


# Advancing transparent algorithmic governance. A Case study in bias auditing

**José Luis Bustelo Gracia**  
ESERP, Digital Business and Law School 

<https://dx.doi.org/10.5209/cgap.97604>

Received: 20/08/2024 • Accepted: 26/03/2025

**EN Abstract.** This article examines the necessity of implementing bias audits in algorithmic systems, particularly within the context of public governance. With the increasing use of artificial intelligence (AI) in governmental decision-making, concerns about the transparency and fairness of these systems have arisen. The paper analyzes various European regulations, such as the Artificial Intelligence Act (AI Act), and case studies that highlight the persistence of biases in algorithms, emphasizing the importance of a proactive regulatory approach and systematic audits. Additionally, it discusses the ethical and social implications of algorithmic governance and proposes solutions to mitigate associated risks.

**Keywords:** Algorithmic governance, transparency, bias audit, artificial intelligence.

## Hacia una gobernanza algorítmica transparente: auditoría de sesgo. Estudio de caso

**ES Resumen.** Este artículo explora la necesidad de implementar auditorías de sesgo en sistemas algorítmicos, particularmente en el contexto de la gobernanza pública. Con el incremento del uso de la inteligencia artificial (IA) en la toma de decisiones gubernamentales, surge la preocupación sobre la transparencia y equidad de estos sistemas. Se analizan diversas normativas europeas, como el Reglamento de Inteligencia Artificial (AI Act), y estudios de caso que evidencian la persistencia de sesgos en algoritmos, destacando la importancia de un enfoque proactivo en la regulación y la implementación de auditorías sistemáticas. Además, se discuten las implicaciones éticas y sociales de la gobernanza algorítmica y se proponen soluciones para mitigar los riesgos asociados.

**Palabras Clave:** Gobernanza algorítmica, transparencia, auditoría de sesgo, inteligencia artificial.

**Summary:** 1. Introduction. 2. Methodology. 3. Theoretical framework. 3.1. Introduction to algorithmic governance. 3.2. Ethics and human rights. 4. European regulations on algorithmic transparency. 5. Case study and critical analysis: Audit of a job recruitment algorithm. 6. Critical analysis. Discussion. 7. Conclusions. 8. References.

**How to cite:** Bustelo Gracia, J. L. (2025). Advancing transparent algorithmic governance. A Case study in bias auditing, in *Cuadernos de Gobierno y Administración Pública* 12(1), e97604. <https://dx.doi.org/10.5209/cgap.97604>

## 1. Introduction

The increasing deployment of artificial intelligence (AI) systems in public governance has raised significant challenges in terms of fairness, transparency, and accountability. As algorithms become integrated into decision-making, their ability to amplify pre-existing biases and generate new forms of discrimination has become a critical concern. Algorithmic transparency and the implementation of bias audits are, therefore, essential tools for mitigating these risks and ensuring that artificial intelligence is used fairly and ethically.

This article falls within the framework of the discussions on governance and public administration addressed in *Cuadernos de Gobierno y Administración Pública*, particularly in relation to the development of new governance models, bureaucratic capacity, and the impact of technology on public management. Previous literature has explored various dimensions of governance (Aguilar Villanueva, 2014), the different types and approaches that exist (Arellano Gault, 2014), as well as the challenges faced by modern society in implementing effective models (García Magariño, 2016). More recently, the debate has broadened to include bureaucratic capacity and its relationship with institutional effectiveness (Guerrero García, 2023) and the integration of artificial intelligence into administrative processes (Crespo-González et al., 2024). In this sense, this paper seeks to contribute to this line of research by analyzing how algorithmic governance redefines decision-making and the role of co-responsibility in the public sphere.

The concept of "algorithmic governance" refers to the use of algorithms to manage and make decisions in governmental and corporate contexts. However, these systems, while promising efficiency and objectivity, often lack transparency, which makes it difficult to identify and correct inherent biases. Several studies have shown that algorithms can perpetuate inequalities, either due to the nature of the data used or design decisions made by programmers (Dekker et al., 2022).

The development of institutional capacity is a key element in contemporary governance, especially in the context of digitalization and the use of technological tools in decision-making. From a theoretical and conceptual perspective, algorithmic governance should be understood as a dynamic process in which institutions must adapt to new forms of management based on data and automation. This requires not only technological infrastructure, but also clear regulatory frameworks and training strategies that allow governmental actors to interpret, monitor, and correct automated decisions when necessary. Strengthening these institutional capacities is essential to ensure that technology not only optimizes processes but also aligns with democratic and equity principles.

The need for bias audits in AI systems has been recognized as an essential practice for responsible governance. Recent research highlights that incorporating independent audits can significantly reduce bias in AI systems, thereby improving fairness in algorithmic decision-making (Conitzer et al., 2022). Furthermore, the case of the implementation of Local Law 144 in New York highlights the challenges and benefits of establishing bias audit regimes as a legal standard, which demonstrates the importance of such measures in improving transparency and public trust in technology (Groves et al., 2024).

Nevertheless, for these audits to be effective, it is essential to understand that bias in algorithms does not only come from the data, but also from the decisions made at all stages of the system's development and deployment. Current legislation in several countries, including the proposed European Union Artificial Intelligence Regulation, is being adapted to address these issues, proposing the need for broader regulation covering all aspects of the algorithm lifecycle.

It is essential to differentiate between algorithms, programs, and data, as each of these elements can introduce bias into a system. An algorithm is a sequence of instructions designed to solve a specific problem, while a program is the practical implementation of one or more algorithms in a programming language. Data, on the other hand, represents the input information that an algorithm processes and can be a critical source of bias if it is not representative or contains inherent prejudices. In this sense, bias auditing must address these three components in a differentiated manner for effective evaluation.

Therefore, this article explores the different dimensions of bias auditing in algorithmic governance, examining both existing regulations and case studies that highlight the urgency of adopting a more transparent and responsible approach to the development and implementation of AI.

## 2. Methodology

This study uses a qualitative and documentary approach to analyze the implementation and effectiveness of bias audits in algorithmic systems within the context of public governance. The following methodological stages were carried out:

**Literature review:** An exhaustive review of the academic literature and relevant European regulations, such as the Artificial Intelligence Act (AI Act), was conducted. This review helped to identify the main challenges associated with transparency and fairness in algorithmic systems, as well as current regulatory proposals to mitigate these problems.

**Case study:** Significant case studies were selected and analyzed, such as the implementation of bias audits in automated recruitment platforms. These cases provide empirical evidence on the effects of algorithmic biases on decision-making and the effectiveness of audits in mitigating them.

**Critical analysis:** An analytical approach was applied to assess the ethical, technical, and legal dimensions of algorithmic governance. This analysis included a review of the structure and training of specific algorithms, as well as an evaluation of their impact in terms of fairness and transparency.

**Proposed improvements:** Based on the findings, regulatory and technical improvements were proposed that could contribute to more transparent and fair algorithmic governance. These recommendations are based on the need for proactive regulation and the incorporation of systematic audits as standard practice in the development and deployment of AI systems.

This methodological approach combines theoretical analysis with practical evaluation, allowing for a comprehensive understanding of the challenges and solutions in the governance of algorithmic systems.

### 3. Theoretical framework

#### 3.1. Introduction to algorithmic governance

Algorithmic governance refers to the growing reliance on automated decision-making systems in the field of public administration and government decision-making. These systems fueled by big data and developed using artificial intelligence (AI) models promise to improve efficiency and fairness in public policy management. However, their implementation also raises serious ethical concerns, particularly around transparency, accountability, and the inherent bias in algorithms.

In recent decades, governments have increasingly adopted algorithmic tools to optimize processes ranging from resource allocation to social needs prediction. The ability of these systems to process large volumes of data and generate fast and accurate decisions has been hailed as a revolutionary advance in public governance (Dekker et al., 2022). However, this revolution has also sparked debates about the potential adverse effects of automation on decision-making. Algorithms are used in a wide variety of applications within the government sphere. From the allocation of housing and social resources to law enforcement, automated systems have demonstrated their ability to improve administrative efficiency. For example, in the justice sector, algorithmic systems have been implemented to assist in judicial decisions, such as assessing the risk of recidivism in convicted individuals.

A notable example of algorithmic governance is the use of prediction systems in social services management. These systems analyze historical and current data to predict the future needs of the population, allowing governments to allocate resources more effectively. However, the accuracy of these predictions depends largely on the quality of the data and models used, which can lead to significant errors if not managed properly (Groves et al., 2024). Although algorithmic governance offers clear advantages, it also presents several challenges that must be addressed to ensure its effectiveness and fairness. One of the main problems is algorithmic bias, which can arise at different stages of the algorithm lifecycle, from data collection to model development and implementation.

Bias in algorithms can have serious implications, especially when applied in sensitive contexts such as justice or the allocation of social benefits. A recent study highlighted how algorithms used in the criminal justice system in the United States showed significant bias against racial minorities, leading to unfair and disproportionate decisions in terms of sentencing and parole (Conitzer et al., 2022). This problem is not limited to the United States; around the world, a lack of diversity in development teams and reliance on biased historical data have contributed to the perpetuation of inequalities through technology.

Another major challenge is the lack of transparency in algorithmic systems. Many of these systems operate as "black boxes," meaning that their internal workings are not understandable to users or even to the developers themselves. This raises serious accountability issues, as it is difficult to determine who is responsible when an algorithm makes a wrong or unfair decision (Groves et al., 2024). The opacity of these systems also makes it difficult to implement control and oversight mechanisms, increasing the risk of abuse. Therefore, transparency is a fundamental principle in public governance, and its importance is amplified in the context of algorithmic management. The ability of citizens and regulatory bodies to understand how algorithms work and how decisions are made is crucial to ensure these systems are used fairly and responsibly. In this regard, various measures have been proposed to improve transparency in algorithmic governance.

#### 3.2. Ethics and human rights

One of the most promising measures is the implementation of algorithm audits, which can help identify and correct biases, as well as ensure that systems comply with ethical and legal standards. These audits should be conducted by independent entities and must be made available to the public so that citizens can trust the fairness and justice of the algorithmic systems used by the government (Dekker et al., 2022). In addition, the adoption of open standards and the publication of algorithm source codes are important steps toward greater transparency. This would allow researchers and civil society organizations to examine the systems and propose

improvements, thereby contributing to more inclusive and participatory governance. It is suggested that algorithmic governance not only affects administrative efficiency but also has profound implications for human rights. The adoption of algorithmic systems in government management can threaten fundamental rights such as privacy, equality, and access to justice. For example, the use of algorithms in surveillance and security has raised concerns about invasion of privacy and racial and ethnic discrimination.

It is essential that any implementation of algorithms in governance be carried out with a human rights-based approach that considers both the benefits and potential risks of these technologies. This involves not only conducting impact assessments prior to implementation, but also taking corrective measures when rights violations are detected. Despite the challenges, algorithmic governance has the potential to positively transform public administration. With the development of new technologies, such as explainable artificial intelligence (XAI), which allows users to better understand algorithmic decisions, and the use of blockchain technologies to ensure data transparency and integrity, it is possible to imagine a future in which algorithms are used fairly and efficiently to improve the lives of citizens. It is a promising area, it is co-creation of algorithms, where developers work collaboratively with government actors, civil society organizations, and citizens to design systems that better reflect the needs and values of the community. This participatory approach not only improves transparency and trust in systems but also helps to ensure that algorithms are developed and used in an ethical and responsible manner (Dekker et al., 2022).

The implementation of algorithmic tools in governance raises ethical and moral dilemmas that require detailed analysis. The objectivity sought through the systematization of decision-making is not without risks, especially if the co-responsibility of the different actors involved in these processes is disregarded. Co-responsibility implies that system designers, public officials, and citizens must actively participate in monitoring and controlling the impact of these tools. Without a solid ethical framework, algorithmic decisions could reinforce existing inequalities or generate unwanted effects on the affected population. An example of this is the use of algorithms in the allocation of social resources: while they can optimize distribution, they can also overlook contextual factors that require human judgment.

Algorithmic management in governance represents both an opportunity and a significant challenge for modern governments. While these systems can improve efficiency and accuracy in decision-making, they also raise serious concerns about bias, transparency, and the protection of human rights. It is essential that governments take a proactive approach to addressing these challenges by implementing measures that ensure transparency, accountability, and fairness in the use of algorithms. Only then can the full potential of artificial intelligence in public governance be leveraged, while ensuring that the fundamental principles of democracy and human rights are respected. It is further noted that the growing incorporation of algorithmic systems into governmental and commercial decision-making has brought with it the challenge of ensuring the transparency and fairness of these technologies. The need for ethically centered algorithmic governance has underscored the importance of audit mechanisms to identify and correct potential biases present in algorithmic models. This approach not only improves fairness in decision-making but also promotes public trust in these technologies. It is true that one of the main political challenges in implementing algorithms in governance is the inherent opacity of many of these systems, which makes it difficult to identify biases and errors. Algorithmic transparency involves making the processes and decisions made by algorithms visible, allowing for their analysis and audit. According to Giunchiglia et al. (2021), diversity is a key factor in driving transparency, as it allows for a better understanding and addressing of biases that may arise in algorithmic systems. Furthermore, transparency is also seen as a prerequisite for building trust and legitimacy in algorithmic systems, especially when they affect fundamental human rights such as privacy and non-discrimination. Bias auditing has become an essential practice to ensure that algorithms do not perpetuate or amplify inequalities. Lam et al. (2024) proposes a criteria-based audit framework to ensure that algorithmic systems comply with ethical and governance standards, adapting these principles to emerging legislative regulations and those that may come, such as bias audits in both public and private hiring algorithms. These audits allow for the systematic identification of critical points where biases may manifest, providing a solid basis for making corrective decisions and preventing harm to vulnerable groups. Explaining and differentiating between explainability and auditability in algorithmic systems becomes one of the major challenges. Although both concepts are related to transparency, they serve different functions. Explainability focuses on providing end-users with a clear understanding of how and why an algorithm makes specific decisions, while auditability is geared towards internal and external analysis by experts to evaluate the accuracy and fairness of a system. Springer and Whittaker (2021) emphasize that these two elements cannot be addressed simultaneously through a single implementation of transparency, as they serve different purposes.

Globally, governments and organizations have begun to establish policies to regulate the use of algorithms, with a particular focus on transparency and bias auditing. Policies in the United States, although still limited, have begun to incorporate transparency and auditing requirements to address bias issues, with initiatives including temporary bans and mandatory assessments. However, there is still a long way to go to implement more

legislation that can effectively mitigate the risks associated with the use of algorithms in critical sectors such as finance and labor recruitment (Qureshi et al., 2024). In these cases, the audit of algorithmic systems is not only a detection tool, but also a mechanism for continuously correcting and improving models. Conitzer et al. (2022) present a study that shows how independent auditing can reduce bias in algorithmic systems, positively impacting fairness in decision-making and providing a basis for future innovations in algorithmic governance. This approach becomes a crucial element to ensuring that systems operate within ethical limits and respect fundamental rights. To move toward more transparent and fair algorithmic governance, it is necessary not only to improve technology, but also to strengthen regulatory frameworks and promote collaboration among multiple actors. The active participation of civil society organizations, alongside rigorous evaluations by independent auditors, can ensure that systems are designed and operated in a way that minimizes bias and maximizes fairness and justice. This also implies the need for continuous training for developers and policy makers, ensuring they understand the complexities and inherent risks of algorithmic systems.

The role of the user in interpreting the results generated by the artificial intelligence system is crucial. Depending on the context, the user may act as a mere recipient of information or as a critical evaluator who validates the results before their application. For example, in the field of justice, a recidivism prediction algorithm should be used as a support tool and not as a definitive, unsupervised human decision. In contrast, in more operational tasks, such as the automatic classification of email, human intervention may not be necessary. To illustrate this, consider an AI-assisted medical diagnosis system: although the algorithm can generate a probability of disease based on the patient's symptoms, the final decision must be made by a healthcare professional who contextualizes the information.

Technological tools applied to governance encompass a wide variety of systems, from predictive models to real-time data analysis platforms. A key aspect of their operation is their ability to process large volumes of information and detect patterns that can be used to optimize decision-making. However, these tools also present challenges, such as the opacity in their decision-making criteria, the possibility of replicating pre-existing biases, and the difficulty of establishing effective supervision mechanisms. For example, in the administration of justice, risk assessment algorithms can influence the determination of sentences or precautionary measures, which raises questions about their impartiality and transparency. To mitigate these risks, it is essential to establish clear criteria for validation and human review in their implementation.

These tools used in governance processes include artificial intelligence systems for predicting trends in public policy, as well as blockchain platforms for transparent data management and smart contracts. All of them operate using algorithmic models that analyze large volumes of information to optimize decision-making. One example of this is predictive analysis systems in urban management, which make it possible to anticipate demand for public services and optimize resource allocation in real time. Another relevant case is the use of citizen participation platforms based on collective intelligence, which facilitate interaction between citizens and governments through online voting and deliberation mechanisms. However, for their proper implementation, these technologies require adequate data infrastructure, interoperability protocols, and regulatory frameworks to ensure their proper functioning and ethical use.

#### 4. European regulations on algorithmic transparency

The Artificial Intelligence Act (AI Act) is a pioneering legislative proposal within the European Union's regulatory framework, aimed at regulating the use, development, and commercialization of AI systems in Europe. It is currently under review and negotiation in various legislative bodies, including the European Parliament and the Council of the EU, where amendments are being debated and proposed. Although the regulation has not yet been definitively implemented, it is expected to come into force in 2025, establishing a binding regulatory framework for all Member States. This framework seeks to ensure that AI systems are safe, ethical, and respect fundamental rights such as privacy and non-discrimination (Busuioc et al., 2022; Gstrein et al., 2024).

The AI Act is positioned as the cornerstone of the EU's regulatory strategy on artificial intelligence. This regulation classifies AI systems according to their level of risk, ranging from low to unacceptable, prohibiting those that could compromise fundamental rights or public safety. Special attention is paid to applications considered high risk, such as facial recognition in public spaces, decision-making systems in critical sectors such as labor, finance, or education, and those related to critical infrastructure (Engelmann, 2023; Sovrano et al., 2022). In this regard, the regulations impose obligations on AI providers and users to ensure transparency through mechanisms such as adequate documentation, data traceability, and the explainability of the algorithms used (Turksen et al., 2024). These requirements not only seek to prevent bias and discrimination, but also to guarantee privacy in automated decision-making (Gyevnar et al., 2023; Calero Valdez et al., 2024).



Algorithmic transparency is positioned as a central pillar within the proposed regulations. The regulations focus particularly on high-risk AI systems that include applications in health, education, employment, and public administration, where not only transparency is required, but also independent audits to verify compliance with established standards. In addition, significant penalties are imposed on organizations that fail to comply with these rules, reinforcing the EU's commitment to the protection of rights and the promotion of public trust. Zharova (2023) emphasizes that transparency is not limited to the accessibility of information, but must also ensure comprehensibility for end-users, which is essential to avoid bias and enable informed decisions. In this context, explainability tools, such as logic-based explanation mechanisms, are crucial for complying with regulatory requirements.

Looking ahead, the AI Act represents a comprehensive effort to address current and future challenges of algorithmic governance across multiple sectors. This regulatory framework not only anticipates the ethical and social risks of the massive deployment of AI, but also proposes solutions that harmonize technological innovation with the protection of fundamental rights (Wörsdörfer, 2023; Laux et al., 2023). With the implementation of the AI Act and the laws planned for 2025, Europe is positioning itself as a leader in AI regulation, establishing a model of transparent and equitable algorithmic governance for the future.

The European regulation on algorithmic transparency is not limited to imposing technical rules but also establishes ethical and social principles that seek to align the development of artificial intelligence with the fundamental values of the European Union. In this regard, the AI Act becomes a key tool for ensuring that emerging technologies are implemented responsibly, considering their impact on society and human rights. This multidimensional approach involves not only compliance with technical requirements, but also the creation of a framework of shared responsibility among developers, users, regulators, and auditors, which reinforces trust in AI systems.

In addition to high-risk systems, the AI Act also proposes measures to manage emerging technologies whose application has not yet been widely explored, such as generative AI and autonomous systems. These technologies present unique challenges in terms of transparency and traceability, which has led the EU to include specific provisions to ensure that their use remains under control and that risks are managed appropriately. The regulation also provides for the constant updating of requirements and standards, recognizing the rapid pace at which technology evolves and the need to adapt regulation to new challenges.

Another important aspect is interoperability between regulatory frameworks. The EU seeks to ensure that the AI Act is compatible with other international regulations and with the principles established in agreements such as the General Data Protection Regulation (GDPR). This alignment facilitates global cooperation in AI governance and allows Europe to exert significant influence on global regulation, promoting standards that combine innovation with the protection of rights. This integrative vision not only strengthens the EU's position in the technological sphere but also sets precedents for the development of public policies that balance progress and responsibility.

As a negative aspect to highlight within the European regulations on algorithmic transparency, although well-intentioned, is that it faces significant challenges in terms of implementation and effectiveness. While it establishes the need for algorithms to be comprehensible and auditable, the inherent complexity of many artificial intelligence models makes this transparency difficult to achieve in practice. In addition, the regulations may impose additional burdens on companies, especially small and medium-sized ones, which may not have the necessary resources to comply with the strict transparency requirements. This could create an environment where only large corporations can comply with the regulations, which could limit innovation and competition in the sector. Therefore, the regulation does not sufficiently address the challenges related to the interpretation and use of the information provided by algorithms, which can lead to a false sense of security regarding the reliability and fairness of automated decisions.

Finally, the AI Act promotes a culture of transparency at all levels of the AI system lifecycle. From the design phase to implementation and use, the regulation requires comprehensive documentation that allows authorities and the public to understand how and why certain automated decisions are made. This approach reduces the gap between technology and citizens, empowering users with clear and accessible information. In short, European regulations on algorithmic transparency not only respond to the need to regulate a rapidly developing technology but also lay the foundations for algorithmic governance based on ethics, transparency, and inclusion, ensuring that technological innovation advances in parallel with respect for fundamental rights and the promotion of a more just and equitable society.

It is important to note that for an auditor to effectively evaluate an algorithm, it is necessary to define clear criteria and specific methodologies. Auditing can be carried out through various strategies, such as the analysis of results across different population subgroups (to detect discriminatory biases), the inspection of the source code, and the review of the data collection and selection process for training. A recommended approach is black box auditing, in which the auditor analyzes the inputs and outputs of the system without access to the internal code, comparing them with human decisions to evaluate inconsistencies. Alternatively, white box auditing allows for the inspection of the internal structure of the model and its decision parameters. A practical example would be the audit of a personnel selection system: it could be analyzed whether certain demographic groups systematically receive lower scores, suggesting possible bias.

## 5. Case study and critical analysis. Audit of a job recruitment algorithm

A case study is presented focusing on the audit of a recruitment algorithm, an area in which inherent biases can have significant effects in terms of equal opportunities and fairness. Algorithmic auditing on recruitment platforms is crucial for mitigating inherent biases that can perpetuate inequalities by prioritizing certain characteristics over others. Research has shown that automated systems, such as the recruitment algorithms of a large technology company, have replicated gender biases by favoring men over women in the technology field, contributing to the underrepresentation of women in this field. Furthermore, recent studies highlight the need for both internal and external audits to identify and correct these discrepancies in automated hiring systems (Ajunwa, 2019). The algorithm analyzed in this study is used in selection processes to recommend candidates based on their resumes and online profiles. The objective of the audit was to evaluate the fairness and transparency of the system, especially in terms of gender and ethnicity. Such tools have become common on recruitment platforms, but recent research has revealed that they can perpetuate biases by prioritizing certain characteristics over others.

In this case, it was discovered that the algorithm showed a systematic preference for male candidates for technical roles, reflecting biases in both the training data and the model structure. According to Albaroudi et al. (2024), biases in hiring algorithms can arise due to historical data that reflects patterns of discrimination and stereotypes in labor recruitment (Albaroudi et al., 2024). Despite efforts to eliminate this type of biases using techniques such as vector space correction in deep learning models, the study found that technical interventions are often insufficient to completely eradicate these problems. Additionally, the audit identified that the algorithm penalized candidates with names associated with ethnic minorities, resulting in a lower probability of being recommended for interviews. This algorithmic discrimination illustrates how systemic biases can be amplified when decisions are automated on a large scale. In their research, Hickok et al. (2022) highlight the importance of conducting comprehensive audits that address not only the final results, but also the way data is processed and transformed throughout the algorithmic pipeline (Hickok et al., 2022). Based on these findings, it is important to reiterate that the problem of bias in recruitment algorithms is not limited to equity in access to job opportunities but also raises fundamental questions about the ethical responsibility of the organizations that implement these systems. The growing reliance on automated tools in hiring decisions has sparked a debate about the need for a more robust regulatory framework that addresses the ethical and legal implications of using artificial intelligence in selection processes. As suggested by Binns (2018) and Mittelstadt et al. (2016), it is essential that companies implement algorithmic governance practices that include continuous audits and the participation of various stakeholders to ensure transparency and accountability in the use of these systems.

Similarly, recent literature highlights the importance of incorporating interdisciplinary approaches into algorithm auditing. Collaboration between experts in ethics, technology, and law is crucial for designing audits that not only identify technical biases but also assess the social and ethical impact of algorithms in specific contexts. These holistic approaches allow for a deeper understanding of how algorithms can affect different social groups and facilitate the creation of solutions that address not only the immediate causes of bias but also the structural dynamics that perpetuate it.

Another key consideration is the role that transparency plays in mitigating algorithmic bias. The opacity of black box algorithms— i.e., those whose internal processes are neither transparent nor understandable to users— makes it difficult to detect and correct bias. According to Diakopoulos (2016), the lack of transparency in automated systems could erode trust in hiring processes, especially if those affected do not have access to mechanisms for appealing or correcting unfair decisions (Diakopoulos, 2016). In this regard, initiatives that promote the explainability and accessibility of algorithms, such as those advocated by Wachter, Mittelstadt, and Floridi (2017), are essential to ensure that candidates can understand and question automated decisions that affect their access to job opportunities (Wachter, Mittelstadt, & Floridi, 2017).

It is important to consider that algorithmic auditing should not be seen as a single or definitive solution, but rather as part of a broader approach to algorithmic governance. As Lepri et al. (2018) suggest, effective algorithmic governance should include not only technical audits, but also the implementation of organizational policies that promote diversity and inclusion at all stages of the hiring process, from data collection to the final evaluation of candidates (Lepri et al., 2018). Furthermore, it is crucial that these policies be backed by a long-term organizational commitment to equity, which requires the active participation of all levels of the organization, from algorithm developers to strategic decision-makers.

For all these reasons, auditing hiring algorithms, as presented in this case study, highlights the need for multifaceted and collaborative approaches to address inherent biases in automated systems. Organizations must take a proactive stance, not only in identifying and correcting biases, but also in creating more inclusive and equitable work environments that reflect a genuine commitment to social justice in the digital age.

Another example or case study is provided by Mensah (2023), which analyzes how hiring algorithms can perpetuate biases in candidate selection. This study highlights that even with multiple technical interventions, gender and ethnic biases prevail due to historical patterns and algorithmic opacity. The research emphasizes the importance of regular audits to identify these biases, suggesting that independent external audits are essential to ensure transparency and fairness in automated processes. This case offers a critical insight into how, despite technical advances, auditing and transparency are fundamental components in addressing discrimination in automated hiring.

Mensah (2023) provides a comprehensive framework for understanding the challenges and limitations faced by these technological tools. His study examines the effects of persistent biases in algorithms used by companies for candidate selection. Despite efforts to mitigate these biases through correction techniques, such as modifying vectors in deep learning models, the results show that gender and ethnic biases continue to manifest themselves due to historical patterns reflected in the data.

A significant issue that stands out is the need to address algorithmic opacity, known as the "black box problem." This term refers to the difficulty of understanding the internal decision-making processes of algorithms, which complicates the identification and correction of biases. In the studied case, it was observed that the algorithm systematically favored male candidates for technical roles, which is directly related to the historical predominance of men in such positions. This structural bias is exacerbated when AI systems replicate discriminatory patterns by prioritizing certain demographic characteristics over others.

Furthermore, the study shows that algorithmic discrimination is not limited to gender, but also affects candidates from ethnic minorities. In this context, it was found that candidates with names associated with minorities had a lower probability of being selected for interviews, even when their qualifications were comparable to those of other candidates. This highlights how systemic prejudices in society can be amplified when large-scale decisions are automated without adequate monitoring. Mensah (2023) argues that the solution lies not only in technical adjustments to the algorithm, but also in the implementation of regular and comprehensive audits, both internal and external. These audits must go beyond a superficial review of the final results and include a critical analysis of how data is processed and transformed throughout the algorithmic lifecycle. This includes evaluating the quality of training data, transparency in modeling, and the system's ability to provide comprehensible explanations for its decisions.

An important conclusion drawn from the study is that transparency, while essential, is not sufficient unless accompanied by explainability and accessibility. The ability to explain automated decisions clearly and comprehensively is crucial to building trust and allowing both candidates and regulators to understand and question the results. Without these mechanisms, transparency risks of being superficial, limited to the disclosure of data and codes that are inaccessible to most users.

The case study also highlights the importance of an interdisciplinary approach to algorithm auditing. Mensah suggests that collaboration between experts in technology, ethics, and law is essential to designing audit frameworks that address not only the technical aspects but also the social and ethical implications of AI systems. This approach allows for a more holistic understanding of how algorithms impact different social groups, which is critical for developing solutions that address the deep-rooted causes of bias, beyond the immediate symptoms.

Mensah's article (2023) exemplifies how auditing of hiring algorithms should not only focus on correcting technical biases, but also on creating an algorithmic governance environment that integrates transparency, explainability, and fairness at every stage of the process. The implementation of regular audits and the adoption of a holistic and multidisciplinary approach are essential to ensure that AI systems in labor recruitment do not perpetuate or amplify existing inequalities in society.

## 6. Critical analysis. Discussion

Analysis of this case reveals a series of ethical and technical issues that need to be addressed to improve the transparency and fairness of AI systems in labor recruitment. First, it is evident that transparency is not limited to the explainability of the algorithm. As Mashhadi et al. (2022) point out, incorporating fairness visualization tools can help system designers and auditors identify patterns of bias that might otherwise go unnoticed. However, transparency does not always guarantee fairness, as a system can be transparent in its operations and still produce biased results.

A critical issue is the limited access to information on how these algorithms are developed and trained. Percy et al. (2021) argues that accountability in AI systems must be supported not only by internal audits, but also by formal external accreditation processes that can guarantee impartiality. This level of scrutiny is particularly relevant in sectors such as hiring, where algorithmic decisions can affect individuals' lives directly and over long-term. The case highlights the need to design models that are not only accurate, but also that they are also fair.



According to Xiang et al. (2022), it is essential to balance accuracy and fairness in predictive models, especially when used in sensitive contexts such as hiring or education. This involves not only adjusting models to reduce bias, but also rethinking the metrics used to evaluate their performance.

On the other hand, the ability of AI systems to correct biases depends greatly on the quality of the data used. Biased training data can lead to discriminatory results, and correcting these biases after the model has been implemented is extremely difficult. This problem is intensified when algorithms are trained using historical data that reflects existing biases in society. As shown in the study by Albaroudi et al. (2024), technical interventions such as synthetic data generation or bias correction in vector space can reduce the impact of these biases, but not eliminate them entirely. This audit highlights the importance of a more comprehensive approach to algorithmic governance. The results suggest that, in addition to technical corrections, vigorous regulatory oversight is required to demand transparency and accountability at every stage of the algorithm's life cycle. Bartley et al. (2021) demonstrate how algorithmic audits can be an effective tool for identifying and correcting biases in digital platforms, but they underscore the need to implement continuous correction mechanisms to adapt to changes in usage patterns and available data.

The case study presented here highlights the complexity of ensuring transparency and fairness in AI systems. Despite advances in auditing techniques and bias mitigation, algorithmic biases remain a persistent challenge. To address this problem, it is essential to combine technical solutions with stronger governance and regulatory approaches that promote transparency and accountability. This includes not only implementing comprehensive audits, but also constantly adapting regulations to ensure that systems remain aligned with ethical values and fundamental rights.

The analysis of this case reveals a series of ethical and technical issues that need to be addressed to improve the transparency and fairness of AI systems in labor recruitment. First, it is evident that transparency is not limited to the explainability of the algorithm. As Mashhadi et al. (2022) emphasize, incorporating fairness visualization tools can help system designers and auditors identify patterns of bias that might otherwise go unnoticed. However, transparency does not always guarantee fairness, as a system can be transparent in its operations and still produce biased results.

A critical issue is the limited access to information on how these algorithms are developed and trained. Percy et al. (2021) argue that accountability in AI systems must be supported not only by internal audits, but also by formal external accreditation processes that can guarantee impartiality. This level of scrutiny is particularly relevant in sectors such as hiring, where algorithmic decisions can directly and long-term affect individuals' lives. The case highlights the need to design models that are not only accurate but also fair. According to Xiang et al. (2022), it is essential to balance accuracy and fairness in predictive models, especially when used in sensitive contexts such as hiring or education. This involves not only adjusting models to reduce bias, but also rethinking the metrics used to evaluate their performance.

Another critical point in this analysis is the role of ethics in the development and deployment of AI systems. Despite efforts to mitigate bias, the lack of a robust ethical framework can lead to partial solutions that fail to address the root causes of the problem. Cath (2018) argues that algorithmic ethics should be integrated into the early stages of system development, with an approach that combines technical and normative considerations to effectively address issues of fairness (Cath, 2018). This integration of ethics into the AI development lifecycle could help mitigate the risks associated with unfair algorithmic decisions.

The incorporation of technological tools into government decision-making significantly modifies traditional public management processes. Among the main implications are reduced decision-making times, access to more accurate information, and the possibility of simulating scenarios to evaluate policies prior to implementation. However, these advances also pose challenges, such as excessive dependence on automated models, the possible loss of human control over critical decisions, and the need to ensure the traceability of the decision-making processes. In this context, it is essential to establish oversight and validation mechanisms to assess the real impact of these technologies on government management, ensuring that their use is aligned with democratic principles and citizens' rights.

Finally, the analysis highlights the importance of adaptability in AI systems to ensure that solutions to bias problems do not become obsolete as data and social contexts evolve. According to Holstein et al. (2019), AI systems must be designed with mechanisms that allow for continuous adjustment in response to changes in data or usage patterns, which is essential for maintaining long-term fairness (Holstein et al., 2019). This suggests that, in addition to correcting existing biases, it is vital that algorithmic systems be inherently flexible and capable of evolving in parallel with the society they serve.

## 7. Critical analysis. Discussion

This article analyzes the crucial importance of implementing bias audits within the framework of algorithmic governance, addressing the technological, regulatory, and ethical dimensions involved. In a context where automation and algorithmic decisions directly affect citizens' lives, it is essential not only to consider the benefits of artificial intelligence (AI), but also to assess the risks and inequalities that these systems may perpetuate or exacerbate.

The growing reliance on algorithms in public administration and business decision-making proposes challenges that go beyond technical efficiency. The opacity characteristic of many AI systems, particularly those based on deep learning, makes it difficult to trace and detect biases. As demonstrated in this analysis, algorithms are not neutral tools. They reflect and amplify social dynamics and biases inherent in the data with which they are trained. In this sense, bias audits should not be considered optional, but rather essential to ensuring fair and equitable AI.

A central aspect of algorithmic governance is the need for proactive and preventive regulation. Regulations such as the Artificial Intelligence Act (AI Act) in Europe seek to establish a comprehensive framework for managing the risks associated with the use of AI, especially in critical areas such as health, education, and justice. However, the effective implementation of these regulations faces considerable challenges, especially due to the heterogeneity of national contexts in the European Union, which complicates the harmonization and uniform application of regulations.

Despite their importance, bias audits have limitations. As noted in the case studies presented, their effectiveness depends on data quality, algorithm transparency, and collaboration among stakeholders. Biases can arise at different stages of the system lifecycle, from data collection to implementation, requiring a holistic approach that considers both technical and social aspects.

A key lesson learned from this analysis is that transparency, while essential, is not sufficient on its own. To be effective, it must be accompanied by explainability mechanisms that allow to understand the decisions made by AI systems. Furthermore, transparency must be accessible and actionable for the general public, which implies not only opening access to algorithms and data, but also ensuring that the information is understandable. Explainability tools and citizen participation mechanisms can play a crucial role in promoting inclusive and democratic governance.

This article also highlights the importance of shared responsibility among the various actors in algorithmic governance. From developers to policymakers and auditing bodies, everyone has a role to play in building ethical AI systems. This collaborative approach not only fosters accountability, but also innovation and continuous improvement in the design of inclusive algorithms.

Regarding future regulations, European legislation is expected to move towards regulating emerging aspects such as synthetic data generation and the integration of explainable AI (XAI). These developments are essential to ensure the auditability and comprehensibility of systems. Regulation will also need to address the challenge of fully autonomous systems in critical sectors such as banking. The adoption of common standards and the promotion of interoperability between systems will be key to the success of algorithmic governance at the European and global levels.

A critical aspect to highlight is the dynamic nature of algorithmic biases. As data evolves and models are adjusted, biases also change, which underscore the need for continuous audits and real-time correction mechanisms. This approach is particularly relevant in contexts where algorithmic decisions significantly impact people's lives, such as in personnel selection, the administration of justice, or the granting of credit. Algorithmic governance, therefore, must be a dynamic process that is adaptable to technological and social changes.

The case study presented shows that even in sectors with strict regulations, such as labor recruitment, biases persist, highlighting the need for rigorous and multidisciplinary approaches in the design and evaluation of algorithms. It is not just a matter of adjusting models, but of reconsidering success metrics to include dimensions of equity and social justice. In this sense, developers must adopt an ethical perspective from the earliest stages of development.

At a global level, the EU has the opportunity to lead the creation of international standards in algorithmic governance, influencing how other regions approach transparency and fairness in AI systems. However, this leadership will require an inclusive dialogue that integrates governments, civil society, the private sector, and academia. Only through strong international cooperation will it be possible to establish inclusive regulatory frameworks that are adapted to the diversity of global contexts.

The incorporation of technological tools into government decision-making significantly modifies traditional public management processes. Among the main implications are reduced decision-making times, access to more accurate information, and the possibility of simulating scenarios to evaluate policies prior to their implementation. Nevertheless, these advances also pose challenges, such as excessive dependence on automated models, the possible loss of human control over critical decisions, and the need to ensure the traceability of decision-making processes. In this context, it is essential to establish oversight and validation mechanisms that

allow to evaluate the real impact of these technologies on government management, ensuring that their use is aligned with democratic principles and citizens' rights.

For all these reasons, effective algorithmic governance must go beyond technical review and consider a multidimensional approach that incorporates continuous evaluation of the social and contextual impact of automated decisions. This approach must be iterative and adaptive, responding to technological changes and emerging social dynamics.

The path to a real transparent and equitable algorithmic governance requires the implementation of rigorous audits, the adoption of dynamic regulations, and collaboration among multiple actors. While the challenges are significant, so are the opportunities to build AI systems that respect fundamental rights and promote a fairer society. In an increasingly digitized world, fairness and accountability must be at the center of every stage of the algorithmic process.

## 8. Critical analysis. Discussion

- Aguilar Villanueva L. F. (2014). Las dimensiones y los niveles de gobernanza. *Cuadernos de Gobierno y Administración Pública*, 1(1), 11-36. [https://doi.org/10.5209/rev\\_CGAP.2014.v1.n1.45156](https://doi.org/10.5209/rev_CGAP.2014.v1.n1.45156)
- Ajunwa, I. (2019). "The Auditing Imperative for Automated Hiring," *Harvard Journal of Law & Technology*, 34 (forthcoming 2021). <https://ssrn.com/abstract=3437631> or <http://dx.doi.org/10.2139/ssrn.3437631>.
- Albaroudi, E., Mansouri, T., and Alameer, A. (2024). "A comprehensive review of AI techniques for addressing algorithmic bias in job hiring," *AI*, 5(1), pp. 383–404. <https://doi.org/10.3390/ai5010019>.
- Arellano Gault D. (2014). ¿Uno o varios tipos de gobernanza? Más allá de la gobernanza como moda: la prueba del tránsito organizacional. *Cuadernos del Gobierno y Administración Pública*, 1(2), 119-137. [https://doi.org/10.5209/rev\\_CGAP.2014.v1.n2.47538](https://doi.org/10.5209/rev_CGAP.2014.v1.n2.47538)
- Bartley, N., Abeliuk, A., Ferrara, E., and Lerman, K. (2021). "Auditing algorithmic bias on Twitter," *Proceedings of the 2021 Web Conference*. DOI: 10.1145/3447535.3462491.
- Beduschi, A. (2024). "Synthetic data protection: Towards a paradigm change in data regulation?", *Big Data & Society*, 11(1), p. 20539517241231277. DOI: 10.1177/20539517241231277.
- Binns, R. (2018). "Fairness in machine learning: Lessons from political philosophy", *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency*; <https://proceedings.mlr.press/v81/binns18a/binns18a.pdf>
- Bouderhem, R. (2024). "Shaping the future of AI in healthcare through ethics and governance," *Humanities and Social Sciences Communications*, 11, p. 416. DOI: <https://doi.org/10.1057/s41599-024-02894-w>.
- Busuioc, M., Curtin, D., and Almada, M. (2023). "Reclaiming transparency: contesting the logics of secrecy within the AI Act," *European Law Open*, 2(1), pp. 79–105. DOI: 10.1017/elo.2022.47.
- Cath, C. (2018). "Governing artificial intelligence: ethical, legal and technical opportunities and challenges", *Philosophical Transactions of the Royal Society A*, 376, p. 20180080. DOI: <http://dx.doi.org/10.1098/rsta.2018.0080>.
- Calero Valdez, A., Heine, M., Franke, T., Jochems, N., Jetter, H.-C., and Schriels, T. (2024). "The European commitment to human-centered technology: The integral role of HCI in the EU AI Act's success," *arXiv*, 2402.14728. DOI: <https://dx.doi.org/10.48550/arXiv.2402.14728>.
- Conitzer, V., Hadfield, G.K., and Vallor, S. (2022). "Technical perspective: The impact of auditing for algorithmic bias," *Communications of the ACM*, 65(12), pp. 40–44. DOI: 10.1145/3571152.
- Crespo-González J., Vicente Vicente M. J., Saudade Baltazar M., and Valmorisco Pizarro S. (2024). Profesorado universitario e IA: ¿una pareja bien avenida?: Reflexiones a partite las expectativas y percepciones del estudiantado. *Cuadernos del Gobierno y Administración Pública*, 11(2), e95972. <https://doi.org/10.5209/cgap.95972>
- Dekker, R., Koot, P., Birbil, S., and van Embden, A.M. (2022). "Co-designing algorithms for governance: Ensuring responsible and accountable algorithmic management of refugee camp supplies," *SAGE*. <https://doi.org/10.1177/20539517221087855>.
- Diakopoulos, N. (2016). "Accountability in algorithmic decision making", *Communications of the ACM*, 59(2), pp. 56-62. DOI: 10.1145/2844110.
- Engelmann, A. (2023). "Algorithmic transparency as a fundamental right in the democratic rule of law," *Brazilian Journal of Law, Technology and Innovation*, 1(2), pp. 169–188. DOI: <https://dx.doi.org/10.59224/bjlti.v1i2.169-188>.
- García Magariño S. (2016). Un cuestionamiento de los supuestos que vertebran la Sociedad moderna en búsqueda de nuevos modelos de gobernanza: el gobierno, la Sociedad civil y la empresa nadando en un mar de complejidad. *Cuadernos del Gobierno y Administración Pública*, 2(2), 187-211. [https://doi.org/10.5209/rev\\_CGAP.2015.v2.n2.51572](https://doi.org/10.5209/rev_CGAP.2015.v2.n2.51572)
- Giest, S. and Grimmekhuijsen, S. (2020). "Introduction to special issue algorithmic transparency in government: Towards a multi-level perspective", *Information Polity*, 25(4), pp. 413-416. DOI: 10.3233/IP-200010.
- Giunchiglia, F., Otterbacher, J., Kleanthous, S., Batsuren, K., Bogina, V., Kuflik, T., and Shulner Tal, A. (2021). "Towards algorithmic transparency: A diversity perspective," *arXiv preprint arXiv:2104.05658*. DOI:10.48550/arXiv.2104.05658.
- Groves, L., Metcalf, J., Kennedy, A., Vecchione, B., and Strait, A. (2024). "Auditing work: Exploring the New York City algorithmic bias audit regime," *ACM*. DOI: <https://doi.org/10.1145/3630106.3658959>.

- Gstrein, O., Haleem, N., and Zwitter, A. (2024). "General-purpose AI regulation and the European Union AI Act," *Internet Policy Review*, 3, pp. 1–15. DOI: <https://dx.doi.org/10.14763/2024.3.1790>.
- Guerrero García A. L. (2023). Cuando hablamos de capacidad burocrática ¿de qué hablamos? Una revisión teórica y empírica del concepto. *Cuadernos de Gobierno y Administración Pública*, 10(2), e83296. <https://doi.org/10.5209/cgap.83296>
- Gyevnar, B., Ferguson, N., and Schafer, B. (2023). "Bridging the transparency gap: What can explainable AI learn from the AI Act?", *Frontiers in Artificial Intelligence*, 10(3), pp. 233-245. <https://dx.doi.org/10.3233/FAIA230367>.
- Hickok, M., Dorsey, C., O'Brien, T., Baur, D., Ingram, K., and Chauhan, C. (2022). "Case study: The distilling of a biased algorithmic decision system through a business lens," *Open Science Framework*. DOI:10.31235/osf.io/t5dhu.
- Holstein, K., Wortman, V., Daumé III, H., Dudik, M., and Wallach, H. (2019). "Improving fairness in machine learning systems: What do industry practitioners need?", *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, p. 600. DOI: 10.1145/3290605.3300830.
- Kharitonova, Y., Malik, N.S., and Yang, T. (2023). "The legal issue of deterrence of algorithmic control of digital platforms: The experience of China, the European Union, Russia, and India," *BRICS Law Journal*, 10(1), pp. 147-170. DOI: 10.21684/2412-2343-2023-10-1-147-170.
- Lam, K., Lange, B., Blili-Hamelin, B., Davidović, J., Brown, S., and Hasan, A. (2024). "A framework for assurance audits of algorithmic systems," *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW2), pp. 1-21. DOI: 10.1145/3630106.3658957.
- Laux, J., Wachter, S., and Mittelstadt, B. (2023). "Trustworthy artificial intelligence and the European Union AI Act: On the conflation of trustworthiness and the acceptability of risk," *Regulation & Governance*, 17(2), pp. 1-21. <https://dx.doi.org/10.1111/rego.12512>.
- Lepri, B., Oliver, N., Letouzé, E., Pentland, A., and Vinck, P. (2018). "Fair, transparent, and accountable algorithmic decision-making processes," *Philosophy & Technology*, 31(4), pp. 611–627. DOI:10.1007/s13347-017-0279-x.
- Mashhadi, A., Zolyomi, A., and Quedado, J. (2022). "A case study of integrating fairness visualization tools in machine learning education," *CHI Conference on Human Factors in Computing Systems*. DOI: 10.1145/3491101.3503568.
- Mensah, G.B. (2023). *Artificial intelligence and ethics: A comprehensive review of bias mitigation, transparency, and accountability in AI systems*. [https://www.researchgate.net/publication/375744287\\_Artificial\\_Intelligence\\_and\\_Ethics\\_A\\_Comprehensive\\_Review\\_of\\_Bias\\_Mitigation\\_Transparency\\_and\\_Accountability\\_in\\_AI\\_Systems](https://www.researchgate.net/publication/375744287_Artificial_Intelligence_and_Ethics_A_Comprehensive_Review_of_Bias_Mitigation_Transparency_and_Accountability_in_AI_Systems)
- Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., and Floridi, L. (2016). "The ethics of algorithms: Mapping the debate," *Big Data & Society*, 3(2), pp. 2053951716679679. DOI: 10.1177/2053951716679679.
- Percy, C., Dragicevic, S., Sarkar, S., and Garcez, A. (2021). "Accountability in AI: From principles to industry-specific accreditation," *AI Communications*, 34(4), pp. 431-445. DOI: 10.3233/AIC-210080.
- Qureshi, N.I., Choudhuri, S.S., Nagamani, Y., Varma, R., and Shah, R. (2024). "Ethical considerations of AI in financial services: Privacy, bias, and algorithmic transparency," *Proceedings of the 2024 IEEE International Conference on Knowledge Engineering and Computer Science (ICKECS)*. DOI:10.1109/ICKECS61492.2024.10616483.
- Sovrano, F., Sapienza, S., Palmirani, M., and Vitali, F. (2022). "Metrics, explainability, and the European AI Act proposal," *Journal of Future Studies*, 5(1), pp. 1–15. DOI: <https://dx.doi.org/10.3390/j5010010>.
- Springer, A. and Whittaker, S. (2021). "Making transparency clear: The dual importance of explainability and auditability", *Proceedings of the 24th International Conference on Intelligent User Interfaces (IUI '21)*. DOI: 10.1145/3447872.3447938.
- Turksen, U., Benson, V., and Adamyk, B. (2024). "Legal implications of automated suspicious transaction monitoring: Enhancing integrity of AI," *Journal of Banking Regulation*. DOI: 10.1057/s41261-024-00233-2.
- Wachter, S., Mittelstadt, B., and Floridi, L. (2017). "Transparent, explainable, and accountable AI for robotics," *Science Robotics*, 2(6), pp. ean6080. DOI: 10.1126/scirobotics.aan6080.
- Wörsdörfer, M. (2023). "Mitigating the adverse effects of AI with the European Union's artificial intelligence act: Hype or hope?", *Journal of Organizational Ethics*, 10(4), pp. 1-13. <https://dx.doi.org/10.1002/joe.22238>.
- Xiang, F., Zhang, X., Cui, J., Carlin, M., and Song, Y. (2022). "Algorithmic bias in student success prediction models: Two case studies," *IEEE TALE Conference*. DOI: 10.1109/TALE54877.2022.00058.
- Zharova, A. (2023). "Achieving algorithmic transparency and managing risks of data security when making decisions without human interference: Legal approaches," *Journal of Digital Transformation and Law*. DOI: 10.21202/jdtl.2023.01.0014.

## José Luis Bustelo Gracia

Dr. José Luis Bustelo Gracia is a university professor specialized in Finance and Accounting, with an outstanding track record in higher education. He currently teaches at Isabel I University, Miguel de Cervantes European University, and ESERP Digital Business and Law School, where he participates in Master's programs in Legal Advice, Management, and Planning. His academic background includes studies at the University of Barcelona, Rey Juan Carlos University, Camilo José Cela University, and La Salle Open University, specializing in financial management, banking, and marketing. His research focuses on artificial intelligence and its impact on business management.