IMPACTO CIENTÍFICO EDITORA

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN MODERN SOCIETY

INTELIGÊNCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA NA SOCIEDADE MODERNA

INTELIGENCIA ARTIFICIAL Y APRENDIZAJE AUTOMÁTICO EN LA SOCIEDAD MODERNA

10.56238/edimpacto2025.034-001

Hermenegildo Woropo Albino Paiva[1]

ABSTRACT

In recent decades, Artificial Intelligence (AI) and Machine Learning (ML) have gone from being mere computer science concepts to becoming fundamental in today's society. They are present in artificial assistants, healthcare diagnostics, security systems, and industrial automation processes. However, this rapid expansion of AI also raises new ethical, social, and regulatory challenges. With algorithms taking on roles that were previously exclusive to humans, concerns arise about data protection, accountability for automated decisions, biases in systems, and the potential replacement of the human workforce. Thus, this paper aims to analyze the use of Artificial Intelligence and Machine Learning in today's society, focusing on the areas of healthcare, security, and automation. It is concluded that the ability to quickly analyze large amounts of data and identify useful patterns for decision-making has enabled progress that previously seemed impossible, improving medical diagnostics, increasing real-time surveillance, and automating industrial processes more efficiently. However, these advantages do not come without ethical, political, and legal issues. The potential of AI, although promising, also presents several challenges that require careful analysis and attentive regulation.

Keywords: Artificial Intelligence. Machine Learning. Ethics in AI. Automation. AI in Healthcare. AI in Public Safety.

RESUMO

Nas últimas décadas, a Inteligência Artificial (IA) e o Aprendizado de Máquina (AM) deixaram de ser meramente conceitos da área de ciência da computação, tornando-se fundamentais na sociedade atual. Elas estão presentes em assistentes artificiais, diagnósticos na área da saúde, sistemas de segurança e processos de automação industrial. Entretanto, essa rápida expansão da IA também levanta novos desafios éticos, sociais e regulatórios. Com algoritmos assumindo papéis antes únicos dos humanos, aparecem preocupações sobre a proteção dos dados, a responsabilidade por decisões automatizadas, os preconceitos nos sistemas e a potencial substituição da força de trabalho humana. Assim, este trabalho tem como objetivo analisar a utilização da Inteligência Artificial e do Aprendizado de Máquina na sociedade atual, com foco nas áreas de saúde, segurança e automação. Conclui-se que, a

[1]Master's degree Computer Science. Federal University of Lavras (UFLA).
E-mail: hermen.velho02@gmail.com Orcid: https://orcid.org/0009−0002−5172−2487

habilidade de analisar grandes quantidades de dados rapidamente e identificar padrões úteis para a tomada de decisões tem permitido progresso que antes parecia impossível, aprimorando diagnósticos médicos, aumentando a vigilância em tempo real e automatizando processos industriais de maneira mais eficiente. No entanto, essas vantagens não vêm sem questões éticas, políticas e legais. O potencial da IA, apesar de promissor, também apresenta vários desafios que demandam análise cuidadosa e regulação atenta.

Palavras-chave: Inteligência Artificial. Aprendizado de Máquina. Ética em IA. Automação. IA na Saúde. IA na Segurança Pública.

RESUMEN

En las últimas décadas, la Inteligencia Artificial (IA) y el Aprendizaje Automático (ML) han dejado de ser meros conceptos informáticos y se han convertido en fundamentales en la sociedad actual. Están presentes en asistentes artificiales, diagnósticos sanitarios, sistemas de seguridad y procesos de automatización industrial. Sin embargo, esta rápida expansión de la IA también plantea nuevos desafíos éticos, sociales y regulatorios. A medida que los algoritmos asumen funciones que antes estaban reservadas a los humanos, surgen inquietudes sobre la protección de datos, la responsabilidad por las decisiones automatizadas, el sesgo en los sistemas y el posible reemplazo de la fuerza laboral humana. Así, este trabajo pretende analizar el uso de la Inteligencia Artificial y el Machine Learning en la sociedad actual, centrándose en las áreas de salud, seguridad y automatización. Se concluye que la capacidad de analizar rápidamente grandes cantidades de datos e identificar patrones útiles para la toma de decisiones ha permitido avances que antes parecían imposibles, mejorando los diagnósticos médicos, aumentando la vigilancia en tiempo real y automatizando los procesos industriales de manera más eficiente. Sin embargo, estas ventajas no están exentas de problemas éticos, políticos y jurídicos. El potencial de la IA, aunque prometedor, también presenta varios desafíos que requieren un análisis cuidadoso y una regulación atenta.

Palabras clave: Inteligencia Artificial. Aprendizaje Automático. Ética en IA. Automatización. IA en La Atención Sanitaria. IA en Seguridad Pública.

1 INTRODUCTION

In recent decades, Artificial Intelligence (AI) and Machine Learning (ML) are no longer merely computer science concepts, becoming fundamental in today's society. They are present in artificial assistants, healthcare diagnostics, safety systems, and industrial automation processes, causing a significant change in the way people interact with technology, make choices, and perform complex activities. The development of neural networks, natural language processing, and predictive analytics is revolutionizing entire industries, bringing promise for greater efficiency, personalization, and agility.

However, this rapid expansion of AI also raises new ethical, social, and regulatory challenges. With algorithms taking on roles once unique to humans, concerns arise about data protection, accountability for automated decisions, biases in systems, and the potential replacement of the human workforce. Therefore, it is essential to understand the different dimensions of the use of artificial intelligence today, in order to ensure that its progress respects the principles of justice, security, and equity.

The choice to examine the application of Artificial Intelligence and Machine Learning in contemporary society is justified by the increasingly central role that these technologies play in various areas of life. The impact already evident in sectors such as health, public safety, and automation demonstrates that intelligent systems are not only tools, but transformers of social realities. Thinking critically about these changes is key to ensuring that the benefits of AI are widely shared and that its risks are controlled through ethical and responsible governance.

Despite technological innovations, society still has a lack of clear guidelines to address the ethical, social, and legal problems that arise with the use of AI. Many sectors implement intelligent systems without fully taking into account their repercussions on human rights, equity and transparency. In this scenario, the following concern arises: what are the main uses, challenges, and ethical implications linked to the use of artificial intelligence and machine learning in the areas of health, safety, and automation, and how can these issues be addressed in a critical and responsible manner?

Thus, this work has the general objective of analyzing the use of Artificial Intelligence and Machine Learning in today's society, focusing on the areas of health, safety and automation. Specifically, it aims to identify the main applications of these technologies, debate the challenges faced by their users and developers, and reflect on the ethical dilemmas that arise from their large-scale adoption.

The type of research used is qualitative, centered on a literature review. According to Gil (2008), the bibliographic review is carried out using material that has already been published, including books, academic articles, digital content and technical reports.

2 DEVELOPMENT

2.1 FUNDAMENTALS OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

According to Ludermir (2021), artificial intelligence is an area of computing dedicated to the creation of systems that can emulate human intelligence functions, such as reasoning, learning, perception, decision-making, and understanding of natural language. Although this idea emerged in the 1950s with figures such as Alan Turing and John McCarthy, it is only in recent times that advances in computing and the large amount of data available have made it possible for AI to become one of the most impactful innovations in the digital world. AI can be divided into two main categories: weak AI, which focuses on specific tasks (such as chatbots or recommender systems), and strong AI, which is still in the testing phase, aiming to replicate human-like general and autonomous intelligence.

Machine learning, or Machine Learning (ML), is a subdivision of AI that is dedicated to techniques and algorithms that enable systems to trigger learning from data, modifying their performance without the need for human intervention, as pointed out by Bigonha (2025). Instead of being programmed with fixed guidelines, these systems receive large amounts of data and, from it, recognize patterns and make predictions or decisions. This learning can occur in a supervised manner, where the system is trained with labeled examples (such as x-rays identified as healthy or sick), or unsupervised, where the system tries to discover groupings and shapes within the data without clear guidance. There is also reinforcement learning, in which the machine acquires knowledge through trial and error, based on rewards and punishments.

It is relevant to highlight the difference between AI and ML, although both complement each other: while AI seeks to create machines with intelligence, ML serves as one of the most efficient ways to achieve this goal. Other areas within AI include natural language processing, computer vision, expert systems, and autonomous robotics. Progress in these fields is mainly influenced by data quality, computational processing capacity, and algorithms that are becoming increasingly advanced, which transforms AI into a science that interacts with areas such as statistics, mathematics, linguistics, neuroscience, and even philosophy (LUDERMIR, 2021).

According to Gabriel Filho (2023), the main strategies used in Artificial Intelligence are based on mathematical and computational models designed to replicate human reasoning

and learning. Among the famous models, we find artificial neural networks, which simplistically mimic the functioning of neurons in the human brain. Deep neural networks, which use multiple layers of processing, have the ability to identify complex patterns in images, sounds, and text in an extremely precise manner. An example of this technology is facial recognition systems, which examine thousands of features in an image to recognize a person clearly.

Another very common approach for Bigonha (2025) is the decision tree, which organizes data in the format of flow graphs, making it simpler to analyze the choices made by the algorithm. This approach excels in industries such as finance, healthcare, and law, where it is essential that the reason behind each decision is understandable. Clustering algorithms, such as k-means, along with size reduction techniques such as PCA, are used to identify hidden patterns in large data sets, and are applicable in areas such as marketing, biology, and information security. In addition, logistic regression methods, support vector machines, and Bayesian models are also part of the Artificial Intelligence toolkit, each offering benefits in particular contexts.

Reinforcement learning has gained prominence recently, especially in contexts such as games, such as DeepMind's AlphaGo, and in autonomous control systems, such as vehicles and industrial robots. In this type of approach, the intelligent agent engages with the environment and learns to make decisions based on continuous feedback, both positive and negative, until its tactics are perfected. This learning model holds promise for situations that require flexible and adaptable decision-making, and is increasingly being adopted in robotics, logistics, and personalized recommendation systems (SCHUSSLER et al. 2018).

In addition to the techniques themselves, Gabriel Filho (2023) points out that it is crucial to highlight the importance of programming tools and libraries, such as TensorFlow, PyTorch, Scikit-learn, and Keras, which have made the construction of Artificial Intelligence models more accessible. These platforms provide ready-made modules for various tasks, such as classification, regression, natural language processing, and computer vision, reducing the technical difficulty for the application of AI in various fields of knowledge. With these resources, academic institutions, startups, and even government agencies have been able to create intelligent solutions that meet their demands.

Although Artificial Intelligence and Machine Learning are often considered equivalent, Bazeia (2023) mentions that it is important to understand their distinctions and how they relate to other emerging technologies, such as Big Data, Internet of Things, Cloud Computing, and Blockchain. AI fundamentally relies on large volumes of data to improve its algorithms, while IoT facilitates the continuous generation of this information through sensors in mobile

devices, vehicles, and industrial equipment. In turn, cloud computing offers the necessary, high-performance and scalable infrastructure to process and store the data that is used in the training of models.

For Arão (2024), another technology that is increasingly integrating with artificial intelligence is Blockchain, particularly in relation to security and tracking of decisions made by algorithms. The fusion of AI with Blockchain has been applied to develop more secure systems that can be audited, being able to document decisions in real time and ensure the integrity of automatic processes. This union proves especially valuable in areas such as healthcare and logistics, where clarity and permanence of information are essential.

In addition, Morais and Branco (2023) point out that AI and Robotics complement each other, where artificial intelligence offers the "brain" and algorithms for decision-making, while robotics operates in physical space with sensors, actuators, and control systems. Examples of this type of integration include surgical robots, drones for delivery, and vehicles that operate autonomously. Often, AI not only makes it possible to perform tasks, but also to adapt them to the context, through continuous learning based on the surrounding conditions.

Finally, understanding these differences is key to avoiding conceptual misunderstandings and to encouraging a critical and well-informed incorporation of digital technologies. According to Arão (2024), AI does not work alone, but in collaboration with other technological innovations, forming an interconnected ecosystem that transforms economic structures, labor relations, modes of consumption, and even the way people perceive reality. By understanding the principles of AI and Machine Learning, the essential foundation is laid to examine their applications more accurately, identify their limitations, and reflect on the consequences of their use without proper regulations.

2.2 APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN STRATEGIC SECTORS

According to Bernasiuk and Sarlet (2024), the health sector is one of the areas that benefits the most from the use of Artificial Intelligence, especially with regard to early diagnosis and adaptation of treatments. AI-based tools have been shown to be highly effective at detecting patterns in imaging tests, such as X-rays, CT scans, and MRIs, often with accuracy equal to or even better than that of doctors. These systems are fed with a large volume of scans and can notice subtle changes that may go unnoticed by a professional, helping in a crucial way in the diagnosis of conditions such as cancer, pneumonia, heart disease, and neurological disorders.

In addition to image analysis, Amaro Jr et al. (2024) cite that AI has also been widely applied in predictions and tailored medicine. By combining clinical, genetic, and behavioral

data, these intelligent systems are able to predict a patient's likelihood of developing certain conditions, indicate more appropriate medications, and suggest treatments based on similar patient profiles. This represents a considerable advance towards personalized medicine, in which treatment is adjusted according to the individual characteristics of each patient, which tends to increase success rates and minimize unwanted effects.

Hospital administration is another sector that is seriously influenced by AI. For Leite et al. (2024), through optimization algorithms and machine learning, it is possible to anticipate the need for beds, monitor the stock of medicines, schedule surgeries more effectively, and even predict disease outbreaks based on epidemiological information. During pandemics, such as COVID-19, AI-powered systems were employed to monitor the spread of the virus, estimate the contagion curve, and inform public health actions, proving their value in public health as well.

Additionally, chatbots and virtual assistants have been integrated into health services to provide basic information, initial screenings, and psychological support. While these systems still have limits, they mark progress in expanding access to healthcare, especially in remote areas or where there is a shortage of professionals. However, these advances also bring to light ethical and legal issues, such as the security of patient data and liability in cases of diagnostic errors, requiring careful regulation in the use of these technologies (AMARO JR et al. 2024).

Still, it is essential to emphasize that the use of artificial intelligence in healthcare should serve as a support for human intervention, and not as a substitute. According to Bernasiuk and Sarlet (2024), medical decision-making encompasses not only technical analysis, but also empathy, ethical discernment, and communication, aspects that systems are not yet able to reproduce in depth. Thus, the greatest value of artificial intelligence in healthcare lies in expanding the effectiveness of professionals, making them more informed and efficient, without removing the human element that is crucial in care.

In the field of public security, Pinheiro (2024) mentions that Artificial Intelligence has been used as a resource to prevent and combat crime. In several cities around the globe, intelligent monitoring systems, based on computer vision and facial recognition, have been implemented to identify suspects in real time, observe anomalous behavior in public spaces, and supervise risk areas. These systems use vast databases and integrated cameras, making it possible to intersect data and issue automatic alerts to the competent authorities.

Another significant area of AI in this field, according to Nagata (2024), is revealed in predictive analytics systems, which evaluate crime records, socioeconomic information, and behavior patterns to estimate the probability of criminal occurrences in specific regions. This

tool has been applied to guide police patrols, distribute resources strategically and even predict possible episodes of urban violence. Despite its promising potential, this use raises important ethical questions, especially about algorithmic bias and the risk of perpetuating social stereotypes or discriminatory practices.

In addition, AI has played an important role in the area of criminal investigations. Big data analysis and text mining programs can process large amounts of information collected from cell phones, social networks, security cameras, and official documents to uncover connections between suspects, map criminal organizations, and speed up the resolution of difficult cases. In addition, natural language processing algorithms can assist in the interpretation of police statements and reports, contributing to the identification of inconsistencies or important patterns (PINHEIRO, 2024).

In prisons, Vieira and Barbuda (2024) point out that AI has been applied to observe the behavior of inmates, prevent disagreements, and improve resource management. Smart sensors and video analysis systems are able to identify suspicious movements, crowds or risk scenarios, providing a faster response from security teams. However, as in other sectors, the intense use of monitoring technologies calls for a balance between public safety and respect for fundamental rights, including privacy and the presumption of innocence.

In the face of these innovations, it is essential that public policies involving AI-based security are formulated with clarity, human oversight, and social control. For Nagata (2024), the application of highly invasive technologies without proper democratic regulation can result in abuses of authority and accentuate social inequalities. Therefore, it is vital to regulate these activities, ensuring that technology benefits the citizenry, and not arbitrary repression or institutionalized prejudice.

According to Oliveira et al. (2024), the industry stood out as one of the leaders in the use of Artificial Intelligence as a driver of the new era known as Industry 4.0, which stands out for the combination of cyber-physical systems, internet of things, and autonomous decisions. In this new production scenario, AI is applied to simplify processes, anticipate machine breakdowns, control inventories, and adapt production immediately, based on requirements and operating conditions. This new model promotes higher efficiency, resource savings and more agile adaptations in production chains.

One of the most notable uses of AI in the industry according to Gonçalves et al. (2023), is in predictive maintenance, which uses sensors and algorithms to constantly monitor the operation of machines and predict failures before they happen. This prevents unexpected interruptions, reduces corrective maintenance expenses, and extends the durability of

equipment. Systems also have the ability to learn from past health data, continuously and autonomously adjusting their predictive models.

Another promising area for Estrela et al. (2024) is the use of collaborative robots, or cobots, which work together with human operators in repetitive, dangerous or precise activities. Equipped with sensors, computer vision, and AI algorithms, these robots are able to adapt to their environment and respond safely to human commands. This not only increases productivity but also improves working conditions, reducing accidents and operator fatigue.

Logistics has also been significantly transformed by the use of AI, through intelligent routing systems, demand forecasting, and automated inventory management. Large retail chains and transportation companies apply algorithms to determine efficient routes, anticipate periods of high demand, and automatically adjust distribution centers. This results in a supply chain that is more effective, cost-effective, and less vulnerable to human error (GONÇALVES et al. 2023).

However, these innovations bring with them the challenge of replacing traditional work, especially in operational and administrative roles. For Oliveira et al. (2024), intelligent automation can lead to the loss of millions of jobs, requiring an immediate requalification of professionals, in addition to public policies that ensure a transition to new job opportunities. As in other sectors, the responsible use of AI in industry requires a balance between technological innovation and social inclusion, ensuring that the benefits of automation are widely distributed.

2.3 CHALLENGES AND RISKS ASSOCIATED WITH THE USE OF ARTIFICIAL INTELLIGENCE

Despite remarkable progress, Artificial Intelligence systems are still far from infallible. According to Bernasiuk and Sarlet (2024), one of the most significant technical risks associated with the use of AI is the possibility of errors or incorrect decisions, which can arise because of code flaws, training with biased data, or limitations of the models in facing unforeseen situations. Such failures can have severe repercussions, especially in sensitive areas such as health, aviation, justice and transportation. For example, in automated medical diagnoses, an error in classification can delay the treatment of serious diseases; In self-driving cars, an error in reading the sensors can lead to lethal accidents. In this way, the reliability of AI is based on the solidity of the algorithms and the quality of the data used.

Another crucial point is the challenge of the "black box", common in deep learning systems, which prevents the understanding of algorithms' decisions. This lack of clarity not

only diminishes users' trust, but also restricts the ability to audit and correct errors when they arise. According to Amaro Jr et al. (2024), in fields such as law and medicine, the justification of decisions is essential. Thus, the lack of explanations raises serious ethical and legal concerns regarding the delegation of complex decisions to automated machines. The need for algorithmic explanations and the emergence of interpretable AI (XAI) techniques appear as solutions to this technical problem.

Additionally, Barroso (2024) mentions that many AI systems are extremely vulnerable to the environment in which they operate. Subtle changes in the input data—such as a slightly modified image or incomplete information—can result in unexpected or outright incorrect responses. This highlights the structural fragility of these systems in contexts that require flexibility and common sense, characteristics that are still quite human. Therefore, reducing technical risks depends on continuous validation of systems, human oversight, and the development of strategies that allow algorithms to deal with uncertainties more effectively.

The massive adoption of artificial intelligence solutions causes significant changes in the labor market, fueling one of the main fears of modern society: unemployment generated by technology. According to Carvalho (2021), repetitive and standardized work, especially in the operational, administrative, and customer service areas, is gradually being replaced by automated systems, which offer greater speed, accuracy, and economy. This directly affects the employment opportunities of millions of workers, especially those with low education or limited access to technological training. As a result, social and economic inequalities tend to intensify, especially in developing nations.

In addition to the issue of unemployment, the dissemination of artificial intelligence highlights the problem of digital exclusion according to Peixoto and Paiva (2024). Access to technological tools, digital skills training, and opportunities for reskilling do not occur fairly across different regions, social classes, and ages. The most vulnerable communities, such as rural dwellers, the elderly, and people in poverty, are often excluded from the innovation process, which creates additional obstacles to inclusion and citizen participation. The centralization of technical knowledge and computing power in the hands of a small number of economic groups and developed countries contributes to widening the global technological gap.

Addressing these problems requires the creation of public policies that prioritize digital education, technological inclusion, and professional retraining. According to Barroso (2024), governments, companies, and educational institutions must collaborate to ensure that the benefits of artificial intelligence are accessible to more than just a technical and economic elite. The transition to an increasingly automated economy must be accompanied by

investments in training, social safety nets, and mechanisms that redistribute opportunities. If this does not occur, artificial intelligence, instead of fostering collective progress, can reinforce scenarios of exclusion and social inequality that are even more pronounced.

For Aranha and Costa (2023), one of the most important and discussed points about the use of Artificial Intelligence is algorithmic biases, that is, the tendency of systems to maintain or even intensify the biases that already exist in the data used for their training. These biases may manifest themselves in subtle and unintentional ways, but they have serious and real effects on decisions about credit, admission, judgments, health care diagnoses, and employee selection. When historical data contains race, gender, class, or other discriminations, algorithms learn and replicate these patterns, reinforcing structural inequalities under the façade of technical neutrality.

Notable examples such as the COMPAS system, as cited with Peixoto and Paiva (2024), used in the United States courts, showed how algorithms can classify black defendants as more likely to reoffend compared to white defendants who have the same history. In the health area, research shows that medical screening models tend to prioritize fewer black patients due to biased associations between ethnicity and clinical risk. These distortions show that AI is not free from the social problems present in the data that underpin it. Even worse, because they are automated decisions, finding and correcting these biases becomes even more complicated.

The reduction of prejudices requires, according to Carvalho (2021), firstly, a critical analysis in the development and validation of AI systems, with the collaboration of teams with different specialties that take into account technical, social, and ethical aspects. It is essential to conduct regular audits on the data and algorithms and implement equity metrics that consider the impact of automated decisions on various social groups. In addition, legislation should accompany these initiatives, establishing clear guidelines for accountability in situations of algorithmic discrimination. Addressing bias in AI is not only a technical challenge, but also a commitment to social justice and human rights.

The growing incorporation of Artificial Intelligence in the productive, administrative and social sectors also raises concerns related to technological dependence and the dangers to digital sovereignty. According to Barroso (2024), nations that lack their own technological infrastructure or the ability to develop AI systems tend to become mere consumers of the solutions developed by large technological nations, such as the United States and China. This causes an inequality in terms of power and knowledge that undermines the decision-making autonomy of peripheral countries and accentuates their geopolitical fragility.

The centralization of control of AI in a small number of large companies, known as "Big Techs", such as Google, Amazon, Microsoft and Alibaba, results in a concentration that affects not only the economy, but also the political and information field. According to Rocha et al. (2024), these groups dominate large volumes of data, have the most advanced algorithms, and establish technological norms that influence social behaviors, consumption choices, and even public policies. When states, companies, and citizens come to depend on these private structures to function, the risk of undermining national digital sovereignty and security increases.

To overcome this dependence, it is necessary to make strategic investments in research, innovation, and national technological infrastructure, as pointed out by Barroso (2024). Nations that seek to preserve their autonomy in the digital age need to encourage the formation of local AI ecosystems, support academic institutions and research centers, promote technological startups, and create regulatory frameworks that favor data sovereignty. The implementation of open standards, the promotion of interoperability, and the strengthening of digital governance are essential actions to ensure that AI favors the strengthening, not the weakening, of the sovereignty of peoples and nations.

2.4 ETHICAL AND REGULATORY IMPLICATIONS

According to Dourado and Aith (2022), privacy stands out as one of the most critical concerns in the ethical debate around Artificial Intelligence, mainly because machine learning algorithms require large volumes of data to work effectively. Much of this data is personal and often delicate, covering information such as health, location, consumption habits, browsing history, and even biometric characteristics. Use without proper consent constitutes a direct transgression of the right to privacy and informational self-determination, guaranteed by various laws around the world. In Brazil, the General Data Protection Law (LGPD) represents significant progress in establishing guidelines such as purpose, necessity, and security, limiting data processing by public and private entities. However, the application of this law to the context of AI still needs specific interpretations and adaptations that take into account the complexities of automated decision-making.

In addition to the collection and use of data, Costa and Aranha (2023) cite that the transfer and storage of sensitive information also generate serious ethical dilemmas, especially when this data is shared with third parties or used for purposes other than those initially informed to the holder. Digital platforms that use AI often gather and connect data from various sources, which can result in the formation of extremely detailed and invasive behavioral profiles, without the user being fully aware of what is being done with their

information. This reality requires not only the strengthening of clear and informed consent mechanisms, but also the development of independent audits and governance systems that ensure the traceability and control of personal data used in AI systems.

The automation of decisions that were previously made only by humans brings to light a significant challenge in the legal and ethical sphere: who should bear responsibility in cases of error, failure, or damage caused by an artificial intelligence system? This issue becomes even more complicated when it comes to a system that adapts autonomously, that is, that learns over time and makes choices based on data that not even those who developed it can fully predict. According to Rocha et al. (2024), responsibility becomes diffuse, covering programmers, data providers, companies that use the technology and, in certain situations, even the end users themselves. This complexity indicates the need to revisit traditional concepts of civil law, such as fault, intent and causal nexus, to make them compatible with the new technological realities.

In this scenario, Costa and Aranha (2023) point out that some experts suggest the creation of new specific legal figures, such as the "digital personality", which could attribute strict liability to certain types of autonomous systems. Others argue that both manufacturers and operators should be held jointly and severally liable, similar to liability for defective products. The crucial aspect is to ensure the presence of clear and effective mechanisms for reparation of damages, especially in sensitive sectors such as health, finance and justice. The lack of a specific regulatory framework related to accountability for automated decisions can result in a dangerous gap, where victims of artificial intelligence errors are left without assistance and those responsible avoid accountability.

Clarity in artificial intelligence systems is an essential ethical demand for public trust in automatic choices. However, many algorithms, particularly those that use deep neural networks, function as true "black boxes", generating results without us being able to understand the logical criteria behind them. This lack of transparency undermines not only institutional oversight but also the rights of affected citizens, who are unable to contest or understand decisions involving them. Explainability, therefore, becomes a vital aspect of algorithmic justice, allowing individuals impacted by AI decisions to understand how and why they were made (FORNASIER AND KNEBEL, 2020).

However, according to Fornasier and Knebel (2020), creating modes of explanation is not a simple task. There is a constant tension between the technical efficiency of the model and its ease of understanding. Models that are more accurate tend to be less understandable, while clearer ones may perform less. However, it is imperative to seek a balance that provides some level of auditing and understanding of automated decisions, especially in high-impact

areas. "Explainable AI" (XAI) projects are already being developed, but they still need appropriate standards and regulations. Sound ethical governance should ensure that AI systems are, at a minimum, auditable by regulators or external experts to ensure compliance with fundamental rights.

The lack of specific regulatory standards for Artificial Intelligence is one of the biggest challenges for the development of AI that is ethical, safe, and accessible to all, as cited by Dourado and Aith (2022). Although there is a variety of legislation, such as the Consumer Protection Code, the Civil Rights Framework for the Internet, and the LGPD, these do not consider the complexity of AI and, therefore, have significant gaps. AI regulation must extend beyond data protection and consumer rights; It should include principles such as algorithmic fairness, transparency, human oversight, and anticipatory accountability. The European Union is already moving down this path with its proposed AI regulation, which defines risk categories and proportional requirements, functioning as a model for other nations.

In addition, Costa and Aranha (2023) mention that, in Brazil, there is still much to be done to create a clear and comprehensive normative system on the issue. There are legislative proposals in progress in the National Congress, but they lack adequate interconnection and a broader public debate. AI governance must go beyond just legal guidelines: it must incorporate interdisciplinary committees, societal engagement, technical assessments, ethical principles, and educational initiatives. Only with a collaborative and integrated strategy will it be possible to ensure that artificial intelligence serves as a means for human advancement, rather than becoming a focus of uncontrollable risks. Thus, the development of an effective regulatory framework is an urgent need, both political and ethical and technical, of modern society.

3 CONCLUSION

The rise of Artificial Intelligence and Machine Learning represents a profound transformation in the way society operates, directly affecting crucial areas such as health, safety, and industry. The ability to analyze large amounts of data quickly and identify useful patterns for decision-making has enabled progress that once seemed impossible, improving medical diagnostics, increasing real-time surveillance, and automating industrial processes more efficiently. However, these advantages do not come without ethical, political, and legal issues. The potential of AI, while promising, also presents several challenges that require careful analysis and careful regulation.

The most significant contemporary dilemma is the need to strike a balance between technological innovation and the fundamental rights of individuals. Problems such as the lack

of transparency in algorithms, the possibility of errors in decisions, the increase in social inequalities and the invasion of privacy are real and growing concerns. The complexity of AI technologies often makes it difficult to hold them accountable for potential damages, creating legal uncertainty and making it urgent to develop specific standards that meet the different levels of risk associated with each use. AI management should therefore be considered not only as a technical aspect, but as a process that involves a comprehensive dialogue between science, legislation, ethics and civil society.

Another crucial aspect is the need to adopt an inclusive and democratic approach in the creation of public policies related to the topic. The advancement of AI cannot be monopolized by large companies or technologically superior nations, at the risk of accentuating both global and local inequalities. It is vital that there is public investment in research, training of professionals, support for regional startups and encouragement of technological solutions that meet social needs. At the same time, it is essential to ensure that the benefits of AI are distributed fairly, tackling the digital divide and promoting the empowerment of citizens to interact with new technologies.

Finally, the future of Artificial Intelligence is more related to the way society decides to incorporate it into everyday life than to its technical ability. This requires not only innovations, but also a collective commitment and adherence to democratic principles. To develop ethical and safe AI, it is essential to promote digital education, ensure clarity in algorithms, carry out strict supervision and, above all, establish regulations that ensure that technology serves the collective good. Only with these foundations will it be feasible to establish a healthy coexistence between humans and machines, converting the power of artificial intelligence into a tool for justice, inclusion and sustainable development.

REFERENCES

Amaro Jr., E., Nakaya, H., & Rizzo, L. V. (2024). Artificial intelligence in health. Revista USP, (141), 41–50.

Arão, C. (2024). Behind artificial intelligence: An analysis of the epistemological underpinnings of machine learning. Continuous Flow, 47(3).

Barroso, L. R. (2024). Artificial intelligence: Promises, risks and regulation. Revista Direito e Práxis, 15(4). https://doi.org/10.1590/2179-8966/2024/123456

Bazeia, G. P. (2023). Fundamentals of artificial intelligence: The case of generative networks (GANS): Ethical dilemmas [Unpublished article]. Information Systems, UNISUL.

Bernasiuk, H. L. B., & Sarlet, G. B. S. (2024). Artificial intelligence and health: The impacts of generative AI on health in Brazil. Revista de Direitos e Garantias Fundamentais, 25(2), 51–84. https://doi.org/10.22478/ufpb.2317-8558.2024v25n2.123456

Bigonha, R. S. (2025). Fundamentals of machine learning (1st ed.). Author.

Carvalho, A. C. P. L. F. (2021). Artificial intelligence: Risks, benefits and responsible use. Estudos Avançados, 35(101), 123–140. https://doi.org/10.1590/s0103-4014.2021.35101.008

Costa, F. D. C., & Aranha, E. (2023). Current issues on artificial intelligence: Risks and regulatory emergency. Revista Acadêmica, 96(2), 45–62.

Dourado, D. A., & Aith, F. M. A. (2022). The regulation of artificial intelligence in health in Brazil begins with the General Law for the Protection of Personal Data. Revista de Saúde Pública, 56, 80. https://doi.org/10.11606/s1518-8787.2022056004680

Estrela, H. D., Araujo, F. R. C., Moreira, K. O., & outros. (2024). The influence of artificial intelligence (AI) on industrial automation processes. EditoraES, 12(3), 1–15.

Fornasier, M. O., & Knebel, N. (2020). Artificial intelligence: Ethical-legal challenges and risks. Law & Peace, 14(43), 25–40.

Gabriel Filho, O. (2023). Artificial intelligence and machine learning: Theoretical aspects and applications. Blucher.

Gil, A. C. (2008). Methods and techniques of social research (6th ed.). Atlas.

Gonçalves, L. S., Fantazia, G. S., Oliveira, D. S., & outros. (2023). Artificial intelligence in industry 4.0. e-Acadêmica, 4(2), e2642485. https://doi.org/10.29327/123456.4.2-5

Leite, K. R. V., Pasquali, M. A. B., Silva, T. A. S. N., & outros. (2024). Healthcare transformation: The impact of artificial intelligence on modern medicine. Contribuciones a las Ciencias Sociales, 17(5), 1–20.

Ludermir, T. B. (2021). Artificial intelligence and machine learning: Current state and trends. Estudos Avançados, 35(101), 141–158. https://doi.org/10.1590/s0103-4014.2021.35101.009

Morais, F. D. B., & Branco, V. R. C. (2023). Artificial intelligence: Concepts, applications and controversies. In XX International Symposium on Integrated Sciences of UNART, Guarujá, SP, Brazil.

Nagata, S. V. (2024). Use of artificial intelligence in public security and its contribution in the Military Police. Brazilian Journal of Development, 10(6), 123–140.

Oliveira, V. S. de, Santos, L. de F., & Ferreira, M. P. (2024). Artificial intelligence in the automation of industrial processes and its impacts. Mackenzie Journal of Economics, 21(1), 162–182.

Peixoto, F. G., & Paiva, E. L. A. (2024). Ethical challenges of the use of artificial intelligence in basic education. Revista Caderno Pedagógico, 21(13), 1–15.

Pinheiro, F. M. L. (2024). The role of technology and artificial intelligence in public security: Challenges and convergences with individual and social fundamental rights. Revista do Ministério Público do Estado do Rio de Janeiro, (94), 45–60.

Rocha, U. B., Santos, W. P. C. dos, & Nano, R. M. W. (2024). Use of artificial intelligence: Advances, risks and challenges related to intellectual property. Multidisciplinary Scientific Journal Knowledge Core, 6(3), 137–149.

Schussler, P. J., Bastiani, E., & Bussler, N. R. C. (2018). Artificial intelligence and machine learning: Using the understanding of human intelligence for reproduction in computing. Hall of Knowledge, Unijuí.

Vieira, P. H. B., & Barbuda, A. S. (2024). The application of artificial intelligence in the activity of public security intelligence. Legal Journal of the Northeast of Minas Gerais, 8(1), 1–15.