

Perspective

Rule of law in the AI era: addressing accountability, and the digital divide

Antonios Kouroutakis¹

Received: 21 August 2024 / Accepted: 5 November 2024

Published online: 27 December 2024

© The Author(s) 2024 [OPEN](#)

Abstract

The rule of law is a dynamic and evolving concept that adapts to the changing needs and values of society. In light of the information technology revolution and the widespread use of AI applications, a fundamental question arises: how will these advancements influence the concept and application of the rule of law? The answer is twofold. Firstly, the ‘black box’ problem and the general lack of understanding regarding how automated decisions are reached present multiple challenges to the rule of law. This can potentially lead to decisions influenced by systemic biases without a transparent and accountable decision-making framework. Addressing these challenges requires upholding the rule of law through human involvement in automated decision-making processes and possibly enforcing an obligation for reason and explainability. However, the extent of human involvement and the need for explainability would vary based on the nature and function of the AI applications. Secondly, the unprecedented benefits derived from the use of AI applications carry the risk of exacerbating the digital divide, significantly impacting equality. Consequently, it will be argued that the rule of law necessitates both governments and private entities utilizing AI applications to implement measures aimed at preventing and narrowing the digital divides.

Keywords AI · Rule of law · Transparency · Explainability · Digital divide

1 Introduction

Rule of law is everywhere today and there is a consensus among policy and law makers, legal scholars and “we the people” that it is a blessing.¹ Historically, the concept of rule of law in theory and as a philosophical idea evolved differently in the UK, USA, France and Germany [19]. Although rule of law as a term is similar to “Rechtsstaat”, or “Etat de droit”, it is important to note at the outset that these notions were not always synonymous (Venice [64]). In particular about the historical development of the rule of law, two models [archetypes] of rule of law are historically identified: the first associated with Aristotle and the rule of reason, and the second with Montesquieu and the rule of the institutional restraint [60]. Closer to the modern conception of the rule of law, Bingham outlines the most important historical events in the development of the rule of law, such as habeas corpus and the abolition of torture [10].

¹ Several articles and theories have been fairly devoted to analysis of the rule of law. (e.g. [10],[2],[56]). Despite the large amount of literature, and while rule of law is subject to differing interpretations, the orthodoxy in theory is that the concept is either perceived from the formal or the substantive perspective (Craig 1997) and, alternately, from the thick or thin dimension [62].

✉ Antonios Kouroutakis, akouroutakis@faculty.ie.edu | ¹IE University, Madrid, Spain.



Discover Artificial Intelligence

(2024) 4:115

| <https://doi.org/10.1007/s44163-024-00191-8>

However, the concept of rule of law is a dynamic and evolving idea, not inherently static that adapts to the changing needs and values of society. Therefore, the key question is how the emergence of the information technology revolution, and the widespread use of AI applications would influence the concept and the application of the rule of law. A multitude of articles address challenges to the rule of law arising from the rapid advancement of technology [13, 21, 32, 34, 36, 37, 63], driven by increased computational power and network connectivity. Various devices are now permanently linked to the internet, continuously collecting and processing vast amounts of data related to virtually every conceivable aspect of life. This trend is further compounded by the advent of transformative technologies such as AI, machine learning, blockchain, and quantum computing.

It might be claimed that the surge of technological advancements, such as AI, will likely bring significant opportunities. Daly points out that AI applications in the tax system have the potential to enhance rule of law with the improving of consistency, transparency and compliance with tax administration [21]. These technologies have the potential to enable more informed decision-making, improve the quality of public and private services, enhance transparency, and promote stronger civic engagement. However, the impact will depend on how these technologies are implemented and regulated. Most importantly, the widespread use of AI technologies also presents significant challenges for the fundamental components of the rule of law, ² key among which are concerns related to the ‘black box’ nature of the autonomous systems, ³ and the detrimental impact on the society from the widening digital divides [28].

In relation to the black box problem two significant consequences emerge concerning the rule of law, as AI, machine learning, and deep learning systems operate based on intricate algorithms. Firstly, the internal decision-making processes are often opaque; the means by which decisions are reached and the critical factors involved remain unknown to humans in most cases. This lack of transparency has detrimental effects, eroding trust in the decision-making process and undermining accountability on both legal and political fronts. ⁴ If individuals disagree with automated decisions, they find themselves unable to challenge or comprehend the outcomes.

Secondly, even when transparency exists, the degree of human comprehensibility of a given “black box” model or decision may lack interpretability. ⁵ While the algorithm’s operation is transparent, detailing the critical elements for autonomous decisions, humans may struggle to grasp the internal workings and logic of the decision-making process. This lack of understanding poses multiple challenges to the rule of law, potentially resulting in decisions influenced by systemic biases without a transparent and accountable decision-making framework. ⁶ To address these issues, upholding the rule of law would necessitate human involvement in the automated decision-making processes and possibly an obligation for reason and explainability. However, as the first part will show, the degree of human involvement and the requirement for explainability would depend on the nature and the function of the AI applications.

In relation to the digital divide, also referred as new digital divide or algorithmic divide [72], while technologies offer unparalleled benefits to users, there’s a risk of widening the digital divide in society, profoundly impacting equality. To illustrate the transformative potential of the information technology revolution, consider the analogous benefits of the industrial revolution. ⁷ The industrial revolution exponentially increased human power as machines mass-produced goods, and transportation innovations eliminated geographical distances. Similarly, the information technology revolution promises to elevate human intelligence to unprecedented levels, profoundly influencing the quantity and quality of intellectual work.

However, information technology operates based on special infrastructure which is not given in many areas of the world. This inevitably creates access disparities between those who have access to that infrastructure and those who don’t. On the top of that, such technology, for instance AI productive tools, require technical skills acquired by special education and constant training. The lack of such digital literacy has a detrimental impact on the employment opportunities and the societal participation and benefit from such technologies widening the economic disparities and the

² Broadly speaking rule of law is described as an “essentially contested concept” [67], or as a cross cutting issue and an open-ended concept [35]. However, it seems that there is consensus on six elements of rule of law. These are: “(1 Legality, including a transparent, accountable and democratic process for enacting law,(2 Legal certainty; (3 Prohibition of arbitrariness; (4 Access to justice before independent and impartial courts, including judicial review of administrative acts; (5 Respect for human rights and (6 Non-discrimination and equality before the law”. For more details see Venice Commission [64].

³ On the black box problem see von Eschenbach [66].

⁴ On the correlation between transparency and accountability see Desai and Kroll [22].

⁵ For more analysis see Ananny and Crawford [4].

⁶ About how systemic bias might produce discrimination, see Barocas and Selbst [7].

⁷ For more detailed analysis about the benefits and the perils of the industrial revolution see Polanyi [53]

equality of opportunities. The concerns about widening of the digital divide was emphatically expressed in a Joint Statement issued by more than 50 countries, member states of the UN General Assembly on the Proposed UN General Assembly Resolution titled 'Seizing the Opportunities of Safe, Secure, and Trustworthy Artificial Intelligence Systems for Sustainable Development'.⁸

Classified these challenges in one way, the former is about the formal aspect of rule of law, and the latter is about the substantive aspect. Classified in another way, the former is about the due process and the latter is about the right to equality and in particular the right to equal opportunities.

The article uniquely connects these challenges to the foundational principles of the rule of law, first emphasizing that human involvement and the requirement of explainability are not universally necessary for every AI application. The level of involvement and the requirement of explainability required should depend on the nature of the AI system and its legal impact. For AI applications with significant legal consequences, humans should have a supervisory role and retain final decision-making authority. In essence, algorithms should be restricted in cases where governments use AI for non-executive actions, such as the coordination of traffic lights, or they should serve in an advisory capacity during the decision-making process. Secondly, the successful deployment of AI technologies risks widening the digital divide. To mitigate this, governments must establish a comprehensive legal and ethical framework that promotes AI literacy and ensures equitable access to and understanding of these technologies, helping to bridge the divide. This article adds to the ongoing debate not by focusing solely on the transparency and explainability of AI applications or the issue of the digital divide, but by examining these concerns through the lens of the rule of law. It aims to analyze how AI applications can be aligned with and uphold the standards of the rule of law.

Within this context, the first part of the article examines the black box problem of AI applications from the perspective of the rule of law and argues that rule of law standards would necessitate that AI technology should be developed with the interests and needs of people at its core ensuring AI systems are more transparent, interpretable, and ultimately more human-friendly. The analysis demonstrates that such requirements in the automated decision process would vary, contingent upon the nature of the AI system and its legal impact. The second part of the article centers on the potential risks associated with the widening digital divide and its impact on the right to equality. It examines the provisions of the recently adopted AI Act [57] that address AI literacy and emphasizes the need for both governments and private entities utilizing AI applications to take proactive measures to prevent and reduce digital divides.

2 The challenge of ai in decision-making: from transparency to the rule of law

2.1 Understanding AI models and the challenge of interpretation

Although AI, machine learning, and deep learning technologies have been developing for over 40 years, their commercial applications are still in their early stages. Despite this commercial deployment, their potential to bring about transformative change is immense, with the capacity to revolutionize decision-making processes across both the private and public sectors. This technology is already employed in the hiring process, where it aids in the selection of candidates for interviews, and in the banking sector it is utilized to evaluate the credit score of consumers.⁹ In the public sector, AI applications are employed to allocate social benefits and facilitate judicial processes.¹⁰

A prime example of AI in action is the Hawk-Eye ball tracking system, which is used in tennis matches to monitor and assist with human decisions, as well as to correct errors. This AI application stands out because its impact can be quantified, demonstrating that it has led to more accurate decisions and fewer mistakes by referees.

The Hawk-Eye ball tracking system was introduced in 2005 as a pretrial measure to present a three-dimensional representation of the ball trajectory. This three-dimensional representation of the ball trajectory is very useful in borderline cases when the human eye has difficulty identifying with accuracy whether the ball was in or out of the court. During the pretrial period, the Hawk-Eye system was not used to influence decisions; however, in 2006, the pretrial period ended, and since then, players could use the AI application to overturn decisions. Recent research comparing data from 700

⁸ See Joint Statement on the Proposed UN General Assembly Resolution on Seizing the Opportunities of Safe, Secure, and Trustworthy Artificial Intelligence Systems for Sustainable Development (March 14, 2024).

⁹ About the different uses of algorithms in relation to hiring and loan decisions see Pasquale [51].

¹⁰ About the prospects of AI applications in the public sector see Kouroutakis [41]. In particular about criminal procedures see Quattrocchio [54].

matches played both during the pretrial period and after the trial period has shown that the use of AI can substantially improve the decisions reached by humans. Hawk-Eye oversight has motivated human officials to enhance their performance, resulting in an 8% reduction in errors [3].

This improvement in decision-making underscores the benefits of employing ‘white box’ AI applications like Hawk-Eye. AI applications built on patterns, rules, or decision trees that can be easily understood by experts, aligning more closely with human language are termed as the ‘white box.’ The benefit of these white box AI applications is that they are designed for enhanced comprehension, allowing for effective human oversight and intervention when necessary to override or correct the model’s logic.¹¹ In theory white box algorithms are transparent in relation to the data used to train the algorithm, their gravity and their relevance for the output, and the logic employed. This transparency ensures that human officials can trust the system’s recommendations and intervene confidently when needed, making Hawk-Eye a robust and reliable tool in the realm of sports officiating.

Given the transparency and interpretability of AI applications built on a white box model, lead us to ponder whether such applications can deliver decisions for instance at the standard of judges [61].¹² Others remark that ‘the question about the future of public administration is not whether digital algorithms are perfect. Rather, it is a question about what will work better: human algorithms or digital ones!’ [18].

However, the design configuration of a white-box algorithm has inherent limits. For instance, there might be limits to performing more complex problems, or its predictive accuracy might be lower when facing interconnected, multi-faceted, or nonlinear data.

2.2 Aligning AI with the rule of law principle of due process

Whereas we often refer to AI as one thing, one technology, in practice AI is a broad field encompassing various technologies and approaches aimed at creating machines or systems that can perform tasks that typically require human intelligence. AI applications with high complexity, such as those consisting of multiple decision trees or with a large number of parameters and layers (e.g., Neural Networks—Deep Learning), often provide high predictive accuracy. However, they are considered black-box models due to the difficulty in understanding how they reached their output.

In relation to black box models, serious concerns have been raised regarding the lack of transparency, which inevitably results in our inability to understand why and how automated decisions are reached. Berman accurately highlights a paradox, stating that ‘the more complex and powerful an algorithm, the more opaque it is likely to be’ [8]. The absence of transparency¹³ has significant detrimental effects. On the one hand it erodes trust in the decision-making process, and on the other hand, it undermines the accountability.¹⁴ When individuals disagree with automated decisions, they find themselves unable to comprehend the reasoning behind those decisions, making it challenging for them to dispute or challenge the outcomes.

What complicates the situation is that even if AI applications operate transparently, both ordinary people and experts may struggle to comprehend the internal processes of the algorithms, the gravity and the significance of each data point, the casual links, and, in general, the logic of the decision-making process. This challenge is often referred to as the problem of interpretability,¹⁵ where the process is visible, but difficult to interpret, (referred to as ‘grey box’ for the purpose of this article).

To illustrate this, consider the utilization of AI applications in detecting precancerous cells. These applications, employing deep learning techniques, are trained on datasets comprising both precancerous and healthy cells. By analyzing various patterns such as cell shape, color, and shades, etc. they are capable of distinguishing between precancerous

¹¹ For more details about white boxes see Loyola-González [44].

¹² Furthermore, Volokh argues that for machines to replace judges it is necessary that they would be able to pass the Turing test and generate decisions that would convince expert lawyers [65]. In contrast Tasioulas argues that ‘the use of AI tools in place of human judges undermines, in three ways, the kind of respect for rational autonomy that the desiderata of the rule of law are intended to secure. Specifically, it undermines the explainability, the accountability, and the reciprocity or community that is intimately bound up with the ideal fulfilment of the rule of law, and – in particular – the desideratum of congruence between law and official decision’ [63]. In 2013, in the annual report Chief Justice Roberts rejected the idea that AI would replace key players in the legal system such as judges. See John G. Roberts, Jr., Chief Justice’s Year-End Reports on the Federal Judiciary (2023) 6.

¹³ Transparency is defined as ‘the ability for a specific model to be understood. In the strictest sense, a model is transparent if a person can contemplate the entire model at once’ [43].

¹⁴ About the need for accountability, transparency and accuracy in autonomous decision-making processes see Citron [17].

¹⁵ Interpretability is defined as the degree of human comprehensibility of a given “black box” model or decision [9].

and healthy. Despite the complexity involved, AI applications can determine whether a cell is precancerous based on a combination of factors that may not be readily interpretable by human experts, thereby surpassing the limits of causal interpretation. Even if an AI model is transparent, it may still lack interpretability, which creates significant challenges for the rule of law. Humans cannot effectively challenge AI applications, which, while theoretically transparent may, in practice, result in decisions influenced by systemic biases due to the absence of a transparent and accountable decision-making framework.

In light of the considerable influence AI applications can have on decision-making processes, the inquiry is directed towards understanding how and under what circumstances these applications align with the foundational principles of the rule of law.

The response to this question hinges on two key factors: the nature of the model, and the nature of the decisions they render.

Firstly, regarding the nature of the AI model, regardless of whether it is a white box or the black box model, human involvement, often referred to as 'human in the loop' or 'AI by the people' remains essential for two primary reasons: ensuring the proper functioning of the AI application and averting malfunctions. Humans play a crucial role in verifying the model's proper functioning, particularly regarding the quality of explanations when necessary, such as in cases where automated decisions reject requests for social benefits or loans. Additionally, humans are vital to prevent malfunctions, such as overseeing AI applications to mitigate issues like data poisoning, which could lead to system malfunctions.

However, concerning black box or grey box AI applications there is a fundamental requirement for transparency and explainability in each system [16, 39, 46]. In addressing the 'transparency' requirement, it is crucial to emphasize the following. Transparency should not be confined solely to the internal workings of AI applications, rather, it encompasses three dimensions. First, transparency pertains to the acknowledgment of an automation process.¹⁶ Second, it involves clarity regarding the internal functioning of the AI application, including the training data. Third, it extends to providing transparency on the process of challenging automated decisions. Consequently, transparency necessitates that individuals are not only aware that they are utilizing an AI application but also possess an understanding of its operations, and importantly, have the means to challenge its decisions.

On the other hand, the quality of the explanation is crucial. Explanation should be clear, appropriate and sufficient. As Tasioulas remarks 'even if an explanation exists and is accessible to a minimally adequate degree, there is still a further question as to whether it is an explanation of the right kind, in the sense of being one that justifies the decision that has been made' [63]. A clear explanation leaves no doubt as to the real reason behind the decision, while an appropriate and sufficient explanation is tailored to specific facts on which the decision was based. However, Bordt et al. argue that explanations do not work especially in adversarial relationships because different algorithms have the potential to produce inconsistent explanations or because there is inherent ambiguity in explanations [11].

The requirement for transparent AI applications coupled with explainability may pose challenges. Specifically, such requirements could have a detrimental effect on the quality of the AI applications, as they might impose design requirements that restrict design choices. Berrada aptly remarks, that 'explainability is an essential property; however, it is not always a necessity. In fact, requiring every AI system to explain every decision could result in less efficient systems, forced design choices, and a bias towards explainable, but less capable and versatile outcomes. Furthermore, making AI systems explainable is undoubtedly expensive; they require considerable resources both in the development of the AI system and in the way it is interrogated in practice.' (Adadi, Berrada, [1]).

This presents a policy dilemma, between the level of transparency and explainability and the compatibility of such applications with rule of law standards requiring the obligation to provide decisions to people in compliance with their due process rights. It could be argued that human agency, and specifically human oversight, might serve as a counterbalance to the risks associated with the lack of transparency or explainability in government algorithms.¹⁷ However, Green contends that evidence suggests individuals are often unable to fulfill the necessary oversight functions effectively. As a result, policies that rely on human oversight to mitigate these risks may legitimize such AI Applications without the

¹⁶ Article 22 of the GDPR requires an explicit consent if a person is subject to a decision based solely on automated processing. See European Parliament [29].

¹⁷ An OECD report for instance in the section titled 'Respect for the rule of law, human rights and democratic values, including fairness and privacy' recommends that 'AI actors should implement mechanisms and safeguards, such as capacity for human agency and oversight, including to address risks arising from uses outside of intended purpose, intentional misuse, or unintentional misuse in a manner appropriate to the context and consistent with the state of the art.' See OECD [49] [1.2].

necessary safeguards [33]. In line with the above, it is argued that human oversight monitoring complex AI systems is ineffective as it not possible to have meaningful control [40].

On the other hand, it is argued that AI applications are not necessarily more opaque than human decision-making Brožek et al. [14]. Daly accurately remarks that ‘Determining the “real motivation” behind a decision, however, is a perennial problem in legal systems’ [21]. Indeed, Chief Justice Roberts Jr. of the US Supreme Court confesses that judges rely on subtle, non-verbal cues—such as body language, tone of voice, and facial expressions—to assess a defendant’s sincerity and emotions during proceedings, such as during sentencing (Roberts, Jr. 2023). These nuanced human interactions are often not explicitly reflected in formal decisions, making judgments based on them appear subjective or arbitrary. Still, human-based decisions uphold a fundamental principle of a fair trial—that individuals should be judged by their peers, a concept rooted in the Magna Carta.

In relation to the nature of the decisions, white box AI applications have the potential to assist by automating repetitive tasks, by reducing human error in administrative processes, by scaling the modern administration to handle the growing volume of data, and by delivering faster decisions. In general AI applications in the public administration can improve the services to the citizens and at the same time can make the bureaucracy more efficient.

However, the intricacies associated with deploying AI applications with a black box model have led many to advocate for their limitation to a supporting role in human decision-making, rather than serving as a sole determinant replacing human judgment [38]. The problem of the lack of transparency and explainability is exacerbated in cases where the administration exercises discretion. A case study in ETIAS, an AI application designed to assess risks posed by visa-exempt third-country nationals entering the EU, has revealed the when the administration exercises wide discretion, it is inevitable that the design and the development of such AI application will be based on arbitrary rules [25].

Furthermore, it has been accurately pointed out that the inner working of AI applications is based on data correlations. However, data correlation is not the same as cause and effect relationships which are the inner working of the legal reasoning [6]

In essence, algorithms are proposed to be restricted in cases where governments employ AI applications for non-executive actions, or they may serve as an advisory role in the decision-making process. Regarding the non-executive utility of AI applications, governments may successfully employ them, for instance, in controlling traffic lights, where a detailed explanation of the system’s workings may not be deemed necessary.

In terms of the advisory role of AI, this approach suggests that AI applications offering input can be advantageous, potentially expediting decision-making with reduced bureaucratic obstacles, thereby enhancing the overall effectiveness of the decision-making process. In such scenarios, existing legal provisions govern the circumstances under which administrations or private entities are obligated to provide justifications for the outcomes of their decisions.

3 Rule of law and the challenge of the digital divide

Artificial Intelligence is anticipated to wield a substantial impact on human civilization. To comprehend its potential, let’s draw a parallel with another transformative technological advancement—electricity. Since its widespread adoption, electricity has not only enriched our lives with convenience and comfort by powering various applications and devices but has also played a pivotal role in materializing the industrial revolution, enhancing productivity [20]. Electricity has been an essential feature of modern communication systems, including radios, telephones, televisions, and the internet, and has modernized many aspects of human activity such as agriculture, education, and healthcare.

Similarly, AI is poised to have a profound and transformative influence.

It can readily assist in automating repetitive tasks, such as data input or contract auditing for specific provisions, thereby freeing up more time for creative and productive work.¹⁸ Furthermore, AI can optimize resource management, contribute to sustainable practices, and is expected to drive innovation in fields such as medicine, entertainment, and education. Despite short-term disruptions in the labor market as it transitions to an AI-driven ecosystem, individuals with access to these technologies and the digital skills to use them stand to benefit significantly. The synergy of artificial intelligence and human intelligence will enhance the quality of intellectual endeavors and will revolutionize manufacturing, services and products.

¹⁸ In relation to the impact of AI technology in the legal profession see S. Migliorini, and J. Moreira ‘The Case for Nurturing AI Literacy in Law Schools’ [2024] Asian Journal of Legal Education 1.

Countries and regions stand to gain significantly from the new AI ecosystem, as AI has the potential to transform manufacturing, services, products, public sector functions, whole industries and governance, driving both economic growth and job creation. AI can also be a powerful driver of cross-border collaboration as industries such as in technology and pharmaceuticals may benefit from supply chains with better logistics and synchronized operations. Mikhaylov et al. [47] argue that international collaboration in AI brings several benefits that align with incentives for adopting AI broadly. By pooling resources and knowledge, countries can improve predictive accuracy for service demand and enable faster, more automated responses to fluctuations in demand globally [47].

A recent report from the World Bank (Digital Progress and Trends Report 2023) highlights that AI is poised to be a critical “competitiveness differentiator” for companies, impacting both product and process innovation [71]. For example, AI-based companies will be able to deliver enhanced services at lower costs, giving them a competitive edge in the market.

However, a serious concern has arisen regarding the potential widening of digital divides, posing a significant impact on equality. These concerns revolve around two main fronts: first, the worry that the use of AI applications may exacerbate inequality, and second, the fact that not everyone has access to such technology, potentially leaving those without access excluded from its benefits.

O’Neil raised concerns about the discriminatory impact of algorithms against certain groups in areas such as hiring, lending and law enforcement [50]. In like manner, Eubanks outlined the impact of AI applications in three areas of public service in the US, relating to welfare, homelessness and child protection and found that such applications failed to support the poorest and the most vulnerable [28]. The failure of AI applications exacerbating inequalities is caused by the bias in the system, which might be in relation to sampling of data, in the design and implementation of the algorithm.¹⁹

In relation to the digital divide, AI’s digital nature fosters an environment where first-mover advantages can be substantial, potentially widening the gap between early adopters leading in technology and others. [12] which could widen the income gap both within countries and between countries [26].

Pinheiro et al. [52] suggested that AI technologies may intensify the digital divide given that lower-income regions with less advanced technology tend to adopt simpler technologies, while higher-income regions with more advanced technology are more likely to embrace complex technologies and industries [52]. In line with the above, if patenting activities and scientific publications on AI signal innovation on AI and the wider use of such technology, it seems that already there is a significant concentration in a few regions such as US, China, Japan and South Korea [42].

Various factors contribute to the emergence of a digital divide—a gap between individuals and groups with access to information technology tools like AI and those without.²⁰ It’s noteworthy that information technology relies on specific infrastructure, which is not universally available. Additionally, not everyone possesses the digital literacy skills necessary to effectively utilize digital resources. Furthermore, those excluded from the information technology revolution are more likely to be unaware of its benefits or lack access to the tools needed to utilize it.

The consequence of this digital divide is expected to be profound, depriving some people of the benefits of such technology. To a lesser extent, even individuals with access to AI applications, which may not be well-designed or fed with appropriate data, may not optimize their benefits from the technology [72]. While the widening of the digital divide poses political, social, economic, and technological challenges, it is also a rule of law issue.

The widening of the digital divide could have a severe impact on the right to equality, as it denies ‘have-nots’ equal opportunities to benefit from the information technology revolution. Equality of opportunities, rights, and obligations constitutes the concept of equality before the law [23], a significant component of the rule of law.

In theory, the ideal of equality of opportunity has garnered attention from scholars in legal and political philosophy [5, 24, 45, 58, 68–70], Wilkinson III [69]; [59]. However, scholars approach the concept of equality of opportunity from different perspectives, leading to the emergence of various conceptions, such as the distinction between formal and substantive equality of opportunity [27]. In the formal dimension of equality of opportunities, all positions and posts are considered open, and their allocation is based on merit. Westen asserts that ‘equal opportunity exists wherever two or more people fall within a class of agents who are all free from the same obstacle to attain the same goal’ [68]. Conversely, the formal version of equality of opportunity has faced criticism from several scholars. Specifically, they argue that the formal dimension perpetuates inequality unless equality of opportunity encompasses fair access to the qualifications required for success [55, 70].

¹⁹ For more details about different forms of bias in AI Applications, see Ferrara [31], Chapman et al. [15]

²⁰ Yu argues that the digital divide has 5 features: awareness, access, affordability, availability and adaptability [72].

In this context, it is imperative to prioritize efforts aimed at bridging the digital divide. The rule of law, guided by its fundamental principle of equality, necessitates that governments take measures to level the playing field concerning technical infrastructure and AI literacy.

Traditionally, literacy has focused on the skills of reading and writing. However, with the widespread use of AI technologies, particularly generative AI, the scope of literacy must expand. First, since AI applications can create and process text, audio, images, and video, users must critically engage with this content to identify potential errors or inaccuracies. Second, understanding the mechanics of AI—how it operates based on algorithms and data—has become essential. On one side, familiarity with algorithms helps users interpret and evaluate AI-driven decisions, supporting informed trust in these systems. On the other, AI outputs depend heavily on the training data, making it crucial for users to recognize the limitations and potential biases in AI results; for example, biased input data will naturally produce biased outcomes. Lastly, given AI's vast potential and its profound societal impacts, AI literacy requires an ethical dimension to encourage responsible and informed use. Ng et al. define AI literacy as encompassing four competencies: the knowledge and skills required to (i) understand, (ii) use, (iii) evaluate, and (iv) ethically apply AI technologies (Ng et al [48]). This expanded definition combines both technical understanding and ethical awareness and highlights the importance of critical engagement, technical insight, and ethical considerations in a well-rounded approach to AI literacy.

Consequently, AI literacy involves making these technologies widely accessible and equipping everyone with the skills to use them effectively, responsibly, and to their full advantage. Ideally, AI applications should be designed to actively promote and enhance digital literacy among end-users, empowering them to engage thoughtfully with AI-driven tools and fostering a culture of informed and ethical technology use. This model offers a win-win scenario for both AI applications and users. A larger population would engage with and benefit from the technology, while end-users would acquire the knowledge to maximize these resources and safeguard themselves from potential risks.

To its credit, the European Parliament, in the recently agreed Artificial Intelligence Act, proposed an amendment aimed at bolstering AI literacy (European [30]) as the concept of AI literacy was absent from the original bill submitted to the EU Parliament.²¹ Specifically, Article 4b.1 of Amendment 214 outlines that both the Union and the Member States bear the responsibility of championing measures to cultivate a satisfactory level of AI literacy across various sectors. Furthermore, Article 4b.2 mandates that providers and deployers of AI systems must implement measures to ensure adequate AI literacy among their staff and individuals involved in operating and utilizing AI systems.

These measures are to be tailored to the technical knowledge, experience, education, and training of the personnel, considering the contextual aspects of AI system usage and the impacted individuals or groups. To elaborate, Article 4b.3 specifies that these literacy measures particularly encompass teaching basic concepts and skills related to AI systems, covering their functionality, diverse product types, uses, risks, and benefits. The overarching objective, as articulated in Article 4b.4, is to attain a satisfactory level of AI literacy enabling providers and deployers to ensure compliance and enforcement of the Regulation. Amendment 214 emerges as a notable addition, emphasizing the significance of widespread AI literacy. This amendment underscores that the EU and Member States are mandated to actively promote educational and training measures for fostering AI literacy across all industries and demographic segments of society.

Having stated this, the standards of the rule of law, particularly its component of equality, necessitate both governments and private entities to actively undertake measures to enhance the AI literacy of the populace. By doing so, not only can the widening of the digital divide be prevented, but there is also the potential for it to narrow.

4 Conclusions

The concept of the rule of law has undergone historical evolution, with distinct models emerging in the UK, USA, France, and Germany. While rooted in similar terms like “Rechtsstaat” and “Etat de droit,” these notions were not always synonymous. The historical development of the rule of law identifies two archetypes: one associated with Aristotle and the rule

²¹ The AI Act defines ‘AI literacy’ as the ‘skills, knowledge and understanding that allows providers, users and affected persons, taking into account their respective rights and obligations in the context of this Regulation, to make an informed deployment of AI systems, as well as to gain awareness about the opportunities and risks of AI and possible harm it can cause and thereby promote its democratic control. AI literacy should not be limited to learning about tools and technologies, but should also aim to equip providers and users with the notions and skills required to ensure compliance with and enforcement of this Regulation. It is therefore necessary that the Commission, the Member States as well as providers and users of AI systems, in cooperation with all relevant stakeholders, promote the development of a sufficient level of AI literacy, in all sectors of society, for people of all ages, including women and girls, and that progress in that regard is closely followed.’ (European [30]).

of reason, and the other with Montesquieu and the rule of institutional restraint. The rule of law is dynamic, adapting to changing societal needs and values. With the information technology revolution, driven by AI, machine learning, and other transformative technologies, challenges to the rule of law arise.

While these advancements offer opportunities for informed decision-making, service delivery, transparency, and civic engagement, they also pose challenges such as the 'black box' nature of autonomous systems and the widening digital divide. The 'black box' problem in AI systems, operating on intricate algorithms, raises transparency concerns. Decision-making processes are often opaque, eroding trust and accountability. Even when transparent, the lack of human comprehensibility of these 'black box' models poses challenges to the rule of law, potentially leading to decisions influenced by biases. Human involvement and explainability are crucial to address these issues, contingent on the nature and function of AI applications.

The digital divide, a new form of inequality, is exacerbated by limited access to information technology infrastructure and the need for specialized education. This divide widens economic disparities and hampers equality of opportunities. The article classifies these challenges as formal (due process) and substantive (right to equality) aspects of the rule of law. Finally, the article emphasizes the need for positive measures to enhance AI literacy, advocating for a rule of law framework that calls for collective efforts by governments and AI applications to bridge the digital divide and safeguard the right to equality in the era of rapid technological advancement.

The phrase "AI by the people, for the people" suggests an ideal of artificial intelligence that is created by and for the benefit of humanity as a whole, reflecting principles of transparency, inclusivity and accessibility. It further encapsulates the idea that AI technology should be developed with the interests and needs of people at its core, and that it should comply with the rule of law, an important pillar of our legal culture.

Author contributions AK wrote the whole article and conducted the research.

Funding MCIU/AEI /<https://doi.org/10.13039/501100011033> / FEDER, UE, PID2023-149184OB-C43

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Adadi A, Berrada M. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE Access*. 2018;6:52138–60.
2. Allan TRS. *Constitutional justice: a liberal theory of the rule of law*. Oxford: Oxford University Press; 1995.
3. Almog D, Gauriot R, Page L, Martin D. AI oversight and human mistakes: evidence from centre court. *Arxiv*. 2024. <https://doi.org/10.48550/arXiv.2401.16754>.
4. Ananny M, Crawford K. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media Soc*. 2018;20(3):973–89. <https://doi.org/10.1177/1461444816676645>.
5. Anderson E. What is the Point of Equality? *Ethics*. 1999;109:287–337. <https://doi.org/10.1086/233897>.
6. Bayamlioglu E, Leenes R. The "Rule of law" implications of data-driven decision-making: a techno-regulatory perspective. *Law Innov Tech*. 2018. <https://doi.org/10.1080/17579961.2018.1527475>.
7. Barocas S, Selbst AD. Big data's disparate impact. *Calif L Rev*. 2016. <https://doi.org/10.1577/Z38BG31>.
8. Berman E. A government of laws and not of machine. *B U L Rev*. 2018;98:1277.
9. V Bernardo. *TechDispatch: explainable artificial intelligence. technology and privacy unit of the european data protection supervisor (EDPS); 2023*
10. Bingham T. *The rule of law*. London: Allen Lane; 2010.

11. S Bordt, M Finck, E Raidl, U Luxburg von, 'Post-hoc explanations fail to achieve their purpose in adversarial contexts' in (ACM 2022) FAccT '22 891
12. Brynjolfsson E, Wang J, McElheran K. The power of prediction: predictive analytics, workplace complements, and business performance. *Bus Econ.* 2021;56:217–39.
13. Brownsword R. Technological management and the rule of law. *Law Innov Technol.* 2016;8(1):100–40. <https://doi.org/10.1080/17579961.2016.1161891>.
14. Brożek B, Furman M, Jakubiec M, Kucharzyk B. The black box problem revisited real and imaginary challenges for automated legal decision making. *Artif Intell Law.* 2023. <https://doi.org/10.1007/s10506-023-09356-9>.
15. Chapman J, Johnson K, Pasquale F. Artificial intelligence, machine learning, and bias in finance: toward responsible innovation. *Fordham L Rev.* 2019;88:499.
16. Choi J, Gunning D, et al. XAI—explainable artificial intelligence. *Sci Robotics.* 2019;4:37.
17. Citron DK. Technological due process. *Wash U L Rev.* 2008;85:1249.
18. Coglianese C, Lai A. Algorithm V. *Algorithm Duke L J.* 2022;71:1281.
19. Costa P, Zolo D. The rule of law history. In: *Theory and criticism.* Cham: Springer; 2007.
20. Crafts N. Productivity growth in the industrial revolution: a new growth accounting perspective. *J Econ Hist.* 2004;64:521.
21. Daly S. Artificial intelligence, the rule of law and public administration: the case of taxation. *Cambridge Law J.* 2024. <https://doi.org/10.1017/S0008197324000448>.
22. Desai DR, Kroll JA. Trust but verify: a guide to algorithms and the law. *Harv J L Tech.* 2017;1:39.
23. Dworkin R. *Taking rights seriously.* Harvard University Press; 1978.
24. Dworkin R. What is equality? II. Equality of resources. *Philos Public Aff.* 1981;10:283–345.
25. Eklund AM. Rule of law challenges of “algorithmic discretion” & automation in eu border control: a case study of ETIAS through the lens of legality. *Eur J Migr Law.* 2023. <https://doi.org/10.1163/15718166-12340152>.
26. Ernst E, Merola R, Samaan D. Economics of artificial intelligence: implications for the future of work. *IZA J Labor Policy.* 2019;9:1.
27. Elford G (2023) Equality of Opportunity. *The Stanford Encyclopedia of Philosophy.* <https://plato.stanford.edu/archives/fall2023/entries/equal-opportunity/>
28. Eubanks V. *Automating inequality: How high-tech tools profile, police, and punish the poor.* New York: St. Martin's Press; 2018.
29. Parliament E. 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 data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official J Eur Union L.* 2016;119:1–88.
30. European Parliament (2023) Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html
31. Ferrara E. Fairness and bias in artificial intelligence: a brief survey of sources, impacts, and mitigation strategies. *SSRN.* 2023. <https://doi.org/10.4855/arXiv.2304.07683>.
32. Gowder P. Transformative legal technology and the rule of law. *Univ Toronto Law J.* 2018;68:82–105. <https://doi.org/10.3138/utlj.2017-0047>.
33. Green B. The flaws of policies requiring human oversight of government algorithms. *Comput Law Secur Rev.* 2022. <https://doi.org/10.1016/j.clsr.2022.105681>.
34. Greenstein S. Preserving the rule of law in the era of artificial intelligence (AI). *Artif Intell Law.* 2022;30:291–323. <https://doi.org/10.1007/s10506-021-09294-4>.
35. R Grote. The rule of law, Rechtsstaat and Etat de Droit. In: Starck C (eds). *Constitutionalism, universalism and democracy a comparative analysis.* Baden-Baden; 1999
36. Huq AZ. A right to a human decision. *Va L Rev.* 2020;106:611.
37. Huq AZ (2021) Artificial intelligence and the rule of law. *routledge handbook on the rule of law, University of Chicago, public law working paper No.* <https://doi.org/10.2139/ssrn.3794777>
38. Kahneman D, Sibony O, Sunstein CR. *Noise: a flaw in human judgment.* London: William Collins; 2021.
39. Kaminski ME. The right to explanation. *Explained Berkeley Tech L J.* 2019;34:189.
40. Koulu R. Proceduralising control and discretion: human oversight in artificial intelligence policy. *Maastricht J Eur Comp Law.* 2020. <https://doi.org/10.1177/1023263X20978649>.
41. Kouroutakis AE. Legaltech in public administration: prospects and challenges. In: DiMatteo LA, editor. *The Cambridge handbook of lawyering in the digital age.* Cambridge: Cambridge University Press; 2021.
42. Kitsara I. Artificial intelligence and the digital divide from an innovation perspective. In: Bounfour A, editor. *Platforms and artificial intelligence the next generation of competences.* Cham: Springer; 2022.
43. Lepri BO, Nuria O, et al. Fair, transparent, and accountable algorithmic decision-making processes: The premise, the proposed solutions, and the open challenges. *Philos Technol.* 2018;31(4):611–27.
44. Loyola-González O. Black-box vs. white-box: understanding their advantages and weaknesses from a practical point of view. *IEEE Access.* 2019;7:154096–113.
45. Mason A. *Levelling the playing field: the idea of equal opportunity and its place in egalitarian thought.* Oxford: Oxford University Press; 2006.
46. Miller T. Explanation in artificial intelligence: Insights from the social sciences. *Artif Intell.* 2019;267:1–38. <https://doi.org/10.1016/j.artint.2018.07.007>.
47. Mikhaylov SJ, Esteve M, Campion A. Artificial intelligence for the public sector: Opportunities and challenges of cross-sector collaboration. *Philos Trans Royal Soc A.* 2018. <https://doi.org/10.1098/rsta.2017.0357>.
48. Ng DTK, Leung JKL, Chu SKW, Qiao MS. Conceptualizing AI literacy: an exploratory review. *Comput Educ.* 2021;2: 100041. <https://doi.org/10.1016/j.caeai.2021.100041>.

49. OECD, 'Recommendation of the Council on Artificial Intelligence' (2019) OECD/LEGAL/0449, <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449> Accessed 25 Oct 2024
50. O'Neil C. Weapons of math destruction: how big data increases inequality and threatens democracy. New York: Crown; 2016.
51. Pasquale F. The black box society: the secret algorithms that control money and information. Cambridge: Harvard University Press; 2015.
52. Pinheiro FL, Balland PA, Boschma R, Hartmann D. The dark side of the geography of innovation: relatedness, complexity, and regional inequality in Europe. *Reg Stud.* 2022. <https://doi.org/10.1080/00343404.2022.2106362>.
53. Polanyi K. The great transformation. Boston: Beacon Press; 2001.
54. Quattrocolo S. Artificial intelligence, computational modelling and criminal proceedings a framework for a European legal discussion. Cham: Springer; 2020.
55. Rawls J. A theory of justice. Harvard University Press; 1999.
56. Raz L (1979) The Authority of Law: Essays on Law and Morality. Clarendon Press
57. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)
58. Rosenfeld M. Substantive equality and equal opportunity: a jurisprudential appraisal. *Calif L Rev.* 1986;74:1687.
59. Scheffler S. What is egalitarianism? *Philos Public Aff.* 2003;31:5–39.
60. Shklar J. Political theory and the rule of law. In: Hutchinson AC, Monahan PJ, editors. The rule of law; ideal and ideology. Toronto: Carswell; 1987.
61. Susskind R. Online courts and the future of justice. Oxford: Oxford University Press; 2019.
62. Tamanaha BZ. On the rule of law: history, politics. Theory: Cambridge University Press; 2004.
63. Tasioulas J. The rule of algorithm and the rule of law Vienna lectures on legal philosophy. *Legal Reason.* 2023. <https://doi.org/10.2139/ssrn.4319969>.
64. Venice Commission (2011) Report on the rule of law. Council of Europe.
65. Volokh E. Chief justice robots. *Duke L J.* 2019;68:1135.
66. von Eschenbach WJ. Transparency and the black box problem: why we do not trust AI. *Philos Technol.* 2021;34:1607–22. <https://doi.org/10.1007/s13347-021-00477-0>.
67. Waldron J. Is the rule of law an essentially contested concept (In Florida). *Law Philos.* 2002;21(2):137–64. <https://doi.org/10.1023/A:1014513930336>.
68. Westen P. The concept of equal opportunity. *Ethics.* 1985;95(4):837–50.
69. Wilkinson JH III. The supreme court, the equal protection clause, and the three faces of constitutional equality. *Virg L Rev.* 1975;61:945–1018.
70. Williams B. The idea of equality. In: Laslett P, Runciman WG, editors. *philosophy, politics and society, series II.* Oxford: Blackwell; 1962.
71. World Bank (2024). Digital progress and trends report 2023
72. Yu PK. The algorithmic divide and equality in the age of artificial intelligence. *Fla L Rev.* 2020;72:354–61.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.