REVISTA
CES DERECHO

INVESTIGATION ARTICLE

Artificial intelligence, algorithmic discrimination and the European Union's regulatory response

Inteligencia artificial, discriminación algorítmica y la respuesta normativa de la Unión Europea

Juan Alejandro Martínez Navarro[1]

[1] Doctor en Derecho. Docente e Investigador de la Universidad Tecnológica Atlántico-Mediterráneo (UTAMED), Málaga. España.

How to cite: Martínez-Navarro, Juan Alejandro. "Artificial intelligence, algorithmic discrimination and the European Union's regulatory response" En: *Revista CES Derecho*. Vol. 16. No. 3, septiembre a diciembre de 2025. pp. 4-20. https://doi.org/10.21615/cesder.7943

Abstract
This paper analyses the risks that artificial intelligence (AI) poses to fundamental rights, with particular emphasis on the principle of equality and the emergence of algorithmic discrimination. It shows that AI systems can reproduce and amplify biases embedded in data or design choices, disproportionately affecting vulnerable groups. By examining *Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonized rules on artificial intelligence* and *Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (GDPR)*, the study assesses the EU's legal capacity to address these risks. It highlights the need to reinforce oversight mechanisms, ensure algorithmic transparency, and promote a model of digital governance that effectively safeguards human rights.

Keywords: artificial intelligence; algorithmic discrimination; fundamental rights; equality; European Union.

Resumen
Este estudio examina los riesgos que plantea la inteligencia artificial (IA) para los derechos fundamentales, en particular el principio de igualdad y la aparición de discriminación algorítmica. Se demuestra que los sistemas de IA pueden reproducir y amplificar sesgos presentes en los datos o en las decisiones de diseño, afectando especialmente a colectivos vulnerables. A través del análisis del *Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial*, y del *Reglamento (UE) 2016/679 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (RGPD)*, se evalúa la capacidad normativa de la Unión Europea frente a estos desafíos. El trabajo subraya la necesidad de fortalecer los mecanismos de control, asegurar la transparencia de los algoritmos y promover una gobernanza digital que garantice el respeto efectivo a los derechos humanos.

Palabras clave: inteligencia artificial; discriminación algorítmica; derechos fundamentales; igualdad; Unión Europea.

Introduction

Our society is undergoing an unprecedented process of technological revolution, commonly referred to as the «Fourth Industrial Revolution». This revolutionary process is characterized by an intense concentration of technologies that blurs the boundaries between the physical, digital, and biological spheres (Torres, 2021: 215). It appears evident that through Revolution 4.0, the way we live, work, and relate to one another is being fundamentally transformed, and this will affect the functioning of public administration with equal intensity

Date of correspondence:
Received: June 11, 2025.
Reviewed: August 8, 2025.
Accepted: October 20, 2025.

DOI: 10.21615/cesder.7943
ISSNe: 2145-7719
https://revistas.ces.edu.co/index.php/derecho

OPEN ACCESS
CC BY NC SA

(Schwab, 2015).

The aforementioned technological revolution is leading to a deficient constitutionalism, where the analog Constitution must regulate a new digital world (Sánchez, 2016: 239).

Depending on the context in which they are applied and the level of technological development, artificial intelligence (AI) systems can generate risks that compromise both public interests and protected fundamental rights. These threats can manifest in various forms, both tangible and intangible.

A critical aspect to consider is the possibility that AI technologies may produce biased outcomes. These inaccuracies can give rise to discriminatory effects, particularly based on factors such as age, ethnicity, race, gender, or disability. Furthermore, the immediate nature of decisions made by these systems in real-time, combined with limited opportunities to verify or correct errors, considerably elevates the risk of violating rights and freedoms, particularly in the realm of law enforcement.

The relevance of this study lies in the fact that, within a context of increasing digitalization and automation, AI systems are being progressively integrated into decision-making processes in both the public and private sectors. These automated decisions, often made in real-time and without sufficient possibility for human review, can significantly influence individuals' fundamental rights, including the right to equality and non-discrimination. Given the risk that AI systems may operate as «black boxes», where the opacity of automated decision-making processes hampers the identification and correction of biases, a rigorous examination of the algorithms and data upon which they are based becomes crucial. This research is essential for developing public policies and regulatory frameworks that ensure responsible, transparent, and equitable use of artificial intelligence, promoting algorithmic governance that respects human rights and strengthens social justice.

In this context, this work proposes to examine the various types of biases that may emerge throughout the lifecycle of AI systems, from data collection and preparation to training processes and automated decision-making. Likewise, the current regulatory framework will be analyzed, with particular attention to the Artificial Intelligence Act, to evaluate its effectiveness in mitigating biases and preventing algorithmic discrimination. Additionally, recommendations will be offered for policy formulation that strengthens algorithmic transparency, accountability, and the interpretability of decisions made by these systems, with particular emphasis on protecting fundamental rights. Through this analysis, we seek to contribute to the academic and regulatory debate on how to ensure that artificial intelligence is developed and applied in ways that promote equity, diversity, and inclusion, avoiding the reproduction of structural inequalities and guaranteeing human dignity and the rights of all citizens.

Artificial intelligence and humans

The history of artificial intelligence (AI) is rich in milestones that have transformed our understanding of machines and their capacity to simulate human intelligence. In recent years, AI has reached new levels of sophistication and practical application, notably the resurgence of deep learning and its applicability in image recognition. The advent of large datasets and advances in hardware have driven machine learning, significantly improving algorithms such as Support Vector Machines and Artificial Neural Networks.

AI systems are those «systems that display intelligent behaviour by analysing their environment and taking actions—with some degree of autonomy—to achieve specific goals» (European Commission, 2018: 1). In this way, the system encompasses a set of methods, theories, and techniques aimed at attempting to reproduce human cognitive capabilities in a machine (CEPEJ, 2018: 69).

The recent Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), in its Article 3.1, defines an AI system as «a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit

adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that can influence physical or virtual environments».

An artificial intelligence (AI) system is characterized by several key elements that enable it to perform complex tasks and simulate human intelligence (European Commission, 2018: 1). The fundamental components of these systems include algorithms and models, data, learning and adaptation, as well as ethical and safety considerations. These elements work together to enable the efficient and responsible functioning of AI (CEPEJ, 2018: 69).

As Barrios Andrés appropriately explains, the current social context is configured around the so-called «algorithmic rule of law», which, in turn, is built upon the following core technological elements: Internet, cloud computing, digital platforms for providing public services, Big Data, process automation through artificial intelligence, advertising systems based on blockchain technologies (Merchán, 2019) and distributed ledger technology (DLT), as well as the ubiquitous connectivity of 5G networks (Barrios, 2020: 3).

At the core of any AI system are algorithms and models, which constitute the mathematical and computational foundations that guide their behavior (Special Committee on Artificial Intelligence in the Digital Age, 2022: 13). Algorithms are an ordered and finite set of operations that allow finding the solution to a problem (Dictionary of the Royal Spanish Academy).

In turn, algorithms are sequences of mathematical equations that interweave to generate a result. Their objective is not to validate the veracity of a proposed hypothesis, but to identify deterministic correlations between different data. The algorithm uses a vast database organized in a comprehensible manner (smart data) (De Salvador, 2023, p. 148). Subsequently, a mathematical model employs these data randomly until it identifies correlation patterns among them, which allows it to make precise estimations.

There exist more advanced algorithms known as machine learning algorithms (De Salvador Carrasco, 2023: 135). These algorithms are essential for pattern identification and data-based decision-making[1] and are more effective at predicting behaviours, but they present a complex challenge: their functioning, upon reaching a certain level of data and variable inclusion, becomes incomprehensible to the human mind and even to their own developers[2].

Certainly, machine learning algorithms, especially those based on deep learning, exhibit an increasing level of autonomy as they are fed with massive volumes of data (Arellano, 2019). These models, in their processing capacity, can handle and correlate thousands or even millions of factors simultaneously, generating a complex network of interrelationships that escapes detailed comprehension by human programmers.

Indeed, deep learning models, such as deep neural networks, are frequently considered «black boxes» due to their lack of transparency in decision-making (Cotino, 2022: 31)[3]. Programmers and data scientists can understand the model's architecture and input data; however, the specific internal processes that lead to certain predictions may remain opaque and difficult to decipher. Additionally, it becomes complicated for developers to

[1] Among these algorithms, Support Vector Machines (SVM), decision trees, and deep neural networks stand out, such as convolutional neural networks (CNN) and recurrent neural networks (RNN). These models are trained to perform specific tasks, from image classification to natural language processing, allowing AI to simulate complex human capabilities.

[2] The importance of these variables can fluctuate as the model adjusts internally during its training, making use of optimization techniques to adjust its parameters to the input data (Borges Blázquez, 2020: 58).

[3] The learning and adaptation capacity is a central characteristic of AI systems. There are three main types of learning: supervised, unsupervised, and reinforcement learning. In supervised learning, models are trained with labeled data, where the desired output is known. In unsupervised learning, models must identify intrinsic patterns in unlabeled data. Reinforcement learning, on the other hand, involves models that make decisions through interaction with an environment, seeking to maximize accumulated reward. These methods allow AI systems not only to learn from past experience but also to adapt to new data and situations (Sánchez, 2020).

identify with precision which specific variables are critical for the model at a given moment[4].

The complex nature of algorithms has led some prominent authors to consider their value as a source of law. Thus, authors such as Balaguer Callejón in his interesting work «The Constitution of the Algorithm», raises the problematic possibility of the legal value of algorithms, especially in the field of public administration (Balaguer, 2023: 38-41)[5].

The reality is that artificial intelligence (AI) feeds on data that represents variables and contains inherent social information. We rely on data, that is, on information derived from daily life. As Miró Llinares aptly argues, despite AI depending on mathematical analyze, these are applied to social variables, and it is the task of social scientists to design adequate explanatory frameworks for their application (Miró, 2019: 103).

The quality and quantity of available data are crucial for training and evaluating AI models (Special Committee on Artificial Intelligence in the Digital Age, 2022: 13). Data provides the necessary raw material for algorithms to learn and improve their performance.

In the field of artificial intelligence, respecting human dignity requires that all people be treated with the inherent respect they deserve as moral agents, and not as mere objects that can be filtered, classified, scored, directed, conditioned, or manipulated. Therefore, AI systems must be designed and developed in ways that respect, protect, and promote the physical and mental integrity of human beings, as well as their personal and cultural identity, and that contribute to satisfying their fundamental needs (High-Level Expert Group on Artificial Intelligence, 2019: 13).

The recognition of fundamental rights was made possible through the dissemination of the value of human dignity. Thus, human rights find their foundation in the essential value that human dignity represents[6]; which is, in truth, «the founding value, the matrix of all values» (Panea Márquez, 2005: 18). As a value that gives meaning to the rest of human rights, it could be said that dignity is at a higher level than these (Gómez Sánchez, 2004: 88). In turn, these rights have been constituted as a universal ethical framework to address complex challenges such as those posed by AI systems (Pérez, 2005: 53)[7].

This not only implies freedom from interference by sovereign power but also requires active intervention by governmental and non-governmental organizations to ensure that all people, especially those at risk of exclusion, have equal access to the benefits and opportunities offered by artificial intelligence. Furthermore, AI systems must be designed with a firm commitment to ensuring that their operation does not undermine the fundamental principles of the rule of law, as well as strict compliance with existing laws and regulations, ensuring respect for procedural guarantees and equality before the law.

In fact, there exists extensive doctrine, among which authors such as De Asís Roig stand out, who defend the

[4] By way of example, in 2016, the Arkansas Department of Social Services decided to automate needs assessment for disabled residents through an algorithm instead of conducting home visits by nurses. This measure sought to reduce costs, but resulted in litigation due to system problems. The algorithm assigned care hours based on complex calculations, but did not allow adjusting beneficiaries' needs if they changed. During the 2017 trial, it was demonstrated that the algorithm produced arbitrary and erroneous results, such as ignoring the need for assistance in cases of amputations. Additionally, it assigned the same level of care to people with serious and less serious conditions, violating federal laws. The federal judge prohibited the state from automating these decisions until the process could be adequately justified. Despite the problems, neither the state nor the system providers/developers managed to identify and correct the problem in the algorithm (Barrios Andrés, 2020: 4).

[5] It should be specified that previously authors such as Boix Palop, Arroyo Jiménez, and De la Cueva already raised the possibility that algorithms materially perform the same functions as regulations (Boix Palop, 2020; Arroyo Jiménez, 2020; De la Cueva, 2018).

[6] In the words of our Spanish Constitutional Court, the principle of human dignity is the «portico of other values or principles (STC 337/91 FJ 12º)». The Spanish Constitutional Court has determined that «projected onto individual rights, the rule of art. 10 [Spanish Constitution of 1978] implies that, as 'spiritual and moral value inherent to the person', dignity must remain unaltered whatever the situation in which the person finds themselves». It continues indicating that it constitutes «an invulnerable minimum that every legal statute must ensure so that, whatever the limitations imposed on the enjoyment of individual rights, they do not entail contempt for the esteem that, as a human being, deserves respect» (STC 120/1990, FJ 4).

[7] However, the emerging question of whether rights constitute a sufficient ethical framework to address all the challenges that new technologies may pose represents a complex debate. Conversely, there are authors who defend the need to recognize and regulate new digital rights (although this debate exceeds the scope of the present study) (De Asís Roig, 2023: 602-603).

need to take human rights as a reference when regulating emerging technologies (De Asís, 2015, p. 62). Therefore, in direct relation to the challenges presented by AI systems, De Asís Roig highlights the absence of a rights-based approach (De Asís Roig, 2024: 28).

Regarding this complex context, it is relevant to mention the reflections of Pérez Ugena, who precisely analyze the system of fundamental rights in the face of challenges posed by artificial intelligence systems. First, he highlights the principle of equality, which constitutes the central axis of this study. Second, he examines the protection of private life. Third, he analyze the risks that AI represents in relation to rights in the field of communication, such as freedom of expression and freedom of information. Finally, he addresses the right to effective judicial protection (Pérez Ugena, 2024: 327-333).

In the following sections of this work, the first of the mentioned points will be addressed exclusively: the principle of equality. An analysis will be conducted on how this fundamental principle is impacted by the development and implementation of artificial intelligence systems, considering the legal and ethical implications that arise in

Algorithms and discrimination: algorithmic discrimination

Our society has transitioned from being predominantly exposed to damages caused by natural phenomena to primarily facing harm derived from the very technology we develop. This observation constitutes a central axis in scientific research, where the benefits that a technology provides are evaluated in contrast to the risks it entails. While in most cases, significant technological advances surpass the risks that these same advances generate, it is undeniable that technological development stands as the essential foundation of social evaluation.

This process, while providing numerous advantages, also harbors multiple dangers, encapsulated in what doctrine has termed the «risk society» (Beck, 2008). In this context of technological damages, law acquires a crucial role. As Esteve Pardo points out, any risk derived from technology must be subjected to legal control (Esteve Pardo, 2003: 138).

The remarkable advancement of science and technology that defines our era imposes upon law the necessity to evolve, in order to encompass an increasingly broad and complex spectrum of situations (Montano, 1995: 81). However, it is imperative that law determines who has the authority to make decisions and under what responsibilities such authority is exercised (Esteve Pardo, 2003: 138). While science has the task of identifying and evaluating risks, it is law that must make decisions in contexts of scientific uncertainty. In summary, science provides information about possible risks, and it corresponds to law to adopt pertinent decisions based on such information[8].

It might be assumed that artificial intelligence systems are immune to human limitations, such as prejudices, and that, therefore, they do not incur discrimination based on race, nationality, gender, or culture. However, the situation is considerably more complex[9]. The mitigation of algorithmic biases—whether they manifest inadvertently or are intentionally introduced—is a very productive sphere of research in artificial intelligence development (Herranz, 2019).

In the field of artificial intelligence, equality implies that the functioning of these systems should not produce unfairly biased results. Likewise, it is necessary to guarantee adequate respect toward potentially vulnerable individuals and groups.

[8] Regarding the rapid technological development to which our society is being subjected due to the implementation of AI systems and their impact on fundamental rights, one issue stands out both theoretically and practically: the protection of digital data. However, as has already been noted, there are other areas of vulnerability that warrant particular attention from legal scholarship (Rallo Lombarte, 2020: 108).

[9] According to UNESCO's 2019 report, it is no coincidence that virtual assistants such as Siri, Alexa, and Cortana have female names and voices. Gender bias is also embedded in algorithm design, as 78 percent of AI professionals are men. This means that male experiences tend to dominate the development and structuring of these algorithms. In short, the companies that design these services are perpetuating an existing social reality (Deva, 2020).

Indeed, data-based artificial intelligence algorithms do not intrinsically generate biases, but they can reproduce them without adequate human intervention, manifesting in three main phases: in data collection, where collected data may reflect pre-existing prejudices; in training data preparation, when selecting and processing the attributes provided to the algorithm; and in decision-making, through proposals and decisions adopted throughout the entire lifecycle of intelligent development (Belloso Martín, 2022: 48).

Regarding this, the proposition that biases present in algorithms simply reflect reality and that algorithmic design is restricted to conforming to such reality lacks coherence with constitutional principles. This perspective denotes a persistent lack of attention to constitutional values by technology companies (Balaguer, 2023: 42).

A bias is a «systematic error that can be incurred when making samples or tests by selecting or favoring some responses over others» (Dictionary of the Royal Spanish Academy). For its part, algorithmic discrimination refers to the situation in which biases inherent to artificial intelligence systems generate or may generate adverse effects that disproportionately affect certain social groups (IALAB, 2021).

As Flores Anarte accurately explains, «the particular characteristics of AI systems' functioning, such as the opacity of mechanized decision-making procedures or the difficulty in controlling bias reproduction in machine learning, make pejorative treatments caused by algorithmic discrimination a type of discrimination much more difficult to identify and tackle and with great potential to contribute decisively to deepening the gender gap» (Flores, 2023: 106).

This occurs in a context where an increasing number of decisions are mechanized, both in the public and private sectors, indicating a trend toward total automation of decision-making processes.

The public sector, for its part, is experiencing growing digitalization and the incorporation of artificial intelligence tools in its operations. This process has given rise to what is described as a paradigm shift in governance models at all territorial levels[10]. In this new framework, a significant portion of public authorities' decisions are made or based on algorithms, representing a transition from traditional control structures based on bureaucracy toward systems supported by intelligent algorithms. This emerging paradigm of Algorithmic Governance promises advances and optimization in public administration activity, and tangible benefits derived from applying AI tools in the public sector have already been observed.

The legal response to algorithmic discrimination in the European Union

Artificial intelligence has emerged as a disruptive technology with undeniable strategic implications across all sectors (Padilla Ruiz, 2019: 103). However, while its transformative potential is widely acknowledged, it is misguided to view artificial intelligence as the ultimate solution for eradicating all forms of discriminatory bias. Despite the efficiency and capabilities offered by AI and other tools such as Big Data, it cannot be assumed that these technologies are capable of completely eliminating the biases inherent in automated processes.

[10] The judiciary is no exception. In the case of State v. Loomis, the Wisconsin Supreme Court examined the use of the artificial intelligence tool COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) in the judicial sentencing process. COMPAS is an algorithm designed to assess an individual's risk of recidivism, and its use in the judicial system has sparked controversy due to several issues in its application.
One of the main concerns identified is the lack of transparency and explainability of the algorithm, as its internal workings and the factors used in its assessments are not fully accessible or comprehensible to defendants. This opacity significantly limits defendants' ability to understand how their risk scores were generated, which in turn restricts their ability to challenge or contest the evaluation.
Another critical issue is the potential reinforcement of biases and discrimination. In the case at hand, the algorithm asserts that "all available scientific evidence indicates that males have higher rates of recidivism and violent crime than females."
Moreover, the use of COMPAS in judicial decision-making imposes constraints on the right to appeal. Due to the automated and opaque nature of the system, defendants face greater difficulties in appealing decisions based on algorithmic outcomes, as they cannot easily access the underlying logic informing those decisions.
The Wisconsin Supreme Court allowed the use of COMPAS but emphasized that judges must not rely solely on the tool when issuing sentences. It also stressed the importance of informing defendants about the inherent limitations of using such an algorithm in their case, thereby safeguarding a fairer and more equitable trial process (Borges Blázquez, 2020: 61–63).

As has become evident, artificial intelligence has demonstrated significant value in enhancing efficacy and efficiency, facilitating the detection—and even prediction—of problems, advancing the automation of procedures, and more (Arellano Toledo, 2023: 34). Nevertheless, it simultaneously presents considerable challenges yet to be resolved, including the perpetuation of biases, lack of transparency, and the complexity of algorithms, which, to some extent, may adversely impact equality policies[11].

Non-discrimination is established as a fundamental guiding and structuring principle. It is enshrined in the Universal Declaration of Human Rights, adopted by the United Nations General Assembly in 1948. Within the United Nations framework, two significant legal instruments adopted in 1966 are worth highlighting: the International Covenant on Economic, Social and Cultural Rights and the International Covenant on Civil and Political Rights, the latter of which articulates non-discrimination as a right under Article 26.

In Spain, the right to equality originates in and is granted constitutional status under Articles 9.2[12] and 14[13] of the Spanish Constitution (CE). This right has subsequently been developed in more specific legal texts, notably Organic Law 3/2007, of March 22, on effective equality between women and men. Article 1 of this law provides that:

> «Women and men are equal in human dignity and equal in rights and duties. The purpose of this Law is to make the right to equal treatment and opportunities between women and men effective, particularly by eliminating discrimination against women, regardless of their circumstances or status, in all areas of life, especially in the political, civil, labour, economic, social, and cultural spheres […]»

Beyond the regulatory framework, the right to equality is operationalized through the Strategic Plan for the Effective Equality of Women and Men (PEIEMH), which is currently embodied in the Third Strategic Plan covering the period 2022–2025 (Ministry of Equality, 2022). This Strategic Plan outlines and implements action measures to promote gender equality and, ultimately, to give practical effect to the right to equality.

The following sections will examine three recently adopted regulatory instruments that are particularly relevant for establishing non-discrimination safeguards in the effective application of AI: Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024, laying down harmonized rules on artificial intelligence; and Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016, concerning the protection of natural persons with regard to the processing of personal data and the free movement of such data.

The Artificial Intelligence Act

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024, establishing harmonized rules on artificial intelligence (AI Act), is structured around the need to safeguard fundamental rights in a context where AI systems play an increasingly central role in everyday life.

This regulation provides a clear framework that obliges AI system operators to act with transparency and accountability. It imposes strict requirements on high-risk systems and aligns with other EU legal frameworks—most notably Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such

[11] As an example, Amazon implemented an artificial intelligence system to optimize its hiring process. To train the system, historical data from employees hired over the previous decade was used. During that period, technical positions had been predominantly filled by men. As a result, the algorithm began to reject female applicants, interpreting their underrepresentation in the training data as an indication that they were not suitable candidates for such roles (Rubio, 2018).

[12] Spanish Constitution, Article 9.2: «It is the responsibility of public authorities to promote the conditions under which the freedom and equality of individuals and the groups to which they belong are real and effective; to remove obstacles that hinder or prevent their full realization; and to facilitate the participation of all citizens in political, economic, cultural, and social life».

[13] Spanish Constitution, Article 14: «Spaniards are equal before the law, and no discrimination shall prevail on grounds of birth, race, sex, religion, opinion, or any other personal or social condition or circumstance».

data (GDPR)—to ensure the proper protection of individuals' rights.

The AI Act takes both a preventive and corrective approach: it seeks to avert potential rights violations through stringent regulation and provides mechanisms for affected individuals to seek redress. This reflects a clear commitment to the fundamental values of the European Union, ensuring that technological advancement does not come at the expense of human rights.

The AI Act addresses the regulation of AI systems comprehensively, focusing on the protection, monitoring, and prevention of fundamental rights infringements. It acknowledges that AI systems may pose significant adverse effects—not only physical, but also psychological and social—arising from misuse or malfunction.

To this end, the AI Act categorizes AI systems according to different levels of risk, paying particular attention to those that may negatively impact fundamental rights (Barrio Andrés, 2022: 64). These classifications are essential for determining the level of regulation and oversight applicable to each type of AI system[14].

The regulation establishes four risk categories (Pérez Ugena, 2024: 322), ranked from most to least severe. First are prohibited AI systems, deemed unacceptable because they violate EU values and fundamental rights. These systems are banned due to the serious risks they pose to individuals' safety and rights—for instance, real-time biometric identification systems in public spaces used by public authorities, except in narrowly defined and authorized situations[15]. In terms of fundamental rights, these systems pose direct threats to human dignity, freedom of thought, privacy, and non-discrimination. By explicitly banning such systems, the regulation provides robust protection for these rights.

Prohibited systems are those that, by their very nature, are considered inherently dangerous to fundamental rights and are therefore completely barred. This reflects a strong normative commitment to the protection of EU values and fundamental rights, drawing clear lines against unacceptable uses of AI[16].

Among the prohibited practices regulated by the AI Act (Pérez Ugena, 2024: 321), the following are worth highlighting: «the placing on the market, putting into service or use of an AI system that exploits any of the vulnerabilities of a specific natural person or a group of persons due to their age, disability or a specific social or economic situation, with the objective or the effect of substantially distorting the behaviour of that person or a person belonging to that group in a manner that causes or is likely to cause significant harm to that person or another person» [Art. 5.1(b)], and «the evaluation or classification of individuals or groups of individuals over a certain period based on their social behaviour or personal or personality characteristics, whether known, inferred or predicted, where the resulting social score leads to […] detrimental or unfavorable treatment of certain individuals or groups» [Art. 5.1(c)].

Second are high-risk AI systems, which, while not prohibited, pose a significant risk to individuals' health, safety, and fundamental rights. These systems are subject to strict obligations prior to their commercialization or deployment (e.g., AI systems used in employment contexts to make decisions about hiring, promotion, or dismissal). The most relevant fundamental rights in this category include privacy, data protection, non-

[14] In 2021, a total of 114 civil society organizations issued a statement calling for the Artificial Intelligence Act to prioritize the protection of fundamental rights and democratic values in light of the growing risks associated with the use of AI systems. The statement criticized the current risk-based approach of the regulation as inadequate and overly rigid, proposing several amendments to enhance the legislation. These included prohibiting AI practices that infringe upon fundamental rights, ensuring accountability from users of high-risk systems, and improving public transparency. The statement also emphasized the need for inclusive and accessible standards and stressed the importance of considering the needs of the most vulnerable individuals and groups. As previously discussed, these concerns were ultimately incorporated into the final text of the regulation (A Civil Society Statement, 2021).

[15] In this context, the case of *Mercadona* is particularly relevant. In 2021, the company was fined over 3 million euros for deploying a facial recognition system intended to prevent individuals with criminal records—specifically those involving offenses against the company's premises, property, or staff—from entering its stores. The fine was imposed as a result of violations of existing data protection regulations (Sánchez-Yara, 2021).

[16] «The Act also pays particular attention to the impact of AI systems in specific contexts, such as the workplace or health insurance, as well as in essential services and public benefits. In the latter case, fundamental rights are primarily addressed through social rights, dignity, and non-discrimination. In all these situations, the individuals involved are either already at a disadvantage (as is the case with workers), or are vulnerable due to their health or socio-economic status» (Laykite, 2024: 163).

discrimination, the right to education, the right to work, and access to justice. High-risk systems must comply with rigorous requirements to minimize the risk of rights violations[17].

Third are limited-risk AI systems, which pose lower risks and are therefore subject to fewer restrictions. The focus of regulation here is on transparency and providing adequate information to users (e.g., chatbots used in customer service contexts). Even when risks are limited, they may still impact privacy and the right to information. It is essential that users be informed they are interacting with an AI system and understand how their data is being used.

Lastly, minimal-risk AI systems are those with negligible risk to fundamental rights and are thus subject to minimal regulation (e.g., email spam filters)[18].

The AI Act underscores the importance of assessing how AI systems may affect individuals' fundamental rights, including privacy, non-discrimination, and freedom of expression. To prevent discrimination, the regulation mandates continuous review of algorithms and datasets to ensure fairness and justice in the outcomes generated by AI systems. It also acknowledges the direct impact of AI, particularly on communication platforms and social networks, on freedom of expression, and establishes safeguards to prevent unjustified infringements of this right[19].

In this regard, the AI Act requires that high-risk AI systems be overseen by human supervisors, independent from any automated control mechanisms[20].

At this point, it is worth noting certain contentious aspects of the regulation. It imposes strict conditions on the use of biometric identification systems, ensuring their proportionality and legitimacy, and explicitly prohibits their use for mass or indiscriminate surveillance. The regulation emphasizes the necessity of operating within a narrowly defined dataset obtained legally and avoids any form of mass or indiscriminate collection of biometric data[21].

Nevertheless, scholars such as Laykyte and Simón Castellano (Simón Castellano, 2023) identify vulnerabilities in the regulation regarding impact assessment. For instance, this assessment is conducted at the system's initial

[17] Article 6.1 of the Artificial Intelligence Act provides that «[…] an AI system shall be considered high-risk where the following two conditions are fulfilled: a) the AI system is intended to be used as a safety component of a product that falls under one of the Union harmonisation legislation acts listed in Annex I, or the AI system itself constitutes such a product; and b) the product of which the AI system is a safety component pursuant to point (a), or the AI system itself as a product, is subject to a third-party conformity assessment under the Union harmonisation legislation acts listed in Annex I».

[18] The category of high-risk AI systems is most closely linked to fundamental rights, as these systems can have a direct and significant impact on areas such as privacy, equality, non-discrimination, access to justice, education, and employment. These systems are subject to strict evaluation, transparency, and oversight requirements aimed at preventing violations of these rights. In contrast, while limited- and minimal-risk systems may also have implications for fundamental rights, their impact is less direct, and the applicable regulation is more flexible, focusing mainly on transparency and the provision of clear information to users.

[19] With respect to the protection of fundamental rights, the AI Act requires a detailed impact assessment to be carried out before the deployment of any high-risk AI system. This assessment must identify all potential risks to individuals' fundamental rights and propose specific measures to mitigate those risks. Moreover, public participation and consultation are encouraged, involving affected communities and human rights experts to ensure a comprehensive and inclusive assessment process.

Article 27.1 of the Artificial Intelligence Act states: «[…] For this purpose, deployers shall carry out an assessment consisting of: a) a description of the deployer's processes in which the high-risk AI system will be used, in accordance with its intended purpose; b) a description of the time period during which each high-risk AI system is intended to be used, and the frequency of its use; c) the categories of individuals and groups potentially affected by the use of the system in the specific context; d) the specific risks of harm that may affect the identified categories of individuals and groups under point (c), taking into account the information provided by the provider pursuant to Article 13; e) a description of the application of human oversight measures, in line with the instructions for use; f) the measures to be adopted in the event that these risks materialise, including internal governance arrangements and complaint mechanisms».

[20] Article 14 of the Artificial Intelligence Act provides: «1. High-risk AI systems shall be designed and developed in such a way that they can be effectively overseen by natural persons during the period they are in use, including by providing appropriate human-machine interface tools. 2. The purpose of human oversight shall be to prevent or minimise risks to health, safety or fundamental rights that may arise when a high-risk AI system is used in accordance with its intended purpose or in a reasonably foreseeable misuse scenario, particularly where such risks persist despite the implementation of other requirements set out in this section».

[21] Thus, Article 27.1 makes it clear that «prior to the deployment of a high-risk AI system […] deployers that are public sector bodies or private entities providing public services, as well as all deployers of high-risk AI systems […] shall carry out an assessment of the potential impact of such systems on fundamental rights».

deployment, and its subsequent update is left to the discretion of the deployer, who decides whether changes in the system are sufficiently relevant to expose individuals and/or groups to new risks[22].

Indeed, the regulation could be strengthened by introducing requirements for continuous monitoring and periodic audits of AI systems to ensure that discriminatory biases do not emerge over time—particularly once systems are deployed and begin interacting with real-world data. Additionally, it would be beneficial to introduce an explicit obligation for developers and users of high-risk AI systems to conduct a detailed discrimination impact assessment prior to system implementation.

Regarding the oversight of fundamental rights, the Artificial Intelligence Act grants broad powers to supervisory authorities to ensure compliance with the regulation. These authorities may request and review all documentation related to high-risk AI systems, including impact assessments, data logs, and any internal audit reports. If the documentation does not provide sufficient evidence of compliance, the authorities may conduct independent testing and evaluation of the AI systems[23].

To ensure an effective impact assessment, it is crucial to highlight that the legislator establishes, on the one hand, a network of shared responsibilities among providers, system deployers, and market surveillance authorities. On the other hand, these responsibilities are reinforced through the implementation of measures, procedures, instructions for use, and other tools designed to ensure that such responsibilities are effectively fulfilled (Laykyte, 2024: 168).

In any case, although the regulation refers to the need for transparency, it is essential to emphasise the explainability of algorithms, particularly regarding how they make decisions that affect vulnerable groups. Companies and public entities using AI should be required to provide understandable explanations of how their systems make decisions, enabling affected individuals to comprehend and, if necessary, challenge those decisions.

Finally, while the Artificial Intelligence Act addresses the rights of vulnerable groups—focusing on protection against bias and discriminatory effects—it devotes only limited attention to issues of accessibility and inclusion. In truth, the drafting adopts a defensive and preventive stance. However, AI systems have the potential to be powerful tools for the integration and empowerment of these groups. Failing to position AI as a means to guarantee fundamental goods such as autonomy, social and professional integration, and active participation in community life is a strategic oversight.

In this regard, the High-Level Expert Group on Artificial Intelligence of the European Commission understands AI not as «an end in itself, but a promising means to foster human flourishing and thus improve individual and collective well-being and the common good, while also contributing to progress and innovation» (High-Level Expert Group on Artificial Intelligence, 2019: 5). Therefore, AI not only can—but must—be used to promote these objectives, given its potential to positively transform the lives of excluded or marginalised individuals. In this sense, applying AI as a tool to achieve these ends is not merely desirable, but imperative to move towards a more inclusive and equitable society.

[22] On the one hand, Recital 96 requires a new evaluation when «[…] the deployer considers that any of the relevant factors has changed». Likewise, Article 27.2 of the Artificial Intelligence Act establishes that «[t]he obligation [to carry out the impact assessment] shall apply at the first use of the high-risk AI system. In similar cases, the deployer may rely on previously conducted fundamental rights impact assessments or on existing assessments carried out by providers. If, during the use of the high-risk AI system, the deployer considers that any of the elements listed in paragraph 1 has changed or is no longer up to date, they shall take the necessary steps to update the Information». As Laukyte rightly points out, the regulation does not clarify what should be understood by «in similar cases» (Laykyte, 2024: 167).

[23] Article 77.1 of the Artificial Intelligence Act provides that: «National authorities or public bodies responsible for supervising or enforcing obligations under Union law regarding the protection of fundamental rights, including the right to non-discrimination, in relation to the use of high-risk AI systems referred to in Annex III, shall have the power to request and access any documentation created or maintained pursuant to this Regulation, in an accessible language and format, when such access is necessary for the effective performance of their mandates, within the limits of their jurisdiction. The relevant authority or public body shall inform the market surveillance authority of the corresponding Member State of any such request».

The General Data Protection Regulation and Automated Decision-Making Systems

Concerning the prominent aspects that link the right to personal data protection with artificial intelligence systems, the General Data Protection Regulation (GDPR) establishes specific restrictions applicable to automated decision-making processes.

> «The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her» (Art. 22).

According to the GDPR, individuals possess the right not to be subjected to decisions based exclusively on automated data processing when such decisions may have relevant legal effects or impact them in comparable ways. This right ensures that no transcendental or significant decision affecting a person's rights or daily life is made solely through automated processes, such as algorithms or computer systems, without direct human intervention (Pérez Ugena, 2024: 329).

Automated data processing also encompasses profiling[24], which involves evaluating personal data to analyze or predict aspects related to work performance, economic situation, health, personal preferences, reliability, behavior, location, or movements of an individual, when such evaluations may have significant legal consequences or similar impact (Article 29 Data Protection Working Party, 2018: 40).

Profiling has the potential to perpetuate pre-existing stigmas and foster social fragmentation, as it may pigeonhole a person into a specific category, limiting their options to predetermined suggestions (Jiménez-Castellanos Ballesteros, 2023: 191-215). This phenomenon can compromise freedom of choice, affecting decisions regarding products and services, such as books, music, or news[25]. Furthermore, profiling may lead to inaccurate predictions or, in some cases, denial of services and goods, thereby perpetuating unjustified discrimination[26].

Regarding the aforementioned, the excessive power of technology companies vis-à-vis the surveillance state itself is evident, not to mention the limited participation attributed to civil society (Cotino Hueso, 2023: 17-63). These companies possess the data, infrastructure, technological capacity, and expertise, compared to a public administration that is barely beginning to approach this phenomenon (Laykite, 2024: 157).

The inclusion of provisions on automated decision-making systems in the General Data Protection Regulation (GDPR) represents a legislative success that demonstrates remarkable foresight on the part of the legislator. Seven years were necessary for the Court of Justice of the European Union (CJEU) to issue the first judgment

[24] Among many other possible scenarios, decisions such as the denial of a loan, a job offer, the acceptance or rejection of an insurance application, or the allocation of public resources should not be determined solely by algorithms without meaningful human intervention influencing the process.

[25] In its ruling of February 5, 2020, the European Court of Human Rights declared the algorithmic system Systeem Risico Indicatie (SyRI) unlawful. This system, used by the Dutch government to assess the risk of social security and public revenue fraud, was found to violate the principles of proportionality, transparency, and respect for private life as enshrined in the European Convention on Human Rights.Key issues identified by the Court included the system's lack of transparency, which prevented individuals from understanding how their data were being processed and on what basis risk assessments were made. Moreover, the Court concluded that the system failed to strike a fair balance between the objective of fraud prevention and the protection of citizens' privacy, ultimately deeming it disproportionate. This judgment highlights the necessity for AI systems to uphold fundamental rights, especially in contexts where they may have direct and significant impacts on individuals' lives (Fernández, 2020).

[26] A significant case concerning automated decision-making is *Arkansas Department of Human Services v. Ledgerwood*, 530 S.W.3d 336 (Ark. Sup. Ct. Nov. 9, 2017). The Arkansas Supreme Court examined the legality of the Department of Human Services' (DHS) use of an automated system to determine eligibility and benefit levels under the Medicaid program. The plaintiff, Ledgerwood, challenged the DHS decision, alleging that the AI system used had made errors that resulted in an improper reduction of her benefits. The ruling identified several critical flaws in the AI system. First, it highlighted errors in data collection and processing, as the system failed to adequately handle the specifics of Ledgerwood's case, leading to a misinterpretation of her financial and medical circumstances. Second, the system was found to lack transparency and explainability, providing no clear rationale for its decisions and thereby hindering beneficiaries' ability to understand or appeal them. Additionally, the system exhibited deficiencies in personalization, applying generic rules automatically without sufficiently accounting for complex or exceptional cases that required a more nuanced analysis. Finally, the Court emphasized that these system failures affected the fundamental rights of beneficiaries by improperly reducing their access to essential healthcare services. The Court concluded that the use of the automated system, in the absence of adequate oversight and robust mechanisms for error correction, was insufficient to ensure fairness and accuracy in benefit determinations. It therefore ordered the DHS to review and improve its automated decision-making processes.

specifically addressing automated decisions within the GDPR framework, with particular focus on the interpretation and application of Article 22.

In the examined case, the CJEU judgment of December 7, 2023, involves SCHUFA, a German entity that employs algorithms to calculate a probabilistic value regarding individuals' creditworthiness. This value is subsequently utilized by financial institutions, such as banks, to make decisions related to credit granting. A German court requested that the CJEU clarify whether such a process constitutes an «automated decision» as stipulated in Article 22 of the GDPR, which regulates individuals' rights against decisions adopted exclusively through automated data processing.

The CJEU adopts a broad interpretation of the concept of «automated decision», determining that SCHUFA's generation of a probability value falls within this category. The underlying rationale is that such a value substantially influences third-party decisions, which can considerably affect the lives of the individuals involved (Cotino Hueso, 2024: 3).

This interpretation carries great significance, as it extends the protection offered by Article 22 of the GDPR. It thus ensures that the contemplated safeguards apply not only to completely automated decisions but also to those in which algorithms play a determining role, even if the final decision is made by a human being.

It is necessary to concur with Cotino Hueso's observations, who emphasizes that, despite Article 22 being part of the GDPR, its principles will play a fundamental role, in general terms, in the regulation of artificial intelligence in the European Union (Cotino Hueso, 2024: 4) and, as a personal addition, in the influence of algorithmic biases.

The CJEU judgment establishes a highly relevant precedent by expanding the scope of protection under Article 22 of the GDPR, ensuring that the right to non-discrimination of individuals or groups is safeguarded against decisions influenced by automated processes, even when human intervention occurs in the final phase. This approach reinforces safeguards against potential adverse effects of automation and lays the groundwork for more robust and protective regulation in the field of artificial intelligence and automated decisions.

Within the framework of our analysis, the VioGén system holds particular interest in Spain—an automated mechanism developed in 2007 by the Ministry of Interior to automatically assess the risk that a woman reporting gender-based violence may be assaulted again. The VioGén system is activated when a woman files a complaint at the police station, where a standardized questionnaire is applied and processed by an algorithm designed to determine the victim's personal risk level, classifying it into categories ranging from «not appreciated» to«extreme». The resulting assessment is used to decide the necessary police measures to protect the victim.

On September 30, 2020, the Administrative Chamber of the National Court ruled against the Ministry of Interior for insufficient protection provided by the Civil Guard to a woman who requested a protection order. The initial questionnaire assessment established a «low» risk, and without additional investigations, the agents determined that the risk was «not appreciated». This assessment influenced the judicial decision to deny the protection order, which culminated in the victim's death. The Chamber emphasized that police intervention in gender-based violence cases should not be reduced to formal aspects, such as initial attention, information about rights, and court summons, but must include effective and priority protection for victims to prevent the consequences of abuse (Borges Blázquez, 2020: 66).

The judgment also points to systemic failures in the Civil Guard and judicial police actions, by relying exclusively on automated data without adequate risk reassessment by agents specialized in gender-based violence. It was concluded that the State failed to fulfil its positive obligation of protection, denying assistance to a victim due to a system insensitive to gender particularities (Borges Blázquez, 2020: 67).

Conclusions

The algorithmic society is exerting a significant impact on human rights, which enshrine the principles of equality and prohibit any form of discrimination. In an environment where virtually all our actions are transformed into data—later processed and managed by algorithms influencing critical decisions—the effective protection of human rights hinges on the appropriate regulation of these technological developments. It is, therefore, imperative to prevent the emergence of a harmful cycle of discrimination, one that manifests both in digital environments and in everyday life.

In this context, the legal system bears an unavoidable responsibility to confront and mitigate the dynamics of inequality, bias, and discrimination that may arise. This requires not only proactive intervention, but also a redefinition of the current legal framework, as human rights theory finds itself undergoing an unprecedented paradigm shift (Belloso Martín, 2022: 67–68). A comprehensive reorganization of the fundamental rights system is essential in order to establish a direct linkage between constitutional rights and emerging digital technologies. Algorithms must not be permitted to serve as a justification for infringing upon these rights.

The phenomenon of algorithmic discrimination has emerged as one of the most pressing concerns in contemporary artificial intelligence research, fundamentally challenging our assumptions about technological neutrality and raising profound questions about the protection of fundamental rights in an increasingly automated society. While AI systems are frequently portrayed as objective arbiters of decision-making, mounting evidence suggests that these technologies often serve as amplifiers of existing societal biases rather than neutral computational tools.

Recent investigations have revealed a troubling pattern: the very data used to train these systems carries within it the historical prejudices and systematic inequalities that have long plagued human institutions. Consequently, when developers design algorithms based on such datasets, they inadvertently encode these biases into the system's core functionality. This creates a feedback loop where discriminatory practices become not just perpetuated but systematically reinforced through technological means.

The development of artificial intelligence systems cannot proceed without careful consideration of ethical implications and safety measures. Transparency has become a watchword in this field, yet achieving genuine algorithmic transparency remains elusive. Users must be able to comprehend how decisions affecting their lives are made, yet the complexity of modern machine learning models often renders such understanding practically impossible.

Privacy concerns add another layer of complexity to this challenge. The collection and processing of personal data—essential for training effective AI models—must be balanced against individuals' rights to privacy and data protection. Meanwhile, the prevention of discriminatory outcomes requires constant vigilance and sophisticated technical solutions that many organizations struggle to implement effectively.

These considerations extend beyond mere technical requirements; they represent fundamental questions about the kind of society we wish to create through technology. Public trust in AI systems depends not only on their effectiveness but on their fairness and accountability.

The European Union's Regulation (EU) 2024/1689 on artificial intelligence represents perhaps the most comprehensive attempt to address these challenges through legislative means. This regulation introduces a risk-based approach to AI governance, categorizing systems according to their potential for harm and imposing increasingly stringent requirements on higher-risk applications.

However, even this ambitious regulatory framework reveals significant gaps when subjected to practical scrutiny. The regulation's focus on pre-deployment assessment, while necessary, fails to adequately address the dynamic nature of AI systems. Algorithms continue to learn and evolve after deployment, potentially developing new

biases or amplifying existing ones in ways that initial testing could not anticipate.

The General Data Protection Regulation (GDPR) provides complementary protections, particularly through its restrictions on purely automated decision-making. Yet the intersection between GDPR requirements and AI capabilities creates new challenges for organizations seeking to comply with both the letter and spirit of data protection law while leveraging the benefits of algorithmic decision-making.

Current approaches to AI governance suffer from a fundamental mismatch between the static nature of regulatory frameworks and the dynamic characteristics of AI systems. Traditional impact assessments, conducted once during the development phase, cannot capture the evolving nature of algorithmic behavior over time.

What emerges from this analysis is the need for continuous monitoring systems that can detect emerging biases and discriminatory patterns as they develop. Such systems would require not only technical sophistication but also institutional commitment to ongoing oversight—a challenge that many organizations, particularly in the private sector, have been reluctant to embrace.

The integration of bias detection and mitigation tools into AI systems themselves represents one promising approach, though it raises new questions about who controls these safeguards and how their effectiveness can be verified independently.

When governments deploy AI systems, the stakes become even higher. Public administration has always been subject to democratic oversight and accountability mechanisms, but AI systems can obscure the decision-making process in ways that undermine traditional forms of democratic control.

Citizens have a right to understand how their government makes decisions that affect their lives, yet complex algorithmic systems can make such understanding practically impossible even for experts. This creates a fundamental tension between administrative efficiency and democratic transparency that has yet to be satisfactorily resolved.

Some jurisdictions have begun experimenting with algorithmic auditing requirements for government AI systems, but these efforts remain fragmented and often lack the technical rigor necessary to provide meaningful oversight.

The challenges posed by algorithmic discrimination cannot be solved through technical fixes alone, nor can they be adequately addressed through regulation that fails to account for the dynamic nature of AI systems. Instead, what is required is a fundamental rethinking of how we approach AI governance—one that recognizes the need for ongoing vigilance, multi-stakeholder collaboration, and institutional innovation.

This means moving beyond the current paradigm of one-time assessments and static regulatory compliance toward systems of continuous monitoring and adaptive governance. It requires not just technical expertise but also deep engagement with the communities most likely to be affected by algorithmic decision-making.

Perhaps most importantly, it demands that we resist the temptation to treat AI bias as merely a technical problem to be solved through better algorithms. The biases embedded in AI systems reflect deeper structural inequalities in society, and addressing them effectively will require confronting these underlying issues rather than simply trying to engineer them away.

The future of AI governance will likely depend on our ability to create institutions and processes that can evolve alongside the technology itself, maintaining democratic oversight and protecting fundamental rights while allowing for beneficial innovation. This represents one of the defining challenges of our technological age.

References

A Civil Society Statement. (2021). "An EU Artificial Intelligence Act for Fundamental Rights". En: *Civil Society Statement* No. 1. Versión digital disponible en: https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf

Arellano Toledo, W. (2019). "El derecho a la transparencia algorítmica en Big Data e inteligencia artificial". En: *Revista General de Derecho Administrativo*. Vol. 50. Madrid: Iustel, p. 1-25.

Arellano Toledo, W. (2023). *Derecho, ética e inteligencia artificial*. Valencia: Tirant lo Blanch.

Arroyo Jiménez, L. (2020). "Algoritmos y reglamentos". En: *Almacén del Derecho* No. 25. Versión digital disponible en: https://almacendederecho.org/algoritmos-y-reglamentos

Balaguer Callejón, F. (2023). *La Constitución del algoritmo*. Zaragoza: Fundación Manuel Giménez Abad de Estudios Parlamentarios y del Estado Autonómico.

Barrio Andrés, M. (2020). "Retos y desafíos del Estado algorítmico de Derecho". En: *Análisis del Real Instituto Elcano*. Vol. 82. Madrid: Real Instituto Elcano, p. 1-6.

Barrio Andrés, M. (2022). *Manual de Derecho digital*. Valencia: Tirant lo Blanch, 2.a edición.

Beck, U. (2008). *La Sociedad del riesgo mundial. En busca de la seguridad Perdida*. Barcelona: Paidós.

Belloso Martín, N. (2022). "La problemática de los sesgos algorítmicos (con especial referencia a los de género). ¿Hacia un derecho a la protección contra los sesgos?". En: *Inteligencia Artificial y filosofía del Derecho*. Murcia: Laborum, p. 45-78.

Boix Palop, A. (2020). "Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la administración para la adopción de decisiones". En: *Revista de Derecho Público: teoría y método*. Vol. 1, No. 1. Madrid: Centro de Estudios Políticos y Constitucionales, p. 223-269.

Borges Blázquez, R. (2020). "El sesgo de la máquina en la toma de decisiones en el proceso penal". En: *Ius et Scientia. Revista electrónica de Derecho y Ciencia*. Vol. 6, No. 2. Sevilla: Universidad de Sevilla, p. 54-71.

CEF – Laboral Social. (2022). "Derecho antidiscriminatorio. Ley 25/2022, una ley de garantías, general e integral". En: *CEF Laboral Social* No. 1. Versión digital disponible en: https://www.laboral-social.com/derecho-antidiscriminatorio-ley-15-2022-garantias-general-integral.html

Comisión Especial sobre Inteligencia Artificial en la Era Digital. (2022). "Informe sobre la inteligencia artificial en la era digital". En: *Parlamento Europeo* No. 2020/2266(INI). Versión digital disponible en: https://www.europarl.europa.eu/doceo/document/A-9-2022-0088_ES.pdf

Comisión Europea para la Eficiencia de la Justicia. (2018). "European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment". En: *CEPEJ* No. 1. Versión digital disponible en: https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c

Comisión Europea. (2018). "Comunicación de la Comisión: Inteligencia artificial para Europa". En: *COM* No. 237. Versión digital disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=COM%3A2018%3A237%3AFIN

Comisión Europea. (2021). "Proposal for a Regulation of the European Parliament and of the Council. Laying down harmonized rules on artificial intelligence (Artificial Intelligence Act)". En: *COM* No. 206. Versión digital disponible en: https://artificialintelligenceact.eu/wp-content/uploads/2024/01/AIA-Final-Draft-21-January-2024.pdf

Cotino Hueso, L. (2022). "Transparencia y Explicabilidad de la Inteligencia Artificial y Compañía (Comunicación, Interpretabilidad, Inteligibilidad, Auditabilidad, Testabilidad, Comprobabilidad, Simulabilidad...). Para qué, para quién y cuánta". En: *Transparencia y explicabilidad de la Inteligencia Artificial*. Valencia: Tirant Lo Blanch, p. 29-70.

Cotino Hueso, L. (2023). "Qué concreta transparencia e información de algoritmos e inteligencia artificial es la debida". En: *Revista Española de Transparencia*. Vol. 16. Madrid: Fundación Hay Derecho, p. 17-63.

Cotino Hueso, L. (2024). "La primera sentencia del Tribunal de Justicia de la Unión Europea sobre decisiones automatizadas y sus implicaciones para la protección de datos y el Reglamento de inteligencia artificial". En: *Diario La Ley*. Vol. 80. Madrid: Wolters Kluwer, p. 1-12.

De Asís Roig, R. (2015). *Una mirada a la robótica desde los derechos humanos*. Madrid: Dykinson.

De Asís Roig, R. (2023). "Sobre los derechos evidentes". En: *Cuadernos Electrónicos de Filosofía del Derecho*. Vol. 49. Valencia: Universitat de València, p. 602-611.

De Asís Roig, R. (2024). "De nuevo sobre Inteligencia Artificial y Derechos Humanos". En: *Derechos y Libertades*. Vol. 51, Época II. Madrid: Universidad Carlos III de Madrid, p. 25-40.

De la Cueva, J. (2018). "Código fuente, algoritmos y fuentes del Derecho". En: *El Notario del siglo XXI: revista del Colegio Notarial de Madrid*. Vol. 77. Madrid: Colegio Notarial de Madrid, p. 24-27.

De Salvador Carrasco, L. (2023). "Aprendizaje automático y protección de datos". En: *Inteligencia artificial: los derechos humanos en el centro*. Madrid: Dykinson, p. 131-154.

Esteve Pardo, J. (2003). "Ciencia y Derecho ante los riesgos para la salud. Evaluación, decisión y gestión". En: *Documentación Administrativa*. Vol. 265-266. Madrid: INAP, p. 137-149.

Fernández, C. B. (2020). "Primera sentencia europea que declara ilegal un algoritmo de evaluación de características personales de los ciudadanos". En: *Diario La Ley* No. 1. Versión digital disponible en: https://diariolaley.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAEAMtMSbH1czUwMDAyNDa3NDJUK0stKs7Mz7M1MjACC6rl5aekhrg425bmpaSmZealpoCUZKZVuuQnh1QWpNqmJeYUp6qlJuXnZ6OYFA8zAQCfSdkrYwAAAA==WKE

Flores Anarte, L. (2023). "Sesgos de género en la inteligencia artificial: el Estado de Derecho frente a la discriminación algorítmica por razón de sexo". En: *Revista Internacional de Pensamiento Político – I Época*. Vol. 18. Huelva: Universidad de Huelva, p. 95-120.

Gómez Sánchez, Y. (2004). "La dignidad como fundamento de los derechos: especial referencia al derecho a la vida". En: *Bioética: La cuestión de la dignidad*. Madrid: Publicaciones de la Universidad Pontificia de Comillas, p. 75-104.

Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial. (2019). "Directrices éticas para una IA fiable". En: *Comisión Europea* No. 1. Versión digital disponible en: https://digital-strategy.ec.europa.eu/es/library/ethics-guidelines-trustworthy-ai

Grupo de Trabajo sobre Protección de Datos del artículo 29. (2018). "Directrices del Comité Europeo sobre Protección de Datos referentes a decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679". En: *Comisión Europea* No. 1. Versión digital disponible en: https://www.aepd.es/documento/wp251rev01-es.pdf

Herranz, A. (2019). "De qué serán capaces la inteligencia artificial y el machine learning en 10 años: los mayores expertos nos responden". En: *Xataca* No. 1. Versión digital disponible en: https://www.xataka.com/robotica-e-ia/que-seran-capaces-inteligencia-artificial-machine-learning-10-anos-mayores-expertos-nos-responden

IALAB. (2021). "Sesgos algorítmicos de Género". En: *Laboratorio de Innovación e Inteligencia Artificial de la Universidad de Buenos Aires* No. 1. Versión digital disponible en:https://www.researchgate.net/publication/359115499_Sesgos_algoritmicos_de_genero/link/6228cf3697401151d20a1a41/download?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19

Jiménez-Castellanos Ballesteros, I. (2023). "Decisiones automatizadas y transparencia administrativa: nuevos retos para los derechos fundamentales". En: *Revista Española de Transparencia*. Vol. 16. Madrid: Fundación Hay Derecho, p. 191-215.

Laykite, M. (2024). "Reflexión sobre los derechos fundamentales en la nueva Ley de la Inteligencia Artificial". En: *Derechos y Libertades*. Vol. 51, Época II. Madrid: Universidad Carlos III de Madrid, p. 151-175.

Merchán Murillo, A. (2019). "Inteligencia artificial y Blockchain: retos jurídicos en paralelo". En: *Revista General de Derecho Administrativo*. Vol. 50. Madrid: Iustel, p. 1-20.

Ministerio de Asuntos Económicos y Transformación Digital. (2020). *Estrategia Nacional de Inteligencia Artificial*. Madrid: Ministerio de Asuntos Económicos y Transformación Digital.

Ministerio de Igualdad. (2022). *Plan Estratégico para la Igualdad Efectiva de Mujeres y Hombres 2022-2025*. Madrid: Ministerio de Igualdad.

Miró Llinares, F. (2019). "El modelo policial que viene: mitos y realidades del impacto de la inteligencia artificial y la ciencia de datos en la prevención policial del crimen". En: *Libro Blanco de la Prevención y Seguridad Local Valenciana*. Valencia: Instituto Valenciano de Seguridad Pública y Emergencias, p. 98-113.

Montano, P. J. (1995). "La responsabilidad de médicos y científicos ante las nuevas tecnologías biogenéticas". En: *Cuadernos de Bioética*. Vol. 1. Madrid: Asociación Española de Bioética y Ética Médica, p. 81-91.

Padilla Ruiz, P. (2019). "Inteligencia artificial y Administración pública. Posibilidades y aplicación de modelos básicos en el procedimiento administrativo". En: *Consultor de los ayuntamientos y de los juzgados: Revista técnica en administración local y justicia municipal*. Vol. 10. Madrid: Wolters Kluwer, p. 96-104.

Panea Márquez, J. M. (2005). "La imprescindible dignidad". En: *Bioética y derechos humanos. Implicaciones sociales y jurídicas*. Sevilla: Universidad de Sevilla, p. 17-28.

Pérez De la Fuente, O. (2005). *Pluralismo cultural y derechos de las minorías una aproximación iusfilosófica*. Madrid: Dykinson.

Pérez Ugena, M. (2024). "La inteligencia artificial: definición, regulación y riesgos para los derechos

fundamentales". En: *Estudios de Deusto*. Vol. 72, No. 1. Bilbao: Universidad de Deusto, p. 307-337.

Rallo Lombarte, A. (2020). "Una nueva generación de derechos digitales". En: *Revista de Estudios Políticos*. Vol. 187. Madrid: Centro de Estudios Políticos y Constitucionales, p. 101-135.

Rubio, I. (2018). "Amazon prescinde de una inteligencia artificial de reclutamiento por discriminar a las mujeres". En: *El País* No. 1. Versión digital disponible en: https://elpais.com/tecnologia/2018/10/11/actualidad/1539278884_487716.html

Sánchez Barrilao, J. F. (2016). "El derecho constitucional ante la era de ultrón: la informática y la inteligencia artificial como objeto constitucional". En: *Estudios de Deusto*. Vol. 64, No. 2. Bilbao: Universidad de Deusto, p. 225-258.

Sánchez-Yara, G. (2021). "La AEPD impone a Mercadona una multa de 3,15 millones de euros por varias infracciones del RGPD relacionadas con su sistema de reconocimiento facial". En: *RGPD Blog* No. 1. Versión digital disponible en: https://rgpdblog.com/la-aepd-impone-a-mercadona-una-multa-de-315-millones-de-euros-por-varias-infracciones-del-rgpd-relacionadas-con-su-sistema-de-reconocimiento-facial-parte-1-2/

Simón Castellano, P. (2023). *La evaluación de impacto algorítmico en los derechos fundamentales*. Navarra: Aranzadi.

Torres Jarrín, M. (2021). "La UE & la gobernanza ética de la Inteligencia Artificial: Inteligencia Artificial & Diplomacia". En: *Cuadernos Salmantinos de Filosofía*. Vol. 48. Salamanca: Universidad de Salamanca, p. 213-234.