A Management Control oriented Governance Framework for Artificial Intelligence

Richard Sentinella, Maël Schnegg and Klaus Möller

Summary: In an age of increasing access to and power of artificial intelligence (AI), ethical concerns, such as fairness, transparency, and human well-being have come to the attention of regulators, standard setting bodies, and organizations alike. In order to build AI-based systems that comply with new rules, organizations will have to adopt systems of governance. This study develops, based on existing frameworks and a multiple case study, a governance framework specifically designed with these challenges in mind: The St. Gallen Governance Framework for Artificial Intelligence focuses on identifying stakeholder concerns and strategic goals, building a management control system, assigning roles and responsibilities, and incorporating dynamism into the system of governance.

Keywords: Artificial Intelligence Governance, Management Control System, Levers of Control, AI Ethics,

Managementkontrollen orientiertes governance Framework für künstliche Intelligenz

Zusammenfassung: Im Zeitalter des zunehmenden Zugangs zu und der Macht von künstlicher Intelligenz (KI) rücken ethische Belange wie Fairness, Transparenz und menschliches Wohlergehen in den Fokus von Regulierungsbehörden, Normungsgremien und Organisationen gleichermaßen. Um KI-basierte Systeme zu entwickeln, die den neuen Regeln entsprechen, müssen Organisationen Governance-Systeme einführen. Diese Studie entwickelt auf der Grundlage bestehender Frameworks und mehrerer Fallstudien ein Governance-Framework, das speziell auf diese Herausforderungen ausgerichtet ist: Das St. Galler Governance Framework für Künstliche Intelligenz konzentriert sich auf die Identifizierung von Stakeholder-Belangen und strategischen Zielen, den Aufbau eines Management-Kontrollsystems, die Zuweisung von Rollen und Verantwortlichkeiten und die Einbeziehung von Dynamik in das Governance-System.

Stichworte: Governance künstlicher Intelligenz, Managementkontrollsystem, Kontrollhebel, Ethik künstlicher Intelligenz

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

1 Introduction

Artificial Intelligence (AI) has been hailed in recent years to be a panacea for organizations plagued by rising costs and faltering productivity, while others claim it is a doomsday technology that will dismantle society and replace workers with computer programs and robots (Butcher and Beridze, 2019; Harari, 2017; Jobin et al., 2019; Quattrone, 2016). Currently, AI is causing changes in the economy, driving organizations in a wide range of sectors to adopt it (Davenport & Ronanki, 2018). The promises and forewarnings of AI have not played out, although organizations are looking for new ways to use AI to their advantage.

Despite being aware of the benefits and promises of AI, its use has not spread quickly and evenly across the economy. AI uses three factors to mimic human intelligence: large high-quality datasets, advanced mathematical models, and large amounts of computing power to train the models on the datasets (Collins et al., 2021). While in the past AI might have been shackled by lack of technical capabilities, now organizations are faced with different challenges of using AI. While research in Information Systems and Technologies brought understanding on the technology itself the deployment of AI in organizations presents challenges beyond technical aspects. Organizations will have to define which of their activities can be enhanced by AI. Organizations will have to acquire specific human capital as well as new processes, governance structures, and operating models (Afiouni & Afiouni-Monla, 2019; Alsheibani et al., 2020; Brynjolfsson et al., 2019; Mikalef & Gupta, 2021; Papagiannidis et al., 2021). Ethical issues, such as fairness, transparency, and human well-being, drive the need for governance structures (Butcher & Beridze, 2019; Jobin et al., 2019). This study addresses the ethical challenges of AI by reviewing existing AI governance frameworks and case studies of Swiss organizations to understand what organizations need. This knowledge is summarized in a new governance framework. Such a framework can benefit two audiences: (1) organizations building ill-conceived projects without framing it into a full organizational concept and (2) organizations reluctant to attempt to leverage on technology due to their fear of ethical misbehavior.

Researchers showed that increasing transparency in AI models mitigated trust issues and increased adoption (Chowdhury et al., 2022). The "black box" problem, where humans cannot verify or understand an AI-based system, can cause the fear that an AI-based model is more biased than previous methods and create issues of fairness (Ntoutsi et al., 2020). The third common ethical concern pertains to the changing role of humans. Harari (2017) claims that widespread automation via AI could create a "useless class," changing the relationship between employer and employee as well as the structure of the labor market. Together, the ethical concerns of transparency, fairness, and human well-being might prove to be enough for organizations to delay adopting AI. Other researchers temper this fear, such as Davenport and Ronanki (2018), who argue that AI will be a complement to workers rather than a replacement, as there are unique advantages to humans and AI.

The tensions mentioned above all stem from the challenge to integrate human- and machine-driven processes in an organization. The objective of this paper is to identify whether governance mechanisms can be identified and implemented to reduce the various gaps. To do so, the existing literature and framework are analyzed and confronted with the reality of four case studies, leading to the proposition of a management control-oriented governance framework for AI. The paper is organized with the following structure. The following section identifies relevant literature on AI Governance, in which several

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

existing AI Governance frameworks are compared. The methodology section explains how the four case studies are selected and analyzed. The findings are presented, which, together with the analysis of existing frameworks, inform the development of an original management control-oriented AI governance framework. The last section is reserved for the conclusion.

2 Development of AI Governance

The relationship between a principle and its agent is sensitive to the balance of power and the need for trust. Governance is a formalized system of checks and balances intended to ensure that the agent is working in the best interest of the principle (Daily et al., 2003; Eisenhardt, 1989). These checks and balances can be formal or informal within the organizational unit (Malmi & Brown, 2008). This system helps the principle achieve organizational goals and maintain formal and informal boundaries by ensuring that the agent is correctly motivated by the system of governance. Examples of governance can be seen in national and international standards, such as the Swiss Code of Best Practices for Corporate Governance (Böckli et al., 2014). Over time, the field of governance has expanded, touching other fields, such as risk management (Beasley et al., 2005), innovation (Sharif, 2012), and the use of information systems (Tonn & Stiefel, 2012). New technologies can affect governance practices (Brennan et al., 2019) and AI is a disruptive technology that will require new governance practices, additional rules, and allocation of responsibilities (Afiouni & Afiouni-Monla, 2019; Alsheibani et al., 2020; Brynjolfsson et al., 2019; Mikalef & Gupta, 2021; Papagiannidis et al., 2021). The literature of management control systems (MCS) and the levers of control (LOC) provides a lens to define and organize the system of governance within an organization (Simons, 1994).

Following the understanding of Schneider et al (2022), AI governance for organizations is defined as "the structure of rules, practices, and processes used to ensure that the organization's AI technology sustains and extends the organization's strategies and objectives." (Schneider et al., 2022, p. 5).

Following the first development of AI and its application in various organizational contexts, practitioners realized the challenge of successfully designing and implementing AI applications within organizations compared to other IT projects. Contrary to a traditional statistical model, the output of an AI model is hardly predictable as the model aims to learn from an evolving dataset (Collins et al., 2021). In the name, machine learning, lies the idea that the model can learn continuously after being released. Even for the software industry, which has been working with regular release of updates, AI could mean a regularly evolving product. In other fields, such as medical treatment, such paradigms are brand new: the release of treatment is meant to evolve over time, which contradicts the current understanding of a – once discovered – continuously valid treatment (Vokinger et al., 2021; Vokinger & Gasser, 2021). This new paradigm created by AI generates the need for new systems of governance.

Early calls into AI governance have focused on governance outside of the organization, for example for research on AI governance and ways to devise norms, policies, and institutions globally (Dafoe, 2018). This call has been taken up by myriad governments (European Commission, 2020; European Parliamentary Research Service, 2022; Personal Data Protection Commission, 2020) intergovernmental organizations (OECD, 2019), standard setting bodies (Cihon et al., 2021) and private groups, such as by the World

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

Economic Forum (2020). The research field of AI governance looking at organizational or corporate governance is relatively young, although organizations play an important role in the AI ecosystem, from research and development to end use (Cihon et al., 2021). Holistic AI governance frameworks have been looking for ways to fill the gap of publicly available knowledge for how organizations can manage the use of AI.

The European Commission's High-level Expert Group on Artificial Intelligence first released the AI Ethics Guidelines for Trustworthy AI in 2019 (High-Level Expert Group on Artificial Intelligence, 2019), which informed its White Paper on AI (European Commission, 2020) and the proposal for the Artificial Intelligence Act (European Parliamentary Research Service, 2022). The EU's desire is to set a standard for "ethical, secure, and cutting-edge AI made in Europe" (High-Level Expert Group on Artificial Intelligence, 2019, p. 4). The European Commission sets rules around which systems have an unacceptable risk (e.g., no social scoring systems would be allowed), which are high risk (e.g., applicant tracking systems for recruiting) and require strict governance and reporting, which systems are low risk (e.g., chatbots) and require transparency, and which systems are minimal risk (e.g., forecasting) and do not have any mandatory governance requirements (European Commission, 2020).

The OECD has their own guidance and recommendations based on previously released guidelines, marking in 2019 the first intergovernmental standard (OECD, 2019). These recommendations are partly technical recommendations (e.g., AI should be robust, secure, and safe) and partly ethical (e.g., AI should foster inclusive growth, sustainable development, and well-being). ISO and IEEE are involved in setting standards for AI use. Although IEEE is more focused on ethical implications and unlikely to be adopted as a mandatory standard, the organization's standards will likely be adopted as a voluntary standard or be expected in certain industries (Cihon et al., 2019). The ISO has had a history of being adopted into national regulation and thus is likely to be mandatory once their standards are further refined (Cihon et al., 2019).

Two recently developed tools for organizations to show they conform to these regulation and standards are the Conformity Assessment Protocol for AI (capAI) and the Responsible AI Institute (RAII) certification. capAI is a structured auditing procedure by which an organization can assess its conformity to the EU's Artificial Intelligence Act. The procedure is composed of three components: the internal review protocol, which is a tool for organizations for quality assurance and risk management; a summary datasheet to be compliant with EU reporting requirements; and an external scorecard, which provides internal and external stakeholders with information about the AI-based system (Floridi et al., 2022). The RAII certification is another way that stakeholders can be assured that the AI-based system complies to various regulations and standards, including those from the EU, OECD, IEEE, and ISO (The Responsible AI Institute, 2022). Currently, the program is being tested in the US and UK on three use cases: AI Procurement, automated lending, and skin imaging. Like capAI, the RAII process involves answering a series of questions, whose response indicators inform a score which will in turn inform the certification (The Responsible AI Institute, 2022). Both approaches enable organizations and their stakeholders to have a reasonable assumption that they conform to the applicable standards and regulations represented by the auditing or certification process.

A governance framework should help the organization understand and design its system of governance while giving suggestions to fill any potential gaps to pass a quality check,

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

such as the capAI or RAII certification. An example is the Implementation and Self-Assessment Guide for Organisations published by the Personal Data Protection Commission of Singapore (2020). Various frameworks have popped up as the field of AI governance has matured (Baquero et al., 2020; EY, 2019; Gasser & Almeida, 2017; Personal Data Protection Commission, 2020; Rolls Royce, 2021; Schneider et al., 2022; Sivakumar et al., 2020; Sokalski et al., 2019; van Giffen et al., 2020). Both academic and practitioner frameworks are presented and analyzed in the following to understand what the state of the art is. The choice to include both types of frameworks reflects their different aims: practitioner frameworks are more proscriptive and offer ways to patch leaky systems, while academicians develop theoretical and descriptive models, mostly stemming from an Information Systems (IS) perspective.

Source	Name	Audience	Main Takeaways
Gasser and Almeda, 2017	A Layered Model for AI Governance	Academia	▪ Governance occurs at different layers (Social and legal/Ethical/Technical). ▪ The "black box" nature of AI creates information asymmetries between the different stakeholders: developers, consumers, and policymakers.
van Giffen et al., 2020	St. Gallen Management Model for AI (SGMM-AI)	Organizations	▪ Organizations need to adapt their management model to gain value from AI-based technologies. ▪ Organizations need to modify their management and organizational models to meet the new challenges presented by AI.
Schneider et al., 2022	Conceptual Framework for Data Governance	Academia	▪ The framework fosters collaboration across functions, structuring and formalizing AI management. ▪ Businesses should define how and who makes decisions, develop supporting artefacts (policy, standards, and procedures), monitoring compliance.
EY, 2019	Building the right governance model for AI/ML	Financial Services	▪ Preparing for coming regulatory hurdles and aligning AI use with organizational strategy will help build stakeholder trust and accountability.
Sokalski et al., 2019	Controlling AI: The imperative for transparency and explainability	Organizations	▪ Building trust around AI is a key goal of business leaders. ▪ Self-regulation will soon end with new policy initiatives. ▪ Companies are not sure how to approach AI governance.
Sivakumar et al., 2020	OmniaAI: Building trust in AI	Organizations	▪ To manage AI risks, business leaders must consider three key questions: when to enact governance mechanisms, who is accountable for them, and how to operationalize governance and enable the organization.

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

Source	Name	Audience	Main Takeaways
Baquero et al., 2020	Derisking AI by design: How to build risk management into AI development	Organizations	▪ Businesses should rethink risk management in the face of new compliance and reputational risks from AI. ▪ Risk management should be embedded into the life cycle of AI.
Personal Data Protection Commission, 2020	The Model Artificial Intelligence Governance Framework	Organizations	▪ AI should be (1) explainable, transparent, and fair and (2) human centric. ▪ The four areas the Model Framework focuses on are: internal governance structures and measures, human involvement in AI-augmented decision-making, operations management and stakeholder interaction and communication.
Rolls Royce, 2021	Aletheia Framework	Organizations	▪ The framework focuses on Social Impact, Accuracy/Trust, and Governance. ▪ AI should be designed to act ethically and produce unbiased results faithful to the designed purposed.

Figure 1: Summary table for existing AI Governance frameworks

2.1 AI Governance Frameworks

Gasser and Almeida (2017) created the layered model for AI governance as a framework to understand the different layers of AI governance in organizations and society. The researchers are concerned with the "black box" problem creating information asymmetries between developers, consumers, and policymakers (Gasser & Almeida, 2017). In the model, the outermost layer is the social and legal layer, which enfolds the norms, regulations, and legislations that apply to AI-based systems. The ethical layer is the middle layer and could be understood as incorporating the stakeholder's interests. The technical layer is the layer of governance that directly affects the AI-based system, such as standards, data governance, and algorithm accountability.

While the Gasser and Almeida model does not explicitly call out organizations as the focus of their framework, van Giffen et al. (2020) published a framework with a focus on helping organizations adapt their management model to gain value from AI-based technologies. The model focuses on "(1) Management of artificial intelligence, (2) organization of business operations, (3) legal, (4) regulation and compliance, (5) life-cycle management, (6) management of technology infrastructure, and (7) cyber security" (van Giffen et al., 2020, p. 11). The authors argue that organizations can achieve great value with ML and AI but have to modify their management and organizational models due to the new challenges presented by AI.

Schneider et al. (2022) did not base their model and recommendations on empirical data, but rather on existing literature. They "emphasize the following six parts of AI governance: fostering collaboration across functions, structuring and formalizing AI management through a framework, focusing on AI as strategic asset, defining of how and who makes decisions, developing supporting artefacts (policy, standards, and procedures),

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

[and] monitoring compliance" (Schneider et al., 2022, p. 5). As one of the most thorough of the governance frameworks analyzed, their work is informative for this article.

Many AI governance frameworks emerged from leading consulting firms. These are published online with the goal of helping organizations tackle the issue of using AI, often with a nod towards ethical issues that organizations might face. The consulting firms agree that AI governance will help organizations overcome ethical and organizational issues that AI technology can bring. *EY* (2019) focuses on preparing for coming regulatory hurdles, especially in industries like banking and financial services, and aligning AI use with organizational strategy. Others, like *KPMG* (Sokalski et al., 2019) or *Deloitte* (Sivakumar et al., 2020), argue in favor of building trust with the technology rather than focusing on regulatory challenges. The AI governance framework by *McKinsey* focuses on how to manage risks to enable greater adoption within the organization (Baquero et al., 2020). These models are freely available for organizations to use, although they are lacking in depth compared to other frameworks.

The Model Artificial Intelligence Governance Framework created by the Personal Data Protection Commission of Singapore stands in contrast to the consultant models regarding the depth of materials available online for organizations to use, since they publish materials on their framework, case studies, and guides on their website. The two guiding principles of the framework are that AI should be (1) explainable, transparent, and fair and (2) human-centric (Personal Data Protection Commission, 2020). The four areas the Model Framework focuses on are: "internal governance structures and measures, human involvement in AI-augmented decision-making, operations management and stakeholder interaction and communication." (Personal Data Protection Commission, 2020, p. 11) The framework is one of the first comprehensive, flexible frameworks to help organizations align themselves to the principles and regulations that are being drafted and implemented in Europe (European Commission, 2020; European Parliamentary Research Service, 2022) as well as voluntary standards and principles around the world (Cihon et al., 2019; Personal Data Protection Commission, 2020). Many organizations have used the first edition of the framework after its release in 2019, the feedback from which has been incorporated into the second edition.

Although many frameworks are currently available for organizations to draw inspiration from, early movers created their own governance frameworks and shared their best practices. The Aletheia Framework (Rolls Royce, 2021) was developed by Rolls Royce, an industrial technology company who faced the need for a governance framework to ensure stakeholders of their ethical and safe use of AI. They later decided to publish it for other practitioners. The framework is aimed at practitioners, foregoing theory to lay out a 32-step process for organizations to follow. The 32 steps are divided into the categories of Social Impact, Accuracy/Trust, and Governance. The steps build a checklist that ensure that the AI is designed to act ethically and produces unbiased results faithful to the intended design.

2.2 Levers of Control as Theoretical Analysis Lens for Governance Frameworks

Corporate governance has focused on solving the collective agency and action problem caused by non-owner managers and its literature on the efficacy of governance mechanisms (Becht et al., 2003; Jensen et al., 1976). AI governance and many of the frameworks for corporate governance of AI evolved, however, organically from the field they originate

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

in: IT and IS literature, where the focus on the mechanisms put in place to ensure compliance and stakeholder interests are safeguarded, such as in De Haes and Van Grembergen (2004). This framework categorizes controls into the categories of structural, procedural, and relational mechanisms. This categorization can be seen in AI governance frameworks from the IS field such as Schneider et al (2022).

For more than 50 years, the field of strategic planning and management control has provided managers with the appropriate control system to drive their organization (Anthony, 1965; Lauzel & Cibert, 1962; Learned et al., 1965). To steer the organization through its internal and external challenges, the need for a balanced set of measures became more salient (Kaplan & Norton, 1992, 1996). Soon, the discussion evolved to include the interactions and complementarity of these systems (Grabner & Moers, 2013; Merchant & Otley, 2020). The need for a lens to look at the whole set of controls present in an organization was influenced by the LOC framework, proposed by Simons (1994), and used to understand how an organization can implement a (change of) strategy through a variety of levers. The use of LOC-based MCSs has been expanding since the framework was first published in 1994. Malmi and Brown (2008) and Ferreira and Otley (2009) both use their variation of the LOC framework to be able to analyze and understand how organizations use management controls for their strategic goals. Tessier and Otley (2012) adapted the LOC framework to understand governance practices and categorize them based on different characteristics, i.e., social/technical, performance/compliance, interactive/diagnostic, etc.

There is a need to govern over AI better, starting from the identification of a need to the implementation and use of AI, encompassing the various impacts it will have on the organization. As a mature field, management control can provide an appropriate lens to review the proposed framework and shore up their potential shortcomings.

3 Methodology

MCSs offer structured perspectives to analyze and understand how an organization uses governance mechanisms to implement strategic goals. The LOC framework in this study is a lens to analyze the governance frameworks in the literature review and the governance needs in the case studies. The analysis is structured along three key aspects of the LOC management control framework: the scope of the management control system, the roles and responsibilities used in the framework, and interactions in the MCS bolster dynamism. The in-depth comparison of the various AI governance frameworks from scientific publications and practical sources allows to derive common features and shortfalls. Thus, in the framework analysis some propositions for effective and efficient AI frameworks are developed using the analytical lens of the LoC.

A multiple-case study design was chosen for its robustness and to investigate the "how" and "why" pertaining to organizational use of AI governance (Yin, 2003). Interviewing multiple employees at each organization allowed for a triangulation of data in the analysis (see Carter et al., 2014; Gibbert & Ruigrok, 2010). The interviews were conducted in person or via an online conference platform. Seven interviews were recorded, while three interviewees could not be recorded for confidentiality reasons. The interviewees were informed about the research goals in advance and could speak German or English, which increased the participants' willingness to talk and the quality of information (Kurz et al.,

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

2007). As algorithmic fairness loomed large in media, the interviewees were assured of anonymity.

The case study interviews were conducted between August and October 2019 with four large and well-established Swiss organizations in telecommunication, transportation, and finance industries (see table 1). The four organizations chosen for the case study were all active in testing or using AI-based systems. The interviewees were chosen from employees involved with AI, but they were spread out among the respective organization's hierarchy, including the perspectives of both employee and manager.

Organization	No. Employees	Revenue	Industry	Interviewee
TelCo	10,000 – 25,000	CHF 10 – 20b	ICT	Head of AI 1[1,4]
				Data Scientist 1[1,4]
TransCo	25,000 – 50,000	CHF 5 – 10b	Transportation	Head of IT[1,5]
				Data Scientist 2 and Data Scientist 3[2,5]
FinCo1	5,000 – 10,000	CHF 1 – 5b	Financial Services	Data Governance Officer 1[1,5]
				Data Governance Officer 2[1,3,5]
FinCo2	50,000 – 75,000	CHF 20 – 50b	Financial Services	Head of AI 2[2,5]
				Head of AI 3[2,4]
				Information Governance Officer[2,3,5]

Notes: [1]the interview took place online; [2]the interview took place on site; [3]the interview could not be recorded for confidentiality reasons; [4]the interview took place in English; [5]the interview took place in German

Table 1: Organizations and interviewees included in the sample

The interviews were semi-structured around an interview guide, although interviewees were allowed to talk about other topics not covered by the questions. The interview guide was composed of open-ended questions around the organizations' AI projects, ethical concerns, and use of governance mechanisms. After each interview, a systematic interview report was written in English with the help of notes or recordings. The insights gathered during the interviews informed the development of propositions that follow.

4 Findings

In the following case studies, an empirical basis is examined to understand the issues facing organizations and what tools they use or plan to use to address these issues. As the organizations varied in their use and governance of AI, trends were able to be gleaned from the interviews, avoiding a specific framework bias but rather providing content-oriented insights. In the following, these trends are highlighted using the MCS lens.

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

4.1 Stakeholders' impact

The organizations care about how their employees view AI and address their concerns over AI use. The employees hold varying attitudes towards AI, but their attitudes are determined by several factors: work experience, skill level, and how the organization uses AI. New hires are positive towards AI, expecting to be provided with AI-support systems, while less-skilled, veteran employees fear job loss. When the focus is on building AI-based systems for unliked tasks, employees are more accepting.

Employees' ethical concerns can be alleviated by increasing explainability, transparency, and trust; *FinCo2* prioritizes systems that give reasons why a decision was made. The organizations refrain from letting AI make important decisions but allow less consequential decisions to be made algorithmically. *FinCo2* and *TelCo* ensure fairness by reducing biases compared to status quo and identifying which biases are acceptable, although *TelCo* believes automated decision making can be more fair than human decision making, as the biases are more easily quantified and alleviated. The levels of transparency and fairness in their AI-based systems are two areas where the organizations are actively investing time and resources.

Customers factor only indirectly into the ethical considerations of the system. There are concerns about handling customer data considering the General Data Protection Regulation (GDPR). Since GDPR article 22 provides the right to customers for an explanation of why a decision was made, the organizations were reticent when building decision-making algorithms that handled customer data. Customer data being used for AI was in long-term plans, but organizations wanted to have regulatory clarity and more experience building AI-based systems. *TelCo* sees complying with regulation as more important than creating value or employee concerns. *TransCo, FinCo1,* and *FinCo2* believe they would benefit from regulation of AI that defines general conditions but still gives room to innovate. The influence of regulators is leading the organizations to self-regulate until concrete regulation is available.

4.2 Scope of AI governance

The organizations had long-term and short-term goals for their use of AI tied to strategic goals and economic pressures. In the long-term at *TelCo*, AI "*will transform any single aspect of what we do*," while in the short-term automation through AI helped achieve two goals of the organization: saving time and increasing customer satisfaction. AI strategy at *FinCo2* is to automate increasingly sensitive systems because AI use is mandatory to remain competitive. Their short-term strategic goals were to enhance their service, improve data quality, and strengthen the protection of the customer. For the organizations, the goal of AI was not to reduce costs, but to solve a short-term (lack of workers) or a long-term (competitive pressure) challenge.

Aligning AI strategy with corporate strategy covered two critical goals. First, this ensured that the use of AI furthers strategic goals, such as increasing competitiveness or easing the effects of a worker shortage. Second, top managers support AI when it becomes a tool for accomplishing strategic goals. Often, top managers lack the technical knowledge and focus on AI's effects on productivity and profitability.

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

4.3 Existing and new roles and responsibilities for AI governance

Responsibilities around AI use (i.e., regarding ethical bias, risks, and decision making) are delegated to individuals and committees. The organizations created new roles or adapted existing roles. *FinCo2* has a risk officer, a compliance officer, and a legal officer, so they did not see an immediate need for additional roles and new responsibilities went to existing roles. The other organizations found use of new roles when figuring out where to assign responsibilities.

TelCo, *TransCo*, and *FinCo1* all had a version of an oversight committee for AI application and use. Its composition varied among the organizations; *TelCo's* committee entailed managers, data scientists, and experts. The committees did not make technical decisions, but made strategic decisions, such as on investments, workforce planning, and risk management. As its committee's purview was limited, *Telco* has a data governance office to review legal issues.

Technical decision making is left to technical employees in the data science or AI department, who have a greater understanding of how the systems work. Management is regularly informed about AI projects, but the decision-making authority remains with the technical project teams, likely due to management having limited technical understanding. While top managers have decision making authority in these organizations, they rely on technical employees for guidance and delegate day-to-day decisions.

The organizations defined roles when it comes to developing and managing AI in their organizations. Many of the organizations have taken advantage of existing resources, such as legal and compliance employees, while others see the need for new roles, such as ethics or oversight committees. The responsibilities associated to these roles were, however, not always clear. Moreover, the evolution of roles and responsibilities is reactionary and uncoordinated. There was a contradiction between two interviews at *FinCo1* whether fully automated decision making was used in the organization, leading to the assumption that there were not clear channels of communication within the organization. In the future, the organizations would benefit from clearer delineation of duties and a more formal organizational structure around AI use and responsibility.

4.4 Dynamic AI governance

Efficient controls are used continuously to identify and correct potential disruptions in the organizational activities or in its environment (Simons, 1994). Sets of indicators are periodically monitored and compared to elicit actionable insights. The organizations used KPIs to monitor AI-based systems as well as AI use overall. KPIs at *FinCo2* are used overall (productivity, customer response time, savings) and for each model (precision, recall). Productivity metrics are used to measure time saved, for example with an NLP system to extract information from unstructured data. *FinCo1* regularly surveys employees for a KPI to measure ethical implications of AI use. Generally, KPIs are used at two different levels, for the individual system to measure effectiveness and technical quality and overall to gauge the impact of AI on the organization's ethical or organizational goals.

Auditing and testing AI-based systems are important before and after deployment. Testing pre-deployment at *TransCo* and *TelCo* ensures that AI is an improvement compared to status quo. Auditing post-deployment for issues such as unintended bias was cited as an important goal for the organizations, but only *Telco* managed to employ a technical

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

auditing tool to test for group fairness. Since *TelCo* could not find a satisfactory auditing tool on the market or open source, they created it themselves. Overall, the tools for conducting AI audits were not well developed or standardized amongst the organizations. This leads to two issues: the auditing process relies on homemade software and does not lend itself to easy comparison across multiple organizations.

The risk management strategies of the organizations revolve around the use of human oversight and the exclusion of human-generated data. Humans are kept in the loop at *FinCo2* selectively based on the "sensitivity" of the use case: AI to help with pricing in B2B relationships supports humans, while text processing is fully automated. *TransCo* feels that their abstention from using human data allows them to avoid many ethical issues. Rules that match the level of risk to restrictions on AI use allow organizations to streamline less risky use cases and outline the risk levels the organization accepts.

To increase transparency at *TransCo*, information on AI use is available internally and they host informational workshops about AI. Nevertheless, few employees read the technical documentation or attend the workshops. Two organizations take stock of their AI use annually; *TransCo* reports on the state of affairs, while *FinCo2* states in their annual report that the organization is committed to transparency. At *TransCo,* data scientists request feedback when they present project results to employees. The organizations feel that there is value in clear communications with stakeholders, but the result is often perfunctory. Targeted trainings, outreach, and promotion of the AI system of governance could be effective ways of influencing stakeholders.

KPIs and other indicators were used to monitor different aspects of AI use. Auditing detects discrepancies between the expected outcome and the real outcome of AI-based systems and risk management matches appropriate restrictions to AI-based systems. Communication, although lacking, could be a valuable tool for influencing stakeholders. These aspects of dynamism ensure that the AI system of governance can adapt to changing circumstances.

5 Framework Development

Building on insights from the literature review and the case studies, a theory-based, practice-oriented holistic framework for AI governance in organizations (which is referred to as the sg-GFAI – St. Gallen-Governance Framework for Artificial Intelligence, see figure 1) has been developed. This framework incorporates four layers, which can be used as guiding steps when designing an organization's AI governance. The layers follow four propositions, which are substantiated in the following sections of this chapter:

A holistic AI governance framework should

1. be oriented towards the ethical concerns of the organization's stakeholders and incorporate the organizational strategy,
2. use all control levers to drive, monitor, and train ethical actions and behaviors,
3. clearly lay out roles and responsibilities needed in the system of governance, and
4. adapt dynamically to changing circumstances and environments.

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

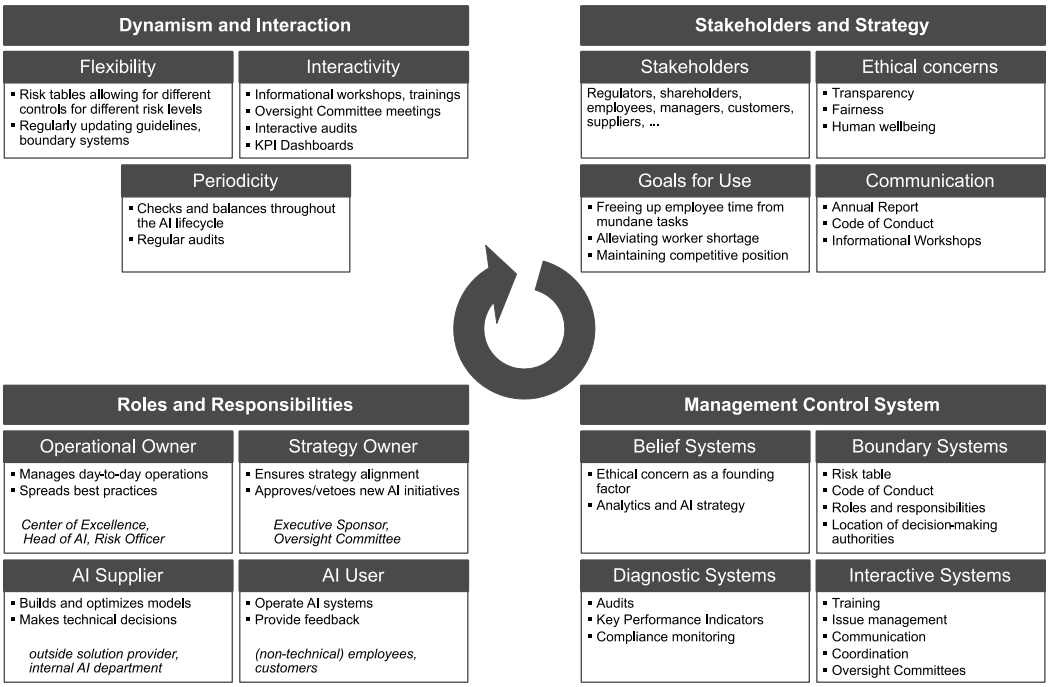

Figure 2: St. Gallen Governance Framework for AI (sg-GFAI)

5.1 Stakeholder and Strategy Layer

Although organizations rely on external forces when it comes to standards and regulations, identifying the concerns fitting to their development of AI can hardly be standardized. Stakeholder involvement is important for successful AI strategy and implementation (Crockett et al., 2021; Lee et al., 2019). The definition of fairness in AI use can vary across cultures, locations, and industries, so the issue of fairness in one organization might differ from another (Ntoutsi et al., 2020). A critical source of information is an organization's stakeholders analysis, including shareholders, board members, management, government regulators, employees, and members of their supply chains. By listening to them, organizations will be able to addresses their ethical concerns in the system of governance.

The literature review brought forward the ethical issues of fairness and transparency. Bias in the data can cause issues from illegal discrimination to flawed organizational decisions taken on flawed outputs. Both were the case when Amazon used an AI-based recruiting software that, trained on biased data, was found to unfairly discriminate against women (Dastin, 2018). The "black box" issue, where advanced AI models, such as neural networks, are indecipherable to humans, causes a lack of trust on a part of the stakeholders (van Giffen et al., 2020). Although techniques for deciphering the outputs or detecting unwanted bias in neural networks are being developed, these need to be included in the AI projects (De-Arteaga et al., 2022).

The St. Gallen Management Model for the Operational Use of AI mention issues that could occur, such as job loss or bias, but do not incorporate these as an ethical base for

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

their framework (van Giffen et al., 2020). The Conceptual AI governance Framework for Business is likewise focused on providing organizations with guidance and a framework for managing AI, but do not include ethical considerations in their framework (Schneider et al., 2022). The Layered AI Governance Model, in contrast, calls for ethical concerns to be included in AI governance, citing issues such as justice and equality, use of force, safety, privacy, displacement of labor, and taxation (Gasser & Almeida, 2017). The authors of this framework do not, however, address other stakeholder interests and strategic goals.

The Model Artificial Intelligence Governance Framework builds its framework by "translating ethical principles into practical recommendations" (Personal Data Protection Commission, 2020, p. 6). Among their ethical principles are that AI should be explainable, transparent, and fair and human centric. The first are in line with the consensus of the other frameworks. The suggestion of a human-centered ideology is echoed by *Deloitte* (Sivakumar et al., 2020) and *EY* (2019). Stating that the framework should be built on human centricity might have the effect of dissuading organizations, if they wish to pursue another ethical foundation.

The Model Artificial Intelligence Governance Framework encourages stakeholder communication and feedback (Personal Data Protection Commission, 2020), which is important when trying to understand which ethical concerns and strategic goals exist among the stakeholders. *KPMG* claims that their AI governance framework will help organizations bridge the "trust gap" and build confidence between the opaque AI-based system and the organization's stakeholders, implying that their framework is anchored in the concerns of stakeholders (Sokalski et al., 2019). The need for stakeholder interaction is not reflected in all frameworks, instead it is often limited to only a few stakeholders.

Strategic goals are an important factor in how the system of governance is shaped. For example, if the strategic goal is to address a worker shortage by automating tasks through AI, reducing the need for additional workers, then a framework that gives recommendations on the premise of human-centricity is not compatible with this strategic goal. Some frameworks do not address that AI governance also needs to enable organizations to innovate and drive forward strategic goals.

The ethical concerns of employees and managers are discussed in both the literature review and case studies. The three main themes are transparency, fairness, and human well-being. The lack of transparency in AI-based systems can be mitigated by avoiding black box AI models like neural networks and keeping humans in the loop as the final decision makers. Unintended or unwanted bias can be avoided or managed through careful data management, testing before deployment, and regular audits throughout the AI-based system's lifecycle. The employer can combat employee fear through policies of human-AI collaboration and clear communication.

There are multiple ways to solicit stakeholder concerns towards AI use in the organization. The organization should keep an eye out for coming AI regulation, such as the EU AI Act in Europe, as well as industry or product specific regulations (European Parliamentary Research Service, 2022). Standard setting bodies, such as the ISO or IEEE, are developing standards that will influence best practices and become industry or national standards(Cihon et al., 2019). Within the organization, the attitudes of employees and managers can be gathered via workshops, surveys, or other direct means. In practice, organizations are reluctant to pursue this information, as asking questions might draw unwanted attention to their practices.

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

In addition to the interests of their stakeholders, the organization should also keep in mind its strategic goals. In the case studies, the main goals were to remain competitive and free up employee time. Due to perceived pressure from outside the organization, AI was a way to remain competitive. Automating away unwanted tasks was seen as a way to free employees from specific processes. Although the organizations were reluctant to say that they wanted to reduce current headcount, their end goal was to reduce the need to hire additional employees, as the tight labor market had already led to employee shortages. Both goals were present in the case studies and influenced how the organizations used AI governance, including their approaches to communication, training, and human-AI interaction.

Once an organization has decided on ethical and strategic goals for their use of AI, the next step will be to codify and communicate these. Communication of ethical goals should be simple, flexible, and not just a written checklist (Crockett et al., 2021). Organizations should add these to their existing codes of conduct, hold informational workshops, or publish information on their use of AI governance in investor-oriented materials, such as their Annual Report.

Organizations should understand the ethical concerns held by stakeholders, elicited directly or indirectly, and match them to their strategic goals. Many examples of concerns and goals have been mentioned but they will be unique to the organization or industry and would likely change over time. In addition to using these to guide the implementation of their system of governance, the organization should think of the avenues to communicate these priorities to their stakeholders.

5.2 Management Control System Layer

One of the most widespread understandings of a management control system distinguish four levers of control to secure appropriate governance: belief system, boundary system, diagnostic system, and interactive system (Simons, 1994). Each system has a different purpose and is composed of various control mechanisms, and usually a practical control system is a mixture of all levers with different degrees of granularity and intensity in design and use.

Derived from the stakeholder and strategy layer, the belief system provides the understanding of why the organization uses AI. The beliefs of the organization set the direction of the ethical use of AI, as well as the organization's AI strategy. For example, *FinCo2* was clear in their belief that AI should augment humans instead of replacing them and that AI should focus on automating unenjoyable tasks. If properly communicated and adhered to within the organization, the belief system would alleviate employee fears. A belief system has the benefit of clearly communicating a stance on AI in a positive sense to internal stakeholders.

The boundary system shows the explicit conditions under which the organization is implementing and operating its AI-based systems. Among these are the formalized and non-formalized boundaries, the roles, and the responsibilities, as well as the location of decision-making authorities. A common way to set a boundary is to rely on a "risk table", identifying different levels of risk and the measures to be taken at each level. For example, the organization might identify the risk of losing personal customer data and will restrict itself from using them in its AI-based systems. This risk table is included in the EU AI Act (European Parliamentary Research Service, 2022) and the Singaporean Model Artificial

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

Intelligence Governance Framework (Personal Data Protection Commission, 2020). The Code of Conduct referenced in the case study identifies the practices the organization wants to avoid. At *FinCo2*, the Code of Conduct identifies the actions and outcomes that it should avoid. Boundary systems allow the organization to work within a confined set of parameters, helping to standardize the approval processes while incorporating regulatory requirements.

The diagnostic system is where the organization's AI use is coordinated and monitored. Audits are a popular tool to detect any variance and analyze its cause. The field of AI auditing is young, although many of the frameworks referenced its importance. *TelCo* stated that they built their own tool to audit their AI-based systems, as shelf-ready tools did not fit their requirements. Performance management, especially in the form of KPIs, was elaborated on in all case studies, where organizations found it useful to use a mix of soft and hard KPIs for each AI-based system and the overall AI program. Although this system is important, it was one of the least fleshed-out categories of management controls.

The interactive system is where stakeholders gather, discuss, and plan their AI use. Training and workshops were used to build technical capabilities or share information about the AI use. For example, *TransCo* hosted informational workshops to increase transparency about their AI use. Many frameworks discussed how an oversight committee, a group of stakeholders interacting at set intervals, might look with varying responsibilities. Interactive control systems help organizations reinvent the way AI is used in the organization and the way the organization can leverage a competitive advantage out of AI. In both the literature review and the case studies, a key observation was that organizations should favor discussion to constantly rethink how to use the technology. It is not clear, however, who should monitor and lead such tasks and at which moments and frequency these discussions should happen. This is discussed in the next two layers of the framework.

5.3 Roles and Responsibilities Layer

Roles and responsibilities are discussed in most of the governance frameworks, except for the Layered AI Governance Model (Gasser & Almeida, 2017). Some frameworks are more explicit than others in their call for roles and responsibilities. For example, van Giffen et al (2020) offers advice for organizations, but does not identify the roles and responsibilities for AI governance. In the other frameworks, the responsibilities are brought up in the context of a specific job title. In such cases, the practitioners relying on such frameworks lack the flexibility regarding the various responsibilities and how to split and distribute them among the various roles. The titles themselves do not matter, as long as the responsibilities exist in the organization (Schäfer et al., 2022).

Often these responsibilities already partially exist within the organization. *KPMG* suggests that organizations should look at what capabilities they are already using (Sokalski et al., 2019). *McKinsey* mentions specifically the need to provide staff with training as does *Deloitte*, so that existing staff can be upskilled to work with AI (Baquero et al., 2020; Sivakumar et al., 2020). For example, training can be provided to an existing project management office so that they can become experts in AI and its ethical issues. To preserve existing structures and not duplicate efforts, it is important to use these structures and roles in an adapted function.

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

The roles and responsibilities should build a system of checks and balances – for example, an executive sponsor wants to increase the accuracy of their sales forecasts, which is then sent to the AI Center of Excellence (CoE) to ensure that it is compliant with internal rules around data security and bias prevention and finally the AI committee approves it as a low-risk project with minimal controls. This allocating of responsibilities among different roles and ideally among multiple employees and stakeholders, allows for the roles and responsibilities to act as a system of checks and balances on the development and use of AI. Expanding on the roles and responsibilities, this layer identifies key roles and where responsibility may be assigned. Examples are given of how the responsibilities could be distributed.

Several of the consultant firm frameworks bring up two roles: the AI Oversight Committee and CoE. The Oversight Committee according to *EY* should "challenge the AI/ML adoption strategy for different aspects like fairness and conduct" (EY, 2019, p. 15). *Deloitte* suggests an Oversight Committee with diverse membership as a forum for decision-making around AI development and deployment (Sivakumar et al., 2020). The recommendation for an independent, critical voice within the organization is clear from these consulting firms' frameworks. The CoE should be staffed by professionals who are well versed in leading AI practices. Davenport (2021) describes the CoE as providing workers with coordination and leadership, helping to make the AI use of the organization coherent and efficient. *EY* suggests that CoEs have varying scopes, from purely advisory to leading AI development; CoEs can bring in leading practices and spread lessons learned through the organization (EY, 2019). *Deloitte* also suggests a CoE as a federated center of subject matter expertise, spreading awareness and training throughout the organization (Sivakumar et al., 2020). The CoE is comprised of full-time employees with a mandate to maintain day-to-day compliance with the strategic goals and ethical mandates, while the Oversight Committee is a diverse group of stakeholders within the organization who make forward-looking AI decisions.

The owner(s) of operational[1] activities take responsibilities for the day-to-day AI implementation and use. They also spread best practices and information about AI within the organization. This role could be covered, fully or in part, by a Chief AI Officer, but also head of AI or program manager. A CoE could also take part in covering the responsibilities of such role. Schäfer et al (2022) details many of the responsibilities for this role, such as external cooperation, managing cooperation between the organization and IT, strategy interpretation, reporting, among others. There should be an executive role that can help the organization adopt and create value with AI (Alkashri et al., 2020). The operational owner(s) take the responsibilities of a risk officer who concentrates on managing AI risks within the organization. The Conceptual AI Governance Framework for Business brings up the AI Risk Manager who can veto AI-based systems (Schneider et al., 2022). Schäfer et al (2022) identify the role of AI Risk Officer, who manages risk identification, reduction, and prevention, as well as audits. Hodge (2020) states that Risk Officers will

1 In the management control and strategic planning literature, a distinction is made between planning made at a strategic or operational level. In Information Systems literature, the use of operational can be confusing since it often refers to the operation of a system (in contrast to the system design and implementation). In this paper, the term operational refers to the planning, decisions, and activities, taking place in the on-going day-to-day implementation and use of AI, in contrast with the strategic activities.

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

need to ensure controls are in place at the right time in the development cycle, as well as that the severity of potential impacts are raised to leadership and safeguards are enacted to prevent negative impact.

At a strategic level, the strategy owner will have to decide which beliefs and boundaries will be followed by the organization and how to diffuse and enforce them. There might already be an executive sponsor for AI, who would be a good source of aligning organizational strategy. Alternatively, this role could also be filled by an interdisciplinary Oversight Committee who would be composed of stakeholders able to steer and oversee the use of AI. This role should be the first to be filled when initiating an AI system of governance, so that they can assign further responsibilities and set the beliefs and boundaries within the organization.

The AI supplier is the provider of the technical aspects of the AI-based system. The decision to buy ready-made systems or develop a system internally is a strategic decision. Ultimately, such decisions will be taken by the strategy owner, while the provider of the system builds and optimizes models for each given use case. In the case studies, this was often the stage that technical decisions were made, as it was specifically pointed out that the strategy owner would not decide technical aspects of AI.

The AI user interacts with the AI-based system. For example, in a system such as an applicant tracking system (ATS), the user would be the candidate applying to an open position, while also being the HR recruiter reading outputs from the ATS. In this role, the user can be the operator or a human whose data is being processed. In either case, the user should provide feedback to the AI supplier and other stakeholders about any technical or ethical issues they have. By involving users in the development process, organizations can increase trust and fairness in the AI-based systems (Lee et al., 2019; Siau & Wang, 2018). In the case studies, the interviewees knew how the users felt about AI and could adjust their AI use to work with their concerns and wishes.

5.4 Interaction and Dynamism Layer

Changing circumstances arise from the quick pace of development of AI-based systems and novel AI methods, driving the need for interaction and dynamism within the organization. Interactive controls can take the form of regular audits, committee meetings, and trainings. They create opportunities for stakeholders to meet and discuss issues.

Periodicity is important in creating a dynamic system of governance; controls, such as audits, are only effective when regular. *KPMG* (Sokalski et al., 2019) and *Deloitte* (Sivakumar et al., 2020) suggest that organizations should look to incorporate governance throughout the lifecycles of their AI-based systems, from development through deployment. The control could be triggered by another event, e.g., audits happening when a complaint has been raised, but organizations should look to which interval or trigger for the given control is appropriate for the model and use case.

McKinsey proposes a series of checks throughout the lifecycle of the AI-based system, as well as audits, to acknowledge, monitor, and abate risks (e.g., Model-robustness review or data-sourcing analysis) (Baquero et al., 2020). *EY* (2019) suggests performance and ecosystem monitoring, while *KPMG* (Sokalski et al., 2019) and *Deloitte* (Sivakumar et al., 2020) suggest that risks be continuously monitored through dashboards. Many of the controls organizations employ will become part of the development process for AI-based systems and manage their lifecycle.

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

The organization can provide for more dynamism through the thoughtful interactive use of its control systems. Periodicity can be introduced through regular audits of the AI-based system or setting specific checks along the lifecycle of an AI-based system, from idea to deployment and throughout its useful life, considering that the output generate by AI will evolve throughout its lifecycle (Vokinger et al., 2021; Vokinger & Gasser, 2021). Several interactive mechanisms have been suggested that should bring different stakeholders together to innovate and work through issues, such as informational workshops, oversight committees, interactive audits, and KPIs. These all serve as ways to balance the organization between maintaining control over its activities and providing room for innovation.

6 Conclusion

AI has been lauded as the next revolutionary technology but comes with ethical risks (Butcher and Beridze, 2019; Harari, 2017; Jobin et al., 2019; Quattrone, 2016). This study focuses on the issues of fairness, transparency, and human wellbeing and how they intersect with organizational decisions. An AI governance management control system, built with ethical concerns as its foundation, can alleviate ethical concerns by building up the necessary processes, structures, and models.

By examining AI governance frameworks and a multiple case-study, an AI governance framework was developed to help organizations build their own management control system. In the framework four distinct areas are identified from the propositions: identifying stakeholders, their needs, and the strategic goals of the organization; building a system of management controls around stakeholder concerns and strategic goals; assigning roles within the organization to ensure specific responsibilities are filled; and incorporating dynamism into the management control system. This framework serves as four steps for organizations to identify the ethical and strategic goals of their AI program and build a system of governance around their specific needs.

This study was exploratory, building up a basis of knowledge and theories of how an MCS would affect an organization's use of AI. MCSs were chosen to analyze and embody the system of governance because they focus on implementing a strategy in the organization, while the interactions between the different systems have the potential to increase beneficial aspects, such as innovation. Further research should confirm the effects of MCSs in managing AI within the organization, looking at both the overall effects of the MCS on innovation and risk, while also the working parts of the MCS. For example, Schäfer et al (2022) are working on detailing best practices for the roles of Chief AI Officer and AI Risk Officer. Research that details the use of MCSs in case studies or further confirmatory studies will add depth to the field.

Literature

Afiouni, R., & Afiouni-Monla, R. (2019). Learning in the Rise of Machine Learning Organizational Learning. In *Organizational*. https://aisel.aisnet.org/icis2019/business_models/business_models/2

Alkashri, Z., Siyam, N., & Alqaryouti, O. (2020). *A detailed survey of Artificial Intelligence and Software Engineering: Emergent Issues*.

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

Alsheibani, S. A., Cheung, Y., Messom, C., & Alhosni, M. (2020). *Winning AI Strategy: Six-Steps to Create Value from Artificial Winning AI Strategy: Six-Steps to Create Value from Artificial Intelligence*. *12*. https://aisel.aisnet.org/amcis2020

Anthony, R. N. (1965). *Planning and control systems: a framework for analysis*. Harvard University Press.

Baquero, J. A., Burkhardt, R., Govindarajan, A., & Wallace, T. (2020, August 13). *Derisking AI by design: How to build risk management into AI development*. McKinsey & Company.. https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/derisking-ai-by-design-how-to-build-risk-management-into-ai-development

Beasley, M. S., Clune, R., & Hermanson, D. R. (2005). Enterprise risk management: An empirical analysis of factors associated with the extent of implementation. *Journal of Accounting and Public Policy*, *24*(6), 521–531. https://doi.org/10.1016/j.jaccpubpol.2005.10.001

Becht, M., Bolton, P., & Roëll, A. (2003). *CORPORATE GOVERNANCE AND CONTROL*.

Böckli, P., Burger, M., Forstmoser, P., Frick, D. P., Hofstetter, K., Stiefel, C., Herzog, E., & Winzeler, C. (2014). *Swiss code of best practice for corporate governance*.

Brennan, N. M., Subramaniam, N., & van Staden, C. J. (2019). Corporate governance implications of disruptive technology: An overview. In *British Accounting Review* (Vol. 51, Issue 6). Academic Press. https://doi.org/10.1016/j.bar.2019.100860

Brynjolfsson, E., Rock, D., & Syverson, C. (2019). Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics. In *The Economics of Artificial Intelligence: An Agenda* (pp. 23–57).

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

Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J., & Neville, A. J. (2014). The Use of Triangulation in Qualitative Research. *Oncology Nursing Forum*, *41*(5), 545–547. https://doi.org/10.1188/14.ONF.545-547

Chowdhury, S., Joel-Edgar, S., Dey, P. K., Bhattacharya, S., & Kharlamov, A. (2022). Embedding transparency in artificial intelligence machine learning models: managerial implications on predicting and explaining employee turnover. *The International Journal of Human Resource Management*, 1–32. https://doi.org/10.1080/09585192.2022.2066981

Cihon, P., Leung, J., Ding, J., Garfinkel, B., Dafoe, A., Maas, M., Zwetsloot, R., Hagebölling, D., Carey, R., Zhang, B., Fischer, S.-C., & Shevlane, T. (2019). *Standards for AI Governance: International Standards to Enable Global Coordination in AI Research & Development*. https://arxiv.org/pdf/1802.07228.pdf

Cihon, P., Schuett, J., & Baum, S. D. (2021). Corporate governance of artificial intelligence in the public interest. *Information (Switzerland)*, *12*(7). https://doi.org/10.3390/info12070275

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management*, *60*. https://doi.org/10.1016/j.ijinfomgt.2021.102383

Crockett, K. A., Gerber, L., Latham, A., & Colyer, E. (2021). Building Trustworthy AI Solutions: A Case for Practical Solutions for Small Businesses. *IEEE Transactions on Artificial Intelligence*, 1–1. https://doi.org/10.1109/tai.2021.3137091

Dafoe, A. (2018). *AI Governance: A Research Agenda*. www.fhi.ox.ac.uk/govaiagenda

Daily, C. M., Dalton, D. R., & Cannella Jr, A. A. (2003). Corporate governance: Decades of dialogue and data. *Academy of Management Review*, *28*(3), 371–382.

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

Dastin, J. (2018, October 11). *Amazon scraps secret AI recruiting tool that showed bias against women*. Reuters. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

Davenport, T. H. (2021). *Enterprise Adoption and Management of Artificial Intelligence* (Vol. 01).

Davenport, T. H., & Ronanki, R. (2018). *Artificial Intelligence for the Real World*.

de Haes, S., & Grembergen, W. van. (2004). *IT Governance and Its Mechanisms*. www.isaca.org.

De-Arteaga, M., Feuerriegel, S., & Saar-Tsechansky, M. (2022). *Algorithmic Fairness in Business Analytics: Directions for Research and Practice*. http://arxiv.org/abs/2207.10991

Eisenhardt, K. M. (1989). *Agency Theory: An Assessment and Review* (Vol. 14, Issue 1). Academy of Management Review.

European Commission. (2020). *On Artificial Intelligence-A European approach to excellence and trust White Paper on Artificial Intelligence A European approach to excellence and trust*. https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf.

European Parliamentary Research Service. (2022). *Artificial intelligence Act*.

EY. (2019). *Building the right governance model for AI/ML*. https://assets.ey.com/content/dam/ey-sites/ey-com/en_us/topics/financial-services/ey- building-the-right-governance-model.pdf?download

Ferreira, A., & Otley, D. (2009). The design and use of performance management systems: An extended framework for analysis. *Management Accounting Research*, *20*(4), 263–282. https://doi.org/10.1016/j.mar.2009.07.003

Floridi, L., Holweg, M., Taddeo, M., Silva, J. A., Mökander, J., & Wen, Y. (2022). *capAI A procedure for conducting conformity assessment of AI systems in line with the EU Artificial Intelligence Act*. https://ssrn.com/abstract=4064091

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

Gibbert, M., & Ruigrok, W. (2010). The '"What"' and '"How"' of Case Study Rigor: Three Strategies Based on Published Work. *Organizational Research Methods*, *13*(4), 710–737. https://doi.org/10.1177/1094428109351319

Grabner, I., & Moers, F. (2013). Management control as a system or a package? Conceptual and empirical issues. *Accounting, Organizations and Society*, *38*(6–7), 407–419. https://doi.org/10.1016/j.aos.2013.09.002

Harari, Y. N. (2017, February 24). *The rise of the useless class*. https://ideas.ted.com/the-rise-of-the-useless-class/

High-Level Expert Group on Artificial Intelligence. (2019). *Ethics Guidelines for Trustworthy AI*. https://ec.europa.eu/digital-

Hodge, N. (2020). The Evolution of the Risk Manager. *Risk Management*.

Jensen, M. C., Meckling, W. H., Benston, G., Canes, M., Henderson, D., Leffler, K., Long, J., Smith, C., Thompson, R., Watts, R., & Zimmerman, J. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. In *Journal of Financial Economics* (Issue 4). Harvard University Press. http://hupress.harvard.edu/catalog/JENTHF.html

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

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard–measures that drive performance. *Harvard Business Review*, *70*(1), 71–79.

Kaplan, R. S., & Norton, D. P. (1996). Using the balanced scorecard as a strategic management system. *Harvard Business Review*, *74*(1), 75–85.

https://doi.org/10.5771/0042-059X-2023-2-182 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

Kurz, A., Stockhammer, C., Fuchs, S., & Meinhard, D. (2007). Das problemzentrierte Interview. In *Qualitative Marktforschung* (pp. 463–475). Gabler. https://doi.org/10.1007/978-3-8349-9258-1_29

Lauzel, P., & Cibert, A. (1962). *Des ratios au tableau de bord*. Entreprise Moderne.

Learned, E. P., Andrews, K. R., Christensen, C. R., & Guth, W. D. (1965). *Business policy : text and cases*. Irwin.

Lee, M. K., Kusbit, D., Kahng, A., Kim, J. T., Yuan, X., Chan, A., See, D., Noothigattu, R., Lee, S., Psomas, A., & Procaccia, A. D. (2019). Webuildai: Participatory framework for algorithmic governance. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW). https://doi.org/10.1145/3359283

Malmi, T., & Brown, D. A. (2008). Management control systems as a package-Opportunities, challenges and research directions. *Management Accounting Research*, *19*(4), 287–300. https://doi.org/10.1016/j.mar.2008.09.003

Merchant, K. A., & Otley, D. (2020). Beyond the systems versus package debate. *Accounting, Organizations and Society*, *86*(xxxx), 101185. https://doi.org/10.1016/j.aos.2020.101185

Mikalef, P., & Gupta, M. (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. *Information and Management*, *58*(3). https://doi.org/10.1016/j.im.2021.103434

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M. E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., … Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(3). https://doi.org/10.1002/widm.1356

OECD. (2019). *Recommendation of the Council on Artificial Intelligence*. http://legalinstruments.oecd.org

Papagiannidis, E., Enholm, I. M., Dremel, C., & Mikalef, P. (2021). *Deploying AI Governance practices: A revelatory case study*. https://www.researchgate.net/publication/351870329

Personal Data Protection Commission. (2020). *Model Artificial Intelligence Governance Framework – Second Edition*.

Rolls Royce. (2021). *The Aletheia Framework 2.0*. https://www.rolls-royce.com/~/media/Files/R/Rolls-Royce/documents/stand-alone-pages/aletheia-framework-booklet-2021.pdf

Schäfer, M., Schneider, J., Drechsler, K., & vom Brocke, J. (2022). *AI GOVERNANCE: ARE CHIEF AI OFFICERS AND AI RISK OFFICERS NEEDED?* https://aisel.aisnet.org/ecis2022_rp/163

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

Sharif, M. N. (2012). Technological innovation governance for winning the future. In *Technological Forecasting and Social Change* (Vol. 79, Issue 3, pp. 595–604). https://doi.org/10.1016/j.techfore.2011.12.004

Siau, K., & Wang, W. (2018). Building Trust in Artificial Intelligence, Machine Learning, and Robotics. In *Get The Cutter Edge free* www.cutter.com (Vol. 31, Issue 2). www.cutter.com

Simons, R. (1994). *Levers of control: How managers use innovative control systems to drive strategic renewal*. Harvard Business Press.

https://doi.org/10.5771/0042-059X-2023-2-162 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.

Sivakumar, N., Vinelli, M., Singh, A., & Viktorova, M. (2020). *Building trust in AI*. https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/financial-services/ca- omnia-ai-operation-trust-pov-aoda-en.pdf

Sokalski, M., Klous, S., & Chandrasekaran, S. (2019). *Controlling AI: The imperative for transparency and explainability*. https://advisory.kpmg.us/content/dam/advisory/en/pdfs/kpmg-controlling-ai.pdf

Tessier, S., & Otley, D. (2012). A conceptual development of Simons' Levers of Control framework. *Management Accounting Research*, *23*(3), 171–185. https://doi.org/10.1016/j.mar.2012.04.003

The Responsible AI Institute. (2022). *A Certification for Responsible AI*.

Tonn, B., & Stiefel, D. (2012). The future of governance and the use of advanced information technologies. *Futures*, *44*(9), 812–822. https://doi.org/10.1016/j.futures.2012.07.004

van Giffen, B., Borth, D., & Brenner, W. (2020). Management von Künstlicher Intelligenz in Unternehmen. *HMD Praxis Der Wirtschaftsinformatik*, *57*(1), 4–20. https://doi.org/10.1365/s40702-020-00584-0

Vokinger, K. N., Feuerriegel, S., & Kesselheim, A. S. (2021). Mitigating bias in machine learning for medicine. *Communications Medicine*, *1*(1). https://doi.org/10.1038/s43856-021-00028-w

Vokinger, K. N., & Gasser, U. (2021). Regulating AI in medicine in the United States and Europe. In *Nature Machine Intelligence* (Vol. 3, Issue 9, pp. 738–739). Nature Research. https://doi.org/10.1038/s42256-021-00386-z

World Economic Forum. (2020). *Empowering AI Leadership*. https://spark.adobe.com/page/RsXNkZANwMLEf

Yin, R. K. (2003). Case study research design and methods third edition. *Applied Social Research Methods Series*, *5*.

Richard Sentinella is Research Associate and PhD Candidate at the Institute of Accounting, Control, and Auditing at the University of St. Gallen.

Address: Institute for Accounting, Controlling, and Auditing, Tigerbergstrasse 9, 9000 St. Gallen, Switzerland, Tel: +41 244 74 00, E-mail: richard.sentinella@unisg.ch

Maël Schnegg is Assistant Professor for Digital Performance Management at the Institute of Accounting, Control, and Auditing at the University of St. Gallen.

Address: Institute for Accounting, Controlling, and Auditing, Tigerbergstrasse 9, 9000 St. Gallen, Switzerland, Tel: +41 244 74 23, E-mail: mael.schnegg@unisg.ch

Klaus Möller is Full Professor for Controlling / Performance Management at the Institute of Accounting, Control, and Auditing at the University of St. Gallen.

Address: Institute for Accounting, Controlling, and Auditing, Tigerbergstrasse 9, 9000 St. Gallen, Switzerland, Tel: +41 244 74 06, E-Mail: klaus.moeller@unisg.ch

https://doi.org/10.5771/0042-059X-2023-2-182 - Generiert durch IP 216.73.216.173, am 23.06.2026, 08:54:19. © Urheberrechtlich geschützter Inhalt. Ohne gesonderte Erlaubnis ist jede urheberrechtliche Nutzung untersagt, insbesondere die Nutzung des Inhalts im Zusammenhang mit, für oder in KI-Systemen, KI-Modellen oder Generativen Sprachmodellen.