Enhancing Artificial Intelligence Control Mechanisms: Current Practices, Real Life Applications and Future Views

Vatousios Antonis, Happonen Ari

DOI: http://dx.doi.org/10.1007/978-3-031-18461-1_19

Enhancing Artificial Intelligence Control Mechanisms: Current Practices, Real Life Applications and Future Views

Usman Ahmad Usmani[1], Ari Happonen[2(✉)], and Junzo Watada[3]

[1] Department of Computer and Information Science, Universiti Teknologi Petronas, 79 LakeVille Seri Iskandar, 32610 Seri Iskandar, Perak, Malaysia
[2] LUT School of Engineering Science, LUT University, Yliopistonkatu 34, 53850 Lappeenranta, Finland
ari.happonen@lut.fi
[3] Waseda University, 1 Chome-104 Totsukamachi, Shinjuku City, Tokyo 169-8050, Japan

Abstract. The popularity of Artificial Intelligence has grown lately with the potential it promises for revolutionizing a wide range of different sectors. To achieve the change, whole community must overcome the Machine Learning (ML) related explainability barrier, an inherent obstacle of current sub symbolism-based approaches, e.g. in Deep Neural Networks, which was not existing during the last AI hype time including some expert and rule-based systems. Due to lack of transparency, privacy, biased systems, lack of governance and accountability, our society demands toolsets to create responsible AI solutions for enabling of unbiased AI systems. These solutions will help business owners to create AI applications which are trust enhancing, open and transparent and also explainable. Properly made systems will enhance trust among employees, business leaders, customers and other stakeholders. The process of overseeing artificial intelligence usage and its influence on related stakeholders belongs to the context of AI Governance. Our work gives a detailed overview of a governance model for Responsible AI, emphasizing fairness, model explainability, and responsibility in large-scale AI technology deployment in real-world organizations. Our goal is to provide the model developers in an organization to understand the Responsible AI with a comprehensive governance framework that outlines the details of the different roles and the key responsibilities. The results work as reference for future research is aimed to encourage area experts from other disciplines towards embracement of AI in their own business sectors, without interpretability shortcoming biases.

Keywords: AI governance · Responsible AI · Real life applications · eXplainable AI · Three lines model

1 Introduction

From business to healthcare, sustainability, product design, industrial and educational context alike, innovations on AI innovation and Industry 4.0 are delivering new opportunities to improve people's lives, all across the globe [1, 3, 10, 19, 23, 26, 32]. This tough does raise problems for fairness inclusions, system security and how to bring privacy, into these systems effectively [2]. User-centered AI systems should consider both general and solutions to machine learning (ML) specific issues [4–8, 37]. Understanding the genuine impact of a system's estimates, suggestions operational models, process principles and decisions depends on how users interact with systems and operations [8, 9, 52, 53]. Also, design attributes such as adequate disclosures, clarity, and control are required for good user experience and are usual parameters explaining estimations for revenue streams and value of AI solutions [11]. Considering augmentation and assistance, a single solution is suitable if designed to service a broad range of users and use cases. In certain circumstances, giving the user a limited set of alternatives is beneficial to the system. Precision across many answers is significantly more challenging achieving than accuracy over a single solution [9]. Think about click-through rate data and customer lifetime value, as well as subgroup-specific false positive and false negative rates, to evaluate overall system performance and short and long-term product health (e.g., click-through rate and customer lifetime value) [12]. It should be ensured that the metrics are suitable for the context and purpose of the system. E.g. a fire alarm system should have a high recall, even if there are some false alarms [13, 14]. We can double-check the comprehension of raw data since ML models reflect the data they are trained on. If this isn't possible, such as with sensitive raw data, it is advised to make the most of your information while maintaining privacy by distributing aggregated summaries. Most model is used that satisfies the performance objectives [15]. Also the users need to be made of aware of system limits. E.g. application utilizing ML, made to identify selected bird species reveals that the model was trained on a tiny sample of images from a particular area of the globe [16]. The quantity of feedback increases by better educating people about the product or application. The most acceptable test processes are learned and quality engineering from software engineering to ensure that the AI system operates as expected and can be trusted. To include a wide variety of consumer needs into development cycles, iterative user testing should be done [17]. Then the quality engineering approach is used to have quality checks into a system, ensuring that unplanned errors are prevented or handled swiftly (for example, if a critical feature is suddenly absent, the AI system will not generate a prediction). The model will be constantly monitored to consider real-world performance and user input (e.g., happiness tracking surveys, HEART framework) [13, 18].

In base definition, all world models are flawed, there is no such thing as 100% perfect system. It is recommended to make time in the product strategy for troubleshooting. Both short- and long-term solutions should be taken into account. While a fast remedy can temporarily solve the issue, it is usually not a long-term answer for the problem. Long-term learning solutions should tough be linked with quick learning solutions. The candidate and deployed model's variation must be considered and how the change will affect overall system quality and user experience before altering a deployed model [20, 21]. The critical point is that AI success is contingent on a group, not a single person or position [22]. According to Collins those disciplines increase scientific process awareness and thinking in complicated, interacting systems. They generally have the critical thinking abilities needed to conduct good experiments and analyze the results of ML applications, says the author. Having a varied staff has several advantages [14].

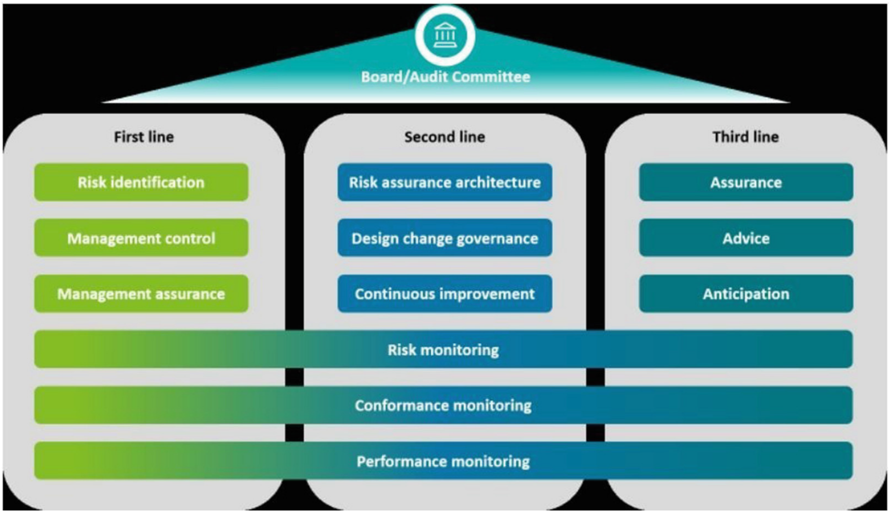

Fig. 1. An overview of the three lines model. [13]

Sitting back and hoping for diversity to come to an individual isn't a realistic team-building approach [7, 24, 25]. The following are the contributions of this manuscript:

- We explain the techniques and moral principles of AI ethics, which is related to responsible use of AI technology and intended to be informative for development.
- AI governance is explained from a perspective of enabling organizations to trust AI-powered outputs to automate current or new business processes for gains in time-to-market advantage at any stage of the real-life application development process.
- A detailed overview of the Three Lines Model (shown in Fig. 1) for AI governance is given to help organizations explain and reinforcing the fundamental principles, broadening the scope, and illustrating how key organizational activities interact to improve governance and risk management, for building a safe AI application.

2 Literature Review

AI's ethical deployment and governance are essential for permitting the large-scale deployment of AI systems required to improve people's welfare and safety. Yet, AI development typically outpaces regulatory development in many areas. The technique also enables AI developers to address typical challenges in automated systems, such as reducing social bias reinforcement, keeping people's jobs and talents, resolving responsibility to ensure confidence in an algorithm's results, and more [27]. Commercial AI systems in radiation clinics have just lately been developed, in contrast to the aerospace sector, with efforts concentrated on showing performance in academic or clinical settings, as well as product approval [28, 29]. Until previously, commercial AI systems for radiation were only available as static goods, allowing cancer specialists to analyze their effectiveness. Although an agile lifecycle management strategy, where AI-based segmentation models are updated with new patient data regularly, sounds appealing, it is unlikely to be accessible anytime soon.

Continuous quality monitoring of linear accelerators and vital software systems for treatment planning and radiotherapy department operations benefit from the same quality assurance and monitoring. In general, it is a good practice to study different fields for their monitoring practices [38, 47, 49], in addition of the specific field where AI is applied to. Before an AI can be deployed, it must first go through a thorough and transparent examination of the ethical implications of its proposed activities, particularly in terms of social impact, but also in terms of safety and bias, before moving on to a five-layer high-frequency checking system to ensure that the AI's decisions are correct and trustworthy. These characteristics are expectation confinement, synthetic data exercise, independence, comprehensiveness, and data corruption assurance. Dose calculation in therapeutic decision support systems, atlas-based auto-segmentation, and magnetic resonance imaging benefit from similar methodologies. [30].

In recent years, (deep) neural networks and machine learning (ML) approaches have complimented and, in some cases, surpassed symbolic AI solutions. As a consequence, its social importance and impact have skyrocketed, bringing the ethical discussion to a far larger audience. The argument has focused on AI ethical principles (nonmaleficence, autonomy, fairness, beneficence, and explainability) rather than acts the "how". Even if the AI community is becoming more aware of potential issues, it is still in the early stages of being able to take steps to mitigate the risks. The purpose of this study is to bridge the gap between principles and practices by developing a typology that can help developers apply ethics at each level of the Machine Learning development pipeline while also alerting researchers to areas where further research is needed. Although the research is limited to Machine Learning, the findings are predicted to be readily transferrable to other AI domains. The following is the difference between ethics via design and pro-ethical design: Because it does not rule out a course of action and requires agents to choose it, the nudge is less paternalistic than pro-ethical design. The nudge is less paternalistic than pro-ethical design since it does not prohibit but rather compels agents to choose a path of action. A simple illustration may help you comprehend the distinction. Because it enables drivers to pay a fine, a speed camera is both pro-ethical and nudging in the event of an emergency. Speed bumps, on the other hand, are a kind of traffic-calming device used to slow down automobiles and increase safety. They may seem to be a good concept, but they need a long-term route adjustment, leaving motorists with few options. This implies that even while responding to emergency, emergency vehicles such as a medical ambulance, police car, or fire engine must slow down.

3 AI Governance Development

While AI moral governance offers promise, it also has limits and risks becoming ineffectual if not used correctly. AI restrictions must be understood and followed. E.g. who has the last say on what constitutes "ethical" AI? Companies are coming up with their ideas and methodologies for establishing what it means to use these technologies ethically and what "ethical" AI implies for society. Those at the top of organizations, mostly white males, set the tone [4, 31]. The AI ethics board should be diverse, reflecting the views of people who AI systems may impact. Bias reduction when the company's and leadership's aims are at odds: Being the first to market is highly prized by businesses. On the other side, eradicating prejudice and building responsible AI conflict with this aim necessitates extra procedures and stop points, lengthening the time it takes to bring a product to market. Traditional goals clash with ethical and responsible AI goals, placing the company in jeopardy and losing money [33].

In addition to the governance in companies the governments and different countries unions should also step in and set their views on table for ethics, governance and methodologies of regulating AI, just as they have done in manufacturing and waste management, to restrict events like Boeing 737 MAX groundings, effectually a result of failed industry self-regulation practices. Examining current goals and making sure that responsible AI is a significant focus is critical. Finally, the interplay between ethical and economic gains demonstrates a fundamental knowledge of market success.

Cutting shortcuts on ethics has long-term and legal ramifications, particularly essential given AI's quickly shifting regulatory framework. In inadequate accountability and training, when it comes to putting ideas into practice, there is sometimes a lack of precise instruction. Furthermore, there is no responsibility when a company's ideals are broken. In many circumstances, corporate culture and the market in general [34] prioritizes efficiency above fairness and prejudice reduction, making it impossible to put the ideas into. When combining bias and fairness criteria, it is critical to what is "fair" for a particular AI system and who defines "fair".

The EGAL brief on ML fairness dives further into the subject, laying out methods and roadblocks. The technical solutions and "ethical washing" should be a priority. Most principles and associated ideas are based on the assumption that technology can solve problems, and they tend to have technical bias. E.g. a variety of qualitative techniques are included in the EGAL brief on ML fairness, which can be helpful [35]. Particularly for higher-risk applications, initial high-level assessments of the technology's potential for damage, as well as a record of decisions made throughout the AI system's construction, are critical. The terms management and governance are not interchangeable. Governance is responsible for supervising how decisions are made, while management is in charge of making them. By extending the same concept to AI governance, we arrive at the following AI Governance for company's definition. The more dangerous an AI application is, for a description of some of these threats, see AI hazards, the more critical AI governance becomes [36, 37]. Because AI enabled robots collect data and information on a continuous basis, biases or unwanted outcomes, such as a chatbot that learns inappropriate or violent language over time, are quite probable. If the data it receives is biased, the system will provide skewed outcomes. Individual topic experts must first inspect the data that enters these computers for biases, and then maintain governance to guarantee that biases do not arise over time.

Companies may use additional visibility, a better understanding of their data, and AI governance to assess the machine's business rules or learnt patterns before adopting and spreading them to staff and consumers. When it comes to ethical AI applications, it's all about trust. Customers, companies, and government authorities all want to know that these smart systems are assessing data and making the best judgments they can. They seek to demonstrate that the business outcomes produced by these machines are in everyone's best interests. Some of the tactics recommended in this article may assist businesses in becoming more trustworthy. They can also enhance how AI offers options to customers, aid with regulatory compliance, improve command and control, and provide total transparency and the ability to make the best decisions possible. In this Section, we explain the Three Lines Model (shown in Fig. 2) where we explain the role of the different execution lines and how the governance can be effective using the end to end governance model for Responsible AI.

3.1 The Three Lines Model

A few of the key'red flag' AI applications include facial recognition, AI for recruitment, and bias in AI-based assessments and recommendations. There should be a way to keep track of everything that's going on. It is desirable to control the significant development and approval method if AI is included in any current law or expert bodies for autonomous self-governance. If there is no precedent to follow, there can be additional risks that have not been considered. Corporate governance and risk management are ideas that have been around for a long time. [44] Standard procedures, norms, and conventions are typical to guarantee that businesses run smoothly. The Federation of European Risk Management Associations (FERMA) and the European Confederation of Institutes of Internal Auditing (ECIIA) developed the three lines of defense concept in 2008–10 as guidance on Article 41 of the 8th EU Company Law Directive.

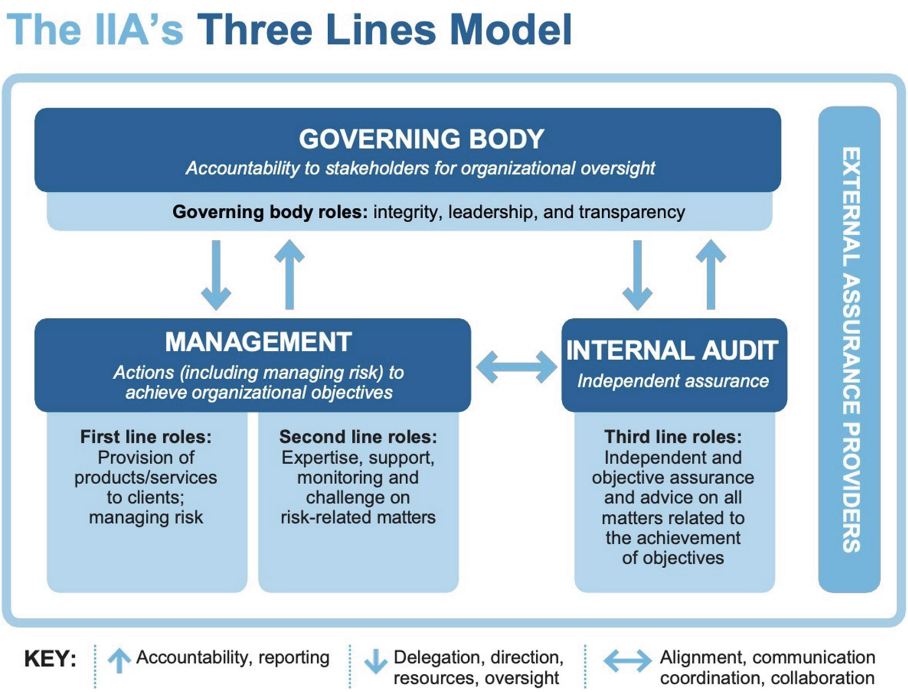

Fig. 2. Three lines model for AI governance [25]

In a position paper titled The Three Lines of Defense in Effective Risk Management and Control, released in 2013, the Institute of Internal Auditors (IIA) endorses this strategy. Risk analysis and management, as well as governance adoption, have all become industry standards [39]. In June 2020, the IIA revised its recommendations to include a position paper on the IIA's Three Line Model. This model covers the six basic governance concepts, essential roles in the three-line model, role connections, and use the three-line model. It includes information on management's tasks and obligations, the internal audit function, and the governing body. The three-line model has been revised to consider a wider variety of financial services organizations, including technology and model risk. Banks often use these three defense lines to manage credit risk, market risk, and operational risk models. We modified this paradigm for AI Governance by creating a new governance organization, methodology, roles, and responsibilities. Managers, Executors, and Inventors: The creators, executors, and operations teams develop, build, deploy, and execute the data, AI/ML models, and software, respectively; the data, software, and models are maintained and monitored by the operations team. Managers, supervisors, and quality assurance staff are responsible for identifying and executing the plan's risks connected to data, AI/ML models, automation, and software. Continuous monitoring is examined as the second line of defense. The second line is also responsible for ensuring that the first line's systems are configured correctly.

The Auditors are responsible for ensuring that the organization's rules, regulations, and goals are followed and that technology is utilized responsibly and ethically [40]. The ethics board consists of a diverse group of corporate leaders and workers. Specific organizations can appoint external members to the Board of Directors. Companies will have to work with external auditors, other assurance providers, and regulators, in addition to their internal duties. In the diagram below, the features of the numerous roles and the critical tasks of each function are shown.

It is critical to monitor the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. The next stage is to devise a strategy [44]. Software and AI models, as previously said, need entirely distinct techniques. While some of the research may not be beneficial, a portfolio strategy increases the likelihood that at least some of them will be correct. At any one-time, established AI adoption and firms have a portfolio of models in different stages of development, including conception, testing, deployment, production, and retirement. The ROI must be monitored and adjusted as required across the portfolio to ensure the best mix of business use cases, efficiency vs. effectiveness initiatives, and so on. Ten human talents and four intelligence are necessary to profit from human-centered AI. The distribution strategy must be carefully evaluated because of the convergence of data, software, and AI models.

To deliver AI-embedded software, or Software 2.0, both waterfall and agile software development methodologies must be updated and interlaced. The essential indicators that must be recorded and monitored for supervision will be identified in the specific delivery plan. The next level of governance is the ecosystem as a whole. The ecosystem in which AI models will be incorporated, as well as the context in which they will be used by personnel both within and outside the company, is what we're talking about. The societal impact of the company's AI should be evaluated. In this area, IEEE's Well-being Measures are a strong contender [41].

3.2 Role of Different Execution Lines

The governing board is in charge of setting the company's vision, mission, values, and organizational appetite for risk. Following that, management is given the duty of fulfilling the organization's objectives and obtaining the necessary resources. The governing body receives management reports on planned, actual, and projected performance and risk and risk management information. The degree of overlap and separation between the governing body's and management's tasks varies depending on the organization. The governing body can be more or less "hands on" when it comes to strategic and operational matters. The strategic plan can be created solely by the governing body or management, or it can be a shared effort. Also, the CEO can be a member of the board of directors and even the chairman. Effective communication between management and the governing body is required in every circumstance. Although the CEO is often the principal point of contact for this communication, other senior executives also interact regularly with the governing body.

Second-line executives, such as a Chief Risk Officer (CRO) and a Chief Compliance Officer (CCO), are sought and required by organizations and authorities. This is entirely compatible with the Three Lines Model's concepts. Management and internal auditing encompassing occupations on the first and second lines. Internal audit's independence from control allow it to plan and carry out its activities without fear of being influenced or interfered with. It has unlimited access to the people resources and information it requires. It makes recommendations to the governing body. Independence, on the other hand, does not imply loneliness. Internal audit and management must for internal audit's work to be relevant and consistent with the organization's strategic and operational goals. Internal audit broadens its expertise and understanding of the firm via all its activities, increasing the assurance and direction as a trusted and strategic partner [26]. Coordination and communication between the first and second lines of management and internal audit are essential to minimize excessive duplication, overlap, or gaps. Because it reports to the governing body, internal audit is frequently referred to as the organization's "eyes and ears".

The governing body is in charge of internal audit oversight, which includes hiring and firing the Chief Audit Executive (CAE), approving and resourcing the audit plan, receiving and considering CAE reports, and providing the CAE with unrestricted access to the governing body, including private sessions without management present. Second-line employment can be created by delegating essential responsibilities and reporting lines to the governing body to give first-line employees and senior management some autonomy. The Three Lines Model allows for as many reporting lines between management and the governing body as are required. For compliance or risk management, for example, as many persons reporting directly to the board are as necessary and organized to give a degree of independence. Second-line roles offer third-line employees with the same advice, monitoring, analysis, reporting, and assurance, but with less discretion. Lower-level employees who make risk management choices—devising and implementing policies, setting boundaries, establishing targets, and so on—are obviously "in the kitchen making sausages" and part of management's actions and responsibilities. Most notably regulated financial institutions, some organizations must have these in place to maintain true independence.

Risk management is still the duty of first-line management in these instances. Risk management monitoring, counseling, directing, testing, assessing, and reporting are examples of second-line responsibilities. Second-line jobs are a component of management's responsibilities. They are never really independent of management regardless of reporting lines since they assist and challenge those in first-line positions. They are critical to management decisions and actions. Third-line occupations are distinguished by their independence from management. Internal audit's independence, which distinguishes it from other activities and gives different assurance and recommendations, is in the Three Lines Model Principles. Internal audit preserves its independence by refusing to make decisions or conduct actions that are part of management's responsibilities, such as risk management, and refusing to assure activities that internal audit is responsible for now or previously. The CAE is expected to take on more decision-making responsibilities for jobs that require similar skills, such as statutory compliance or enterprise risk management (ERM), specially to look for upkeeping the best practices for company performance [42].

3.3 End to End Governance Model for Learning Responsible AI Practices

Other sources of assurance cannot be available. Effective governance requires accurate task assignment and strong activity alignment via cooperation, collaboration, and communication. Internal audit should provide the governing body with confidence that governance structures and processes are well-organized and operating as intended. Organizations are humans that work in a constantly turbulent, multifaceted, interconnected, and chaotic environment. They generally include many stakeholders, all of whom have varying, competing, and often conflicting interests. Stakeholders entrust supervision to a governing body that has empowered management with the resources and authority to make crucial choices like risk management. For these and other reasons, businesses need efficient structures and processes to their objectives while retaining good governance and risk management. The governing body and management rely on internal audits to provide independent, objective assurance and advice on all issues and inspire and promote innovation and growth since the governing body receives management reports on actions, results, and predictions. The governing body is ultimately accountable for governance, which is implemented via the governing body's actions and, management, and internal audit. The Three Lines Model assists businesses in establishing structures and procedures that support good governance and risk management while also helping them achieve their objectives [43, 44, 45].

Any organization uses the model, and it is bolstered by: Using a principles-based approach and customizing the model to the aims and circumstances of the organization. Risk management is considered to achieve the objectives, create value and be "defensive" by safeguarding the assets. The model's representation should be of roles and duties, as well as their interrelationships. Place processes should guarantee that actions and goals are in line with stakeholders' main concerns. The first assumption is that the government exists. Governance is made up of structures and procedures that enable accountability to stakeholders for monitoring by a governing body via integrity, leadership, and transparency. Management operations include risk-based decision-making and resource allocation, including risk management, to accomplish goals [27]. Through thorough investigation and intelligent communication, an independent internal audit position provides assurance and guidance to offer clarity and confidence and encourage and support continual growth [46].

The governing body ensures the establishment of the relevant institutions and processes for efficient governance. The organization's aims and operations sync with its stake- holders' significant objectives. The governing body assigns tasks and resources to management to achieve its goals while adhering to legal, regulatory, and ethical obligations; establishes and maintains an unbiased, objective, and competent internal audit department to provide clarity and confidence in goal progress. Management, as well as first and second-line employees, are covered by Principle 3. Governance is accountable for both first and second line functions to organizational goals. Support responsibilities are included in first-line jobs, which are most directly engaged in providing goods and/or services to the organization's customers. Posts help with risk management. It is possible to combine or split the first and second lines.

Specialists are appointed to specialized second-line occupations to give extra knowledge, supervision, and challenge to first-line activities. Second-line risk management jobs may concentrate on internal control, information, technology security, sustainability, and quality assurance, among other risk management goals. While second-line jobs, such as enterprise risk management (ERM), may have more responsibilities, risk management is still a part of first-line operations and is managed by management. Because of external factors at work, there is always a risk in the retail business. Customer credit is an example of an external factor that has a significant influence on a business' profitability. If a business does a customer credit risk analysis and finds that things aren't going as planned, it is able to lower its risk. This is done by stopping the invoice extensions for clients who are deemed high risk by the organization to decrease risk. Take the manufacturing business, for example. A company wants to develop a new product. They must do a comprehensive risk analysis before beginning production to determine the degree of risk that the firm may face. They may then decide if the advantages of producing a new product exceed the dangers. Internal audit offers unbiased assurance and advice on the appropriateness and effectiveness of governance and risk management. This is accomplished by the skillful use of rigid and disciplined methodologies and knowledge and insight. To encourage and support future development, it informs its findings to management and the governing body. It may evaluate assurance from various internal and external sources throughout this process.

The independence of the third line Internal audit's impartiality, authority, and credibility are all dependent on its independence from management obligations. It includes accountability to the governing body, unrestricted access to the people, resources, and data it needs to perform its job and independence from prejudice or involvement in the design and delivery of audit services. Creating and preserving value is the sixth principle. They work together to develop and maintain value when all functions are coordinated and prioritize stakeholder interests. Communication, cooperation, and collaboration are used to align activities. This maintains the information's consistency, coherence, and transparency, essential for risk-based decision-making. Responsibilities are distributed differently in different organizations. The high-level acts that follow, on the other hand, help to emphasize the Three Lines Model's Principles. Stakeholders understand that the governing body is in charge of organizational monitoring.

Second-line posts help with risk management. It is possible to combine or split the first and second lines. Specialists are appointed to specialized second-line occupations to give extra knowledge, supervision, and challenge to first-line activities. Second-line risk management jobs may concentrate on internal control, information, technology security, sustainability, and quality assurance, among other risk management goals. Internal audit offers unbiased assurance and advice on the appropriateness and effectiveness of governance and risk management. This is accomplished by the skillful use of rigid and disciplined methodologies and knowledge and insight [50]. To encourage and support future development, it informs its findings to management and the governing body. Formulation, execution, and continuous improvement of risk management procedures, including internal control at the process, system, and entity levels are examples of second-line jobs [48].

Risk management goals are achieved through adherence to laws, norms, and accepted ethical behavior; internal control; information and technology security; sustainability; and quality assurance—analyzes and reports on risk management's effectiveness and appropriateness, including internal control. Internal audit is separate from management and reports to the governing body. For example, The President reports to the Audit and Risk Management Committee of the Board of Governors, while they report to the Executive Director, University Governance, and University Secretary, respectively. Assurance and advice on the adequacy and usefulness of governance and risk management with the internal control are communicated to management and the governing body independently and objectively to support the achievement of organizational objectives and to promote and facilitate continuous improvement. Any threats to impartiality or independence are brought to the governing body's attention, which takes appropriate action. Further assurance is provided to meet legal and regulatory duties that safeguard stakeholders' interests, and internal assurance sources are augmented to meet management and governing body needs.

Data architects are also vital in the governance of AI systems. In order to model AI, businesses must have a solid data or metadata pipeline. Keep in mind that the success of AI is contingent on well-organized data architecture devoid of mistakes and noise. There will be a need for data standards, data governance, and business analytics. The development of the AI governance function necessitates the use of human resources. They may, for example, seek for employees who "fit" into the company's present AI framework and provide existing staff training tools to help them learn how to build ethical AI applications. When AI technology is deployed, it is critical to guarantee that no legal boundaries are breached. AI solutions meet organizational and industry-specific regulatory standards. There is no such thing as a one-size-fits-all plan that takes into account all legal and regulatory issues. Customers' perceptions of ethical behavior in the financial services industry, for example, may vary significantly from corporate ethics. The AI governance function's integration of legal and regulatory teams gives a diverse set of decision-making inputs. Marketing, sales, human resources, supply chain, and finance efforts all realize the advantages of AI. As a consequence, subject knowledge is required not just for app creation but also for app evaluation. As a consequence, having a strong business presence on the core AI governance council may aid in improving outcomes. People from various backgrounds should be represented on a company's governing board. It also contributes to inclusive and smooth governance by considering all of the company's issues. Product-based businesses provide a diverse variety of AI-enabled products. When a business purchases a product that isn't primarily based on AI, it often falls beyond the purview of the AI regulatory agency. But what if the business introduces an AI-assisted process, service, or product? Procurement and finance departments should ideally have AI professionals on staff to help with product onboarding. A well-functioning AI governance position will provide a framework for monitoring AI algorithms and products more effectively. In addition, developing an agile and cross-functional AI governance committee would bring a diverse set of perspectives to the table and help in the spread of AI knowledge.

4 Future Views

Despite the development of ethical frameworks, AI systems are nevertheless being quickly implemented across a wide range of vital areas in the public and commercial sectors—including healthcare, education, criminal justice, and many more—with no protections or accountability procedures in place. There are a number of challenges that must be addressed, and no one endeavor, nation, or corporation will be able to do it by themselves. Emerging technologies are becoming more cross-border, and if norms and practices impacting technical development and implementation in various countries do not coincide, significant possibilities can be missed (WTO, 2019) [46]. New conflicts can erupt both inside and between states in a divided globe. In terms of economic prosperity, it is feasible that the development of certain technical systems can grow more costly, postponing innovation. This can lead to injustice and new divides between technologically advanced and technologically disadvantaged nations or regions. Additionally, major differences in how new technologies (particularly AI) are handled and utilized in terms of human rights can make guaranteeing people's equal access to rights and opportunities across borders more difficult. New technologies can be used as new digital surveillance tools, allowing governments to automate citizen monitoring and tracking; they can also help policymakers allocate public goods and resources more efficiently; and they can even be powerful mechanisms for private companies to forecast our behavior [50].

A personal data can be retained and used for AI in an open or hidden manner. It can be voluntarily offered as a kind of remuneration, or it can be taken without the agreement or knowledge of the owner. Overall, arguments about who has access to our data, who has the right to make decisions about it, and who has the instruments to enforce that authority haunt the path to the digital future. This isn't to say that all technological governance should be done at the global level. It is critical for regions, states, and cities to be able to adapt to their residents' social, economic, and cultural needs. While the majority of research has focused on wealthy nations, there is a need for additional information about the geographically specific effect of AI systems on developing countries, as well as how new technology can perpetuate historical inequalities in these areas.

Global processes, on the other hand, are essential even if they do not result in integrated systems, since inequity thrives in the absence of universal laws. To manage the digital transition and achieve social inclusion, it will be required to create internationally identical ethical, humanitarian, legal, and political normative frameworks. Furthermore, while taking into account geopolitical and cultural disparities, there will be a growing need to focus on algorithmic criteria rather than ethical principles. The G20's role in aligning interests and organizing such projects will be critical in the coming years.

The G20 brings together some of the most powerful political and economic forces on the planet. It encompasses the whole globe and includes some of the world's most strong economies. It is the perfect place to examine the future of digital governance and respond to one of the most major contemporary difficulties and concerns facing our world today since it is a crucial venue for dialogue and involvement, both executive and legislative [51]. Right now, there is no one-size-fits-all solution for the best AI technique, but there are lots of options. We must all work together to determine which choice will benefit the most people. By participating in and leading this discussion, the G20 has the potential to become the spinal column of a new architecture for the 21st century, ensuring a brighter future for everyone.

Although recognition and classification aren't the only tasks given to AI systems, they are the most popular. The flexibility of AI approaches might be seen as a sign of variety. When research focuses on small, local gains in well-known, well-suited tasks like identification and classification and then applies the effort to comparable difficulties over a wide variety of domains, however, the predominance of a few activities may pose a risk. It's also feasible that the consequences of a system failure will be completely unexpected. The systems may result in a wide variety of symptoms, from mild discomfort to death. This backs with prior research that shows AI piques people's interest in a broad variety of topics, regardless of whether they are beneficial. On the other hand, mechanisms that protect people's privacy or remove the potential of discrimination are few and few between. Furthermore, not every system requires extensive testing: a system that causes pain is much more forgiving than one that causes death. As a result, the severity of a system failure should be considered while designing a system. Despite their significant dispersion throughout a broad range of categories, the systems are primarily defined by their limited application within those categories. In each facet, just a few, if any, systems represented several categories. For example, several domains were considerably underrepresented and underrepresented in an inconsistent manner.

In a variety of industries, e.g. agriculture and medical applications, robots have lately overtaken humans. The most critical system tasks were recognition and classification. This might be due to researchers' access to resources like robots, the overwhelming popularity of particular applications like self-driving vehicles, or technology's increasing capacity to apply itself to these more difficult fields. However, less well-known challenges should not be overlooked in future research, since less well-known does not always imply lower value, and AI might be useful in situations like crop maturity assessment, disease diagnosis, and MRI scan help. This is particularly significant since practical research in these potentially extremely beneficial areas may aid in the dissemination of findings and should be included in studies. In this discipline, software engineering research hasn't always prospered. Similarly, the majority of the tasks entrusted to the systems are of moderate complexity.

The most significant category, for example, ‘recognition & classification,’ is a difficult task, yet it is often required due to the difficulties it solves: Is there any mold on the product? Is the person in both pictures the same person? Is it now time to reap the benefits of your labor? Assembly, as compared to recognition, is a combination of the two: identifying and actively assembling key components. As a result, it seems that only a small portion of genuine AI system development is concentrated on very difficult challenges. Many model-centered validations, as well as data-driven ML testing in general, are plagued by data problems [61]. If sufficient high-quality data is available, model-centered methods, on the other hand, allow more efficient and maybe faster validation of model-centered systems, or at least the models used in the systems. If a system has several components, we suggest carefully assessing whether the model-centered approach is the optimum validation technique. In general, the study seems to place a higher priority on first validation than ongoing validation. This isn’t surprising; it’s always been done this way: a system is tested before being released when it seems to be functional. Given the demand discrepancies between AI and traditional systems, this may not be the case [62].

The first validation of a self-configuring system often includes self-configurability and first configuration validation. If the system reconfigures itself during deployment, it should almost likely be assessed to ensure that it continues to satisfy the system’s needs. The same can be said for a video streaming platform’s recommendation algorithm, which may have a constantly shifting user base: without continuing validation, the system may fail to meet its requirements, which developers and users may be unaware of. Our validation and continuing validation categories are based on original research. As a consequence, the completeness of the categories will not be examined in this study. Different taxonomies can be beneficial for recognizing and grasping AI validation; as a consequence, researching, expanding, and improving these taxonomies for validity and utility will be a future job. Monitoring the system’s outputs, thresholds, and other factors on a continuous basis may help AI systems improve their accuracy and efficiency. The first step in the oversight process could be to create a list of all AI systems in use at the company, along with their specific uses, techniques used, names of developers/teams and business owners, and risk ratings – such as calculating the potential social and financial risks that could arise if such a system fails. Examining the AI system’s inputs and outputs, as well as the AI system itself, may need a different methodology. Although data quality standards aren’t exclusive to AI/ML, they do have an influence on AI systems that learn from data and provide output depending on what they’ve learned. Training data may be used to assess a data collection’s quality and biases.

If practical and appropriate, benchmarking of other models and existing approaches to improve model interpretability can be incorporated in AI system assessment. Understanding the elements influencing AI system outputs helps to boost AI system confidence. Drift in AI systems might cause a plethora of problems. A shifting link between goal variables and independent variables, can lead to poor model accuracy. As such, drift detection is useful tool in AI problems, e.g. in the security, privacy fairness of a model, as avoidance measures. By evaluating whether input data varies considerably from the model’s training data, monitoring may assist discover “data drift”. Accounting for the model’s data collected in production and analyzing the model’s correctness is one way for acquiring insight into the model's "accuracy drift". In lending institutions, compliance, fair lending, and system governance teams are prevalent, and they seek for signs of bias in input variables and procedures. As a consequence of technology advancements and the deployment of de-biasing AI, a portion, if not the majority, of this labor can be automated and simplified in future. Fair AI, on the other hand, may need a human-centered approach. The generalist knowledge and experience of a well-trained and varied group probing for discriminatory bias in AI systems is unlikely to be completely replaced by an automated procedure. As a result, human judgment might be utilized as a first line of defense against biased artificial intelligence. By the recent research, discrimination-reducing approaches have found to be able to minimize disparities in class-control context and still keeps good predictive quality. In order to reduce inequities, mitigation algorithms design the "optimal" system for a certain degree of quality and discriminating measures.

The algorithms look for alternatives when there isn't another system with a greater degree of quality for a certain level of discrimination. However, no solution has been devised that completely removes bias for any given level of quality. Before such algorithms utilization in production environment, one needs more testing and validation studies. E.g. traditional algorithm searches and feature specifications for valid and less discriminating systems, as well as more modern approaches adjusting input data or the algorithms' optimization functions themselves, are the two sorts of methodologies. To reduce disparate impact, feature selection may be used, which involves removing one or two disparate-effect components from the system and replacing them with a few additional variables. In complicated AI/ML systems, these tactics have been demonstrated to be ineffective. For bias reduction, one needs new strategies in pre-processing and inside the decision-making phase of the algorithm, continuing to the output post-processing phase. The legal context, in which technology is used, as well as how it is used, has an impact on whether specific tactics are allowed in a certain circumstance.

Accuracy drift detection might be useful in the business sector since it can detect a decrease in model accuracy before it has a major effect on the company. Precision drift may lead to a loss of precision in your model. Data drift, on the other hand, aids companies in determining how data quality varies over time. It may be challenging for many businesses to guarantee that AI/ML explanations are both accurate and useful (explainability). AI/ML explanations, like the underlying AI/ML systems, may be poor approximations, wrong, or inconsistent. In the financial services industry, consistency is crucial, especially when it comes to unfavorable action letters for credit lending decisions. To lessen explainability issues, explanatory procedures may be tested for accuracy and stability in human assessment studies or on simulated data, depending on individual implementations. According to a new study, providing explanations and forecasts about how AI systems function may aid criminal actors.

Businesses should only provide information with customers when they directly request it or when it is mandated by law, to prevent security concerns. Traditional security techniques such as real-time anomaly detection, user authentication, and API throttling may be employed to secure AI/ML systems trained on sensitive data and producing predictions available to end users, depending on the implementation and management environment. In AI applications, traditional robust technologies, as well as cyber safeguards, may be effective risk mitigators. As adversarial learning improves, it might be utilized to help construct safe machine learning systems.

Despite the fact that this is a relatively young subject of study, the technology sector is considering a variety of possible mitigation techniques. Differential privacy has been proposed as a means of keeping personal information, including training information, secret. Differential privacy anonymizes data by infusing it with random noise, allowing statistical analysis without revealing personally identifiable information. As a consequence, the system produces the same results even if a single user/data element record is destroyed. Strong technological and cyber controls may be an effective mitigation depending on implementations and context, whereas mitigation methods for the AI/ML threats are still being investigated. Even though effective information security processes may prevent model extraction attacks, watermarking can be used to identify an extracted model. The AI/ML system is taught to give unique outputs for certain inputs in watermarking. If another system delivers the exact unique result for the same precise inputs, it might be a sign of intellectual property theft.

5 Conclusions

Utilization of AI promises new brighter future, especially in context of new generation industries and business context, where data is available in wide scale. These new contexts could be e.g. big fleets with lot of data [56, 57] business models operating in platform levels [58] and wide variety of huge groups contributed data has to be processed like in Digital citizen science activities [59, 60]. Even many traditional industries and processes like accelerating product design [3], sustainability and circularity boosting [19] and e.g. predicting assets maintenance based on collected big data [54]. But for this utilization to work, different actors, asset owners and platform utilizers have to be willing to share the gains and pains in shared development efforts [55] and also in the costs of keeping the system and its intelligent parts developing. In the ethical AI sense, actors who have fair cooperation do give extra incentives (reward and/or punishment) for ethical AI development in addition to the reasons that exist now or may exist by default in the future. Such actors value intrinsically a good compliance with norms and they do prefer reward those who comply with appropriate criteria. On the other hand, one might want to create an incentive for oneself to act in a certain way, as a commitment mechanism; or one can use incentives as a supplement to other governance tools such as behavior monitoring and direct regulation, or one wants to influence the incentives of a large number of people. There are a variety of incentives to consider, like rapid and high return on investment, meeting the needs of social beneficial developments and so on. Creating incentives for essential stakeholders for engagement can assist with all the issues such as public funding, international cooperation, etc. at once.

This research focused on the responsible use of AI, which has lately been highlighted as an essential requirement for ML technology adoption in real-world applications. Our research was about an end-to-end governance model for AI responsible use, emphasizing fairness, and responsibility in large-scale AI technology deployment in real-world organizations. Key deliverables or artefacts were looked upon, which the three lines provide, in order for AI to be used ethically and effectively. We infer that the model generated can aid businesses in the task of identifying the structures, processes, and responsibilities that best support goal attainment while simultaneously ensuring robust governance and risk management. The concept of our proposed model can help manage and regulate risks effectively.

References

1. Collins, C., Dennehy, D., Conboy, K., Mikalef, P.: Artificial intelligence in information systems research: a systematic literature review and research agenda. Int. J. Inf. Manag. **60**, 102383 (2021)
2. Smuha, N.A.: Beyond a human rights-based approach to AI governance: promise, pitfalls, plea. Philos. Technol. **34**(1), 91–104 (2021). Yang, Q
3. Ghoreishi, M., Happonen, A.: New promises AI brings into circular economy accelerated product design: a review on supporting literature. In: E3S Web Conference, vol. 158, pp. 1–10 (2020). https://doi.org/10.1051/e3sconf/202015806002
4. Tigard, D.W.: Responsible AI and moral responsibility: a common appreciation. AI Ethics **1**(2), 113–117 (2020). https://doi.org/10.1007/s43681-020-00009-0
5. Shneiderman, B.: Responsible AI: bridging from ethics to practice. Commun. ACM **64**(8), 32–35 (2021)
6. Berlin, S.J., John, M.: Particle swarm optimization with deep learning for human action recognition. Multimedia Tools Appl **79**(25–26), 17349–17371 (2020). https://doi.org/10.1007/s11042-020-08704-0
7. Rakova, B., Yang, J., Cramer, H., Chowdhury, R.: Where responsible AI meets reality: practitioner perspectives on enablers for shifting organizational practices. Proc. ACM Hum. Comput. Interact. **5**(CSCW1), 1–23 (2021)
8. Wearn, O.R., Freeman, R., Jacoby, D.M.: Responsible AI for conservation. Nat. Mach. Intell. **1**(2), 72–73 (2019)
9. Arrieta, A.B., et al.: Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf. Fusion **58**, 82–115 (2020)
10. Ghoreishi, M., Happonen, A., Pynnönen, M.: Exploring industry 4.0 technologies to enhance circularity in textile industry: role of Internet of Things. In: Twenty-first International Working Seminar on Production Economics, Austria, 24–28 February 2020, pp. 1–16 (2020). https://doi.org/10.5281/zenodo.3471421
11. Metso, L., Happonen, A., Rissanen, M.: Estimation of user base and revenue streams for novel open data based electric vehicle service and maintenance ecosystem driven platform solution. In: Karim, R., Ahmadi, A., Soleimanmeigouni, I., Kour, R., Rao, R. (eds.) IAI 2021. LNME, pp. 393–404. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-93639-6_34
12. Usmani, U.A., Haron, N.S., Jaafar, J.: A natural language processing approach to mine online reviews using topic modelling. In: Chaubey, N., Parikh, S., Amin, K. (eds.) COMS2 2021. CCIS, vol. 1416, pp. 82–98. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-76776-1_6
13. Trocin, C., Mikalef, P., Papamitsiou, Z., Conboy, K.: Responsible AI for digital health: a synthesis and a research agenda. Inf. Syst. Front., 1–19 (2021)
14. Peters, D., Vold, K., Robinson, D., Calvo, R.A.: Responsible AI—two frameworks for ethical design practice. IEEE Trans. Technol. Soc. **1**(1), 34–47 (2020)
15. Clarke, R.: Principles and business processes for responsible AI. Comput. Law Secur. Rev. **35**(4), 410–422 (2019)

16. Usmani, U.A., Watada, J., Jaafar, J., Aziz, I.A., Roy, A.: A reinforcement learning based adaptive ROI generation for video object segmentation. IEEE Access **9**, 161959–161977 (2021)
17. Sambasivan, N., Holbrook, J.: Toward responsible AI for the next billion users. Interactions **26**(1), 68–71 (2018)
18. Butler, L.M., Arya, V., Nonyel, N.P., Moore, T.S.: The Rx-HEART framework to address health equity and racism within pharmacy education. Am. J. Pharm. Educ. **85**(9) (2021)
19. Ghoreishi, M., Happonen, A.: Key enablers for deploying artificial intelligence for circular economy embracing sustainable product design: three case studies. In: AIP Conference Proceedings **2233**(1), 1–19 (2020). https://doi.org/10.1063/5.0001339
20. Usmani, U.A., Watada, J., Jaafar, J., Aziz, I.A., Roy, A.: A reinforcement learning algorithm for automated detection of skin lesions. Appl. Sci. **11**(20), 9367 (2021)
21. Dignum, V.: The role and challenges of education for responsible AI. Lond. Rev. Educ. **19**(1), 1–11 (2021)
22. Leslie, D.: Tackling COVID-19 through responsible AI innovation: five steps in the right direction. Harv. Data Sci. Rev. (2020)
23. Ghoreishi, M., Happonen, A.: The case of fabric and textile industry: the emerging role of digitalization, Internet-of-Things and industry 4.0 for circularity. In: Yang, X.-S., Sherratt, S., Dey, N., Joshi, A. (eds.) Proceedings of Sixth International Congress on Information and Communication Technology. LNNS, vol. 216, pp. 189–200. Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-1781-2_18
24. Wang, Y., Xiong, M., Olya, H.: Toward an understanding of responsible artificial intelligence practices. In: Proceedings of the 53rd Hawaii International Conference on System Sciences, pp. 4962–4971. Hawaii International Conference on System Sciences (HICSS), January 2020
25. Cheng, L., Varshney, K.R., Liu, H.: Socially responsible AI algorithms: issues, purposes, and challenges. J. Artif. Intell. Res. **71**, 1137–1181 (2021)
26. Happonen, A., Ghoreishi, M.: A mapping study of the current literature on digitalization and industry 4.0 technologies utilization for sustainability and circular economy in textile industries. In: Yang, X.-S., Sherratt, S., Dey, N., Joshi, A. (eds.) Proceedings of Sixth International Congress on Information and Communication Technology. LNNS, vol. 217, pp. 697–711. Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-2102-4_63
27. Ashok, M., Madan, R., Joha, A., Sivarajah, U.: Ethical framework for artificial intelligence and digital technologies. Int. J. Inf. Manag. **62**, 102433 (2022)
28. Usmani, U.A., Watada, J., Jaafar, J., Aziz, I.A., Roy, A.: A reinforced active learning algorithm for semantic segmentation in complex imaging. IEEE Access **9**, 168415–168432 (2021)
29. Maree, C., Modal, J.E., Omlin, C.W.: Towards responsible AI for financial transactions. In: 2020 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 16–21. IEEE, December 2020
30. Rockall, A.: From hype to hope to hard work: developing responsible AI for radiology. Clin. Radiol. **75**(1), 1–2 (2020)
31. Constantinescu, M., Voinea, C., Uszkai, R., Vică, C.: Understanding responsibility in responsible AI. Dianoetic virtues and the hard problem of context. Ethics Inf. Technol. **23**(4), 803–814 (2021). https://doi.org/10.1007/s10676-021-09616-9
32. Happonen, A., Santti, U., Auvinen, H., Räsänen, T., Eskelinen, T.: Digital age business model innovation for sustainability in university industry collaboration model. In: E3S Web of Conferences, vol. 211, Article no. 04005, pp. 1–11 (2020). https://doi.org/10.1051/e3sconf/20202110400
33. Al-Dhaen, F., Hou, J., Rana, N.P., Weerakkody, V.: Advancing the under- standing of the role of responsible AI in the continued use of IoMT in health-care. Inf. Syst. Front., 1–20 (2021)

34. McDonald, M.L., Keeves, G.D., Westphal, J.D.: One step forward, one step back: white male top manager organizational identification and helping behavior toward other executives following the appointment of a female or racial minority CEO. Acad. Manag. J. **61**(2), 405–439 (2018)
35. Usmani, U.A., Roy, A., Watada, J., Jaafar, J., Aziz, I.A.: Enhanced reinforcement learning model for extraction of objects in complex imaging. In: Arai, K. (ed.) Intelligent Computing. LNNS, vol. 283, pp. 946–964. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-80119-9_63
36. Lee, M.K., et al.: Human-centered approaches to fair and responsible AI. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–8, April 2020
37. Yang, Q.: Toward responsible AI: an overview of federated learning for user-centered privacy-preserving computing. ACM Trans. Interact. Intell. Syst. (TiiS) **11**(3–4), 1–22 (2021)
38. Hirvimäki, M., Manninen, M., Lehti, A., Happonen, A., Salminen, A., Nyrhilä, O.: Evaluation of different monitoring methods of laser additive manufacturing of stainless steel. Adv. Mater. Res. **651**, 812–819 (2013). https://doi.org/10.4028/www.scientific.net/AMR.651.812
39. Sen, P., Ganguly, D.: Towards socially responsible ai: cognitive bias-aware multi-objective learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 03, pp. 2685–2692, April 2020
40. de Laat, P.B.: Companies committed to responsible AI: from principles to-wards implementation and regulation? Philos. Technol. **34**(4), 1135–1193 (2021)
41. Happonen, A., Tikka, M., Usmani, U.: A systematic review for organizing hackathons and code camps in COVID-19 like times: literature in demand to understand online hackathons and event result continuation. In: 2021 International Conference on Data and Software Engineering (ICoDSE), pp. 7–12 (2021). https://doi.org/10.1109/ICoDSE53690.2021.9648459
42. Wangdee, W., Billinton, R.: Bulk electric system well-being analysis using sequential Monte Carlo simulation. IEEE Trans. Power Syst. **21**(1), pp. 188–193 (2006)
43. Usmani, U.A., Usmani, M.U.: Future market trends and opportunities for wearable sensor technology. IACSIT Int. J. Eng. Technol. **6**(4), 326–330 (2014)
44. Dignum, V.: Ensuring responsible AI in practice. In: Dignum, V. (ed.) Responsible Artificial Intelligence. Artificial Intelligence: Foundations, Theory, and Algorithms, pp. 93–105. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-30371-6_6
45. Amershi, S.: Toward responsible AI by planning to fail. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, p. 3607, August 2020
46. Cath, C.: Governing artificial intelligence: ethical, legal and technical opportunities and challenges. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. **376**(2133), 20180080 (2018)
47. Eskelinen, T., Räsänen, T., Santti, U., et al.: Designing a business model for environmental monitoring services using fast MCDS innovation support tools. TIM Rev. **7**(11), 36–46 (2017). https://doi.org/10.22215/timreview/1119
48. Truby, J.: Governing artificial intelligence to benefit the UN sustainable development goals. Sustain. Dev. **28**(4), 946–959 (2020)
49. Happonen, A, Salmela, E.: Automatic & unmanned stock replenishment process using scales for monitoring. In: Proceedings of the Third International Conference on Web Information Systems and Technologies - (Volume 3), Barcelona, Spain, 3–6 March 2007, pp. 157–162 (2007). https://doi.org/10.5220/0001282801570162
50. Braun, B.: Governing the future: the European central bank's expectation management during the Great moderation. Econ. Soc. **44**(3), 367–391 (2015)
51. Nitzberg, M., Zysman, J.: Algorithms, data, and platforms: the diverse challenges of governing AI. J. Eur. Public Policy (2021)

52. Salmela, E., Santos, C., Happonen, A.: Formalisation of front end innovation in supply network collaboration. Int. J. Innov. Reg. Dev. **5**(1), 91–111 (2013). https://doi.org/10.1504/IJIRD.2013.052510
53. Piili, H., et al.: Digital design process and additive manufacturing of a configurable product. Adv. Sci. Lett. **19**(3), 926–931 (2013). https://doi.org/10.1166/asl.2013.4827
54. Metso, L., Happonen, A., Rissanen, M., Efvengren, K., Ojanen, V., Kärri, T.: Data openness based data sharing concept for future electric car maintenance services. In: Ball, A., Gelman, L., Rao, B.K.N. (eds.) Advances in Asset Management and Condition Monitoring. SIST, vol. 166, pp. 429–436. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-57745-2_36
55. Happonen, A., Siljander, V.: Gainsharing in logistics outsourcing: trust leads to success in the digital era. Int. J. Collab. Enterp. **6**(2), 150–175 (2020). https://doi.org/10.1504/IJCENT.2020.110221
56. Kärri, T., Marttonen-Arola, S., Kinnunen, S-K., Ylä-Kujala, A., Ali-Marttila, M., et al.: Fleet-based industrial data symbiosis, title of parent publication: S4Fleet - service solutions for fleet management, DIMECC Publications series No. 19, 06/2017, pp. 124–169 (2017)
57. Kinnunen, S.-K., Happonen, A., Marttonen-Arola, S., Kärri, T.: Traditional and extended fleets in literature and practice: definition and untapped potential. Int. J. Strateg. Eng. Asset Manag. **3**(3), 239–261 (2019). https://doi.org/10.1504/IJSEAM.2019.108467
58. Metso, L., Happonen, A., Ojanen, V., Rissanen, M., Kärri, T.: Business model design elements for electric car service based on digital data enabled sharing platform, Cambridge. In: International Manufacturing Symposium, Cambridge, UK, 26–27 September 2019, p. 6 (2019). https://doi.org/10.17863/CAM.45886
59. Palacin, V., Gilbert, S., Orchard, S., Eaton, A., Ferrario, M.A., Happonen, A.: Drivers of participation in digital citizen science: case studies on Järviwiki and safecast. Citiz. Sci. Theory Pract. **5**(1), Article no. 22, pp. 1–20 (2020). https://doi.org/10.5334/cstp.290
60. Palacin, V., et al.: SENSEI: harnessing community wisdom for local environmental monitoring in Finland. In: CHI Conference on Human Factors in Computing Systems, Glagsgow, Scotland UK, pp. 1–8 (2019). https://doi.org/10.1145/3290607.3299047
61. Zhang, D., Yin, C., Zeng, J., Yuan, X., Zhang, P.: Combining structured and unstructured data for predictive models: a deep learning approach. BMC Med. Inform. Decis. Mak. **20**(1), 1–11 (2020)
62. Vassev, E., Hinchey, M.: Autonomy requirements engineering. In: Vassev, E., Hinchey, M. (eds.) Autonomy Requirements Engineering for Space Missions. NASA Monographs in Systems and Software Engineering, pp. 105–172. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-09816-6_3