Governance of Artificial Intelligence for Business

Dr. Mahesh Sharma[1], Dr. Satpal Arora[2]
[1]*Associate Professor, Ideal Institute of Management and Technology [GGSIPU, Delhi], Karkardooma, Delhi – 110092*
[2]*Associate Professor, Ideal Institute of Management and Technology [GGSIPU, Delhi], Karkardooma, Delhi – 110092*

Abstract: Although the governance of artificial intelligence (AI) is extensively debated on a philosophical, sociological, and legislative level, few publications specifically target businesses. By drawing a conceptual framework from the literature, we fill this gap. We break down "AI governance" along the axes of who, what, and how "is governed" into the governance of data, machine learning models, and AI systems. This breakdown makes it possible for current governance arrangements to evolve. Measuring the value of data and developing new AI governance roles are fresh, business-specific issues.

Keywords: Machine Learning, Artificial Intelligence, IT Governance, Governance Framework, Data Governance.

INTRODUCTION

With expected spending of about 100 billion US dollars by 2023, up from 38 billion in 2019, artificial intelligence (AI) has become a crucial area of study and application (Shirer & Daquila, 2019). AI demonstrates intelligent behaviour in various ways, opening for a wide range of effective and economical applications. Although AI uses a variety of methodologies, in this study we define AI as systems that learn from samples, i.e., these systems rely on models from a sub branch of AI referred to as "machine learning." Deep learning is one of the machine learning (ML) techniques that uses data to infer decision-making behaviour. Since learning from data rather than extracting and applying domain expert rules produces AI systems of greater performance at comparatively low costs, it is responsible for the majority of AI triumphs. In many application areas, such as job recruitment (Pan et al., 2022), credit scoring (H. Wang et al., 2019), designing floorplans for microchips (Khang, 2021), managing predictive maintenance strategies (Arena et al., 2022), autopilots in aviation (Garlick, 2017), or autonomous driving, AI has demonstrated remarkable success (Meske et al., 2022). (Grigorescu et al., 2020). As a result, AI is a focus for many firms, and research indicates that 90% of CEOs believe that AI presents a business opportunity that is essential to the success of their organization (Ransbotham et al., 2019). At the same time, just 10% of CEOs claim that integrating AI has resulted in a major financial gain (Ransbotham et al., 2020). As a result, there is still a great deal of uncertainty around the efficient use of AI technology to create value in enterprises and the precise way to employ the technology to make profits for organizations.

Despite the fact that AI is not a new concept, its recent technical and legal developments are astounding (Burt, 2021), and they are likely to continue moving quickly for some time. Fast AI development makes it challenging for businesses to stay up with and discover effective governance systems to gain economically from AI. Companies must also abide by a growing number of rules pertaining to data, ML models, and AI systems. Furthermore, AI demonstrates traits that make it both desirable and difficult to rely on, for instance:

(1) Even for applications where AI's models are simple to construct, the output of the technology is frequently challenging to understand (Adadi & Berrada, 2018). (e.g., using AutoML or adapting existing models). It is challenging to understand why AI makes a particular choice or how an AI system functions in general. Meeting regulatory standards and upgrading systems beyond what is already known are made difficult by a lack of understanding.

(2) AI generates unanticipated outcomes that are partially out of an organization's control. It displays unpredictable, "ethics"-illiterate, data-driven behaviour that results in new security,

safety, and fairness problems. The performance of AI is largely guaranteed by statistics. AI could fail in unpredictable situations that could be taken advantage of by bad actors, as has been shown in the case of current self-driving automobiles (Morgulis et al., 2019). An AI model may be built on unbalanced data, such as image data that has more samples of one race than another and is frequently trained using objectives like optimizing overall accuracy. This might result in immoral models that, for instance, show that one race has a considerably higher error rate than another.

(3) Data may cause AI systems' judgments to be biased. As seen, for example, with Amazon's AI-based recruitment tool, systems with biased decision-making represent a reputation risk (Dastin, 2018). This might ultimately have legal repercussions. Even while these arrangements frequently spark public outrage, they are not a surprise. AI systems' brains, ML models, are often tuned for a particular measure (Goodfellow et al., 2016). They frequently ignore other considerations such as the rationale and knowledge on what these decisions should or should not be founded on in favour of focusing solely on performance, for example, by maximizing the number of accurate decisions. This demonstrates the necessity of procedures (and governance) to guarantee sound decision-making.

(4) AI is rapidly developing technologically, opening up new commercial prospects and chances for digital innovation. Many goods and businesses that have not previously used AI technology now contain AI components. It also produces new regulations and standards (Cihon, 2019). (Burt, 2021). Although AI does not currently display "general intelligence," this may change in the future (Ford, 2018). AI could potentially undermine human authority, which is one of the reasons why legislation relating to human agency and oversight are likely to be implemented in the future (European Commission, 2020). Therefore, governance systems are crucial for reducing AI-related problems and increasing the potential of AI within enterprises. The performance of a company is significantly correlated with its governance (Bhagat & Bolton, 2008). Business AI governance is the framework of laws, customs, and procedures utilized to guarantee that the organization's AI technology upholds and furthers its aims and goals. Our mission is to give organizations a governance viewpoint that is concrete, exposes pertinent governance ideas, and establishes guidelines and best practices for the successful application of AI to achieve corporate goals like profitability and efficiency. At the same time, research is still being done on AI acceptance and capabilities (e.g., Jöhnk et al., 2021; Mikalef & Gupta, 2021). This paper's emphasis is on (long-term) governance of AI, not on temporary AI governance during a transitory adoption phase.

In addition to legislative considerations, AI has made technological advancements that have an impact on governance. They enable new forms of governance or at the very least provide some insight into how feasible it is to put in place governance systems. In this work, we focus specifically on model governance, testing, data valuation, and data quality challenges.

While there is still disagreement regarding how disruptive AI will be to governance (Liu & Maas, 2021), it is important to take into account the current governance systems from both a research and practical standpoint. As a result, we see AI governance via the frameworks for data and information technology (IT) governance that are already in place. For instance, there are models for AI governance, such as a layered model for AI governance (Gasser & Almeida, 2017), which focuses on a strong societal and humanistic aspect and consists of a social and legal layer (covering norms, regulation, and legislation), an ethical layer (covering criteria and principles), and a technical layer (data governance, accountability, and standards). However, current methods disregard business and practical considerations. We hope to close this gap with this article. We create a conceptual framework for this goal by combining the literature on AI, including allied topics like ML.

This paper is set up as follows. In the beginning, we present the methodology and lay out the framework. The conclusion of the literature review is then stated together with specifics of the proposed framework. Then, we offer the full framework together with a discussion of significant distinctions from other frameworks resulting from the many characteristics of AI (compared to IT). We then offer a discussion and further research. We demonstrate specifically how our approach may be used to carry out theory development. We wrap off with a summary and suggestions for additional research.

METHODOLOGY

We conducted a thorough literature search before starting our study. While this led to a number of intriguing works, we felt that they fell short in key respects when it came to covering AI governance from a business perspective. Thus, to fill in the gaps not covered by the articles from the systematic literature review, we augmented the systematic search with a narrative search (King & He, 2005). The narrative search was influenced by two complementary approaches to AI governance: I looking at governance frameworks on IT (Tiwana et al., 2013) and data (Abraham et al., 2019) with an emphasis on business-relevant aspects and determining the extent to which each aspect in this framework was covered by literature; and (ii) looking at key characteristics of AI that are likely to have an impact on their governance and determining the degree to which these characteristics have been studied. An overview of the process is provided in the figure 1 as follows:

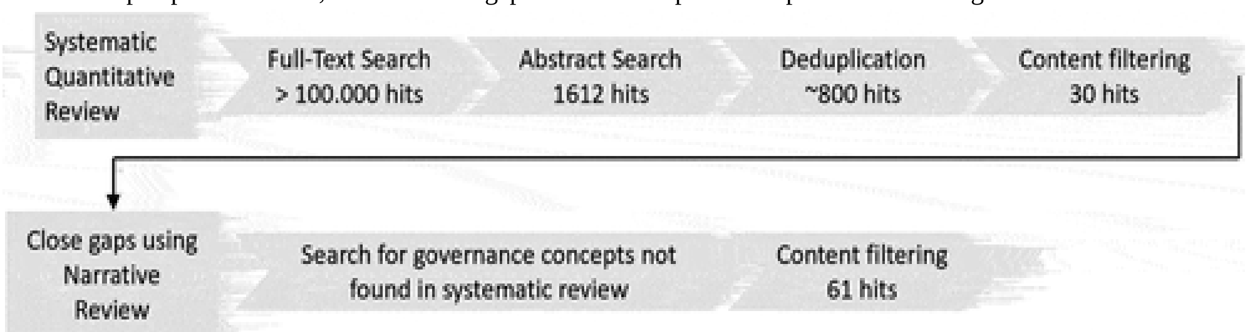

Figure 1: Overview of Literature Review Process

We began by conducting a search using the terms "Machine Learning OR AI OR Artificial Intelligence" AND "governance OR management" AND "compan* OR business* OR firm? OR organisation?" We used all of the databases at our university, including ProQuest, EBSCO, Google Scholar, Science Direct, Springer, and AISnet, and we only included peer-reviewed journals and English articles in our search. Google Scholar has supplanted other well-known databases like Web of Science and Scopus (Martn-Martn et al., 2018). The selected databases also include our main research interests, such as managerial works, computer science works, and works at the junction of these fields, such as information systems. We looked at articles released between May of 2011 and May of 2021. A full-text search on Springer and Google Scholar produced 30,000 and even more hits, respectively. As a result, we just searched abstracts further. This resulted in 1,627 hits overall (1,212 without governance and 415 without management). About half of all the hits, according to our estimation, were duplicates. The title and abstract of the articles were the main criteria for removal. Numerous articles discussed topics other than artificial intelligence (AI), did not focus on a single company (e.g., public policy pieces), did not address AI governance or management, or did not even mention blockchain or Industry 4.0. (e.g., using AI to support management). This left us with 13 of the 415 items for governance. Many management articles focused on particular applications and difficulties in a particular industry (such as manufacturing) (e.g., predictive maintenance). We now have 17 articles, for a total of 30 articles. While experiences of the handling of particular cases are fascinating, they are not really relevant for a comprehensive literature study.

We discovered that these works alone did not address a variety of governance-related topics that are present in the context of IT, data, and AI in a business setting, such as how to value data. Additionally, they lacked concrete ML processes and procedures. Additionally, the recognized works did not provide a deeper comprehension of fundamental AI governance concepts or the establishment of workable, effective governance policies. A significant amount of research has also been done on the practical ML concerns that affect governance, such as ML governance and management, ML model testing, guaranteeing fairness, and explaining "black boxes." Consequently, we also carried out a narrative literature review (King & He, 2005). The onus of conducting an adequate search technique falls on the reviewer of a narrative literature review. In order to find relevant articles, we synthesised works from the systematic literature review and mapped them to current governance frameworks (Abraham et al., 2019; Tiwana et al., 2013), keeping in mind important AI traits as autonomy, learning capacity, and obscurity. The framework's missing pieces were specifically looked for. Thus, we searched both forward and backward using a wide range of keywords, such as "AI design," "AI best practises," "AI safety," "AI strategy," "AI policy," "AI standards," "AI performance measurement," "data quality," "data valuation," "ML management," "ML process," "Model governance," and "ML testing" (Webster & Watson, 2002). We used a single search engine because the narrative literature review contains a lot of search terms. We made use of Google Scholar because Martn-Martn et al. (2018) shown that it was superior to well-known databases like Web of Science and Scopus. Since novel concepts are frequently introduced at academic conferences before they are published in journals and because a sizable body of works, notably in computer science, only exist as conference articles, we lifted the restriction to journals. If there was a wealth of literature, we focused on surveys (for a keyword or during forward search). We also incorporated grey literature, or primary material, in the event that there was a dearth of peer-reviewed articles. This includes preprints from arxiv.org and reviews from MIT and Harvard. Other fields, including computer science, commonly cite preprints on arxiv.org, which lacks a peer review mechanism but uses a moderation procedure for rudimentary quality assurance. Pre-prints were only taken into account when we evaluated their quality. Our key driving forces were to: (1) provide the reader with the most recent information; and (2) address the vast topic of AI and governance from several perspectives. This strategy, in our opinion, is required to guarantee practical relevance and appropriately explain potential areas for future research while taking into account work that will probably be finished shortly. Pre-prints, then, are often scientific papers that are undergoing peer review right now or very soon. Readers who ignore them run the risk of concentrating on research gaps that are already being filled.

We use current IT and data governance frameworks to examine the identified works (as described in section Framework Outline). Figure 2's basic, more general dimensions were employed. After being recognized throughout the assessment process, individual works were first categorized into more general categories, as illustrated in Figure 2. Everything was capable of this. Subcategories within the generic categories were drawn from pre-existing frameworks but were continuously enlarged throughout the review process. For the works that did not fit the framework, we created new categories. As further explained in the discussion of our final framework, our final framework also rejects elements of earlier frameworks (on IT and data governance) if they are not applicable in our environment.

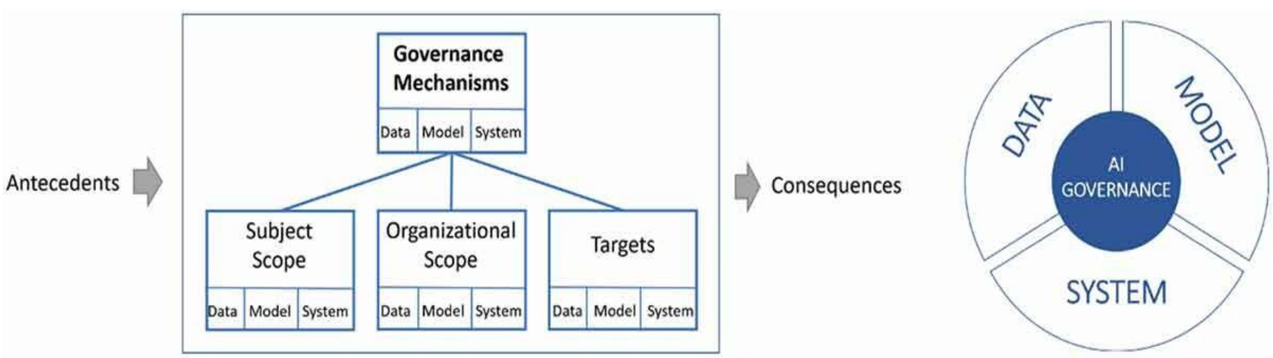

Figure 2. Our perspective on AI governance (left) and the core of our conceptual framework (right).

LITERATURE REVIEW

Based on the publications that were reviewed, this section goes into detail about AI governance. Before applying the structure of the conceptual framework depicted in Figure 2, we must clarify the scope in the form of an AI governance definition. The insights from the literature are used to improve each aspect of the conceptual framework. We begin with a discussion of the governance procedures before moving on to the aims, organizational scope, and subject scope.

AI Governance

We provide a working definition of the crucial term "AI governance for Businesses" in order to make clear the scope of this effort. It combines "corporate governance," a term with a clear definition, and "AI," a term with considerable ambiguity in the literature. We suggest the definition below:

The framework of policies, procedures, and procedures used to ensure that an organization's AI technology upholds and furthers its aims and objectives is known as AI governance for business.

We emphasize the following six aspects of AI governance, in line with Abraham et al. (2019):

1. promoting cross-functional cooperation
2. establishing a framework for organizing and formalizing AI management
3. emphasizing AI as a tactical tool
4. identifying the decision-makers.
5. creating ancillary documents (policy, norms, and procedures), and
6. keeping track of compliance.

Building "machines that can compute how to operate successfully and securely in a wide variety of unexpected scenarios" is what artificial intelligence (AI) intends to do (Russell & Norvig, 2020). The purpose of machine learning (ML), a subset of artificial intelligence, is to learn from data. ML thus refers to the models and procedures for creating and validating models using data-driven learning.

The goal of AI is to create an AI system with an ML model and typically other components, such as a graphical user interface and input handling for an AI (software) system or physical components like those found in a robot. The ML model, which is in charge of making decisions, can be thought of as the AI's brain.

The term "AI" is distinct from "ML." It is also more sophisticated in terms of the methods employed to produce intelligent behaviour. While AI includes methods that are not learning-based, such as logical reasoning, ML focuses on learning from data (Russell & Norvig, 2020).

The exploration of an environment by an agent utilizing reinforcement learning (RL; Russell & Norvig, 2020), such as in "Alpha Go Zero," may be the basis for data indicating experience in the form of samples (Silver et al., 2017). Since there is no data to provide domain knowledge in this situation, the agent must create it through exploration. We don't go into great detail on RL because it's less used in practice and supervised models provide the most benefits for enterprises (Witten & Frank, 2017).

From a governance standpoint, it makes sense and seems desirable to define AI in terms of data, models, and (AI) systems because they are related to already-existing governance domains. They also construct on top of one another. Data and IT governance for AI includes both systems (sometimes a sizable software code base) and models (typically a small code basis). The governance of relatively tiny software codes that contain models and methods for training, evaluating, and testing, as opposed to traditional software, is what is referred to as model governance.

Governance Mechanisms

Businesses can employ a variety of controls to manage AI. They include formal frameworks tying together business, IT, data, model, and machine learning (ML) and system management functions, formal processes and procedures for decision-making and monitoring, as well as procedures that encourage stakeholder engagement and collaboration (De Haes & Van Grembergen, 2009; Peterson, 2004). We use the distinction between (a) structural, (b) procedural, and (c) relational governance mechanisms in accordance with the literature on IT governance (De Haes & Van Grembergen, 2009; Peterson, 2004). Figure 3 gives an overview:

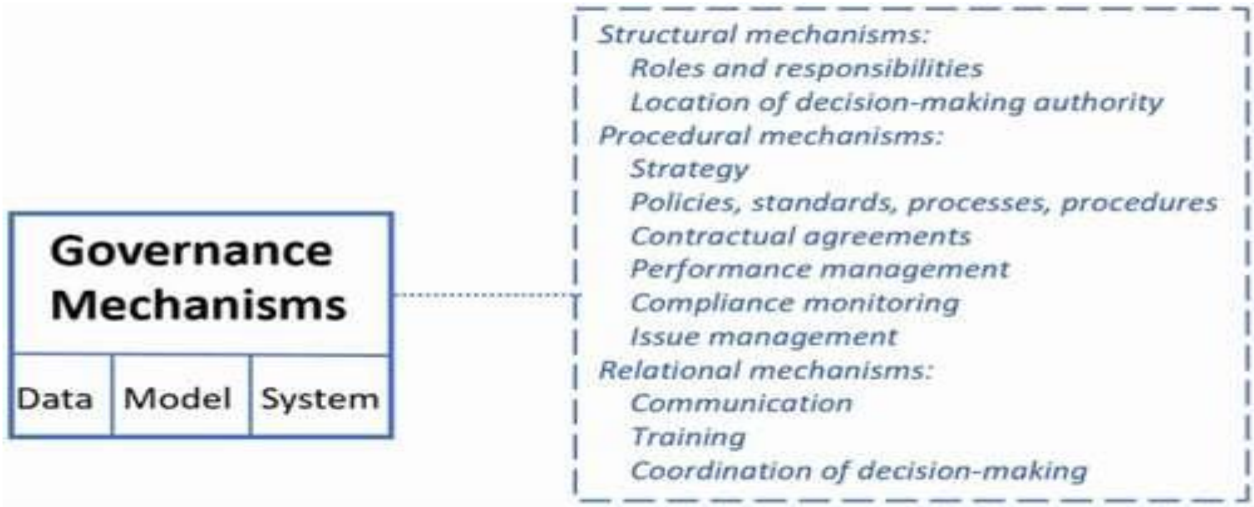

Figure 3. Overview of governance mechanisms

Structural Governance

Reporting frameworks, governing bodies, and accountability are defined by structural governance processes (Borgman et al., 2016). They consist of (2) the distribution of decision-making authority and (1) roles and duties. While there is a wealth of research on data governance (Abraham et al., 2019), the same cannot be said for AI governance. In the context of image analysis in healthcare, Ho et al. (2019) provided a brief overview of the duties of a chief data officer (CDO) and chief information officer (CIO). Standards, procedures, and rules for using a hub that centralizes duties such as personnel acquisition, performance management, and AI have also been established (Fountaine et al., 2019). Establishing a centre of excellence has been recommended throughout the adoption phase (Kruhse-Lehtonen & Hofmann, 2020). The creation of AI systems involves many different fields of study (Tarafdar et al., 2019). In order to manage the complicated interplay between model outputs, training data, regulatory and business needs, it may be necessary to construct the (interdisciplinary) AI governance council that has been pushed for AI in healthcare (Reddy et al., 2020). Executive sponsors are also vital to the process (Pumplun et al., 2019). Depending on the company's adoption level, the executive sponsor's level of power and decision-making over performance objectives may vary. For instance, Pumplun et al. (2019) argued that a dedicated AI budget without any performance requirements would help adoption, at least in the early stages. Investigations are still ongoing into more precise roles connected to model elements. According to Serban et al. (2020), designating an owner for each feature is excellent practice.

Procedural mechanisms

Procedure-based governance mechanisms are designed to make sure that AI systems and ML models are held securely, that they operate correctly and efficiently, and that their operation complies with all applicable laws, regulations, and company internal standards and guidelines regarding explainability, fairness, accountability, security, and safety. Explainability, justice, and accountability are the main concerns of ML models, even though these objectives eventually need to be met for the entire system and the ML model. Contrarily, safety is a more all-encompassing quality. For instance, while a machine learning (ML) model may be regarded unsafe, measures like backup components that are not based on AI and offer limited functionality may guarantee the safety of the overall system. Though its effectiveness is debatable, a "huge red emergency button" may at least allow the AI to be turned off at any time (Arnold & Scheutz, 2018). As seen in Figure 3, procedural mechanisms also seek to guarantee data, model, and system-relevant features and targets. They include (1) a strategy, (2) policies, (3) standards, (4) processes and procedures, (5) contractual agreements, (6) performance measurement, (7) compliance monitoring, and (8) issue management for data, models, and systems both separately and collectively.

Based on strategic corporate objectives, the strategy indicates high-level guiding of actions. The review by Keding (2020) looked at how AI and strategic management interact, with a particular emphasis on two aspects: (a) antecedents, such as data-driven workflows (data value chain and data quality), managerial willingness, and organisational determinants (AI strategy and implementation); and (b) consequences of AI in strategic management on a personal and organisational level (e.g., human-AI collaboration, AI in business models). The work focuses on using AI for business intelligence-related decision-making within an organisation. A number of AI strategy-related components are currently being worked on, including guiding concepts by academics and practitioners (Iansiti & Lakhani, 2020; Kruhse-Lehtonen & Hofmann, 2020; Pichai, 2018; Smit & Zoet, 2020).

High-level standards and norms are provided by AI policies. Key objectives, accountability, roles, and obligations are all communicated by organisations through AI policies. Politicians are actively debating AI policies. For instance, the European Union (EU; European Commission, 2020) has presented policy options on how to promote the use of AI while minimising hazards, and a portion of them are already included in a proposal (European Commission, 2021). Best practises could be a source for policies at the corporate level (Alsheiabni et al., 2020). As an example, Mishra et al. (2020) explore several aspects of how to assess AI's societal and economic impact as well as hazards and threats of AI systems. Measurement in (public) AI policy is also an active area of debate and research.

While there has been significant progress in the development of data standards (DAMA International, 2009), there has been less progress in the development of AI standards. Examples include Wellbeing metrics for ethical AI (IEEE P7010), benchmarking accuracy of facial recognition systems (IEEE P7013), fail-safe design for AI systems (IEEE P7009), and ethically driven AI Nudging methodologies (IEEE P7008). For a more thorough overview, see Cihon (2019). According to our approach, standards should guarantee that data representations, ML models, and the design of AI systems, along with the processes that go along with them, are uniform and standardised across the entire enterprise. They ought to promote interoperability both within and between companies and make sure they are suited for the intended purpose (Mosley et al., 2021).

Final Framework

Our research approached AI governance as an outgrowth of the existing IT and data governance expertise. We specifically based our framework on the three general dimensions relating to what, who, and how is managed by using the IT governance cube of Tiwana et al. (2013). Despite the fact that this appears to support the validity of the current framework, we propose an updated framework shown in Figure 4 and a more streamlined version in Figure 5.

AI governance cube based on IT governance cube, Figure 4. (Tiwana et al., 2013). It can act as a structure for developing theories using rotation and permutation.

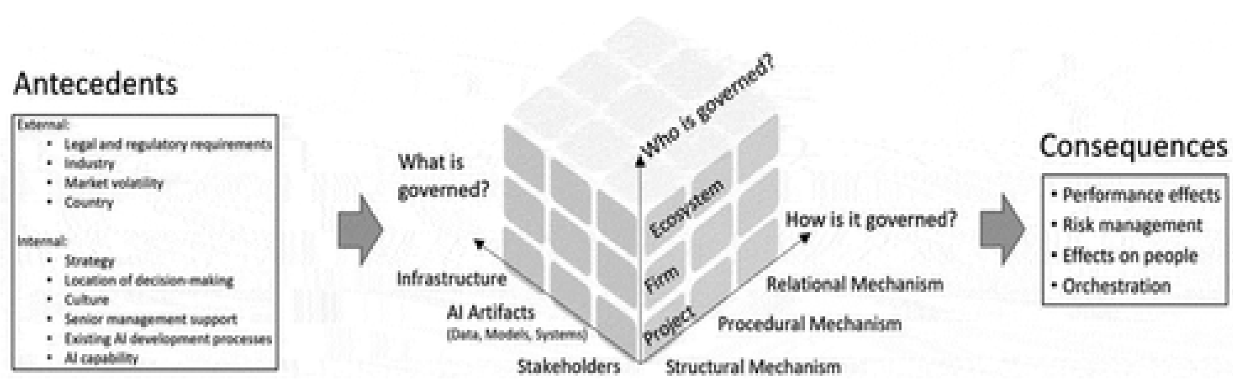

Figure 4: AI governance cube based on the IT governance cube (Tiwana et al., 2013). it can serve as a framework for theory building through rotation and permutation.

Figure 5. Concepts within the conceptual framework for data governance. concepts.

The distinction made by Tiwana between stakeholders, content, and IT artefacts appears to be imperfect. The idea of a model differs from the traditional idea of an IT artefact, which is a piece of hardware or software that frequently contains data or information. Data determine the behaviour of a model, which is more declarative. Due to the black-box nature of AI, it is difficult to determine the precise content of a trained model; it can only be assumed to be a function of the input. Software and hardware are typically static and purchased from outside sources for many companies. In other words, because software must be expressly coded, upgrades are uncommon and expensive. AI, however, has the potential to be self-learning, allowing for the generation of models that perform better over time and are mostly used internally. From our vantage point, Tiwana's definition of IT artefacts in terms of hardware and software could be seen as infrastructure that is used to process and store data, build models, and even deploy AI systems. Although they do not necessarily need to be managed as a single entity and may instead be governed using various targets, we see data, models, and systems as being very intimately connected.

Additionally, Tiwana et al. thought of the IT architecture as a potential tool for governance. Even while corporate architecture is important for AI adoption (Stecher et al., 2020), once adopted, designs are hard to change, making it challenging to deploy a governance tool in a changing context. The architecture of a model is significant in the context of models. Even Nevertheless, the training data and the target they optimise are both extremely important, making the term "architecture" inappropriate as a general phrase. So, we do not include architecture. In the context of AI, relational mechanisms such as decision-making coordination and training—which are less pronounced in Tiwana's cube—are extremely important in addition to structural mechanisms (relating to decision rights) and procedural mechanisms (relating to control, standards, and practises). These three different sorts of mechanisms were first described by De Haes and Van Grembergen and by Peterson (2004). (2009). Concepts included in the conceptual framework for data governance are shown in Figure 8. Concepts from (Abraham et al., 2019) are in italicized, blue, and dashed boxes.

CONCLUSION

This work shed light on the foundations of AI governance. Our research examines the body of work on AI and ML using a conceptual framework. Our assessment emphasizes areas that appear particularly intriguing from a business standpoint, even though the majority of issues in AI governance deserve more attention in general. In addition to a conceptual framework, we offered more specific viewpoints and recommendations prompted by technological advancements, such as those relating to data quality and value, as well as comparisons to related fields, such as responsibilities in software development and data governance. These suggestions can be developed in the future by, for instance, looking at governance roles and data value using our AI governance cube. Selecting a certain cell along the three dimensions to get a range of possibilities. How, by what means, and by whom is anything governed?

REFERENCE

[1] Abraham, R., Schneider, J., & Vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49, 424–438. https://doi.org/10.1016/j.ijinfomgt.2019.07.008 [Crossref], [Web of Science ®], [Google Scholar]

[2] Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable Artificial Intelligence (XAI). *IEEE Access*, 6, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052 [Crossref], [Web of Science ®], [Google Scholar]

[3] Aggarwal, C. C. (2015). *Data mining: The textbook*. Springer. [Google Scholar]

[4] Almeida, P., Santos, C., & Farias, J. S. (2020). Artificial intelligence regulation: A meta-framework for formulation and governance. In *Proceedings Of The 53rd Hawaii International Conference On System Sciences*, https://doi.org/10.24251/hicss.2020.647 [Crossref], [Google Scholar]

[5] Alsheiabni, S., Messom, C., Cheung, Y., & Alhosni, M. (2020). Winning AI strategy: six-steps to create value from artificial intelligence. In *Proceedings of the Americas Conference on Information Systems*. [Google Scholar]

[6] Amabile, T. M. (2020). Creativity, artificial intelligence, and a world of surprises. *Academy of Management Discoveries*, 6(3), 351–354. doi:https://doi.org/10.5465/amd.2019.0075 [Google Scholar]

[7] Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P. et al. (2019). Guidelines for human-ai interaction. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3290605.3300233 [Crossref], [Google Scholar]

[8] Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). Concrete problems in AI safety. *arXiv Preprint arXiv*, 1606.06565 [Google Scholar]

[9] Arena, S., Florian, E., Zennaro, I., Orrù, P. F., & Sgarbossa, F. (2022). A novel decision support system for managing predictive maintenance strategies based on machine learning approaches. *Safety Science*, 146, Article 105529. https://doi.org/10.1016/j.ssci.2021.105529 [Crossref], [Google Scholar]

[10] Arnold, T., & Scheutz, M. (2018). The “big red button” is too late: An alternative model for the ethical evaluation of ai systems. *Ethics and Information Technology*, 20(1), 59–69. https://doi.org/10.1007/s10676-018-9447-7 [Crossref], [Google Scholar]

[11] Ashmore, R., Calinescu, R., & Paterson, C. (2021). Assuring the machine learning lifecycle: Desiderata, methods, and challenges. *ACM Computing Surveys (CSUR)*. doi:https://doi.org/10.1145/3453444 [Google Scholar]

[12] Bellet, T., Cunneen, M., Mullins, M., Murphy, F., Pütz, F., Spickermann, F., Braendle, C., & Baumann, M. F. (2019). From semi to fully autonomous vehicles: New emerging risks and ethico-legal challenges for human-machine interactions. *Transportation Research. Part F, Traffic Psychology and Behaviour*, 63, 153–164. doi:https://doi.org/ 10.1016/j.trf.2019.04.004 [Crossref], [Google Scholar]

[13] Bhagat, S., & Bolton, B. (2008). Corporate governance and firm performance. *Journal of Corporate Finance*, 14(3), 257–273. https://doi.org/10.1016/j.jcorpfin.2008.03.006 [Crossref], [Web of Science ®], [Google Scholar]

[14] Borgman, H., Heier, H., Bahli, B., & Boekamp, T. (2016). Dotting the I and Crossing (out) the T in IT governance: new challenges for information governance. In *Hawaii International Conference on System Sciences (HICSS)*, (pp. 4901–4909). https://doi.org/10.1109/hicss.2016.608 [Crossref], [Google Scholar]

[15] Brodley, C. E., & Friedl, M. A. (1999). Identifying mislabeled training data. *Journal of Artificial Intelligence Research*, 11, 11. 131–

167. https://doi.org/10.1613/jair.606 [Crossref], [Google Scholar]

[16] Brundage, M., Avin, S., Wang, J., Belfield, H., Krueger, G., Hadfield, G. et al. (2020). Toward trustworthy AI development: Mechanisms for supporting verifiable claims. *arXiv Preprint arXiv*, 2004.07213 [Google Scholar]

[17] Burt, A. (2021). New AI regulations are coming. is your organization ready? *Harvard Business Review* [Google Scholar]

[18] Chen, A., Chow, A., Davidson, A., DCunha, A., Ghodsi, A., Hong, S. A. et al. (2020). Developments in MLflow: A system to accelerate the machine learning lifecycle. In *Proceedings of the International Workshop on Data Management for End-to-End Machine Learning*, (pp. 1–4). https://doi.org/10.1145/3399579.3399867 [Crossref], [Google Scholar]

[19] Chui, M., & Malhotra, S., (2018), *"AI adoption advances, but foundational barriers remain" Mckinsey and company*. Retrieved May 01, 2021, from https://www.mckinsey.com/featured-insights/artificial-intelligence/ai-adoption-advances-but-foundational-barriers-remain [Google Scholar]

[20] Cihon, P. (2019). Standards for AI governance: International standards to enable global coordination in ai research & development. In *Future of Humanity Institute*. University of Oxford. [Google Scholar]

[21] Clarke, R. (2019, August). Principles and business processes for responsible AI. *Computer Law & Security Review*, 35(4), 410–422 [Crossref], [Web of Science ®], [Google Scholar]

[22] Corbett-Davies, S., & Goel, S. (2018). The measure and mismeasure of fairness: A critical review of fair machine learning. *arXiv preprint arXiv*, 1808.00023 [Google Scholar]

[23] DAMA International. (2009). *Dama guide to the data management body of knowledge* (1st edition ed.). Technics Publications. [Google Scholar]

[24] Das, A., & Rad, P. (2020). Opportunities and challenges in explainable artificial intelligence (XAI): A survey. *arXiv preprint arXiv*, 2006. 11371 [Google Scholar]

[25] Dastin, J. (2018). Amazon scraps secret ai recruiting tool that showed bias against women. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G [Google Scholar]

[26] Davenport, Thomas, H., & Kirby, J. (2015). Beyond automation. *Harvard Business Review*, 59–65 [Google Scholar]

[27] De Haes, S., & Van Grembergen, W. (2009). An exploratory study into IT governance implementations and its impact on business/IT alignment. *Information Systems Management*, 26(2), 123–137. https://doi.org/10.1080/10580530902794786 [Taylor & Francis Online], [Web of Science ®], [Google Scholar]

[28] Du, M., Yang, F., Zou, N., & Hu, X. (2020). Fairness in deep learning: A computational perspective. *IEEE Intelligent Systems*. 36. 4. doi:https://doi.org/10.1109/MIS.2020.3000681 [Crossref], [Google Scholar]

[29] European Commission. (2020). *White paper on artificial intelligence. A European approach to excellence and trust*., Retrieved November 02, 2020, from https://doi.org/10.2139/ssrn.3706099 [Google Scholar]

[30] European Commission. (2021). *Laying down harmonized rules on artificial intelligence* Retrieved May 01, 2021, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 [Google Scholar]

[31] Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. doi:https://doi.org/10.1007/s11023-018-9482-5 [Crossref], [PubMed], [Web of Science ®], [Google Scholar]

[32] Ford, M. (2018). *Architects of intelligence: The truth about ai from the people building it*. Packt Publishing Ltd. [Google Scholar]

[33] Fountaine, T., McCarthy, B., & Saleh, T. (2019). *Building the ai-powered organization*. Retrieved November 03, 2020, from https://hbr.org/2019/07/building-the-ai-powered-organization [Google Scholar]

[34] Franklin, S., & Graesser, A. (1996). Is it an agent, or just a program?: A taxonomy for autonomous agents. In *International Workshop On Agent Theories, Architectures, And Languages*, (pp. 21 - 35) https://doi.org/10.1007/bfb0013570 [Google Scholar]