Ethics and Information Technology (2022) 24: 37
https://doi.org/10.1007/s10676-022-09634-1

REVIEW PAPER

Characteristics and challenges in the industries towards responsible AI: a systematic literature review

Marianna Anagnostou[1] · Olga Karvounidou[1] · Chrysovalantou Katritzidaki[1] · Christina Kechagia[1] · Kyriaki Melidou[1] · Eleni Mpeza[1] · Ioannis Konstantinidis[1] · Eleni Kapantai[1] · Christos Berberidis[1] · Ioannis Magnisalis[1] · Vassilios Peristeras[1]

Published online: 29 August 2022

Abstract
Today humanity is in the midst of the massive expansion of new and fundamental technology, represented by advanced artificial intelligence (AI) systems. The ongoing revolution of these technologies and their profound impact across various sectors, has triggered discussions about the characteristics and values that should guide their use and development in a responsible manner. In this paper, we conduct a systematic literature review with the aim of pointing out existing challenges and required principles in AI-based systems in different industries. We discuss our findings and provide general recommendations to be considered during AI deployment in production. The results have shown many gaps and concerns towards responsible AI and integration of complex AI models in the industry that the research community could address.

Keywords Artificial intelligence · Responsible AI · Characteristics · Challenges · Business · Ethics

Introduction

Entering a new era of technological advances, Artificial Intelligence (AI) is becoming increasingly ubiquitous in various industries, including healthcare, security, e-government and public sector, transportation, logistics, education, etc. Current advances in the development and applications of AI systems, such as machine learning, deep learning, autonomous or semi-autonomous systems, reshape science and society, creating new opportunities in the way we live, work, travel and do business.

Despite an apparent agreement that AI systems offer tremendous potential in industry and everyday lives, we must still face challenges around their responsible use and development. Given that there is no human intervention in many cases, concerns such as workers being replaced, unfair decisions, misuse by malevolent actors, and lack of privacy and security have been at the forefront of research activity by academia and business organizations.

In order to address these concerns, we need to develop fair and trustworthy systems establishing principles, best practices and values that should guide AI applications for more effective communication and equitable collaboration between the various communities and industries.

To that end, the term of "Responsible AI" has been introduced and defined as a governance framework that documents how a specific organization is addressing the challenges around artificial intelligence (AI) from both an ethical and legal point of view (Dignum, 2019).

There are existing works that propose a set of general principles for responsible AI, like Google's AI principles (Google, 2021), but not at the industry level, as business sectors exhibit different needs and requirements for responsible AI.

In this paper, we conduct a systematic literature review to explore the current status of responsible AI in various industries by identifying the challenges and required characteristics of AI systems. In that way, we introduce new research directions to come across a set of principles that an individual, an organization, or a government could follow in order to create responsible AI models that benefit people and society.

✉ Ioannis Konstantinidis
i.konstantindis@ihu.edu.gr

[1] International Hellenic University, Thessaloniki, Greece

Background

According to Dignum (2019), "Responsible AI is about being responsible for the power that AI brings". Realizing the significant effect of responsible AI systems, academia, international, and other organizations try to first understand and then act to mitigate the potential negative impacts of AI systems. This action takes various forms, including the launch of major ethical guidelines, principles and recommendations (HLEG, 2019; Floridi et al., 2018; Google, 2021; Microsoft, 2021), articulating theoretical and practical approaches, developing tools (PwC, 2019) and examining the possibility of a common AI ethics language when developing and deploying AI-driven products and services (Morley et al., 2020).

Having extensively studied the bibliography on existing ethical guidance and methodologies, we identified a high degree of overlap and repetition among the principles that is also verified by many outputs across academia (Ryan & Stahl, 2021). Floridi (2013) in order to address this principle proliferation, introduced a five principle framework (beneficence, non-maleficence, autonomy, justice and explicability), extending the already existing literature and adopting principles used in bioethics. Since then, the framework was adopted by various works such as Ethics Guidelines for Trustworthy AI published by the European Commission's High-Level Expert Group on Artificial Intelligence (HLEG, 2019) and the OECD's Recommendation of the Council on Artificial Intelligence (Floridi & Cowls, 2019). AI4People presented these five principles along with 20 concrete recommendations to lay the foundations for the establishment of a Good AI Society (Floridi et al., 2018).

The most famous principles excluded from the ongoing academic discussions are Transparency, Accountability, Responsibility and Fairness (Dignum, 2019; Vakkuri et al., 2019a) without composing the core of responsible AI frameworks. The Berkman Klein Center discusses the upsurge of policy activity and introduces eight key themes for deeper commentary and detailed principles: privacy, accountability, safety & security, transparency & explainability, fairness & non-discrimination, human control of technology, professional responsibility, and promotion of human values (Healey, 2020).

Although numerous guidance documents indicate an emerging convergence around the importance of certain ethical principles, an equivalent proportion of documents conclude to a significant divergence among the applied domains (Jobin et al., 2019). Triggered by this complicated picture, our objective was to investigate responsible AI on different industries in the literature in order to identify related challenges and recommendations.

Research methodology

Goal and research questions

The goal of our research work is to identify the challenges and characteristics towards responsible AI in different industries. For this reason, we define the following research questions:

RQ1: What are the challenges of AI models for responsible AI in different business sectors?

RQ2: What are the required characteristics of AI models for responsible AI in different business sectors?

Search process

In order to effectively address our research questions, we conducted a systematic literature review (SLR) based on Kitchenham & Charters (2007) methodological guidelines. Our automatic searching activity to identify relevant publications was conducted on Scopus,[1] Springer[2] and Science Direct[3] databases on December 2020, using the following research query [("machine learning" OR "artificial intelligence" OR "AI") AND ("responsible" OR "ethics") AND (business OR industry)].

Inclusion–exclusion criteria

The inclusion criteria that were used to set the boundaries for the literature review and answer the proposed research questions are the following:

CR1: We included sources on academic theoretical and empirical research (journal articles, chapters from edited books, conference proceedings, etc.) in English, published within a 10-year period (2010–2021).
CR2: We included publications that explicitly mention ethical AI challenges or characteristics.
CR3: We included papers that address specific industries and use cases. Each industry should be referred to at least five papers in order to have adequate information on each domain. For that reason, the final industries had been classified into more broad categories, which are the following: Business Industry, Autonomous Machines, Healthcare Industry, E-Government, Social Industry.
CR4: We excluded sources that focus on the technical implementation of algorithms.

[1] https://www.scopus.com/.
[2] https://link.springer.com/.
[3] https://www.sciencedirect.com/.

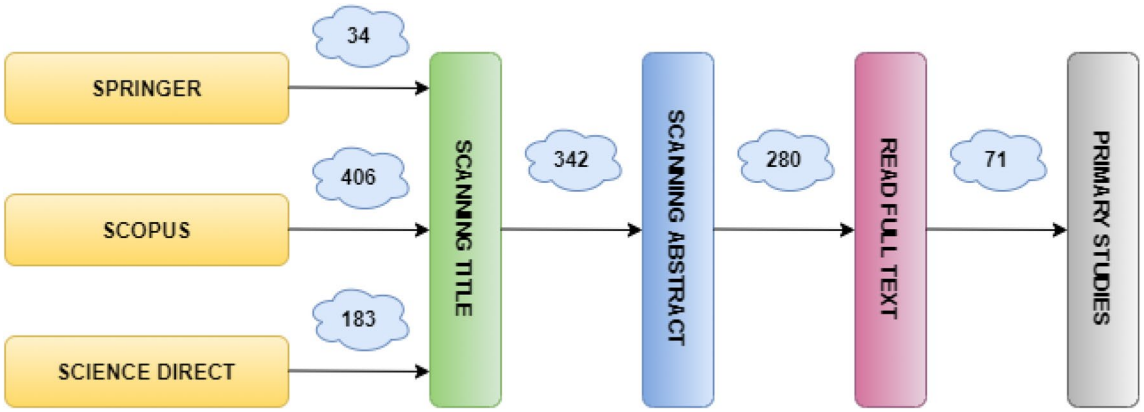

Fig. 1 Process of identifying primary studies

Table 1 Number of related publications in different combinations of challenges and industries

Industry Sector	Ambiguity	Vulnerability	Privacy	Lack of accountability	Morality & bias	Papers
Business management	6		4	3	4	1,8,11,12,13,14,21,24,35,37,44,48,51,52,53,64
Transportation	3	1	1	2	3	7,27,32,40,43,58,59,61
Healthcare industry	3	1	2	2	3	5,6,16,31,36,39,54,60
E-Government & public sector	6	5	2	3	2	1,9,16,20,22,29,30,34,42,50,55,69
Information technology	4	4	1	2	5	2,10,12,18,25,34,38,46,49,61,63,66

Figure 1 illustrates in detail the selection process of the records found through database searching. Our research results identified 71 references as our primary studies that were assessed for eligibility in RQ1 and RQ2.

RQ1: challenges of AI systems for responsible AI in different business sectors

In order to address RQ1 and identify potential challenges of responsible AI in the industries, we reviewed the selected primary studies. We observed challenges related to the use of AI systems and others related to the autonomy of those and classified them according to similar scenarios and characteristics, concluding to the following categories: transparency, security, privacy, responsibility/accountability, morality/bias. The latter were ignored as they appear a rather narrow focus in specific and limited domains. According to the Global Industry Classification Standard (MSCI, 2020), we also identified the following industry sectors.

1. Business management: This sector generally refers to Finance and Consumer Discretionary, identified in the literature without focusing on a specific domain. It consists of AI-driven insurance systems and risk management, consumer analytics and business strategies in light of AI-powered applications and technological systems in order to improve customer's experience.
2. Transportation: It includes literature about self-driving cars and AI-driven transport systems and services.
3. Healthcare: It consists of AI-driven healthcare equipment and services, biotechnology and pharmaceutical industry groups.
4. E-government & public sector: It comprises of e-services and social-political implications of AI technology use in our daily life.
5. Information technology: This sector refers to research studies on a more technical level, including AI-based software services, hardware, robotics and autonomous machines.

Table 1 presents the number of related publications on each challenge and industry.

Ambiguity

Ambiguity is defined as the lack of understanding of how the system works (Vakkuri et al., 2020). It involves the algorithm, the data, as well as the technical aspect and the

process of developing the system. A major challenge is establishing the transparency and explainability of AI systems. A lot of algorithms are incomprehensible to the end-users (Cath et al., 2017). The user should not only be able to understand how the system works but should also have the ability to gather information on who made the system the way it is and why (Vakkuri et al., 2020).

Business management

Emerging technologies can heavily impact business processes and AI systems have proven to be useful in the business industry by helping in the automation of tasks as well as the facilitation of new methods (Mendling et al., 2018). However, for the algorithms to be ethically applied, the parties involved should guarantee their transparency. In the case of business processes, it is mentioned that transparency is crucial to ensure that the end-users are fully aware of the situation they are involved in (Aitken et al., 2020; Braun & Garriga, 2018; Pedersen & Johansen, 2019). Sometimes, businesses and service providers convince their customers to accept their offers and products without disclosing all the necessary information. In addition, companies often use customer data for their AI-based systems without clearly informing them about their practices and by including "terms and conditions" that are not directly accessible. This leads to customers being unaware of how, why, and with whom their data will be shared (Munoko et al., 2020). Lastly, methods based on AI must ensure the explainability of the process. The ambiguity about how an algorithm works could discourage the companies from adopting a new method fast enough, even if it is extremely useful, resulting in losses (Breidbach & Maglio, 2020).

Transportation

In all instances, it is not entirely comprehensible how the algorithm reaches a decision, which creates more issues and ultimately leads to the user blindly trusting the system (Timmers, 2019). In transportation services and especially self-driving cars, technology often takes the role of imitating human behaviour and executing tasks accordingly (Roberts et al., 2020). However, an AI agent might be capable of driving a car but ambiguous on how it reaches a decision from so many data points, yet the driver may blindly trust that decision. A driverless car can face many challenges, ethical and legal, and it will need to decide on its course of action. By ensuring the transparency of the algorithm, the involved parties are able to study the way the machine operates under specific circumstances, make appropriate changes, and utilize the information for the future (Etzioni & Etzioni, 2017).

Healthcare

Ambiguity is considered a critical challenge for industries that directly involve the well-being of the customer (Pitman et al., 2019). One of those instances is healthcare, a domain where the doctor and the parties involved should ensure that the system is not ambiguous and capable of explaining the reasons for their decisions (Loftus et al., 2020). For example, a doctor should be able to request more information on how a diagnosis was reached in order to avoid errors that could be fatal or hinder the healing treatment of the patient (Cath et al., 2017).

e-Government/public sector

A major challenge when it comes to e-Government and the public sector is ensuring explainability and transparency of the solutions used (David et al., 2019; Rahwan, 2017). Most AI systems used are incomprehensible to users, and even though some techniques might focus on providing explanations, those are usually not fully accurate (Cysneiros & do Prado Leite, 2020). For that reason, there is a lack of public involvement and engagement, which means that the service providers cannot easily verify that an AI solution aligns with the public's interest (Aitken et al., 2020; Marri et al., 2019). There are also difficulties associated with low data quality and decision-making, which can be incredibly challenging in the human-centered environment of the public sector (Cath et al., 2017).

Information technology

For computer scientists, it was much easier to understand and manage the first generation of AI systems. However, through the years and with the introduction of much more complicated techniques, such as deep learning, it has become really challenging to understand how the machines come to specific decisions (Breidbach & Maglio, 2020; Turner, 2018). Ambiguity can undermine the importance of this technology by making users afraid of adopting it and disinterested in the results (Kuleshov et al., 2020). For example, ambiguity and lack of explainability can affect search results (Howard & Borenstein, 2017). When someone uses a search engine, the set of criteria on which the results are based is not immediately visible and straightforward. This way, the user is confronted with results that could be biased, either because of society or because of the developer.

Vulnerability

Responsible AI systems offer benefits and improve human lives, but also expose private data to risks and uncertainty. It is important to highlight the need for an ecosystem for

secure and trustworthy AI. AI can be misled or manipulated, creating security issues in monitoring tools, financial systems, autonomous vehicles, and machines. In the related literature, we identified several security issues in responsible AI (Ashrafian, 2014b). With the term "Security" we mean not only security in general but also cyber-security. Cyber-security refers to defending computers, servers, mobile devices and electronic systems in general from cyber-attacks (Gill, 2019; Timmers, 2019).

Transportation

Artificial Intelligence is changing the transport sector as well. Beyond all the benefits they offer, these AI inventions pose security threats in most cases, such as transportation. According to a related study (Hauer, 2019), most drivers would agree that an autonomous car, trained with the goal of minimizing damage in critical situations, is ethically designed. However, many of them would be unwilling to drive such a car and be hesitant to include other people.

Healthcare

Another critical issue is the AI systems' vulnerability in the healthcare sector. An example of this issue is robotic surgical platforms. These platforms with virtual constraints intended to protect anatomic structures could delay or prevent a surgeon from gaining control of an injured blood vessel, harming a patient and pitting human versus machine in assigning liability. Machine learning models can predict risk for several postoperative complications with high accuracy but often lack electronic and clinical workflow integration, limiting their use in routine clinical practice and posing risks related to health or even loss of life (Loftus et al., 2020).

e-Government/public sector

E-government utilizes information and communication technologies in order to improve relations between citizens, businesses and government, aiming at better services, communication and governmental efficiency (Cath et al., 2017; Montes & Goertzel, 2019; Rahwan, 2017). However, despite the adoption of e-Government services, the threats of privacy and security remain major concerns. Other previous research findings stated that AI-based software solutions will demand safety to be mandatory. E-Government services reveal private data and may provoke cyber attacks and exploitation of personal information. This fact may pose several threats to not only privacy but also security for citizens (Cysneiros & do Prado Leite, 2020; Gill, 2019).

Information technology

In many industry fields, advancements in computer science and information technology, in general, have improved computability with AI methods and tools. There is a need for security risk assessment of these technologies to complete tasks in a similar or even a better way than humans (Ashrafian, 2014a). As machine learning systems increasingly affect our lives, it is crucial to ensure that we will be safe under the control of such machines (Chakraborty et al., 2020; Mittelstadt et al., 2016). Designing and training a machine is a complex process and several errors may occur. Machine intelligence should not be equated to human intelligence. So, a machine learning model should be designed properly. An AI machine may harm a person in several ways, such as being exposed to danger or acting in a wrong and unprogrammed way. The most important is the right and entirely programming of machines to complete several activities concerning humans' safety. AI machines should be ready and programmed to react to different situations with real-time protection (McAleenan, 2020).

Privacy

Over the last 30 years, data protection constituted a hot subject of AI ethics in the industry (Mason, 1986). Privacy is a persistent concern, as demonstrated by several AI ethics guidelines and studies. Data protection has become an unusual issue to be inspected from the viewpoint of corporate roles: data privacy officers are the most similar to environmental managers (Juho, 2019). Although these positions seem to be compliance-oriented, the need for a more proactive approach to privacy problems has been recognized (Kleindienst et al., 2017).

Business management

The information revolution and the lack of clear legal boundaries in digital environments are at the root of some of today's most pressing data protection issues. Because of the pervasiveness of information, state action is frequently ineffective (Kleindienst et al., 2017); (Pagallo, 2012). Firms have always been mindful of the growing importance of privacy and regulation when building an AI-based solution (Mason, 1986). According to Campbell et al. (2020), consumers are curious about what kind of data is gathered about them and how advertisers use it. As a result, companies like Apple are proactively opting to limit the types of customer data gathered and how it is used.

Transportation

Self-driving cars collect continuous data about their surroundings using advanced sensors. There is no fair expectation of privacy in public spaces where autonomous vehicles run, and no warning or preference is given, creating privacy concerns. Public interests and complaints must first be investigated to increase public awareness of networked autonomous vehicles and promote the implementation of technological and policy frameworks to safeguard privacy. For example, a survey of 302 participants was conducted using privacy breach scenarios to understand participants' opinions on the various possibilities and comforts of self-driving cars. After analyzing the results, it was concluded that most participants were not comfortable with autonomous vehicles (Bauer et al., 2017).

Healthcare

The privacy issues arising from the use of AI in healthcare cannot be ignored, and it is essential to consider that the boundary between personal data and confidential data is becoming rapidly blurred, as we can now derive patient information from activity habits and other non-sensitive data. However, innovations must adhere to the legislation, rules, and privacy standards in order to ensure that creativity benefits the public interest. That is because deploying AI in the medical sector poses a slew of privacy and ethical issues, including the protection of personal records, the ethical limits of creativity, and the real effect of technology on doctors and patients (Ten Teije et al., 2017). For example, a characteristic challenge is the difficulty of health professionals to be trained on how to work with a machine and the effect that this way will have to employees. To effectively exploit AI in healthcare, we have to answer questions like "How the machine increases doctor's abilities?", "What will happen if the machine makes the decisions and not the doctor?" (Bartoletti, 2019).

e-Government/public sector

Human beliefs and AI are in continuous flux. As a result, technological advancements will drastically modify what culture finds acceptable and consider how privacy standards have shifted due to the convenience offered by mobile phones and the Internet (Rahwan, 2017). Smart cities provide municipalities fresh and unparalleled economic possibilities, however, new developments are often followed by protection and privacy issues. On top of that, a smart city needs a higher level of network access to personal data to support a wide range of disparate devices with varying software and hardware capacities. For example, a smart city with healthcare provisions makes it easier for people to live healthy lifestyles by giving quick access to medical services. Smart cities offer gadgets that connect and help public health experts by integrating medical systems and patient data records, often maintained by e-government (Yang et al., 2019).

Information technology

For information technology, with the vast array of sensors and cameras that IoT (Internet of Things) devices now employ, privacy has become a major concern in the technological world. It is obvious that deep learning requires a large amount of data to produce optimal results, which can only be achieved through continuous data collection and monitoring, which many users may regard as a breach of privacy. The hacking of widely used smart security cameras and doorbells is one example of the consumer and industrial sectors' lack of security procedures for these technologies (Chakraborty et al., 2020).

Lack of accountability

Accountability has to do with the consequences assigned to an actor for specific actions and decisions. The absence of accountability in the case of autonomous systems and their potential misuse is one of the biggest challenges for industry and the end-user. To maintain clarity, decisions must be taken from and explained by the decision-making algorithms used. This requires the need to embody the universal principles and social expectations that the agent employs for deliberation in the action framework. In AI, accountability encompasses both the task of directing behavior (forming values and making decisions) and the purpose of explanation by placing decisions in a broader context and by classifying them along moral values (Dignum, 2019).

Business management

There are some important steps that companies need to embrace in order to progress towards responsible and accountable AI. Explainability and interpretability are essential requirements for responsible AI. Juho (2019) describes on whom the responsibility and accountability have been set in the business industry. The authors present three types of workers that the AI technology age will create: trainers, explainers, and maintainers. Thus, they offer a clear position on the roles related to the management of ethical and responsible AI systems. Data can include a variety of issues that raise ethical questions. The individual responsible for selecting, updating, cleaning and inputting project data will have extensive responsibility for ethical behavior (Dignum, 2019; Juho, 2019). Additionally, in operating processes and challenges for emerging market models, AI innovations are

rapidly at stake in a competitive and socially conscious manner. The stakeholders will play a constructive or marginal role in the value creation for industry, according to their own environmental consciousness. These problems remain 'open' in a number of sectors such as the food system, where rethinking and redesigning of the entire business model involves the introduction of emerging technology (Vaio et al., 2020).

Transportation

In transportation and specifically self-driving cars, the issue of responsibility and accountability is considered critical. The exact role of the AI system in driverless vehicles is still not obvious, which complicates the question of who is responsible for an accident in which a self-driving car was involved. For that reason, many argue that there should be a semi-autonomous approach so that the driver can be allowed to take certain decisions in case of an emergency (Russell & Norvig, 2002). For example, in several areas, including safety and fuel efficiency, self-driving cars are projected to surpass human drivers. There are some claims that AI might make automobiles "emotionally aware" and able to detect the sentiments, emotional states, and health conditions of drivers and passengers. This would allow the driver to be advised, for instance, when tiredness is identified by being proposed to take a break or to make decisions such as slowing down the car (Lugano, 2017).

Healthcare

Furthermore, people are more curious about the ethics of intangible shifts, such as behind-the-scenes technologies. Doctors' decisions will be assisted by electronic technology, but the doctor will still be in charge. The ethical dilemma is one that should be answered by both technology engineers and healthcare practitioners. It should concentrate on how individuals and technologies communicate. For example, many fields of robotics research and development are focused to enhancing patient care, such as robotic wheelchairs or smart walkers. It is believed that assistive robots will significantly improve the mobility and autonomy of the elderly. AI is also used in the creation of high-resolution digital pictures and holograms (Banks, 2018). In addition, AI is progressively a part of digital medicine and will lead to research and practice in the field of mental health (Graham et al., 2019). The authors represent a network of mental health research and therapy professionals who must collaborate to fully comprehend AI capability. They state that it is important that the algorithms used to predict or classify mental health conditions are accurate without exposing patients to risk. They also mention that no clear rules govern the use of AI and other emerging technology in healthcare and that AI systems must supplement, not replace, medical practice, and public information about these algorithms must be useful and contextual. It is evident that ethics must be incorporated into the development of AI through science and education.

e-Government/public sector

Moreover, e-Government and the public sector deal with responsibility and accountability. In the case of the United Arab Emirates (UAE), both the government and the private sector are responsible for the implementation and adoption of AI policy in the e-services sector (Ghandour & Woodford, 2019). This has been achieved through the Ministry of AI, which was created primarily to allow the government to implement AI in its different sectors. The private sector will contribute to AI growth by engaging in research and incorporating AI into various facets of life. This can be achieved by collaborating with the government and thus pooling their resources. One approach to overcoming problems raised by AI adaptation is integrating AI technology into a citizen-centered policy focusing on results. The government should do this in terms of the population's end-to-end process, in which case the citizen is placed first. This technology has the potential to be inclusive as well. As a result, there must be some attention to generational, educational, income, and language differences. The avoidance of ethical risks, that would be critical to driving decision-making by AI, is the next important approach that can be used to solve challenges encountered during the adaptation of AI into government services. This approach would be especially helpful in overcoming the bias question that AI has revealed. Bias can also be avoided by including multidisciplinary and representative teams in the application of AI, as well as the use of ethicists. AI cannot be used to make delicate decisions that would have a huge effect on people's lives before this is accomplished (Marri et al., 2019). According to Rahwan (2017), advances in AI have raised several concerns regarding regulatory and governance mechanisms for autonomous machines and complex AI systems. It is addressed that algorithmic processes are not responsible because they are black boxes whose internal workings are not open to all stakeholders. Other concerns include people inadvertently living in filter bubbles generated by news recommendation algorithms, while others claim that data-driven decision-making processes can exacerbate inequality, and they can often be biased either in their nature or by capturing individual prejudices in their training data. In addition, algorithms may generate feedback loops that perpetuate inequalities, such as the usage of AI in predictive policing or creditworthiness prediction, rendering it impossible for individuals to avoid the vicious cycle of poverty.

Information technology

Responsibility and accountability are equally important challenges in the IT industry. Developers usually fail to create systems that can ensure the AI system's accountability during the decision-making process, which often results in these systems making decisions that are biased and discriminatory. Specifically, Breidbach & Maglio (2020) discuss the issue of "proxy discrimination", which means that certain attributes can be used to indirectly discriminate against groups of people. For example, a ZIP code can be used to profile a consumer racially. In all cases, discrimination against race or gender is illegal. However, when these algorithms come to a biased decision, it is too difficult to identify who is responsible for that act. Overall, there is a growing concern that developers and designers could be careless in their making of AI systems by simply shifting the blame on the algorithm when problems arise (Timmers, 2019).

Morality & bias

Artificial intelligence has many applications in our lives in many different industries. But it is essential to ensure that its appliance is in an ethical manner. This restriction poses many challenges. One of them is morality. How can we allow intelligent machines to make moral decisions about human lives and who should design the algorithms for the decision will be taken by intelligence machines is a question that should be answered? Another critical challenge that AI systems must deal with is bias. Issues of bias are fundamentally related to the human evaluation of results obtained using AI, thus connecting to building AI Systems based on ethical principles (Kuleshov et al., 2020). The algorithms are not smart by themselves but are trained by humans and their explicit biases are observable, with gender and racial biases most found in machine learning algorithms at present. Making the correct decision in the face of an ethical conflict is extremely difficult. Laws are often based on vague concepts such as "safe" and "reckless", which may prove difficult to quantify. As a result, there has been a growing recognition in recent years of the need to ensure that AI systems are aligned with human values.

Business management

It refers to the difficulty that AI systems face in acting and making ethical decisions when they are applied in diverse kinds of businesses. This encompasses a diverse set of moral codes, social norms, moral values, and actions. For example, an AI system can reject an application based on biased data. Moral dilemmas refer to situations in which AI systems must choose between opposing alternatives. Compatibility of the machine and human judgment refers to the discrepancy between machine-based decision-making and human values, which may not be adequately reflected in AI systems (Caner & Bhatti, 2020). Intelligent agents raise new questions about who bears the risk of unintended effects, how liability can be assigned in online business settings, and how the law can respond in situations where technology is intelligent enough to behave autonomously rather than automatically (Dahiyat, 2010). In the 1990s, some researchers started to note expert systems' ethical consequences, such as their lack of cognitive skills (such as intellect, feelings, and values) and bias. Despite these ethical questions, the profession adopted newer AI techniques at the turn of the century. On the other hand, today's AI systems do not seem to have resolved prior ethical issues. It could be difficult for the auditor using the AI to justify the selection of specific samples or processes for testing if the AI is a "black box." In such cases, auditors may exhibit automation bias and complacency, i.e., a lack of skepticism and faith in the AI system's accuracy (Munoko et al., 2020). As a result, machine ethics are important as they will confront the moral choices of semi-autonomous and autonomous machines that communicate with animals of all kinds, giving it a wide and fertile area of application (Bendel, 2017).

Transportation

The challenge of artificial morality has different implementations in each industry. A great concern regarding the morality in AI refers to the industry of transportation and, more specifically, self-driving cars. Although this innovation has many advantages for the end-user, it poses several questions regarding the car's decision and the fact that it is produced in certain ways on behalf of the driver. Researchers are wondering how we can ascribe something like a mental state to a machine or how a machine will decide who will ''kill'' in extraordinary situations (Etzioni & Etzioni, 2017). A set of issues has been arising to find the optimal solution in designing that kind of automated vehicle. The first one is to let the developers of intelligent machines predefine the car's behavior, but they should not exclusively decide about their morally laden behaviors. Moreover, developers are not well-informed and aware of ethics and every person acts differently in the same situation. Another one is to let users decide their actions; the car will receive data and then it will act in the same way. This is called collectively deciding, but still, there are limitations. That case imposes a burden that will reduce the efficiency of the creative processes of these machines or the general pace of innovation (Martin, 2016). Conclusively, as the authors explained, no matter how good the learning algorithm is, it is always as good as the data you give it. The one who controls the data controls the game (Hauer, 2019).

Table 2 Required characteristics of AI systems in industries

Industry	Transparency	Account-ability	Fairness	Empathy	Security	Papers
Business management	5	3	1		1	1,4,12,23,15,21,41,48,51
Transportation	3	2		1		27,28,32,45,47,66
Healthcare Industry	3	3	1	1		5,19,31,36,39,67,68
E-Government & public sector	4	4	1			1,9,16,26,42,55,56
Information technology	5	6	1	1	3	10,17,25,27,33,38,46,57, 61,62,63,66,68

Healthcare

There is a growing concern that bias-influenced AI systems will further perpetuate health inequalities for patients regarding the healthcare industry. The underlying reason why bias is retained in AI models is often related to training data that do not represent the entire population equally (Kang et al., 2021). There are still many obstacles to clearing up for robotics to play a more prominent role in health care. Since nursing depends on human communication, there is a reluctance to use technology in that field, partly for ethical reasons and partly due to the need for personal interaction (Banks, 2018). The Human Touch: Practical and Ethical Implications of Putting AI and Robotics to Work for Patients). The need for study and ethics approval, which is usually needed for further research, must be considered when developing AI tools. The first major challenge is identifying the clinical need, deciding who should make those decisions, and developing goals for algorithm development (Pitman et al., 2019). Artificial intelligence in digital pathology: a roadmap to routine use in clinical practice.)

e-Government/public sector

In the same way, regarding the e-Government sector, it is doubtful that a machine-learning algorithm can equate to ethical theories. Either the system's developers would devise a method for identifying ethical choices, such as machine learning, or a moral principle would be selected for the system and applied before it was deployed. As a result, the developer must be unbiased in this case, which is also a prerequisite for responsible AI (Bjørgen et al., 2018). The Federal Trade Commission reported that algorithms based on big data sets could reproduce existing patterns of discrimination, inherit prior decision-makers prejudice, or simply represent the widespread prejudices that remain in society (Howard & Borenstein, 2017).

Information technology

Potentially negative or unethical consequences can result from the use of AI or machine learning in the industry of information technology (Breidbach & Maglio, 2020). The algorithms used in AI predict future outcomes by searching for correlations. These algorithms are designed by people and in many cases, the designers did not manage to equate correlation to causation and the ability of any predictive algorithm to make correct decisions are contingent on the quality of the data sets used to train the algorithms in the first place, with low training-data quality leading to poor algorithmic performance (Vakkuri et al., 2019b). All the above have, as a result, the problem of moral disagreement (Bogosian, 2017). More specifically, designers of AI systems are also humans with different approaches in morality (Dodig Crnkovic & Çürüklü, 2011). For someone who believes in a particular approach to ethics, the correct system for implementation in artificial moral agents may be obvious, but the path forward for institutions and society remains unclear (McAleenan, 2020).

RQ2: required characteristics of AI models for responsible AI in different business sectors

In our research, we encountered several traits that could be attributed to AI agents and systems. Although not all of them display the same characteristics, we concluded on specific capabilities that an AI system should have for it to be useful, informative, and ethical and divided them into the following categories: transparency, accountability, unbiasedness, empathy, and security. We also identified in the literature other characteristics, including autonomy, that did not meet the required number of mentions in papers, or they did not refer to any of the five industries, so these challenges have been rejected. In the literature, we distinguished the same

business sectors as in RQ1. Table 2 shows the number of publications for each characteristic and industry.

Transparency

A major prerequisite for an algorithm to be considered just and ethical is transparency, and the concept of transparency is multifarious. When referring to transparency in AI systems, we are interested in, for example, why the algorithm reaches a decision, how it was done, what steps can be accounted in the decision-making process, what were the determining factors that led to the choice, what alternatives could have been made (McAleenan, 2020).

Business management

Transparency is an essential requirement in AI systems used in many industries, and especially algorithms and methods used for business processes (Breidbach & Maglio, 2020; Mendling et al., 2018). For business processes to be ethical and maintain transparency, researchers suggest that managers should thoroughly understand AI technology, the way it works, and how it is implemented to be able to use it effectively to resolve issues and encourage data transparency (Luo, 2019). Furthermore, companies are required to enforce regulations, act professionally and transparently in order to build trust with the customers, respect their privacy, and abide by their code of ethics (Balmer et al., 2020; Munoko et al., 2020).

Transportation

The algorithm needs to be explainable and transparent not only to the developer but also to the distributor and the end-user (Etzioni & Etzioni, 2017). For example, the details of how a driverless car operates and reaches a decision should be completely accessible and understandable to the operator to ensure safety and ethical decision-making. This might mean that a driving test should be a requirement for the self-driving car to prove it can navigate, deal with roads, and abide by the codes, law, and cultural habits of the geographical region it will operate in (McBride, 2016). Additionally, if a driverless car gets involved in an accident and we cannot understand why, it is impossible to assess the incident and the ethical implications leading up to the decision if the system is not transparent (Vakkuri et al., 2019b).

Healthcare

Transparency is especially crucial for decision-making progresses that directly affect human lives. Especially in the health sector, researchers suggest that anytime an AI agent is applied, doctors and scientists should note the system's details and provide information on how it was trained to ensure explainability (Banks, 2018). Specifically, it is essential to evidently document where, when, and how AI technology was used in any part of the clinical process (Kang et al., 2021). Without transparency, it is impossible to discover the cause of a medical diagnosis, and the process of discovery is crucial if we want to be able to track the faults that caused an incident and establish accountability (Winfield et al., 2019).

e-Government/public sector

Many authors agree that a required characteristic for an AI system in all industries is to be transparent in order to avoid bias, be ethical, ensure autonomy and freedom, and foster self-responsibility (Rességuier & Rodrigues, 2020). This is especially true for the public sector, where many researchers stress the importance of transparency in e-government services and practices where the public is directly involved (Marri et al., 2019). Specifically, the European Union is a strong advocate of the transparency of AI systems (Cath et al., 2017). For instance, it proposes proper documentation, codification, and labeling of the phases for someone to directly identify why an error was made during the process and, consequently, assist in preventing future mistakes while enabling explainability (Aitken et al., 2020).

Information technology

Transparency is also necessary for systems such as machines and robotics (McAleenan, 2020). Programmers should be transparent during their designing process and make sure that end-users understand what steps are taken when a machine makes a decision and why (Vakkuri et al., 2019b). Autonomous machines are not humans, and they are not governed by emotions, and for that reason, users and developers should be able to identify the ethical principles on which machines base their decisions, rather than try to reason with them or justify the final result (Bogosian, 2017; Etzioni & Etzioni, 2017). Additionally, some researchers suggest that ethical machines should be put under a 'probationary period' during which all their decisions are examined and evaluated to ensure explainability (Winfield et al., 2019).

Accountability

One of the most common characteristics of responsible AI is accountability. Accountability is the ability to determine whether a decision was made in accordance with procedural and substantive standards and to hold someone responsible if those standards are not met. It is an important element of good public and private governance, so efforts about how to create accountable AI are significant (Juho, 2019). On the

contrary, poor efforts may result in regulation that not only fails to truly improve accountability but also stifles the many beneficial applications of AI systems. Day by day, machines are becoming more sophisticated and make people's lives more enjoyable while being autonomous with huge control over everything. In healthcare, business, e-government, information technology, or transport and communication, there are many examples of responsible AI agents who learn by example and act in compliance with other external factors while considering environmental, sensory, and actuation stochasticity (Breidbach & Maglio, 2020; Munoko et al., 2020).

Business management

Accountability mainly concerns the ascription of responsibilities in the business sector. Deep learning, which is also a feature of AI, is a machine learning technique that teaches computers to do what comes naturally to humans, to learn by example from raw information. Computers can complete human tasks with responsibility and with respect to ethical rules. The autonomy of intelligent agents enables them to generate the contractual offer without human intervention or knowledge and consequently to act more like initiators or intermediaries than messengers or instruments (Carter, 2019; Mendling et al., 2018). While some authors assert that software agents cannot be held responsible, others think that such agents might be held responsible once they arrive at a reliable degree of autonomy, intelligence, and sophistication. For example, Dahiyat (2010) mentions that a software agent for searching and roaming the Internet infringes others' rights (such as copyright or privacy right), performs illegal transactions, or operates without the user's authorization and sells rather than buys certain shares. Also, while gathering information, this agent corrupts a third party's database or causes the server to crash. Someone must be accountable for such mistakes and errors.

Transportation

Artificial Intelligence is changing the transport sector with autonomous cars, ships, or trains that make the traffic flow smoother and our lives easier. Such new autonomous vehicles inventions may help in reducing human errors and consequently traffic accidents. Nevertheless, new ethical questions arise regarding liability for the decisions taken by machines in place of humans (Hauer, 2019). A characteristic example is all autonomous driving vehicles (ADVs). Faulhaber et al. (2018) explained that humans, in situations like a car accident, base their decisions mainly on reflexes and instincts rather than deep thoughts. They act in utilitarian ways, trying to cause as little overall damage as possible. This will change with the introduction of ADVs, given that the car's decisions in all kinds of possible traffic scenarios will be programmed beforehand, including guidelines for unforeseen events and even highly unlikely scenarios. However, since there is always the possibility of an ADV causing an accident, the question arises: Who is responsible for this act; the driver, or the car manufacturer? The most effective way to ensure accountability is the assembly of rules from human behavior when they are going to come across this kind of dilemma that will be applicable to all ADVs. An ethical setting that will be implemented in their car to actually use it without a doubt or fear. These settings aim to establish an ethical decision-making framework for moral dilemmas in driving situations that can then serve as a foundation for an obligatory ethical setting to be implemented in ADVs.

Healthcare

AI-based healthcare applications already can reach or even exceed the performance of clinicians for specific tasks. These applications can be beneficial for global challenges, including shortages of clinicians to meet the demands of aging populations and the inequalities in access to health care in low-resource countries (Graham et al., 2019). Healthcare, however, is a complex, safety–critical domain in which technological failures can lead directly to patient harm (Banks, 2018). An example is the use of m-health technology in cardiology (Vervoort et al., 2020). This application is concerned primarily with telephone-based call and text messaging methods to promote and control medication adherence, lifestyle changes, and smoking cessation. But what happens if the application gives incorrect information to the patient? Who will be held accountable for this incident? Proper application of AI should aim to enhance positive social change, sustainability, and responsibility. Particularly in medicine, rules and regulations should be put in place to ensure responsibility and accountability of AI systems, their users, and their appropriate utilization. The most effective way to ensure accountability is explanation. This explanation will be given through the cooperation of AI systems and physicians. Namely, computational scientists may train AI using datasets to make meaningful assessments or predictions and the physicians will evaluate them in order to provide the final information to patients.

e-Government/public sector

AI has lots of benefits in governmental issues as well. AI applications are used for improving both the e-Government systems and citizens' interactions (Cath et al., 2017; Rahwan, 2017). However, there are challenges related to AI accountability in automated e-government applications and services (Marri et al., 2019). An example of the use of artificial intelligence in e-Government and e-services is an AI-enabled lending decision-making system in a bank.

If the system's algorithm does not work properly, then discriminated against customers may have access to the bank system. In this part, someone should be held responsible for all these decisions taken by machines to keep privacy at a high level. Here comes the law. Legal liability for AI systems could originate from either criminal or civil law. Civil liability can be further divided into contractual, tortious, and statutory liability. We take a comparative legal approach, either referring to genuinely transnational sources of law or highlighting common patterns in several jurisdictions' law (Erdélyi & Erdélyi, 2020).

Information technology

Information technology is about software/hardware functions, automated machines, robots, or cyber-related applications. Deep learning is a technique that teaches computers/ robots to do what comes naturally to humans, to learn through real examples. A characteristic example in this sector is the ability to complete automated and repetitive tasks; robots complete everyday tasks, handle commands and reduce the time consumption and the cost related to these activities. (Hoeschl et al., 2018). Autonomous AI agents' ethical aspects have been insufficiently researched until now, among others, based on the misconception that intelligent artifacts do essentially what they have been programmed to do, which is true only for very simple agents. With growing complexity and increasing autonomy, learning and adaptive abilities, ethical challenges are multiplying (Etzioni & Etzioni, 2017; McAleenan, 2020; Timmers, 2019; Tonkens, 2009). Proper AI applications should seek to improve social impact as well as sustainability and environmental responsibility. For this reason, AI systems should include the engineering ethics of designers, manufacturers, and maintenance services, as well as ethical attitudes of users and ethical aspects of the artifacts themselves, in order to focus more on the accountability of these AI agents (Dodig Crnkovic & Çürüklü, 2011).

Fairness

One of the ostensible advantages of using AI or machines in general to make decisions is that they may be impartial, objective, and free of the same biases that humans do, making them more “fair.” Recent research has revealed that AI systems can be biased as well.

Business management

A crucial aspect of trustworthiness is avoiding unfair bias in algorithmic decision-making. This is especially important for companies that use AI to improve decision-making efficiency and accuracy. For instance, Ability, Benevolence, Integrity, and Predictability (referred to as the ABI + model) are the four fundamental traits on which trustworthiness judgments are founded (Aitken et al., 2020). Predictability will support impressions of the trustee's ability, generosity, and integrity. Ability refers to the level that an entity can have the abilities and capabilities to carry out specific duties, suitable to the situation that they would be trusted. Benevolence is the level at which a trustee desire to do good to the trustor. Integrity is that the entity will follow a set of principles, which are suitable with the ethics of trustor.

Healthcare

Reliance on dogma and heuristics: Under the time constraints and uncertainty imposed by acute surgical disease and busy clinic schedules, surgeons frequently engage patients in high-stakes shared decision-making. These conditions encourage people to rely on dogma and heuristics, which can lead to bias, cognitive errors, and harm that could have been avoided. Surgical decision-making innovations may be able to address these issues. Surgeons may be supplemented or replaced by intelligent, autonomous machines. These technologies should be understood and guided by surgeons (Loftus et al., 2020).

e-Government/public sector

It is unlikely that an AI system would compare ethical theories in e-Government. Such a system would either develop a way to discern ethical options, (e.g. through machine learning), or its developers would choose a moral theory for the system and implement it before it was deployed. In this case, the developer must be impartial, which is both a requirement for a responsible AI and a challenge (Bjørgen et al., 2018).

Information technology

AI must avoid bias by conducting thorough and representative research to ensure that incorrect heuristics are not used to discriminate. An indirect way to ensure better data selection is to aim for a broader demographic of programmers, including minorities and women, rather than simply asking programmers to be "more sensitive" to bias. It is thought that issues will be more likely to be identified by encouraging a diversity of viewpoints (Turner, 2018).

Empathy

Empathy is a key part of our lives. It leads to unity and to society, social behavior and prosperity. It transcends culture. An AI system with empathic ability can generate more natural interactions while evaluating our moods or feelings. Based on the sort of information we want technology to

communicate, how we want individuals to react to it, or how we predict they will react to it, we will need to consider their present situation in some way; their emotional state will give us some clue as to how they're likely to respond. It is not as easy as acknowledging someone's tone of voice, posture, or facial expression—there is a lot more to it than that.

Transportation

Furthermore, empathy appears as a characteristic in Transportation & Energy industry. Maxmen (2018) describes in the paper what happens when a driver slams on the brakes to avoid colliding with a pedestrian who has crossed the road unlawfully. A moral decision must be made when the impending danger is being transferred from the pedestrian to the passengers in the vehicle. Self-driving vehicles will soon be forced to make ethical decisions on their own, taking into account some moral values influenced by some factors. For example, according to a survey of 2.3 million individuals worldwide, many of the moral concepts that affect driver's actions differ by nation. People in highly affluent countries with robust institutions, such as law enforcement, were less inclined to spare a pedestrian who unlawfully entered traffic. The findings indicate cultural subtleties that governments and self-driving car manufacturers must consider if the vehicles are to achieve widespread acceptability. In the Moral Machine poll, people in 130 nations with at least 100 respondents indicated a preference for sacrificing older lives to save younger ones (Maxmen, 2018).

Healthcare

Workplace injuries are a widespread and significant public health issue with significant economic repercussions across the world. Cheng et al. (2020) explain how to create a record for work-related injuries. It is about a stable and unified cloud infrastructure that includes a range of data storage, data analytics and machine learning management tools. It employs AI to do in-depth analysis using text-mining techniques to retrieve both dynamic and static data from work accident cases to perform unsupervised and supervised machine learning algorithms. When completely developed, this method can provide a more reliable forecast model for the cost of work accidents. The project consists of three phases. It is essential to mention the two stages. The first step involves identifying human variables, both in terms of facilitators and obstacles, for the return-to-work (RTW) process, by conducting face-to-face interviews and working groups with various stakeholders to gather perspectives about facilitators, barriers, and important RTW interventions for wounded employees. Phase two consists of developing a model of machine learning that uses artificial insight to carry out a detailed study (Cheng et al., 2020).

Information technology

Empathy is considered essential on Information Technology. A theoretical model of ethical decision-making for Autonomous Agents is explained by Cervantes et al. (2015). This concept has the intention to provide Autonomous Agents with appropriate frameworks for making ethical judgment-based decisions. More detailed, the model's design and execution cycle were influenced by current results in disciplines studying brain processes underpinning decision-making. The suggested model mimics the behavior of certain brain regions. Ethical Decision Making is implemented as a three-phase method using a computational model: estimation, implementation, and result evaluation. Each of these steps entails processing data produced by different cognitive functions, including awareness, organizing, and emotions. The author also indicates that emotional input must be incorporated into the decision-making mechanism according to neuroscientific facts. This emotional knowledge affects the rational decision-making mechanism of the agent and Ethical Decision Making. Emotions play a crucial role as people face moral dilemmas by helping them determine which ethical law should be respected.

Security

Artificial intelligence (AI) systems are becoming part of our lives and societies. The more decisions such systems make for us, the more we need to ensure that their decisions have a positive individual and societal impact (Bjørgen et al., 2018). An essential characteristic of AI systems is not only security but also cybersecurity. The term cyber-security refers to the protection of computers, servers, mobile devices, and other electronic systems from cyber-attacks in general (Timmers, 2019).

Business management

When it comes to security, the CIA triad of security, integrity, and availability must be mentioned. This triad is a method for designing an organization's security. Controlling access to information (confidentiality), avoiding data alteration or destruction (data integrity), and ensuring timely access to information(availability) are all essential. These three security criteria are the foundation of security and can be used in a wide range of AI applications, including manufacturing, logistics, building automation, and smart electrical grids (De Las Morenas et al., 2020). A building automation AI-based system, for instance, combines a large number of IoT devices, most of which are physical equipment, raising security concerns even further on the data used to train the AI models. Hacking into databases to steal information is a bad enough threat; however, physical devices can cause

more serious impacts for human beings in the vicinity, making security and privacy high priority issues for those systems. For that reason, building automation systems, due to their distributed design, provides a suitable testing environment for security problems (De Las Morenas et al., 2020).

Information technology

AI also has many applications in the sector of information technology. The top three fundamental human values, according to most reports, are protection and conformity. The basic individual and community criteria for a safe environment lead to security. Security, harmony, and prosperity of the community, relationships, and oneself are the motivational goals. The basic approach to trusted AI considers the system's trustworthiness, which includes technical and functional reliability as well as security (Kuleshov et al., 2020). AI systems that directly affect human security and life can pose a threat to humans. To ensure that artificial intelligence systems are secure, they must be standardized and tested before they are widely used. Adversarial examples assault AI systems in the same way that computer viruses exploit security software weaknesses. A key design aspect is making AI strong and shielding it from attack. Technological solutions have been created that programmers may utilize to discover and decrease the risk of machine learning systems. For instance, a technique that provides security in the information technology sector is automated repetitive tasks. More specifically, AI-assisted risk assessment approaches are being developed to decrease cyber risks and attacks (Turner, 2018). Additionally, Game AI is used to detect negative, abusive, and fraudulent behaviour in online communities and ban members who exhibit this kind of behaviour to ensure environmental security (Riedl & Zook, 2013).

Discussion

After reviewing the existing bibliography, we found that there is a lack of a single ethical framework that will set out our commitment to developing AI technology responsibly in different aspects of our daily life. One potential solution for ensuring that requirements are met in the practical usage of AI systems is to codify those ethical criteria and standards based on every industry sector's needs. Nevertheless, all these standards should have a common basis and direction. As an outcome of this research and in view of the aforementioned characteristics, we believe that the AI application should:

1. *Be explainable*. Transparency helps people to see the way and the purpose for which the AI applications or systems use specific data, whether these systems have been fully checked, as well as if they can understand why such decisions were made. AI applications or systems should provide consent opportunities, promote architectures with privacy protections, and provide adequate clarity and control over data usage.
2. *Be accountable*. AI technologies should be driven and controlled by humans as required.
3. *Be objective*. Bias can be reflected using AI systems and datasets. We must understand that separating the ethical from the unethical is not always easy and varies across cultures and communities. We should aim to avoid any relation with sensitive characteristics, such as region, gender, sexual orientation, etc.
4. *Be empathetic*. Empathy is a crucial component. While assessing our mood or emotions, an AI system with empathic capacity will produce more natural interactions.
5. *Be secure*. To avoid unintended consequences that pose a risk of danger, we should establish and implement good safety and security practices. We should design our AI systems to be sufficiently vigilant, and we should aim to improve them in line with current AI safety research best practices. We should test AI technologies in constrained environments and track their activity after implementation in appropriate cases.

Simultaneously, all of the characteristics mentioned above have a purpose: to resolve the issues we face on a daily basis while using AI-driven applications and services. Furthermore, we identified the following core challenges that could pinpoint possible research directions:

1. *Transparency*. AI is not without flaws, and this is because it is based on human actions. It is capable of making errors, as it does on occasion. We cannot tell if an outcome is the result of a mistake if AI is not transparent.
2. *Security*. Artificial intelligence (AI) technologies collect and use data in order to offer many services. We must ensure that individuals' personal data are only used for specific reasons (which should be expressly and specifically always specified and readily accessible to users—challenge related with transparency), and that they will never use to provoke or encourage harm to persons, and also that they will not breach generally recognized standards of international law and human rights.
3. *Privacy*. The privacy risk is closely linked to possible loss of control of personal information. If this happens, there is a risk of intrusion, processing and dissemination of information, either incorrect, which may lead to misinformation or manipulation that may cause targeting and possible harm to an individual or group of individuals or a society.

4. *Responsibility & Accountability*. According to our research, responsibility and accountability (Table 1) are two of the most important challenges in the industries we examined, for possible improper actions and decisions made through AI-based programs and applications. In our opinion, humans and not the machines or the programs should be considered responsible, as they are the ones who will "train" the machines or the programs to operate according to proper actions and practices and in addition, they are the ones who will judge the result. This challenge is strictly linked with the next one.
5. *Morality & bias*. Machine learning algorithms and AI systems learn from training data and make projections based on input data. Often the training data, input data, or both are morally or politically, or socially motivated. While there can be several advantages of designing a model that works on confidential data, it is vital to consider the possible effects of using unethical or biased data.

However, as our study does not include all industry sectors, the above proposals for ethical and responsible AI systems should not be considered a panacea for all cases of application of AI-driven technology.

Last but not least, every industry should aspire to develop innovations that address critical issues and assist people in their everyday lives. We believe AI and other advanced technologies have tremendous potential to motivate people, support current and future generations, and work for the common good. We understand that this is a complex and dynamic area, and we should approach its development with modesty, a dedication to both internal and external involvement, and a willingness to change our approach as we gain experience over time. If we embrace these truths, we will be able to take a big step against the prejudice that exists about the use of AI.

Conclusion

In this paper was we presented the challenges and required characteristics of AI systems in different industries. As we live in an era where we heavily rely on AI systems to perform a variety of tasks, there is a growing need to ensure that these systems run smoothly and without causing harm to humanity. We have identified possible challenges of the AI systems, as well as characteristics that an AI agent should have to be considered responsible. We concluded our paper by providing insights and solutions that could be useful to combat challenges that may arise.

Even though there is a long way to go, progress has been achieved. There will be more difficulties that we might need to overcome, but with the co-operation of organizations, developers, as well as everyday people, we hope that AI systems will continue to improve and expand in ways that better our lives and assist with everyday issues. We believe our findings shed some light on the situation around AI systems and start a conversation on various aspects.

Appendix

ID	Authors	ID	Authos
1	(Aitken et al., 2020)	36	(Kang et al., 2021)
2	(Ashrafian, 2014a)	37	(Kleindienst et al., 2017)
3	(Ashrafian, 2014b)	38	(Kuleshov et al., 2020)
4	(Balmer et al., 2020)	39	(Loftus et al., 2020)
5	(Banks, 2018)	40	(Lugano, 2017)
6	(Bartoletti, 2019)	41	(Luo, 2019)
7	(Bauer et al., 2017)	42	(Marri et al., 2019)
8	(Bendel, 2017)	43	(Martin, 2016)
9	(Bjørgen et al., 2018)	44	(Mason, 1986)
10	(Bogosian, 2017)	45	(Maxmen, 2018)
11	(Braun & Garriga, 2018)	46	(McAleenan, 2020)
12	(Breidbach & Maglio, 2020)	47	(McBride, 2016)
13	(Campbell et al., 2020)	48	(Mendling et al., 2018)
14	(Caner & Bhatti, 2020)	49	(Mittelstadt et al., 2016)
15	(Carter, 2019)	50	(Montes & Goertzel, 2019)
16	(Cath et al., 2017)	51	(Munoko et al., 2020)
17	(Cervantes et al., 2015)	52	(Pagallo, 2012)
18	(Chakraborty et al., 2020)	53	(Pedersen & Johansen, 2019)
19	(Cheng et al., 2020)	54	(Pitman et al., 2019)
20	(Cysneiros & do Prado Leite, 2020)	55	(Rahwan, 2017)
21	(Dahiyat, 2010)	56	(Rességuier & Rodrigues, 2020)
22	(David et al., 2019)	57	(Riedl & Zook, 2013)
23	(De Las Morenas et al., 2020)	58	(Roberts et al., 2020)

ID	Authors	ID	Authos
24	(Dignum, 2019)	59	(Russell & Norvig, 2002)
25	(Dodig Crnkovic & Çürüklü, 2011)	60	(Ten Teije et al., 2017)
26	(Erdélyi & Erdélyi, 2020)	61	(Timmers, 2019)
27	(Etzioni & Etzioni, 2017)	62	(Tonkens, 2009)
28	(Faulhaber et al., 2018)	63	(Turner, 2018)
29	(Ghandour & Woodford, 2019)	64	(Vaio et al., 2020)
30	(Gill, 2019)	65	(Vakkuri et al., 2020)
31	(Graham et al., 2019)	66	(Vakkuri et al., 2019b)
32	(Hauer, 2019)	67	(Vervoort et al., 2020)
33	(Hoeschl et al., 2018)	68	(Winfield et al., 2019)
34	(Howard & Borenstein, 2017)	69	(Yang et al., 2019)
35	(Juho, 2019)		

Author contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by MA, OK, CK, CK, KM, EM. The first draft of the manuscript was written by MA, OK, CK, CK, KM, EM and all authors commented on previous versions of the manuscript. Ioannis Konstantinidis devised the project, the main conceptual ideas, supervised the findings of this work and contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript. EK, CB, IM and VP contributed to the writing of the manuscript and on addressing the revisions required in the final version. All authors read and approved the final manuscript and provided critical feedback to shape the research, analysis and manuscript.

References

Aitken, M., Toreini, E., Carmichael, P., Coopamootoo, K., Elliott, K., & van Moorsel, A. (2020). Establishing a social licence for financial technology: Reflections on the role of the private sector in pursuing ethical data practices. *Big Data & Society*. https://doi.org/10.1177/2053951720908892

Ashrafian, H. (2014). AIonAI: A humanitarian law of artificial intelligence and robotics. *Science and Engineering Ethics, 21*(1), 29–40. https://doi.org/10.1007/S11948-013-9513-9

Ashrafian, H. (2014). Artificial intelligence and robot responsibilities: Innovating beyond rights. *Science and Engineering Ethics, 21*(2), 317–326. https://doi.org/10.1007/S11948-014-9541-0

Balmer, R. E., Levin, S. L., & Schmidt, S. (2020). Artificial intelligence applications in telecommunications and other network industries. *Telecommunications Policy, 44*(6), 101977. https://doi.org/10.1016/J.TELPOL.2020.101977

Banks, J. (2018). The human touch: Practical and ethical implications of putting AI and robotics to work for patients. *IEEE Pulse, 9*(3), 15–18. https://doi.org/10.1109/MPUL.2018.2814238

Bartoletti, I. (2019). AI in Healthcare: Ethical and privacy challenges. *Lecture notes in computer science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11526 LNAI*, 7–10. https://doi.org/10.1007/978-3-030-21642-9_2

Bauer, L., Bloom, C., Tan, J., & Ramjohn, J. (2017). Self-driving cars and data collection: Privacy perceptions of networked autonomous vehicles. *Thirteenth Symposium on Usable Privacy and Security ($\{$SOUPS$\}$ 2017)*, 357–375.

Bendel, O. (2017). LADYBIRD: The animal-friendly robot vacuum cleaner. *2017 AAAI Spring Symposium Series*. Retrieved September 21, 2021, from https://www.aaai.org/ocs/index.php/SSS/SSS17/paper/view/15277

Bjørgen, E. P., Madsen, S., Bjørknes, T. S., Heimsaeter, F. V, Håvik, R., Linderud, M., Longberg, P.-N., Dennis, L. A., & Slavkovik, M. (2018). Cake, Death, and Trolleys * Dilemmas as benchmarks of ethical decision-making. *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. https://doi.org/10.1145/3278721

Bogosian, K. (2017). Implementation of moral uncertainty in intelligent machines. *Minds and Machines, 27*(4), 591–608. https://doi.org/10.1007/S11023-017-9448-Z

Braun, A., & Garriga, G. (2018). Consumer journey analytics in the context of data privacy and ethics. *Digital Marketplaces Unleashed*. https://doi.org/10.1007/978-3-662-49275-8_59

Breidbach, C. F., & Maglio, P. (2020). Accountable algorithms? The ethical implications of data-driven business models. *Journal of Service Management, 31*(2), 163–185. https://doi.org/10.1108/JOSM-03-2019-0073

Campbell, C., Sands, S., Ferraro, C., Tsao, H. Y., & Mavrommatis, A. (2020). From data to action: How marketers can leverage AI. *Business Horizons, 63*(2), 227–243. https://doi.org/10.1016/J.BUSHOR.2019.12.002

Caner, S., & Bhatti, F. (2020). A conceptual framework on defining businesses strategy for artificial intelligence. *Contemporary Management Research, 16*(3), 175–206. https://doi.org/10.7903/CMR.19970

Carter, D. (2019). New technologies and new data sources: The business information survey 2019. *Business Information Review, 36*(3), 95–105. https://doi.org/10.1177/0266382119871848

Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2017). Artificial intelligence and the 'good society': the US, EU, and UK approach. *Science and Engineering Ethics, 24*(2), 505–528. https://doi.org/10.1007/S11948-017-9901-7

Cervantes, J.-A., Rodríguez, L.-F., López, S., Ramos, F., & Robles, F. (2015). Autonomous agents and ethical decision-making. *Cognitive Computation, 8*(2), 278–296. https://doi.org/10.1007/S12559-015-9362-8

Chakraborty, A., Bagavathi, R., & Tomer, U. (2020). A comprehensive decomposition towards the facets of quality in IoT. *Proceedings—International Conference on Smart Electronics and Communication, ICOSEC, 2020*, 759–764. https://doi.org/10.1109/ICOSEC49089.2020.9215428

Cheng, A. S. K., Ng, P. H. F., Sin, Z. P. T., Lai, S. H. S., & Law, S. W. (2020). Smart work injury management (SWIM) system: Artificial intelligence in work disability management. *Journal of Occupational Rehabilitation, 30*(3), 354–361. https://doi.org/10.1007/S10926-020-09886-Y

Cysneiros, L. M., & do Prado Leite, J. C. S. (2020). Non-functional requirements orienting the development of socially responsible software. *Lecture Notes in Business Information Processing*, *387 LNBIP*, 335–342. https://doi.org/10.1007/978-3-030-49418-6_23

Dahiyat, E. A. R. (2010). Intelligent agents and liability: is it a doctrinal problem or merely a problem of explanation? *Artificial Intelligence and Law, 18*(1), 103–121. https://doi.org/10.1007/S10506-010-9086-8

David, W., Pappalepore, P., Stefanova, A., & Sarbu, B. A. (2019). AI-powered lethal autonomous weapon systems in defence transformation. Impact and challenges. *Lecture notes in computer science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11995 LNCS*, 337–350. https://doi.org/10.1007/978-3-030-43890-6_27

Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer Nature.

DodigCrnkovic, G., & Çürüklü, B. (2011). Robots: ethical by design. *Ethics and Information Technology, 14*(1), 61–71. https://doi.org/10.1007/S10676-011-9278-2

Erdélyi, O. J., & Erdélyi, G. (2020). The AI liability puzzle and a fund-based work-around. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. https://doi.org/10.1145/3375627

Etzioni, A., & Etzioni, O. (2017). Incorporating ethics into artificial intelligence. *The Journal of Ethics, 21*(4), 403–418. https://doi.org/10.1007/S10892-017-9252-2

Faulhaber, A. K., Dittmer, A., Blind, F., Wächter, M. A., Timm, S., Sütfeld, L. R., Stephan, A., Pipa, G., & König, P. (2018). Human decisions in moral dilemmas are largely described by utilitarianism: Virtual car driving study provides guidelines for autonomous driving vehicles. *Science and Engineering Ethics, 25*(2), 399–418. https://doi.org/10.1007/S11948-018-0020-X

Floridi, L. (2013). *The ethics of information*. Oxford University Press.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.3831321

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4people—an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Ghandour, A., & Woodford, B. J. (2019). Ethical issues in artificial intelligence in UAE. *Proceedings—2019 International Arab Conference on Information Technology, ACIT 2019*, 262–266. https://doi.org/10.1109/ACIT47987.2019.8990997

Gill, A. S. (2019). Artificial intelligence and international security: The long view. *Ethics & International Affairs, 33*(2), 169–179. https://doi.org/10.1017/S0892679419000145

Google AI. (2021). *Responsible AI practices*. Google AI. Retrieved September 21, 2021, from https://ai.google/responsibilities/responsible-ai-practices/

Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H.-C., & Jeste, D. V. (2019). Artificial intelligence for mental health and mental illnesses: An overview. *Current Psychiatry Reports, 21*(11), 1–18. https://doi.org/10.1007/S11920-019-1094-0

Hauer, T. (2019). Society caught in a labyrinth: of algorithms: Disputes, promises, and limitations of the new order of things. *Society, 56*(3), 222–230. https://doi.org/10.1007/S12115-019-00358-5

Healey, J. (2020). Principled artificial intelligence: Mapping consensus in ethical and rightsbased approaches to principles for AI. *Artificial Intelligence and Financial Services*. https://doi.org/10.1007/978-3-030-29761-9_6

AI HLEG. (2019). Ethics guidelines for trustworthy AI. In: *European Commission*. Retrieved September 21, 2021, from https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Hoeschl, M. B., Bueno, T. C. D., & Hoeschl, H. C. (2018). Fourth industrial revolution and the future of engineering: Could robots replace human jobs? How ethical recommendations can help engineers rule on artificial intelligence. *Proceedings—2017 7th World Engineering Education Forum*, 21–26. https://doi.org/10.1109/WEEF.2017.8466973

Howard, A., & Borenstein, J. (2017). The ugly truth about ourselves and our robot creations: The problem of bias and social inequity. *Science and Engineering Ethics, 24*(5), 1521–1536. https://doi.org/10.1007/S11948-017-9975-2

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

Juho, V. (2019). Ethics of AI technologies and organizational roles: Who is accountable for the ethical conduct? *Conference on Technology Ethics*.

Kang, J., Thompson, R. F., Aneja, S., Lehman, C., Trister, A., Zou, J., Obcemea, C., & El Naqa, I. (2021). National cancer institute workshop on artificial intelligence in radiation oncology: Training the next generation. *Practical Radiation Oncology, 11*(1), 74–83. https://doi.org/10.1016/J.PRRO.2020.06.001

Kitchenham, B., & Charters, S. M. (2007). Guidelines for performing Systematic Literature Reviews in SoftwareEngineering. EBSE. Retrieved September 21, 2021, fromhttps://www.elsevier.com/data/promis_misc/525444systematicreviewsguide.pdf

Kleindienst, D., Nüske, N., Rau, D., & Schmied, F. (2017). Beyond mere compliance—delighting customers by implementing data privacy measures? *Wirtschaftsinformatik 2017 Proceedings*. Reterived 23 January 2017, from https://aisel.aisnet.org/wi2017/track08/paper/2

Kuleshov, A., Ignatiev, A., Abramova, A., & Marshalko, G. (2020). Addressing AI ethics through codification. *Proceedings—2020 International Conference Engineering Technologies and Computer Science, EnT 2020*, 24–30. https://doi.org/10.1109/ENT48576.2020.00011

De Las Morenas, J., Da Silva, C. M., Funchal, G. S., Melo, V., Vallim, M., & Leitao, P. (2020). Security experiences in IoT based applications for building and factory automation. *Proceedings of the IEEE International Conference on Industrial Technology*, *2020-February*, 322–327. https://doi.org/10.1109/ICIT45562.2020.9067229

Loftus, T. J., Filiberto, A. C., Balch, J., Ayzengart, A. L., Tighe, P. J., Rashidi, P., Bihorac, A., & Upchurch, G. R. (2020). Intelligent, autonomous machines in surgery. *Journal of Surgical Research, 253*, 92–99. https://doi.org/10.1016/J.JSS.2020.03.046

Lugano, G. (2017, July 7). Virtual assistants and self-driving cars. *Proceedings of 2017 15th International Conference on ITS Telecommunications, ITST 2017*. https://doi.org/10.1109/ITST.2017.7972192

Luo, S. (2019). Research on the change of educational management in the era of artificial intelligence. *Proceedings—2019 12th International Conference on Intelligent Computation Technology and Automation, ICICTA 2019*, 442–445. https://doi.org/10.1109/ICICTA49267.2019.00101

Marri, A. Al, Albloosh, F., Moussa, S., & Elmessiry, H. (2019). Study on the impact of artificial intelligence on government E-service in Dubai. *Proceeding of 2019 International Conference on Digitization: Landscaping Artificial Intelligence, ICD 2019*, 153–159. https://doi.org/10.1109/ICD47981.2019.9105866

Martin, D. (2016). Who should decide how machines make morally laden decisions? *Science and Engineering Ethics 2016 23:4*, *23*(4), 951–967. https://doi.org/10.1007/S11948-016-9833-7

Mason, R. O. (1986). Four ethical issues of the information age. *MIS Quarterly: Management Information Systems, 10*(1), 5–12. https://doi.org/10.2307/248873

Maxmen, A. (2018). A moral map for AI cars. *Nature, 562*(7728), 469–470.

McAleenan, P. (2020). Moral responsibility and action in the use of artificial intelligence in construction. *Proceedings of the*

Institution of Civil Engineers - Management, Procurement and Law. *173*(4): 166–174. https://doi.org/10.1680/JMAPL.19.00056

McBride, N. (2016). The ethics of driverless cars. *ACM SIGCAS Computers and Society, 45*(3), 179–184. https://doi.org/10.1145/2874239.2874265

Mendling, J., Decker, G., Hull, R., Reijers, H., & Weber, I. (2018). How do machine learning, robotic process automation, and blockchains affect the human factor in business process management? *Communications of the Association for Information Systems, 43*(1), 19. https://doi.org/10.17705/1CAIS.04319

Microsoft AI. (2021). *Responsible AI principles*. Microsoft AI. https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3Aprimaryr6

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data and Society*. https://doi.org/10.1177/2053951716679679

Montes, G. A., & Goertzel, B. (2019). Distributed, decentralized, and democratized artificial intelligence. *Technological Forecasting and Social Change, 141*, 354–358. https://doi.org/10.1016/J.TECHFORE.2018.11.010

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (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. https://doi.org/10.1007/s11948-019-00165-5

MSCI. (2020). *Global industry classification standard*. MSCI Inc. Retrieved September 21, 2021, from https://www.msci.com/documents/1296102/11185224/GICS+Methodology+2020.pdf/9caadd09-790d-3d60-455b-2a1ed5d1e48c?t=1578405935658

Munoko, I., Brown-Liburd, H. L., & Vasarhelyi, M. (2020). The ethical implications of using artificial intelligence in auditing. *Journal of Business Ethics, 167*(2), 209–234. https://doi.org/10.1007/S10551-019-04407-1

Pagallo, U. (2012). *On the principle of privacy by design and its limits: Technology, ethics and the rule of law*. European Data Protection In Good Health. https://doi.org/10.1007/978-94-007-2903-2_16

Pedersen, T., & Johansen, C. (2019). Behavioural artificial intelligence: An agenda for systematic empirical studies of artificial inference. *AI & SOCIETY, 35*(3), 519–532. https://doi.org/10.1007/S00146-019-00928-5

Pitman, H., Colling, R., Oien, K., Rajpoot, N., Macklin, P., Snead, D., Sackville, T., & Verrill, C. (2019). Artificial intelligence in digital pathology: A roadmap to routine use in clinical practice. *The Journal of Pathology, 249*(2), 143–150. https://doi.org/10.1002/PATH.5310

PwC. (2019). *A practical guide to responsible artificial intelligence*. https://www.pwc.com/gx/en/issues/data-and-analytics/artificial-intelligence/what-is-responsible-ai/responsible-ai-practical-guide.pdf

Rahwan, I. (2017). Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology, 20*(1), 5–14. https://doi.org/10.1007/S10676-017-9430-8

Rességuier, A., & Rodrigues, R. (2020). AI ethics should not remain toothless! *A Call to Bring Back the Teeth of Ethics*. https://doi.org/10.1177/2053951720942541

Riedl, M. O., & Zook, A. (2013). AI for game production. *IEEE Conference on Computatonal Intelligence and Games, CIG*. https://doi.org/10.1109/CIG.2013.6633663

Roberts, H., Cowls, J., Morley, J., Taddeo, M., Wang, V., & Floridi, L. (2020). The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation. *AI & SOCIETY, 36*(1), 59–77. https://doi.org/10.1007/S00146-020-00992-2

Russell, S., & Norvig, P. (2002). *Artificial intelligence: A modern approach*. Prentice Hall.

Ryan, M., & Stahl, B. C. (2021). Artificial intelligence ethics guidelines for developers and users: Clarifying their content and normative implications. *Journal of Information, Communication and Ethics in Society, 19*(1), 61–86. https://doi.org/10.1108/JICES-12-2019-0138

Ten Teije, A., Popow, C., Holmes, J. H., & Sacchi, L. (2017). *Artificial Intelligence in Medicine: 16th Conference on Artificial Intelligence in Medicine, AIME 2017, Vienna, Austria, June 21–24, 2017, Proceedings* (Vol. 10259). Springer.

Timmers, P. (2019). Ethics of AI and cybersecurity when sovereignty is at stake. *Minds and Machines, 29*(4), 635–645. https://doi.org/10.1007/S11023-019-09508-4

Tonkens, R. (2009). A challenge for machine ethics. *Minds and Machines, 19*(3), 421–438. https://doi.org/10.1007/S11023-009-9159-1

Turner, J. (2018). *Robot rules: Regulating artificial intelligence*. Springer International Publishing. https://doi.org/10.1007/978-3-319-96235-1

Vaio, A. . Di., Boccia, F., Landriani, L., & Palladino, R. (2020). Artificial intelligence in the agri-food system: Rethinking sustainable business models in the COVID-19 scenario. *Sustainability, 12*(12), 4851. https://doi.org/10.3390/SU12124851

Vakkuri, V., Kemell, K. K., & Abrahamsson, P. (2019b). Ethically aligned design: An empirical evaluation of the resolved-strategy in software and systems development context. *Proceedings—45th Euromicro Conference on Software Engineering and Advanced Applications, SEAA 2019*, 46–50. https://doi.org/10.1109/SEAA.2019.00015

Vakkuri, V., Kemell, K.-K., & Abrahamsson, P. (2019). AI ethics in industry: a research framework. *ArXiv Preprint*.https://arXiv.org.1910.12695

Vakkuri, V., Kemell, K. K., Kultanen, J., & Abrahamsson, P. (2020). The current state of industrial practice in artificial intelligence ethics. *IEEE Software, 37*(4), 50–57. https://doi.org/10.1109/MS.2020.2985621

Vervoort, D., Marvel, F. A., Isakadze, N., Kpodonu, J., & Martin, S. S. (2020). Digital cardiology: Opportunities for disease prevention. *Current Cardiovascular Risk Reports, 14*(8), 1–7. https://doi.org/10.1007/S12170-020-00644-6

Winfield, A. F., Michael, K., Pitt, J., & Evers, V. (2019). Machine ethics: The design and governance of ethical ai and autonomous systems. *Proceedings of the IEEE, 107*(3), 509–517. https://doi.org/10.1109/JPROC.2019.2900622

Yang, L., Elisa, N., & Eliot, N. (2019). Privacy and security aspects of E-government in smart cities. *Smart Cities Cybersecurity and Privacy*. https://doi.org/10.1016/B978-0-12-815032-0.00007-X

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).
Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.
These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.
We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.
While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.
These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.
To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.
Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.
If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com