Review of The Governance, Risk and Compliance Approaches For Artificial Intelligence

David Lau Keat Jin[1], Ganthan Narayana Samy[2], Fiza Abdul Rahim[3], Mahiswaran Selvananthan[4], Nurazean Maarop[5] & Noor Hafizah Hassan[6]

[1,2,3,5,6]Advanced Informatics Department, Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Kuala Lumpur, 54100, Malaysia
[4]Perdana Department, Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Kuala Lumpur, 54100, Malaysia
davidkeat@graduate.utm.my

Article history

Received:
18 Sep 2023

Received in revised form:
25 Oct 2023

Accepted:
16 Nov 2023

Published online:
18 Dec 2023

*Corresponding author
davidkeat@graduate.utm.my

Abstract

Advancement in the domain of big data, computing power and internet of things continue to spur the development of algorithmic models that morphed into artificial intelligence. Notable achievements have been made in the application of artificial intelligence in image recognition, natural language processing, smart farming, personal learning assistance, and autonomous systems. As its adoption increases and proliferates into every sphere of activities, governments, businesses and organizations begin to formulate strategies and measures to facilitate its adoption even as it is still rapidly progressing. Meanwhile, artificial intelligence's impact to the individual, organizational and societal levels as well as the mechanisms to ensure realization of its benefits and minimization of its drawbacks are actively being pursued by the academic communities. This study endeavours to aggregate the perspectives from multiple review studies to shed light on the approaches pertaining to the governance, risk management and compliance of artificial intelligence. The concepts, elements and practices relevant to the three aspects are presented together with the proposed way forward to facilitate artificial intelligence adoption by the organizations.

Keywords: *Artificial Intelligence, governance, risk management, compliance*

1. Introduction

Artificial Intelligence (AI) is a buzzword used to encapsulate the technology that could process information that mimics human cognitive abilities. As a pioneer in the field, John McCarthy defined AI in 1956 as "the science and engineering of making intelligent machines" [1]. Adding to this definition, Brooks stated that "AI is intended to make computers do things, that when done by people, are described as having indicated intelligence" [2]. AI's definition is still evolving as it gains acceptance and adoption in various industries and spheres of activities. Recent definition has been given by the Institute of Electrical and Electronics Engineers (IEEE) Corporate Advisory Group as "the combination of cognitive automation, machine learning (ML), reasoning, hypothesis generation and analysis, Natural Language Processing (NLP), and intentional algorithm mutation producing insights and analytics at or above human capability" [3]. ML is a subset of AI that allows a

* Corresponding author. davidkeat@graduate.utm.my

system to automatically learn and improve based on feedback while NLP denotes algorithm that can process and analyze human-readable text.

The catalysts of AI development are the advancement made in the Internet of Things (IoT), Big Data and computing capabilities. The algorithms involve in the operation of AI is such that relevant dataset is required to train the algorithmic model to produce the desired output. As AI is data-dependent, inaccurate, biased or intentionally malicious data fed into the algortihmic model may produce inaccurate, biased and erroneous output resulting in adverse or catastrophic consequences, depending on its actual application. For example, the twitter chatbot launched by Microsoft was forced to shutdown after other twitter users trained it with racist information which in turn produced racially offensive and insensitive statements [4]. Other instances where AI produced biased output include classifying higher rate of recividism based on race [5] and gender discrimination when it comes to employment preference [6].

Furthermore, inaccurate results produced by AI used in autonomous system like self-driving vehicle may lead to injury or loss of lives, as in the case of an accident reported when such system is permitted to be used in public roads [7]. Such mishap may be caused by failure to capture all functional requirements during development [8], inadequate verification and validation [9], model misspecification or uncertainty [10], or deliberate attack launched against the system by a malicious actor [11]. When such desirable incidents occur, accountability issues arise as multiple stakeholders are involved in the approval, design, development, deployment, operation, maintenance and oversight of the system [12, 13]. This is exacerbated by the inscrutability of certain AI models which can be considered "black box" and the absence of any formal regulation or standard that could enforce compliance on responsible parties [14, 15].

More recently, AI is used to generate images that can be used for education, marketing, and presentation purposes [16]. This is a reality as AI chatbot is able to understand the meaning of text entered by humans communicating with chatbot [17]. While there is no outstanding issue regarding the usability of the images generated [18], there are however issues related to plagiarism if chatbot generates output for use in academic literature [19]. As the technology is still evolving, the academia is grappling with its permissible use in the context of learning and education [20]. As concern is raised regarding the use of text and image generated by AI, it is more alarming in the case of voice generated by AI as produced by the application called Deepfakes, especially if it is used to impersonate other person for malicious purposes [21].

2. Governance, Risk and Compliance

Due to the issues and concerned pertaining to the use of AI, numerous articles have been published regarding the ethical principles as well as the mechanism to bridge the ethical principle-practice gap [22]. [23] postulated that governance is the central practice that impacts other ethical requirements found from previous studies. In light of this, [24] define organizational AI governance to be a subset of IT governance that intersects with data governance. IT governance in turn is encapsulated within the domain of corporate governance. There are also other governance mechanisms that are beyond the scope of an organization such as ratification of standards and legislation. This relationship is illustrated in Figure 1.

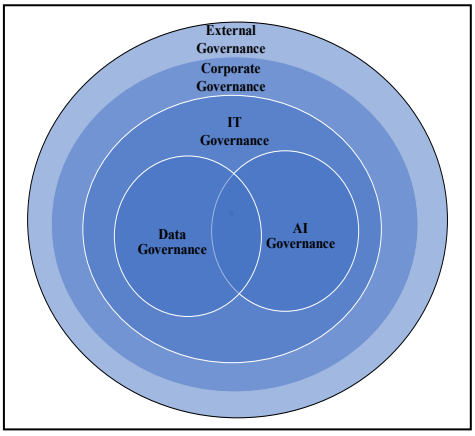

Figure 1. Multilevel Governance of AI [24]

While methods and tools have been proposed by extant literature in order to bridge the gap between ethical principles with actual practices [25, 26], implementation of ethical measures in each phase of AI lifecycle entails resource allocation and utilization such as infrastructure, manpower, expertise, time and cost [27]. In the absence of formal regulations and industry-wide standards, business owners and top management of organizations may not be supportive of those efforts, resulting in developers failing to alleviate the ethical concerns during development phase [28]. Hence, visibility of risks in the AI's context of use should be conceived through formal mechanism within the organization, taking into account the various stakeholders involved as envisaged by National Institute of Science and Technology (NIST) AI Risk Management Framework [29].

As legislations and industry-wide standards are being formulated, organizations would be prudent to ensure that their adoption of AI complies with the governancing laws and standards. Enforcement of these legislations and standards may require expertise in performance of test, evaluation, verification and validation (TEVV) as recommended by [29]. This would include examination by human-before-the-loop, human-in-the-loop as well as human-over-the-loop and executed at the four control points in the AI lifecycle as expounded by [30]. If requisite services pertaining to the use of AI model such as acquisition of data or/and model are necessary, then the countermeasures has to be put in place in the form of contractual terms or specifications required [31, 32]. In light of these propositions, a reproduction of integrated GRC framework is presented in Figure 2 [33].

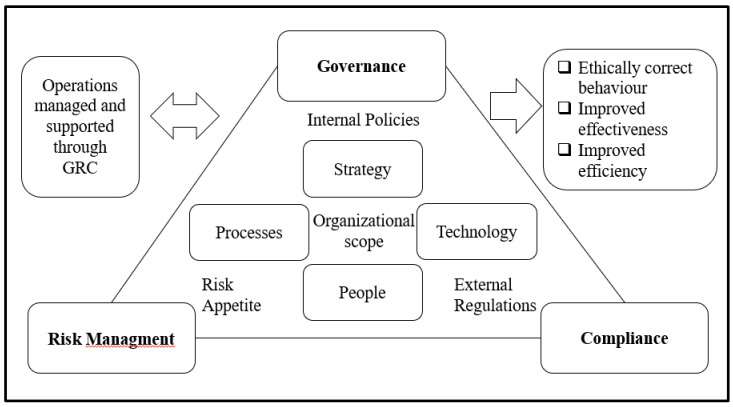

Figure 2. Frame of Reference for Integrated GRC [33]

3. Methodology

This study aims to extract and synthesize insights regarding the GRC elements and practices by previous review studies. Hence, the inclusion criteria as outlined in Table 1 is applied.

Table 1. Article Selection Criteria

No.	Attribute	Inclusion Criteria
1.	Database	Scopus
2.	Type of Article	Review Article
3.	Period of Publication	January 2019 until August 2023
4.	Topic	Governance, risk, compliance, ethical design, framework, tools, safety, trust or impact of AI
5.	Quality Consideration	Peer-reviewed article
6.	Language	English

4. Result

14 review articles are obtained from the search results. A summary of the articles' content is given in Table 2.

Table 2. Summary of Review Articles Obtained

Article	Objective	Key Findings	No. Of References Examined (Period In Year)
[34]	To reveal knowledge gaps in AI Governance (AIG) in the organization.	The gaps identified include: i. Limited understanding of AI governance implementation; ii. lack of attention to the AIG context; iii. uncertain effectiveness of ethical principles and regulation; and iv. insufficient operationalization of AIG processes.	68 (2010-2021)
[35]	Governance process is divided into structural, procedural and relational aspects. The recommended governance practices are: i. fostering collaboration across functions ii. structuring and formalizing AI	Governance mechanism will impact data scope, organizational scope and targets. Each of the impacted area is sub-divided into data, model and system. This framework captures and grouped the requirements in hierarchical order while	61 (2011-2021)

Article	Objective	Key Findings	No. Of References Examined (Period In Year)
	management through a framework iii. focusing on AI as a strategic asset iv. defining how and who makes decisions v. developing supporting artifacts (policy, standards, and procedures), and vi. monitoring compliance	considering the antecedents and consequences.	
[36]	A determine the requirements for responsible artificial intelligence	The responsible AI governance framework is composed of 10 ethical concepts. 75% of the reviewed references are discursive in nature without empirical data	12 (2019-2023)
[37]	To ensure AI systems are contestable by design: responsive to human intervention throughout the system lifecycle.	A framework is proposed which consists of five system features and six development practices that contribute to contestable AI	19 (2016-2021)
[38]	To describe the main ethical themes in the field of data science	Identification of 5 project phases where each phase consist of an ethical theme and multiple ethical considerations.	50 (2009-2019)
[23]	To identify ethical implications for the use of AI in Digital Technology (DT) archetypes.	A conceptual model is developed with 14 ethical aspects mapped to 7 DT archetypes. An ontological framework was also conceived showing the interrelationship among physical, cognitive, information and governance domain.	59 (2000-2022)
[26]	To bridge the gap between principles and practices by constructing a typology that may help practically-minded developers apply ethics at each stage of the Machine Learning development pipeline, and to signal to researchers where further work is needed.	Method of incorporating 5 ethical principles into 7 AI lifecycle phases are compiled from the literature review.	106 (2013-2020)

Article	Objective	Key Findings	No. Of References Examined (Period In Year)
[39]	To present the empirical research regarding human trust in AI	The form of AI representation (robot, virtual, embedded) and the level of AI's machine intelligence are important antecedents to the development of trust. A framework that addresses the elements that shape users' cognitive and emotional trust is proposed. Also, AI's tangibility, transparency, reliability and immediacy behaviors are important in developing trust. The absence of consensus on the suitable scales and behavioural measurement for trust across disciplines may hamper inter-disciplinary cooperation.	150 (1999-2019)
[40]	To analyze the implementation of AI in the organizations and the effects it has on organizations.	The Input, Process and Output (IPO) conceptual framework for understanding AI implementation in organizations is constructed with antecedents, challenges, guidelines and consequences identified and group into clusters. Interestingly, 8 major theories were highlighted as antecedents for application of AI in the organizations.	61 (2017 – 2020)
[41]	To elaborate the different types of adversarial attacks with various threat models and as well as the efficiency and challenges of recent countermeasures against them	Taxonomy of adversarial model for (a) evasion attacks and (b) poisoning attacks with respect to adversarial capabilities and goals. Defence strategies include: i. adversarial training; ii. gradient hiding; iii. defensive distillation; iv. feature squeezing; v. blocking the transferrability; vi. defense-generative adversarial network (gan) vii. magnet; viii. using high-level representation guided denoiser (hgd); and ix. using basis function transformation.	9 (threat model) 16 (attack and application) 9 (defense) (2006-2018)
[42]	To review the methods proposed for AI assurance	10 metrics were proposed in evaluation of the methods	250

Article	Objective	Key Findings	No. Of References Examined (Period In Year)
	which includes testing, verification, validation toward assurance of data-driven, trusworthy, explainable, ethical, unbiased and fair to its users	proposed in previous studies which include: i. specificity to ai' ii. the existence of a formal method; iii. declaration of successful results in testing; iv. availability of datasets; v. system size vi. declaration of successful results in real-world application; vii. existing limitation; viii. generalizable; ix. used in real-world application; and x. contrasted with other methods	(1985-2021)
[43]	To identify the positive and negative impacts to safeguard AI's benefits and avoid its downsides.	Elucidation of purpose, scope, organisational context, expected issues, timeframe, process and methods, transparency and challenges related to AI Impact Assessment (IA). A baseline process of implementing AI-IA is also proposed for AI developers and vendors.	38 (2016-2021)
[44]	To develop a value-based assessment framework that is not limited to bias auditing and that covers prominent ethical principles for algorithmic systems. Used 9 review articles as starting point.	The required values and their manifestation methods for the development team, auditing team, data domain experts and data subjects are identified. A circular-based assessment framework that visualizes closeness and tensions between values are constructed with operational guidelines.	192 (2006-2022)
[30]	To analyze the requirements of fairness, explainability, accountability, reliability, and acceptance to uncover the approaches that can mitigate AI risks and increase trust and acceptance of the systems. It also discusses existing strategies for validating and verifying these systems and the current	Elaboration of fairness, explainability, accountability, privacy and acceptance were given. 4 levels of human involvement in assuring the requirements are met are before-the-loop, in-the-loop, in-the-loop and over-the-loop corresponding to 4 control points in AI lifecycle.	227 (2010-2021)

Article	Objective	Key Findings	No. Of References Examined (Period In Year)
	standardization efforts for trustworthy AI.		

The major elements or practices required in the domain of GRC as extracted from the documents are described as follows:

(a) audit -the examination and inspection implemented to ensure compliance to certain policies, procedures or standards;
(b) certification – the formal recognition that the organization has been audited and complied with the stated policies, procedures or standards;
(c) countermeasures – controls or defense strategies that can be implemented to reduce identified risks;
(d) data governance – revolves on the process carried out on the data to ensure that it is fit for use as well as protecting the privacy of subjects;
(e) ethical training – the training that pertain to ethical use of data, AI models, deployment and assessments to ensure ethical practices in AI lifecycle;
(f) impact assessment – the exercise conducted to determine the consequences of using AI which include risk and benefits to the relevant stakeholders;
(g) Input, Process and Output (IPO) framework – a framework that considers the antecedents, challenges, guidelines and consequences in AI adoption;
(h) metrics – measurements involve in gauging the level of ethical compliance or performance of AI;
(i) Policies, Procedures and Standards (PPS) – the governing mechanism of an organization which ensure all internal employees as well as external parties working with its employees adhere to approved rules and regulations;
(j) process models – the incorporation of ethical practices and controls in the stages of AI lifecycle from planning and designing stage until deployment of AI model.
(k) role of stakeholders – the responsibilities of all those who are involve in the development as well as affected by the deployment and use of AI.
(l) software and tools – the codes, libraries, application that assist in ethical practices of AI development and deployment.
(m) Testing, Evaluation, Verification and Validation/Oversight – these revolves around the practices of assessment and inspection of AI operation in various stages of the lifecycle to ensure that the results fulfill minimum requirements.

Table 3 lists the articles obtained and the main GRC elements as elucidated from (a) to (m) where 'X' denote the main GRC elements elaborated.

Table 3. Mapping of Review Articles to GRC Elements

Article	Main GRC Elements												
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)
[34]	X					X							X
[35]										X			
[36]										X			
[37]											X		
[38]					X								X

Article	Main GRC Elements												
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)
[23]									X				
[26]										X			
[39]								X			X		
[40]									X	X			
[41]			X			X							
[42]								X					X
[43]						X							
[44]	X	X		X							X		
[30]						X							X

Examination of the review articles showed that there are some semantic similarities in their work but with subtle differences in the terms that the authors used. For example, contestable AI [37], governance measures [34-36], value-based assessment [44] and implementation approaches [25, 26, 38, 40] entail recommended processes to deliver assurances of AI [42] , alleviate ethical concerns [23], increase human trust and autonomy [30, 39] as well as reduce risks [41, 43, 45]. It is thus unavoidable that there are some overlapping themes as well as same references used amongst the review articles. Without deduction for duplicate references, the combined total references cited by the reviewed articles are 1,385 from the years 1985 until 2023.

5. Discussion

There are certain concepts that form semantic proximities with one another where only the scope or coverage of the terms is different. For example, in assurance of AI [42], the author defined that it must fulfill its ethical requirements as well as function according to the expectation of stakeholders. Hence, standards that define terminologies used in AI are fitting as reference for organizations. In the same vein, risk assessment can be considered a subset of impact assessment [43] which would include risk identification and evaluation. Organizations can leverage on existing risk management practices for this purpose. Also, measures for assurance [42] hint at activities such as contestation of decisions [37], human oversight [30] or audit exercises which may include testing, evaluation, verification and validation [29]. In short, legislations and industry-wide standards are required to for high risks AI applications to ensure compliance by the organizations involve in development, deployment or maintenance of those applications to safeguard the interests of users impacted by the use of AI applications.

The results directly and indirectly validated the methodology used to collate information and insights regarding the approaches to GRC recommended for AI adoption based on the following assertions:

i. review of review articles affords greater coverage of the knowledge pool in the area of research considering certain review articles like Stahl et. al. [43] extended their list of references to grey literature like web sites of relevant organizations, books and reports.

ii. most articles from academic databases provide critical analysis and synthesize new knowledge or concepts from previous studies [36];
iii. review of other review articles has previously been implemented in tourism industry [46];
iv. all the review articles contributed to GRC approaches as required by this study with two of the review articles which dedicated extensive effort to technical discourse of AI attack and hence provide insights to risk management only [41];
v. AI is a progressive field with conceptual fuzziness and this prompted Shneider et al. [35] and Stahl et al. [43] to begin their studies based on other relevant reviews; and
vi. The insights afforded by previous review articles afford more conclusive evidence regarding the research direction of the topic under consideration [47].

6. Conclusion

Backward snowballing method may be employed to comprehensively list out the elements and practices involved in each of the GRC domain as expounded by the 16 review articles examined [48]. Thereafter, a GRC framework can be constructed. Established standards, framework, models and guidelines available for corporate governance [49], Information Technology (IT) governance [50], data governance [51], procurement [31], risk management [29] and other relevant publications with expert consensus can form the basis for the proposed GRC framework. The acceptance and suitability of the framework can then be validated empirically in different setting and industry where AI is adopted.

Acknowledgments

The author would like to thank the Malaysian Public Services Department that sponsored the research.

References

[1] McCarthy, J., et al. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine, 27*(4), 12-12.
[2] Brooks, R.A., *Intelligence without reason*, in *The artificial life route to artificial intelligence*. 2018, Routledge. p. 25-81.
[3] Group, I.C.A., *IEEE guide for terms and concepts in intelligent process automation*. 2017.
[4] Hunt, E. *Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter*. 2016 Aug 23, 2023]; Available from: https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter.
[5] Jeff Larson, S.M., Lauren Kirchner, Julia Angwin. *How We Analyzed the COMPAS Recidivism Algorithm*. 2016 Aug 25, 2023]; Available from: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.
[6] Dastin, J. *Amazon scraps secret AI recruiting tool that showed bias against women*. 2018 Aug 23, 2023]; Available from: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G.
[7] NTSB. *Tesla Crash Investigation Yields 9 NTSB Safety Recommendations*. 2020 Aug 23, 2023].
[8] Nitta, I., et al. *AI Ethics Impact Assessment based on Requirement Engineering*. in *30th IEEE International Requirements Engineering Conference (RE)*. 2022. Electr Network: Ieee.
[9] Mökander, J. and M. Axente. (2023). Ethics-based auditing of automated decision-making systems: intervention points and policy implications. *AI and Society, 38*(1), 153-171. doi:10.1007/s00146-021-01286-x
[10] Zhang, X.G., et al. (2022). Towards risk-aware artificial intelligence and machine learning systems: An overview. *Decision Support Systems, 159*, 13. doi:10.1016/j.dss.2022.113800
[11] Blauth, T.F., O.J. Gstrein, and A. Zwitter. (2022). Artificial Intelligence Crime: An Overview of Malicious Use and Abuse of AI. *Ieee Access, 10*, 77110-77122. doi:10.1109/access.2022.3191790
[12] Burton, S., et al. (2020). Mind the gaps: Assuring the safety of autonomous systems from an engineering, ethical, and legal perspective. *Artificial Intelligence, 279*, 103201.
[13] Santoni de Sio, F. and G. Mecacci. (2021). Four responsibility gaps with artificial intelligence: Why they matter and how to address them. *Philosophy & Technology, 34*, 1057-1084.

[14] Kouroupis, K. (2022). The AI Act in light of the EU Digital Agenda: A critical approach. *Journal of Data Protection and Privacy, 5*(3), 216-229.
[15] Erdélyi, O.J. and J. Goldsmith. *Regulating artificial intelligence: Proposal for a global solution*. in *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. 2018.
[16] Elasri, M., et al. (2022). Image generation: A review. *Neural Processing Letters, 54*(5), 4609-4646.
[17] Oppenlaender, J. *The creativity of text-to-image generation*. in *Proceedings of the 25th International Academic Mindtrek Conference*. 2022.
[18] AKTAY, S. (2022). The usability of images generated by artificial intelligence (AI) in education. *International technology and education journal, 6*(2), 51-62.
[19] Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. *Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning (February 11, 2023)*.
[20] Sok, S. and K. Heng. (2023). ChatGPT for education and research: A review of benefits and risks. *Available at SSRN 4378735*.
[21] Verdoliva, L. (2020). Media forensics and deepfakes: an overview. *IEEE Journal of Selected Topics in Signal Processing, 14*(5), 910-932.
[22] A. Jobin, M.I.a.E.V. (2019). The global landscape of AI ethics guidelines.
[23] Ashok, M., et al. (2022). Ethical framework for Artificial Intelligence and Digital technologies. *International Journal of Information Management, 62*, 102433.
[24] Mäntymäki, M., et al. (2022). Defining organizational AI governance. *AI and Ethics, 2*(4), 603-609. doi:10.1007/s43681-022-00143-x
[25] Prem, E. (2023). From ethical AI frameworks to tools: a review of approaches. *AI and Ethics*. doi:10.1007/s43681-023-00258-9
[26] Morley, J., et al. (2020). From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Science and engineering ethics, 26*(4), 2141-2168.
[27] Mannes, A. (2020). Governance, Risk, and Artificial Intelligence. *Ai Magazine, 41*(1), 61-69.
[28] Zhang, B., et al. (2021). Ethics and Governance of Artificial Intelligence: Evidence from a Survey of Machine Learning Researchers. *Journal of Artificial Intelligence Research, 71*, 591–666. doi:10.1613/jair.1.12895
[29] Tabassi, E. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0).
[30] Kaur, D., et al. (2022). Trustworthy artificial intelligence: a review. *ACM Computing Surveys (CSUR), 55*(2), 1-38.
[31] WEF. *Guidelines for AI Procurement*. 2019 31 May 2023]; Available from: https://www3.weforum.org/docs/WEF_Guidelines_for_AI_Procurement.pdf.
[32] Dor, L.M.B. and C. Coglianese. (2021). Procurement as AI governance. *IEEE Transactions on Technology and Society, 2*(4), 192-199.
[33] Racz, N., E. Weippl, and A. Seufert. *A frame of reference for research of integrated governance, risk and compliance (GRC)*. in *Communications and Multimedia Security: 11th IFIP TC 6/TC 11 International Conference, CMS 2010, Linz, Austria, May 31–June 2, 2010. Proceedings 11*. 2010. Springer.
[34] Birkstedt, T., et al. (2023). AI governance: themes, knowledge gaps and future agendas. *Internet Research, 33*(7), 133-167. doi:10.1108/intr-01-2022-0042
[35] Schneider, J., et al. (2023). Artificial Intelligence Governance For Businesses. *Information Systems Management, 40*(3), 229-249. doi:10.1080/10580530.2022.2085825
[36] Camilleri, M.A. (2023). Artificial intelligence governance: Ethical considerations and implications for social responsibility. *Expert Systems*. doi:10.1111/exsy.13406
[37] Alfrink, K., et al. (2022). Contestable AI by Design: Towards a Framework. *Minds and Machines*. doi:10.1007/s11023-022-09611-z
[38] Saltz, J.S. and N. Dewar. (2019). Data science ethical considerations: a systematic literature review and proposed project framework. *Ethics and Information Technology, 21*, 197-208. doi:10.1007/s10676-019-09502-5
[39] Glikson, E. and A.W. Woolley. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals, 14*(2), 627-660.
[40] Lee, M.C., et al. (2023). The Implementation of Artificial Intelligence in Organizations: A Systematic Literature Review. *Information & Management*, 103816.
[41] Chakraborty, A., et al. (2021). A survey on adversarial attacks and defences. *CAAI Transactions on Intelligence Technology, 6*(1), 25-45.
[42] Batarseh, F.A., L. Freeman, and C.-H. Huang. (2021). A survey on artificial intelligence assurance. *Journal of Big Data, 8*(1), 60.
[43] Stahl, B.C., et al. (2023). A systematic review of artificial intelligence impact assessments. *Artificial Intelligence Review*, 1-33.
[44] Yurrita, M., et al. *Towards a multi-stakeholder value-based assessment framework for algorithmic systems*. in *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 2022.
[45] Gao, Y., et al. (2020). Backdoor attacks and countermeasures on deep learning: A comprehensive review. *arXiv preprint arXiv:2007.10760*.
[46] Pahlevan-Sharif, S., P. Mura, and S.N. Wijesinghe. (2019). A systematic review of systematic reviews in tourism. *Journal of Hospitality and Tourism Management, 39*, 158-165.
[47] Munn, Z., et al. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC medical research methodology, 18*, 1-7.
[48] Wohlin, C. *Guidelines for snowballing in systematic literature studies and a replication in software engineering*. in *Proceedings of the 18th international conference on evaluation and assessment in software engineering*. 2014.
[49] Calder, A., *ISO/IEC 38500: the IT governance standard*. 2008: IT Governance Ltd.
[50] Khanyile, S. and H. Abdullah. (2013). COBIT 5: an evolutionary framework and only framework to address the governance and management of enterprise IT. *no. September*, 7.
[51] International, D., *DAMA-DMBOK: data management body of knowledge*. 2017: Technics Publications, LLC.