accountability might easily result in the decision not to bring certain GenAI projects into production, as their impact might not be accepted. Also, this might result in the introduction of various governance mechanisms in which the systems evolve, and the actors work together to ensure that GenAI works as intended.

Table 1 provides insight into the breadth of the field and the variety of governance mechanisms. While expansive, it is not exhaustive. Governance mechanisms for organizations can be internally or externally focused. The societal governance mechanisms shown in Table 1 aim to align organizations' efforts with societal expectations, whereas the organizational governance mechanisms focus on governing the internal organization. Also, GenAI projects can have specific governance, which is important, as GenAI might not be used throughout the organization. Technology-embedded governance mechanisms refer to governance mechanisms that are included in the software.

Overall, there is no shortage of governance instruments to be used by organizations, but they differ in effectiveness, and a combination of appropriate mechanisms should be selected. The societal, organizational, project, and technology governance mechanisms should be aligned and complement each other. Multiple levels of defense are often needed, and a single governance mechanism cannot do the job. The type of governance mechanism depends on the type of AI and the context. Key risks include the absence of necessary governance mechanisms or the use of inappropriate mechanisms. Furthermore, governance might be dominated by software companies, politicians, managers, or other stakeholders with limited knowledge of AI. Governance needs to be tailored to their specific needs and challenges. The CAS lens suggests that governance mechanisms are needed for the internal organization and the social network. For governance, multiple lines of defense might be needed in high-risk situations and governance mechanisms should, therefore, complement each other.

Moving toward responsible governance

AI governance often starts at the project level, neglecting organizational realities and not being aligned with the overall organizational governance. This results in fragmentation, missed opportunities, and a lack of clear accountabilities at the organizational level. Hence, governance should move to the organizational level. This might result in a formal and procedural focus in which nobody dares to take the lead. Furthermore, organizations' efforts might not be aligned with society's efforts, which may result in conflicting interests. Therefore, responsible governance is needed. Responsible governance should go beyond general policies, procedures, and guidelines and include actionable policies, controls, and checks and balances, providing multiple levels of defense and joint accountabilities. A broad range of governance mechanisms is needed, from decision-making authorities, procedures and processes, business–IT–user relationships, and communication to risk assessment. Governance evolves with AI systems, data, and other elements. Hence, new governance mechanisms might need to be introduced, and others can be removed over time.

Often GenAI starts as a project within an organization, and the accompanying governance is focused on the project level. Over time, there will be a move to the organizational level, and it will no longer be treated as a separate project but as part of IT governance. Thereafter, responsible governance can be created and focused on shifting the view toward the societal view in which the interacting agents and developments are taken into account. Table 2 shows the move from project to organizational to responsible governance and their main characteristics. During the transition phases, governance systems can exhibit characteristics belonging to multiple stages. The focus shifts from the project to the organizational and societal levels. Subsequently, the scope of the governance shifts from the project to the IT governance level, in which GenAI is governed separately by the organizations, to responsible governance, where GenAI is integrated into the whole organization. In the latter, a CAS view is adopted that considers the interacting agents. Also, the aims shift from making the project a success to balancing risks and returns to taking societal responsibility.

Responsible governance can be defined as joint efforts between stakeholders to establish oversight mechanisms based on societal values and norms, where societal concerns and the complexity of AI governance are taken as a starting point, and where people and society are engaged and empowered.

Table 1. Overview of governance mechanisms.

	Formal structures	Formal procedures	Relational governance	Communication	Risk assessment
Societal governance mechanisms	Ethical committees AI repositories External auditing	Public values identification Complaint procedures	Community engagement Partnerships Support in understanding AI use for outsiders	Algorithm registers Explaining to Journalist Awareness creation	Risks for liability Tracking societal complaints Following the news
Organizational governance mechanisms	Chief AI officer (CAIO) Chief Data Stewards (CDS) Human control of inputs, processing, and outputs Algorithm repositories AI stewardship Data stewardship Partnerships Involvement of experts Internal auditing Contracts and service level agreements	Planning & Control process Data life-cycle management es to deal with exceptions Samples and output checks Risk assessments processes Privacy and ethical risk assessment Anomalies detection Training and Education	Knowledge sharing Escalation and conflict Sharing of practices Shared understanding Job rotations Contracts Liability	Training Best practices sharing	Human-control of outputs External expertise Privacy and security assessment AI red team
Project-oriented governance mechanism	Teams with different fields of expertise Regular meetings to discuss progress, challenges	Data usage processes Involvement of others in decision-making Data auditing and assessment	Team-building Job rotating	Regular meetings Informal team-building Impact awareness sessions	Risk assessment Sandboxing
Technology-embedded governance mechanisms	Governance-by-design	System-level controls Privacy-by-design Transparency-by-design	Automated alerts	Escalation	Bias and discrimination detection and prevention

Table 2. Overview of changing characteristics when moving from project to responsible Governance.

	Project governance	IT governance	Responsible governance
Primarily focus	Project	Organizational	Societal
Governance scope	Project level	Separate governance of GenAI within the organization	Complex adaptive system on GenAI governance
Governance aims	Realization of project benefits of AI and progress	Balancing risks and return	Social responsibility

Responsible governance is a move from an inward organizational perspective to an outward organizational perspective on governance and a move from a static perspective on defining responsibilities toward a complex adaptive perspective, where joint accountability is stressed. CAS suggests having some basic governance principles that guide the effort, including the following:

1. *Public values as a starting point.* Societal values and expectations are taken as a starting point instead of a focus on profit and minimizing risks. Organizations keep track of dynamics and interact with society by taking a CASs approach.
2. *Data provenance.* Know what the original data are and be able to trace this back.
3. *Governance mechanisms for dealing with society.* Ensure governance mechanisms for interacting with society in the areas of decision-making authorities, formal processes, relationships, communication, and benefit/risk assessment. The governance mechanisms should facilitate learning and improvement to deal with the complexity and evolving nature, resulting in patterns of accountability.
4. *Countervailing governance mechanisms.* The risk of joint accountability is that there are no countervailing powers. Ensuring that opposing interests are mobilized by assigning persons responsible for certain concerns by separation of concerns and mobilizing opposing interests. For example, a responsibility for each public value can be allocated to somebody in the organization. These countervailing mechanisms should also result in saying “no” to AI.
5. *Defining joint accountability.* Agree on shared accountability for the use and impact of GenAI on society. This results in a deliberate trade-off being made between using AI or not.
6. *Multiple lines of defense.* One governance mechanism is not enough. Transparency about algorithms alone, for example, is not sufficient, and auditors should also be allowed to scrutinize algorithms.
7. *Enabling learning.* Learn from mistakes and drive continuous improvement. Start small and be critical to foster learning.

In future research, this list can be extended and refined based on new insights, the context, and the type of AI at hand. Nevertheless, CAS suggests that organizations can be pointed in the desired direction with a limited set of governance mechanisms. The AI Act also requires EU Member States to establish at least one regulatory “sandbox” to test AI systems before they are deployed. Furthermore, GenAI systems should disclose that their content or decisions were AI-generated. However, this does not take into account the need to establish governance and looks only at the technological realm. Governance is not easy to arrange, and what is sufficient governance is not always clear. Hence, we suggest establishing effective governance before using AI, testing its ability to deal with undesirable situations, and determining if governance is sufficient to perform the necessary interventions.

Conclusions

GenAI governance by organizations can easily overlook society, fail to consider the dynamics of the many interacting agents, and be overly internally focused. Although there is no shortage of governance mechanisms, the question is how effective these mechanisms are in keeping the focus on society. Often, organizations are too optimistic about the effectiveness that governance mechanisms can entail, as they disregard the complex interactions at play. CAS theory states that systems cannot be understood by looking at their individual parts but by taking the relationships between interacting agents and the underlying complexity as a starting point. High complexity is hard to govern, and the first step for

responsible governance might be to reduce complexity. The CAS lens shows that addressing accountability in a complex network of many stakeholders is fundamental, but the picture is often blurred due to co-evolution, myriad changes, and high complexity. Governance mechanisms should aim to deal with the complexity and co-evolution. The CAS lens suggests that the right directions can be chosen with a relatively limited set of governance mechanisms. The basic principle consists of identifying key societal values and then assigning governance mechanisms such as joint accountability roles, countervailing mechanisms, separating concerns, mobilizing opposing interests, and multiple lines of defense. Together, these mechanisms should enable learning and ensure that GenAI use is properly scrutinized.

Whereas IT governance is focused on performance improvements, the CAS approach sheds light on potential issues stemming from this focus, including depersonalization, deepening distrust, and alienation. The CAS view also emphasizes that governance mechanisms should evolve as AI and other agents evolve. Responsible governance takes the expectation of society as a starting point and enables learning and change over time. Responsible governance embodies a shift from an internal and static organizational view of AI governance to an external and dynamic view where complexity needs to be considered and where organizations and society should share experiences.

Responsible AI governance cannot be an afterthought and should be integrated into the overall governance model. AI, data, systems, and humans are material and interlinked, and governance should consider the whole system instead of having separate governance systems for each element. Current GenAI efforts are still basic and will likely become more advanced, being a first step in the GenAI evolution. Governance should evolve with AI systems. If AI remains at the project level, the assignment of clear accountabilities is challenging, and AI implications go beyond a single project. A major challenge is that no responsibility, and therefore no accountability, is taken due to the high risks of incorrect AI output and a lack of understanding of the complete socio-tech system. Governance is not easy to set up, but we recommend getting governance right before we start using AI. As such, organizations should start off in a sandbox environment, implementing and refining responsible governance, and assigning clear accountabilities, before implementing AI in real life.

Conflict of interest

None declared.

References

- Aaronson, S. A. (2023). Data dysphoria: the governance challenge posed by large learning models. *SSRN Electronic Journal*, SSRN 4554580. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4554580
- Abraham, R., Schneider, J., & Vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49(6), 424–438. <https://doi.org/10.1016/j.ijinfomgt.2019.07.008>
- Almeida, V., Mendes, L. S., & Doneda, D. (2023). On the development of AI governance frameworks. *IEEE Internet Computing*, 27(1), 70–74. <https://doi.org/10.1109/MIC.2022.3186030>
- Brown, A. E., & Grant, G. G. (2005). Framing the frameworks: A review of IT governance research. *Communications of the AIS*, 25(1), 696–712. <https://doi.org/10.17705/1CAIS.01538>
- CapGemini. (2024). *Harnessing the value of generative AI: 2nd edition*. <https://www.capgemini.com/insights/research-library/generative-ai-in-organizations-2024/>
- Carmichael, T., & Hadžikadić, M. (2019). *The fundamentals of complex adaptive systems*. Springer.
- Cave, S., & ÓhÉigeartaigh, S. S. (2018). An AI race for strategic advantage: Rhetoric and risks. *Paper presented at the Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, New Orleans, February 2–3.
- Chen, L., Zaharia, M., & Zou, J. (2023). How is ChatGPT's behavior changing over time? *arXiv preprint arXiv:09009*. <https://arxiv.org/abs/2307.09009>
- Dafoe, A. (2018). *AI governance: a research agenda*. Vol. 1442. Governance of AI Program, Future of Humanity Institute, University of Oxford.
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., & Ahuja, M. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71(4), 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Eisenhardt, K. M., & Sull, D. N. (2001). *Strategy as simple rules*. Harvard Bus Pub.

- Fayyad, U. M. (2023). From stochastic parrots to intelligent assistants—the secrets of data and human interventions. *IEEE Intelligent Systems*, 38(3), 63–67. <https://doi.org/10.1109/MIS.2023.3268723>
- Gasser, U., & Almeida, V. A. (2017). A layered model for AI governance. *IEEE Internet Computing*, 21(6), 58–62. <https://doi.org/10.1109/MIC.2017.4180835>
- Giray, L. (2023). Prompt engineering with ChatGPT: a guide for academic writers. *Annals of Biomedical Engineering*, 51(12), 2629–2633. <https://doi.org/10.1007/s10439-023-03272-4>
- Gonog, L., & Zhou, Y. (2019). A review: Generative adversarial networks. Paper presented at the 2019 14th IEEE conference on industrial electronics and applications (ICIEA), Xi'an, China.
- Hill, P. S. (2011). Understanding global health governance as a complex adaptive system. *Global Public Health*, 6(6), 593–605. <https://doi.org/10.1080/17441691003762108>
- Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy Artificial Intelligence. *Government Information Quarterly*, 37(3), 101493. <https://doi.org/10.1016/j.giq.2020.101493>
- Janssen, M., & Helbig, N. J. G. I. Q. (2018). Innovating and changing the policy-cycle: Policy-makers be prepared!. 35(4), S99–S105. <https://doi.org/10.1016/j.giq.2015.11.009>
- Janssen, M., & Kuk, G. (2016). The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly*, 33(3), 371–377. <https://doi.org/10.1016/j.giq.2016.08.011>
- Khanal, S., Zhang, H., & Taeihagh, A. (2024). Why and how is the power of Big Tech increasing in the policy process? The case of generative AI. *Policy and Society*, 44(1), 52–69. <https://doi.org/10.1093/polsoc/puae012>
- Lewin, R., & Regine, B. (1999). *The soul at work: Unleashing the power of complexity science for business success*. Simon & Schuster.
- Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. *AI and Ethics*, 2(4), 603–609. <https://doi.org/10.1007/s43681-022-00143-x>
- Maragno, G., Tangi, L., Gastaldi, L., & Benedetti, M. (2021). The spread of Artificial Intelligence in the public sector: A worldwide overview. Paper presented at the Proceedings of the 14th International Conference on Theory and Practice of Electronic Governance, Athens, Greece.
- Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, 6(3), 175–183. <https://doi.org/10.1007/s10676-004-3422-1>
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data and Society*, 3(2), 2053951716679679. <https://doi.org/10.1177/2053951716679679>
- Motoki, F., Neto, V. P., & Rodrigues, V. (2023). More human than human: Measuring ChatGPT political bias. *Public Choice*, 198(1), 3–23. <https://doi.org/10.1007/s11127-023-01097-2>
- Nah, F., Cai, J., Zheng, R., & Pang, N. (2023). An activity system-based perspective of generative AI: Challenges and research directions. *AIS Transactions on Human-Computer Interaction*, 15(3), 247–267. <https://doi.org/10.17705/1thci.00190>
- Nan, N. (2011). Capturing Bottom-up information technology use processes: A complex adaptive systems model. *MIS Quarterly*, 35(2), 505–532. <https://doi.org/10.2307/23044054>
- Nfuka, E. N., & Rusu, L. (2011). The effect of critical success factors on IT governance performance. *Industrial Management and Data Systems*, 111(9), 1418–1448. <https://doi.org/10.1108/02635571111182773>
- Perks, C., & Beveridge, T. (2003). *Guide to enterprise IT architecture*. Springer.
- Peterson, R. (2004). Crafting information technology governance. *Information Systems Management*, 21(4), 7–22. <https://doi.org/10.1201/1078/44705.21.4.20040901/84183.2>
- Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. *Policy and Society*, 40(2), 178–193. <https://doi.org/10.1080/14494035.2021.1929728>
- Roberts, N. (2002). Keeping public officials accountable through dialogue: Resolving the accountability paradox. *Public Administration Review*, 62(2), 658–669. <https://doi.org/10.1111/1540-6210.00248>
- Sambamurthy, V., & Zmud, R. W. (1999). Arrangements for information technology governance: A theory of multiple contingencies. *Management Information Systems Quarterly (MISQ)*, 23(2), 261–290. <https://doi.org/10.2307/249754>
- Scherer, M. (2015). Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. *Harvard Journal of Law & Technology*, 29(2), 353–398.
- Schneider, J., Abraham, R., & Meske, C. J. (2024). Governance of generative artificial intelligence for companies. *Information Systems Management*, 40(3), 229–249. <https://doi.org/10.1080/10580530.2022.2085825>
- Sullivan, T. (2011). Embracing complexity. *Harvard Business Review—Business Communications*.

- Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and Society*, 40(2), 137–157. <https://doi.org/10.1080/14494035.2021.1928377>
- Van de Poel, I. (2020). Embedding values in artificial intelligence (AI) systems. *Minds and Machines*, 30(3), 385–409. <https://doi.org/10.1007/s11023-020-09537-4>
- Van Noordt, C., & Misuraca, G. (2022). Artificial intelligence for the public sector: Results of landscaping the use of AI in government across the European Union. *Government Information Quarterly*, 39(3), 101714. <https://doi.org/10.1016/j.giq.2022.101714>
- van Noordt, C., & Tangi, L. (2023). The dynamics of AI capability and its influence on public value creation of AI within public administration. *Government Information Quarterly*, 40(4), 101860. <https://doi.org/10.1016/j.giq.2023.101860>
- Weill, P. (2004). Don't just lead, govern: How best performing organisations govern IT. *MIS Quarterly Executive*, 3(1), 1–17.
- Weill, P., & Ross, J. W. (2004). *IT governance: How top performers manage IT decision rights for superior results*. Harvard Business Press.
- Wirtz, B. W., Weyerer, J. C., & Sturm, B. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration*, 43(9), 818–829. <https://doi.org/10.1080/01900692.2020.1749851>
- Yu, E. (2023). 75% of businesses are implementing or considering bans on ChatGPT. ZDNET. <https://www.zdnet.com/article/75-of-businesses-are-implementing-or-considering-bans-on-chatgpt/>