INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 3, MARCH 2026 | ISSN: 2456-4184 | IJNRD.ORG

Artificial Intelligence and Human Rights: Strengthening Legal Safeguards against Algorithmic Discrimination

Shibanee Acharya[1*], Omkar Acharya[2], Ashish Kumar Mishra[3]
[1]Department Research Fellow, [2]LL.M. Student, [3]Advocate, Orissa High Court
[1]PG Department of Law
[1]Fakir Mohan University, Balasore – 756019, Odisha, India

Abstract: Algorithms and artificial intelligence are now organizing access to jobs, credit, welfare, policing, and justice and are also contemplating entrenching algorithmic discrimination and compromising foundational human rights like equality, non-discrimination, privacy, due process, and human dignity. Objectives include to determine the impact of AI systems on equality, non-discrimination, privacy, dignity, due process and social rights in different sectors; examine the application of the current non-discrimination, data protection and human rights norms to address the issue of algorithmic bias and discriminatory results and find out the particular doctrinal and enforcement gaps and reforms such as hybrid secured ground regimes, compulsory human rights, algorithmic audit, independent oversight organisms. The AI systems that are trained on biased data can be systematically disadvantaged to legally protected groups and black box and opacity architecture can render it hard to refute and show that they are discriminative. However, such regimes are seen to have severe flaws upon application to AI, such as enforcement gaps, information asymmetries. The mixed doctrinal and comparative case study design is an appropriate one and optionally supplemented by limited empirical work is to be done. The argument presented in this paper is the case of the enhanced legal protection based on human rights-grounded like more robust integration of anti-discrimination and data protection laws by means of impelled algorithmic audits, impact analysis and access to the meaningful explanations; sector-specific regulation of high-risk applications like recruitment, credit rating, welfare distribution, predictive policing, and biometric policing; a dignity-based system that instils active anti-discrimination responsibilities and good regulatory solutions both in domestic legislation and new AI tools.

Keywords- Algorithmic, data, privacy, human rights, surveillance, governance

INTRODUCTION

The integration of Artificial Intelligence (AI) into governance, commerce, and daily life has been an accelerated process and has changed fundamentally the decision-making processes in spheres. From recruitment algorithms to credit scoring systems, predictive policing and welfare distribution, AI systems have an increasingly large role in determining access to vital opportunities and rights.[1] These technological systems are commonly represented as impartial and efficient alternatives to human decision making; however, the growing influence of these systems prompts fundamental human rights concerns of their impact. In particular, principles like equality, non-discrimination, privacy, due process and human dignity are increasingly implicated in algorithm-driven decisions that extend our individual life chances.[2]

AI technologies work on complex computational models to analyze huge data sets to find patterns and make predictions. While this ability creates increased efficiency and scalability, it also creates a lot of risk as well. Algorithmic decision-making may not be sensitive to contextual factors, it may not take social inequalities into account, and it may lead to disproportionate impact on vulnerable groups.[3] For example, automated hiring systems might discriminate against women or minority candidates, while predictive policing

[1] Solon Barocas & Andrew D. Selbst, Big Data's Disparate Impact, 104 Calif. L. Rev. 671, 673 (2016).
[2] U.N. High Comm'r for Human Rights, The Right to Privacy in the Digital Age, U.N. Doc. A/HRC/39/29 (2018).
[3] Kate Crawford, Artificial Intelligence's White Guy Problem, N.Y. Times (June 25, 2016).

technology might perpetuate racial profiling.[4] These concerns highlight the importance of critical examination of the implications of AI systems from a human rights perspective.

A key problem in this discussion is the confusion of algorithmic neutrality. Contrary to common assumption, AI systems are not inherently unbiased, but actually often reflect and perpetuate existing inequalities in our society by being built with training data and system design that contain inequalities.[5] Historical data on which machine learning models have been trained may include discriminatory patterns which are then perpetuated and amplified through automated decision-making processes. As a result, marginalized and historically disadvantaged groups may be subject to systemic exclusion or unequal treatment. This phenomenon is problematic especially because it can happen without the explicit intention for discrimination and therefore challenging the traditional legal frameworks based on intent-based analysis.

The problem is being further compounded by the lack of transparency of many AI systems - commonly known as the "black box" problem. Complex machine learning models, especially those that use deep learning, lack transparency and it is often hard to understand how decisions are made.[6] This lack of explainability poses considerable barriers in terms of accountability as individuals who are affected by an adverse decision may be unable to identify, challenge and remedy discriminatory outcomes. Moreover, the asymmetry of information between AI developers and the affected individuals further complicates enforcement of legal rights, undermining procedural fairness and access to justice.[7]

In response to these challenges, several jurisdictions have started to develop regulatory frameworks which are aimed at governing AI and mitigating its risks. The European Union has become a global leader in that regard, through the adoption of General Data Protection Regulation (GDPR) and the proposal of the Artificial Intelligence Act.[8] The GDPR includes some crucial safeguards, such as rights concerned with automated decision-making and data protection, and the proposed AI Act uses a risk-based approach for the regulation of high-risk AI systems.[9] However, despite these improvements, there are large gaps in doctrine and enforcement. Existing legal frameworks sometimes have a hard time addressing forms of indirect discrimination stemming from algorithmic processes, and this is particularly the case when groups affected do not belong to traditional protected categories that are recognized under the anti-discrimination law.[10]

Furthermore, the appearance of "algorithmic groups" categories that are formed based on data-driven classifications, rather than legally-defined categories, presents a special challenge to traditional legal doctrines. These groups may be harmed in similar ways as those that are protected under existing laws but don't have formal recognition, limiting legal remedies for them. The inability to prove causation, and the dearth of access to algorithmic information also undermine the effectiveness of existing regulatory mechanisms.

Against this backdrop, this paper aims to controversially address the intersections between AI and human rights, through the lens of the adequacy of existing legal frameworks in tackling the problem of algorithmic discrimination. It examines the shortcomings of existing non-discrimination and data protection regimes, especially in the European context and highlights some key doctrinal and enforcement gaps. The paper makes the case for a human rights-centered approach to AI governance one that combines anti-discrimination law, the principles of data protection and sector-specific regulation to ensure accountability, transparency and fairness.

[4] Cathy O'Neil, Weapons of Math Destruction 3–10 (2016).
[5] Barocas & Selbst, supra note 1, at 677.
[6] Jenna Burrell, How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms, 3 Big Data & Soc'y 1 (2016).
[7] Lilian Edwards & Michael Veale, Slave to the Algorithm? 16 Duke L. & Tech. Rev. 18, 25 (2017).
[8] Regulation (EU) 2016/679 (General Data Protection Regulation).
[9] Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), COM (2021) 206 final.
[10] Sandra Wachter et al., Why a Right to Explanation of Automated Decision-Making Does Not Exist in the GDPR, 7 Int'l Data Privacy L. 76 (2017).

By recommending mechanisms like algorithmic audits, impact assessments and increased transparency obligations, this study seeks to add value to the creation of a strong legal framework that can address the multifaceted challenges posed by AI systems. Ultimately, the goal would be to ensure that technological advancement does not occur at the expense of fundamental human rights, but rather works in a way that strengthens and encourages the fundamental values of equality, justice and human dignity.

1. LITERATURE REVIEW

Scholarly discussion about the intersection between artificial intelligence (AI) and human rights has grown exponentially in the last decade as concern has increased worldwide about algorithmic bias, systemic inequality, and the consequences of automated decision-making for basic human rights. Early academic interventions had built the basis for the insight that ostensibly neutral technologies could help reproduce and even worsen pre-existing social hierarchies.

One of the most influential contributions in this respect is the work of Solon Barocas and Andrew D. Selbst, who showed how "big data's disparate impact" can lead to discriminatory consequences, even when there is no conscious bias.[11] They suggest that algorithmic systems, which are highly dependent on past data, perpetuate structural inequality imbued in past data to the detriment of already marginalised groups. This understanding laid to rest the traditional legal emphasis on intent in discrimination law and emphasized outcomes and systemic effects.

Similarly, Cathy O'Neil's seminal work Weapons of Math Destruction brings up the dangers of opaque and large-scale algorithmic systems that have no accountability.[12] O'Neil shows how these systems work in important areas such as education, employment, and criminal justice and can frequently perpetuate inequality in the name of objectivity. Her critique highlights the dangers of entrusting important decision-making power to systems that are not transparent and over which meaningful oversight is not provided.

Building on these foundational critiques, legal scholars have asked questions about the adequacy of existing anti-discrimination frameworks to redress algorithmic harms. Sandra Wachter, Brent Mittelstadt and Luciano Floridi believe that the traditional legal regimes are incapable of addressing algorithmic discrimination because they rely on clearly defined protected characteristics such as race, gender or religion.[13] In contrast, AI systems are likely to produce "data-driven" or "algorithmic" groupings that lack an appropriate legal framework of categories but may nonetheless have discriminatory effects. This mismatch has opened up a large doctrinal gap and harmful outcomes can escape legal scrutiny.

From the perspective of data protection, the General Data Protection Regulation (GDPR) of the European Union has become the most important regulatory tool. Scholars have been deeply analysing its provisions, and in particular, Articles 13-15 and Article 22 that provide individuals with rights in relation to automated decision-making, e.g. the right to be informed as well as in some cases the right not to be subject to solely automated decisions.[14] However, the effectiveness of such provisions is still in question. Lilian Edwards and Michael Veale suggesting that the much-discussed "right to explanation" is neither clearly articulated in the GDPR nor sufficiently robust to ensure meaningful accountability; They argue that technical complexity, trade secret and low enforcement capacity make the realisation of transparency an illusion.[15]

In parallel, a growing body of literature calls for using a human rights-based approach to AI governance. International organizations, especially the United Nations, have stressed that AI systems must be created and implemented in a way that is in line with established human rights norms, such as equality, privacy and due process.[16] The idea of "human-centric AI," championed by organizations like the European Commission and the Organisation for Economic Cooperation and Development (OECD), further supports the importance of

[11] Solon Barocas & Andrew D. Selbst, Big Data's Disparate Impact, 104 Calif. L. Rev. 671 (2016).
[12] Cathy O'Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy (2016).
[13] Sandra Wachter, Brent Mittelstadt & Luciano Floridi, Why a Right to Explanation of Automated Decision-Making Does Not Exist in the GDPR, 7 Int'l Data Privacy L. 76 (2017).
[14] Regulation (EU) 2016/679 of the European Parliament and of the Council (General Data Protection Regulation), arts. 13–15, 22, 2016 O.J. (L 119) 1.
[15] Lilian Edwards & Michael Veale, Slave to the Algorithm? Why a 'Right to Explanation' Is Probably Not the Remedy You Are Looking For, 16 Duke L. & Tech. Rev. 18 (2017).
[16] U.N. High Comm'r for Hum. Rts., The Right to Privacy in the Digital Age, U.N. Doc. A/HRC/39/29 (2018).

incorporating ethical principles, such as fairness, accountability, and respect for human dignity, into AI systems.[17]

Recent scholarship has also discussed the issues of indirect and intersectional discrimination as they relate to AI. Unlike traditional forms of discrimination, algorithmic bias is often subtle and diffuse and it is often difficult to prove causation or identify who is affected. Researchers have highlighted the issue of 'proxy discrimination', where apparently neutral variables (token of zip codes or buying behaviours) are used as a substitute for the protected characteristics, thus it would continue to perpetuate inequality, without any explicit mention of prohibited grounds.[18]

Despite these important contributions, a literature review shows that there remains gap between theory and practice. Enforcement mechanisms are still weak, especially when pitted against strong private actors over proprietary algorithms. Transparency requirements are inadequate to deal with the complexity of machine learning systems and current legal frameworks do not easily account for the fluidity and dynamicity of algorithmic classifications. Moreover, there is little agreement on how concepts such as fairness and accountability and explainability can be operationalized in concrete regulatory terms.

2. AI AND HUMAN RIGHTS: CONCEPTUAL FRAMEWORK

The intersection between Artificial Intelligence (AI) and Human Rights requires a strong conceptual framework that reflects the technical basis of algorithmic systems as well as their normative implications. AI systems are not simply neutral computational tools, but socio-technical constructs that are embedded in existing structures of power, inequality and governance. As such, their use within decision-making processes raises important issues of fairness, accountability and the protection of fundamental rights.

This conceptual framework is based on two interrelated dimensions:

(i) the nature and operation of algorithmic decision-making, in particular, the occurrence of bias in AI systems, and
(ii) the impact of such systems on fundamental human rights including equality, privacy, due process and human dignity, and

By combining knowledge from the fields of law, computer science, and human rights theory, this framework aims to offer comprehensive analytical lens to assess the use of algorithmic discrimination.

3.1 Decision-Making Algorithmic and Bias

AI systems, especially those powered by machine learning, are dependent on the analysis of large datasets for the detection of patterns and the making of predictions or decisions. These datasets are often based on historical records, which may be characterized by persistent social, economic and institutional bias. Consequently, AI systems are able to perpetuate and even magnify such biases to produce discriminatory outcomes with no explicit intent on the part of the developers or users.[19]

One of the major hurdles of understanding algorithmic discrimination is the fallacy that AI systems are objective. In reality, there is a series of subjective decisions to be made in the process of algorithmic decision-making, including the selection of data, feature engineering, model design, and optimization criteria. Each of these stages introduces the potential for bias, and therefore compromises the neutrality of the AI systems.

Algorithmic bias can be classified pretty broadly into three all-interrelated forms:

a) ***Data Bias***- Data bias occurs when the datasets used to train AI models are incomplete, unrepresentative or skewness. For example, if a recruitment algorithm is trained on historical data of hiring practices that reflect gender discrimination then the algorithm may learn to favour when hiring male candidates over equally qualified female candidates.[20] Similarly, predictive policing systems

[17] European Comm'n, Ethics Guidelines for Trustworthy AI (2019); OECD, Principles on Artificial Intelligence (2019).
[18] Andrew D. Selbst et al., Fairness and Abstraction in Sociotechnical Systems, 103 Proc. Conf. Fairness, Accountability & Transparency 59 (2019).
[19] Solon Barocas & Andrew D. Selbst, Big Data's Disparate Impact, 104 Calif. L. Rev. 671 (2016).
[20] Cathy O'Neil, Weapons of Math Destruction (2016).

that are trained on biased data on crime may be used to disproportionately target minority communities. Data bias is therefore a reflection of the wider social inequalities contained in historical records.

b) ***Model Bias***- Model bias is the bias that has been introduced during the design and implementation of the algorithm itself. This involves the choice of variables; the weighting mechanisms and assumptions built into the model. For instance, the use of proxy variables, such as zip codes, or educational background, may indirectly encode protected characteristics, e.g. race, or socio-economic status, which may lead to discriminatory results.[21] Additionally, optimization techniques that focus on accuracy but not fairness can also increase the disparities between different groups.

c) ***Outcome Bias***- Outcome bias occurs in the final decisions or predictions that are made by AI systems. Even without stated bias in the design of data or models, the combination of several factors can lead to disparate effects on some groups of people. Outcome bias is especially important from a legal standpoint, as it directly influences individuals' access to the opportunities and resources, such as employment, credit, and public services.[22]

A unique feature of algorithmic decision-making is the "black box" nature of many AI systems, especially those that are based on deep learning. These systems tend to be largely opaque, which makes it hard to see how certain decisions are made. This opacity causes serious difficulties in the detection and proof of discrimination, which affects the effectiveness of existing legal remedies.[23]

Furthermore, AI systems often make classifications according to complex correlations in the data, and this has resulted in the creation of "algorithmic groups" that do not align with traditional legal groups. This makes it difficult to apply anti-discrimination law, which is usually based on pre-defined protected grounds, such as race, gender or religion.[24]

3.2 Effect on Fundamental Rights

The use of AI systems in decision-making processes has far-reaching implications for a variety of fundamental human rights. These impacts go beyond individual cases of discrimination, impacting wider structures and governance mechanisms at society.

a) ***Equality and Non-Discrimination***- The principle of equality is a key element in human rights law, which mandates that individuals be treated equally and without unjustified discrimination. Algorithmic decision-making rebuts this principle by causing disparate outcomes on protected and unprotected groups. Unlike traditional forms of discrimination, algorithmic bias is often indirect so it is harder to identify and address.[25]

Moreover, the use of data to determine classifications can result in the exclusion of people who do not fit neatly within predefined classifications. This begs the question on the adequacy of current legal frameworks in dealing with the new forms of discrimination which are brought about by AI systems. The concept of indirect discrimination therefore has to be reinterpreted to include algorithmic decision-making processes.

b) ***Privacy***- AI systems heavily depend upon collection, processing and analyzing huge amounts of personal data. This raises serious concerns about the right to privacy, especially the topics of surveillance, profiling and data sharing. Advanced methods of AI are capable of inferring sensitive information such as political beliefs, health conditions and behavioural patterns, even from seemingly innocuous data.[26]

The loss of privacy is also made worse with the integration of AI into systems of public and private sectors, such as biometric surveillance and predictive analytics. Such practices may introduce a chilling effect on individual autonomy and freedom and deter democratic values.

[21] Sandra Wachter et al., Why a Right to Explanation of Automated Decision-Making Does Not Exist in the GDPR, 7 Int'l Data Privacy L. 76 (2017).
[22] Joshua A. Kroll et al., Accountable Algorithms, 165 U. Pa. L. Rev. 633 (2017).
[23] Frank Pasquale, The Black Box Society (2015).
[24] Sandra Wachter & Brent Mittelstadt, A Right to Reasonable Inferences, 2019 Colum. Bus. L. Rev. 494 (2019).
[25] European Union Agency for Fundamental Rights, Bias in Algorithms (2018).
[26] U.N. High Comm'r for Human Rights, The Right to Privacy in the Digital Age, U.N. Doc. A/HRC/39/29 (2018).

c) ***Due Process and Fair Trial Rights***- The use of AI in decision-making processes also raises concerns about procedural fairness and due process. Often, individuals who are affected by algorithmic decisions do not have access to meaningful explanations of such decisions or the means to challenge such decisions. This subverts basic principles of law such as the right to be heard, and the right to an effective remedy.[27] The opacity of AI systems makes it hard to hold these systems accountable; it can be unclear who is responsible for a particular decision the developer, the deployer or the data provider. This scattering of responsibility also further undermines legal protections and the impact of existing regulatory frameworks.

d) ***Human Dignity***- At a more basic level, the utilization of AI in decision-making processes raises concerns about human dignity. Automated systems can strip away the human complexity and uniqueness from human experience and reduce individuals to data points. Decisions that have a serious impact on people's lives such as eligibility for welfare benefits or sentencing in criminal courts may be made without human oversight or empathy.[28]
This dehumanization destroys the inherent value of individuals and questions the ethics underlying legal frameworks. A dignity-based approach to AI governance therefore emphasizes the importance of human control, transparency, and accountability in AI decision-making.

3. LEGAL FRAMEWORKS AND THEIR LIMITATIONS

The regulation of algorithmic discrimination as part of artificial intelligence (AI) systems is currently based on a mix between anti-discrimination law and data protection regimes, especially in the European Union. While these frameworks offer some important normative basis for protecting fundamental rights, they were not originally meant to address the complexities of AI-driven decision-making. As a result, their application shows great doctrinal and practical limitations, in particular in situations where opaque, data-driven systems are involved.

4.1 European Non-Discrimination legislation

European anti-discrimination law, based on international instruments such as the Equal Treatment Directives and the Charter of Fundamental Rights of the European Union, bans both direct and indirect discrimination.[29] Direct discrimination occurs in which people are treated less favourably on expressly stated grounds of race, gender, religion or disability. Indirect discrimination, on the other hand, is discrimination where seemingly neutral policies have an unjustifiable negative impact on certain protected groups.[30]

While there is nothing wrong with this framework conceptually, there are limitations to its success in tackling the question of algorithmic discrimination in a number of important ways.

First of all, the framework is very dependent on predetermined protected grounds. AI systems, however, tend to produce classifications according to complex correlations in the data and not according to categories recognized by law. For example, algorithmic profiling can segment people according to some behavioral patterns, geographical indicators or consumption habits, generating some "algorithmic groups" of individuals without the protection of classical legal principles.[31] As a result, people who suffer discrimination on these grounds may be without legal redress despite the fact that the harm they experience may be functionally equivalent to discrimination based on protected characteristics.

Second, there is a great evidentiary difficulty in demonstrating causation. To establish indirect discrimination, it is necessary to prove that a particular practice has a disproportionate negative effect on a protected group. In the case of AI, however, in terms of the complexity of machine learning models and the number of variables involved, that's a pretty difficult task. The lack of transparency in the decision-making

[27] Lilian Edwards & Michael Veale, Slave to the Algorithm? 16 Duke L. & Tech. Rev. 18 (2017).
[28] Luciano Floridi et al., AI4People- An Ethical Framework for a Good AI Society, 28 Minds & Machines 689 (2018).
[29] Charter of Fundamental Rights of the European Union arts. 20–21, 2012 O.J. (C 326) 391.
[30] Council Directive 2000/43/EC, 2000 O.J. (L 180) 22.
[31] Sandra Wachter, Brent Mittelstadt & Chris Russell, Counterfactual Explanations Without Opening the Black Box, 31 Harv. J.L. & Tech. 841 (2018).

process of algorithms also complicates the process of establishing a causal relationship between the design of the system and the discriminatory outcome.[32]

Third, there is lack of access to algorithmic data and decision-making logic. Claimants do not always have the technical expertise or legal entitlement to access the underlying datasets or models that AI systems use. This leads to an imbalance of structure between those who are deploying AI (individuals) and the institutions (state) which may seek to deploy AI; thus, undermining the effectiveness of anti-discrimination law as a tool for redress.[33]

In sum, even if European non-discrimination law offers an important normative framework, it has difficulties to adapt to the dynamic and opaque nature of algorithmic systems, leaving important gaps in the protection.

4.2 Data Protection Law (GDPR)

The General Data Protection Regulation (GDPR) is one of the most detailed pieces of legislation that takes into consideration the challenges of automated decision-making. It sets out a host of rights to safeguard people against the misuse of their personal information, many of which are directly relevant to AI systems.

Among its important protections are the right to information and the right to access, which require data controllers to make meaningful data on the processing of data available to individuals.[34] These rights should be used to promote transparency and empower individuals to be able to understand and challenge decisions that affect them.

Additionally, in Article 22 of the GDPR, individuals are also given protection from the application of solely automated processing upon which decisions are made with legal or similar-produced effects. This provision has been especially relevant in situations that involve credit scoring, recruitment, and the allocation of welfare, as AI systems can have significant impacts on people's lives.

Despite these safeguards, there are a number of limitations on the ability of the GDPR to deal with algorithmic discrimination.

One of the most widely debated issues is that of the ambiguity of the "right to explanation." While the GDPR does state that data controllers should provide "meaningful information about the logic involved," there is no explicit establishment of a full right to a detailed explanation of algorithmic decisions. Scholars have contended that this ambiguity restricts the practical value of the provision especially in more complicated machine learning situations wherein explanations can be hard to generate and interpret.

Furthermore, the GDPR only applies narrowly to some private sector algorithms, especially in the case that such decisions are not deemed to have "legal or similarly significant effects." Many AI systems, which are deployed in targeting advertising, in recommendation engines and in risk profiling, might not fall under the ambit of article 22, despite the potential impact on behaviour and re-enforcement of discrimination.

Another crucial challenge is poor enforcement mechanisms. While supervisory authorities are empowered to fine and make sure that they comply, enforcement has been patchy from jurisdiction to jurisdiction. Resource constraints, technical complexity and the global nature of AI systems further create challenges in conducting proper regulation.[35]

Thus, while the GDPR serves as a significant basis for the governance of AI, the provisions of the regulation are inadequate to fully address the dangers associated with algorithmic discrimination, especially with regards to transparency, scope and enforcement.

[32] Solon Barocas & Andrew D. Selbst, Big Data's Disparate Impact, 104 Calif. L. Rev. 671 (2016).
[33] Tal Zarsky, Transparent Predictions, 2013 U. Ill. L. Rev. 1503 (2013).
[34] Regulation (EU) 2016/679, arts. 13–15 (General Data Protection Regulation).
[35] Paul De Hert & Vagelis Papakonstantinou, The New General Data Protection Regulation, 28 Computer L. & Sec. Rev. 179 (2016).

4.3 Gaps in Existing Frameworks

The interplay between non-discrimination law and data protection law exposes a number of structural gaps that prevent these two bodies of law from working together to combat algorithmic discrimination.

A major concern is one of information asymmetry. People who are victims of algorithmic decisions typically have no access to the data, models, or decision-making processes that have produced those outcomes. This asymmetry not only restricts their capacity to contest the discriminatory practices but also undermines the accountability of the institutions using AI systems.[36]

Closely related is the problem of opacity, which has also been called the "black-box" character of AI. Many machine learning models work in ways that are not easily interpretable, even by the developers. This lack of transparency also makes it difficult to be held accountable under the law or subject to regulation that would enable an understanding of whether a system meets fundamental rights standards.[37]

Finally, there is a major problem of legal regime fragmentation. Anti-discrimination law and data protection law exist in relative isolation, with their own different objectives, standards and methods of enforcement. This lack of integration leaves gaps in regulations, especially in cases where algorithmic discrimination does not clearly fall under either of them.[38]

In order to find a more coherent and integrated approach to the solutions of these challenges, there is a strong need for a more coherent and integrated legal approach, which builds on the strengths of the existing frameworks, while remedying its limitations. Such an approach should focus on transparency, accountability and the preservation of human dignity in the design and deployment of AI systems.

4. JUDICIAL INTERPRETATION

Judicial interpretation has been an important factor in the development of the legal response to algorithmic discrimination and its implications for human rights. Courts of different jurisdictions have begun to tackle the challenges of artificial intelligence, in particular by finding a balance between technological innovation and fundamental rights such as privacy, equality and due process.

One of the most notable judicial interventions into this field is the Dutch decision of NCJM vs. State of the Netherlands (SyRI case) in which the District Court of the Hague examined the compatibility of an algorithmic welfare fraud detection system with human rights requirements The Court held that the use of the SyRI system was in contradiction to Article 8 of the European Convention of Human Rights (ECHR) because it held that the lack of transparency and safeguards meant that the interference with privacy was not necessary or proportionate.[39] This decision had been a landmark recognition that it is possible for opaque algorithmic systems to infringe fundamental rights and that they have to be subject to strict scrutiny.

Similarly, the changing jurisprudence of the Court of Justice of the European Union (CJEU) has helped to reinforce procedural safeguards for automated decision making. In a recent decision in 2025, the Court made it clear that people to whom automated decisions have been made must be given "meaningful and intelligible explanations" of the logic used.[40] The Court also made it clear that such explanations must create the possibility for individuals to understand and effectively verify and challenge decisions and thus to reinforce the procedural fairness and due process. Judicial interpretation has also focused on the inadequacy of existing legal frameworks, in this case GDPR, the General Data Protection Regulation. While Article 22 GDPR does provide for protection against purely automated decisions, courts and scholars have noted on the narrow scope of application of this provision, especially where human involvement is nominal or symbolic.[41] This leaves a regulatory gap for algorithmic decisions that can be free of any meaningful judicial scrutiny.

[36] Mireille Hildebrandt, Smart Technologies and the End(s) of Law (2015).
[37] Frank Pasquale, The Black Box Society (2015).
[38] Karen Yeung, Algorithmic Regulation, 12 Regulation & Governance 505 (2018).
[39] NJCM c.s./De Staat der Nederlanden (SyRI), Rechtbank Den Haag [District Court of The Hague]
[40] Case C-634/21, SCHUFA Holding AG, Judgment of the Court (CJEU)
[41] Regulation (EU) 2016/679 (General Data Protection Regulation), art. 22; Sandra Wachter et al., Why a Right to Explanation of Automated Decision-Making Does Not Exist in the GDPR, 7 Int'l Data Privacy L. 76 (2017).

In the United States, courts have taken the approach of traditional anti-discrimination law to the issue of algorithmic discrimination. Judicial review often is concerned with questions of whether or not a charge of discriminatory intent or disparate impact can be made. However, as noted in emerging jurisprudence, intent is a particularly hard thing to prove in algorithmic systems as biases are often built in indirectly through data or design of models as opposed to explicit decisions made by humans. This has led to the calls for a move to outcome-based assessments as opposed to intent-based frameworks.

Another dimension of judicial interpretation is important, that of the duty to provide reasons. Courts have turned more attention to the notion that automated decision-making should comply with principles of reasoned decision-making, which are a basic concept of administrative law. The obligation of explainability has been seen as not only necessary in order to ensure accountability, but also to ensure access to justice.[42] Without adequate explanations, the affected persons are in effect deprived of the means to challenge unfavorable decisions. Collectively, the judicial approaches paint a picture of a slow crawl towards an understanding of the specific dangers that AI systems pose: Courts are turning away from formalistic interpretation of existing laws to more substantive interpretation with human rights orientation. However, jurisprudence is fragmented and reactionary, and often treats with problems on an individual case by case basis rather than through the development of comprehensive doctrine.

6. Doctrinal Challenges and Emerging Issues

6.1 Groups and Protected Grounds That Are Algorithmic

Traditional anti-discrimination law is based on the identification of harm done to persons belonging to predefined protected categories, such as race, gender, religion or disability.[43] These categories have had the historical role to be the basis of recognizing structural disadvantage and ensuring equality before the law. However, the development of AI-driven systems of decision-making has disturbed this doctrinal basis by the concept of "algorithmic groups."

Algorithmic groups are groups of people that are the result of data analytics based on behavioural patterns, correlations or predictive attributes rather than legally recognized markers of identity. For example an AI system might group people based on the purchasing habits, geographical indicators or online behaviour of individuals and so might end up with groups of people that don't reflect the established protected characteristics. While such classifications may appear to be neutral, they may act as proxies for protected characteristics, and thereby create discriminatory outcomes without necessarily mentioning them.

This leaves a great doctrinal gap. Since the anti-discrimination laws generally require claimants to show that they are disadvantaged based on a protected ground, those who are harmed by algorithmic classification may have a difficult time showing that they have standing to sue. For example, a credit scoring algorithm that discriminates unnecessarily against people from certain postal codes may be discriminating against racial minorities, but the algorithm may not be subject to scrutiny because "postal code" is not a protected category.

Scholars have believed that this limitation warrants rethinking of what is protected either by an expansion of existing categories under notions of the protected grounds or a more flexible approach based on its effects.[44] The European Court of Human Rights and the Court of Justice of the European Union have found the reason to deal with intersectional and indirect forms of discrimination in some instances, but these are yet inadequate in the light of the complexity of algorithms.[45]

A promising solution is in the construction of a "hybrid protected grounds" framework which recognizes both traditional categories and data-driven groupings which have similar harms. Such an approach would steer our minds away from formal classification and towards substantive equality as we do our best to ensure that individuals are protected against discrimination regardless of how it is generated.

[42] Rethinking the Judicial Duty to State Reasons in the Age of Automation, Cambridge Forum on AI & Law (2025).
[43] Council Directive 2000/43/EC, 2000 O.J. (L 180) 22.
[44] Tarunabh Khaitan, A Theory of Discrimination Law 87–90 (2015).
[45] Case C-83/14, CHEZ Razpredelenie Bulgaria AD v. Komisia za zashtita ot diskriminatsia, 2015 E.C.R. I-0000.

6.2 Indirect Discrimination under AI Situations

Indirect discrimination (which is characterized as practices that are facially neutral, but have a disproportionately negative impact on certain groups) is particularly relevant in the context of AI systems. However, the extension of this doctrine to algorithmic decision-making is a special challenge.

First, proof of causation is not easy to make. Traditional legal analysis requires the ability to see a clear link between a policy or practice and the discriminatory effect of a policy or practice. In AI systems the decision-making process is often hidden and involves a number of different variables interacting in complex ways. This makes it difficult to identify the particular factor that is contributing to the discriminatory outcome. For example, a machine learning model used to make a recruitment decision can use hundreds of variables, so it can be almost impossible to demonstrate that a particular input resulted in gender or racial bias.

Second, the discriminatory effects are likely to be obscured in complex models. Many AI systems are "black boxes", in which even the developers do not understand exactly how the decisions are generated. This opacity contributes to the inability of affected people to contest discriminatory results, as well as the success of legal recourse.

Moreover, existing legal standards for indirect discrimination may be less than ideal for handling statistical and probabilistic harms inherent in algorithmic systems. Courts will normally rely on evidence including disparate impact analysis, it is often limited as to the access of relevant data due to trade secrets or privacy implications. This creates an information asymmetry between the individuals and developers of AI, which makes it even more difficult to enforce.

To overcome these challenges, scholars have proposed algorithmic impact assessments and tightened transparency requirements, which would involve organizations assessing and disclosing the potential discriminative effects of their systems.[46] Additionally, shifting the burden of proof in some cases - requiring developers to show the fairness of their systems - could help beef up protections against indirect discrimination.

6.3 Being Accountable to Individuals and Organisations

Determining accountability and liability in AI systems is one of the most important issues in modern law. Unlike traditional decision-making processes, there are various actors involved in AI systems such as developers, developers, data providers and end-users. This diffusion of responsibility makes it difficult to assign the liability when harm happens.[47]

For example, one possible cause of a discriminatory outcome in a hiring system that is powered by AI is biased training data, faulty algorithm design, or incorrect deployment. Identifying the party that is responsible requires a nuanced understanding of the lifecycle of the system that is lacking in many existing legal system frameworks. Furthermore, there can be confounding responsibilities due to the contractual arrangements between stakeholders, rendering the people affected with the problem with no effective remedy.

Another problem is the lack of clear regulatory standards for artificial intelligence systems. While there are some instruments like the GDPR and the proposed EU Artificial Intelligence Act giving some guiding, the issues of liability are not tackled comprehensively.[48] In many jurisdictions the existing laws are applied analogically which leads to uncertainty and inconsistent results.

The idea of strict liability has been suggested as a possible solution to this, particularly with high-risk AI applications.[49] Under this, organizations which implement AI systems would be held accountable for harm regardless of fault which would be an incentive to be more cautious and invest more in risk mitigation. Alternatively, a shared liability model can be used which can allocate responsibility to different actors based on the level of control and contribution to harm.

[46] OECD, Principles on Artificial Intelligence (2019).
[47] European Commission, Proposal for an Artificial Intelligence Act, COM (2021) 206 final.
[48] Regulation (EU) 2016/679 (General Data Protection Regulation).
[49] European Parliament, Civil Liability Regime for Artificial Intelligence (2020).

7. STRENGTHENING LEGAL SAFEGUARDS

The growing application of Artificial Intelligence (AI) to making high-stakes decisions requires a recalibration of existing legal frameworks in order to appropriately make provision for algorithmic discrimination. Current regimes especially anti-discrimination and data protection laws function in relative silos that result in doctrinal fragmentation and lack of enforcement efficiency. Strengthening legal safeguards calls for an integrated legal and transparent human rights-centred regulatory approach that is able to address the technical and normative dimensions of AI governance.

7.1 Anti-Discrimination Laws & Data Protection Laws Integration

A basic step towards good regulation is to apply theories and frameworks of anti-discrimination and data protection into a coherent legal architecture. Traditionally, anti-discrimination law is concerned with outcomes and equality, whilst data protection law is concerned with procedural safeguards such as consent, transparency and data minimization. However, algorithmic discrimination could often be faced at the intersection of these domains which demand a coordinated approach.

A unified framework should ensure that before AI systems are put into use, they are required to undergo Algorithmic Impact Assessments (AIAs), especially in sectors that are high risk. These assessments would measure the potential discriminatory effects, identify the vulnerable groups and suggest ways to mitigate the effects. Inspired by Data Protection Impact Assessments under the GDPR, AIAs can be used as preventive devices for ensuring compliance with equality as well as privacy norms.[50]

Further, the framework must specify a need for meaningful explanations for automated decisions. Whilst there is some protection against purely automated decisions under Article 22 of the GDPR, the extent of this is still controversial.[51] A reinforced regime should go beyond mere formalistic disclosures and towards substantive explanations that will allow individuals to understand, contest and correct algorithmic outcomes.

Equally as important is access to effective remedies. Victims of algorithmic discrimination are often faced with barriers such as lack of awareness, complexity of technology and evidentiary challenges. Integrating procedural rights related to data protection (for example, access and correction of data or data privacy regulations) with substantive rights related to anti-discrimination can help to increase legal recourse and accountability.

7.2 Algorithmic Audits & Transparency

Transparency is a foundation in a subject of accountable AI governance. However, the opacities of machine learning models especially those based on complex neural networks provide great challenges to oversight. In order to overcome this, independent algorithmic audits should be made a regulatory requirement.

Such audits would be conducted on AI systems to see if they are biased, fair, accurate, and in line with legal standards. Independent auditors with technical and legal expertise can be used to examine training data sets, model architecture and output patterns to determine if they have discriminatory impacts. These audits should be conducted both before a system is implemented and periodically throughout the life of the AI system.

In addition to this, there have to be public disclosure obligations for high-risk AI systems. Transparency registers, as proposed under the European Union's Artificial Intelligence Act, can be used to provide information about the purpose, functioning and classification of risks of AI systems. This not only helps build more trust among the public but it also sparks the external check of the civil society and researchers.[52]

Finally, continuous monitoring mechanisms are critical to ensure that AI systems are compliant with time. Given that machine learning models are updated by data inputs, static compliance checks are not good

[50] Regulation (EU) 2016/679 (General Data Protection Regulation), art. 35.
[51] Id. art. 22; Sandra Wachter et al., Why a Right to Explanation of Automated Decision-Making Does Not Exist in the GDPR, 7 Int'l Data Privacy L. 76 (2017).
[52] Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), COM (2021) 206 final.

enough. Regulatory frameworks should mandate real-time monitoring, reporting of incidents, and regular reassessment of emerging biases and systemic risks.[53]

7.3 Sector-Specific Regulation

Given the different levels of risk inherent in different uses of AI, there needs to be a sector specific regulatory approach. High-risk areas include employment, credit scoring, healthcare, welfare distribution and law enforcement as these sectors have to be more closely monitored due to the impact that they have on fundamental rights.

In these sectors, certification requirements should be introduced in order to ensure compliance of the AI systems with the predefined standards of fairness, accuracy and accountability prior to the put-in process. Certification bodies, which are governmental or accredited private organizations, can check for the compliance of the technical and ethical standards.[54]

On top of that there should be compliance standards developed for individual sectors. For example, AI systems deployed for recruitment purposes need to comply with the norms of equality, while those deployed in the healthcare sector need to focus on accuracy and patient safety. Sectoral guidelines can help bring clarity and consistency of regulatory expectations. Regular monitoring and control by the regulatory authorities is also important. This involves regular inspection, enforcement measures and penalties for non-compliance. The establishment of specialised supervisory bodies with technical expertise may make such supervision more effective.

7.4 Regulatory Approach founded on Dignity

On the normative level, AI regulation should be grounded on the principle of human dignity that is the foundation of all fundamental rights. A dignity-based approach has the benefit of shifting the emphasis away from procedural compliance and placing a greater emphasis on substantive justice and fairness. At the centre of this structure is the concept of substantive equality which recognizes that formal equality may be insufficient to combat systemic disadvantages: AI systems need to be designed and tested against real-world impact that they have on marginalized groups so that they don't perpetuate or increase existing inequalities. Human supervision is another important element. Fully automated systems for making decisions run the risk of dehumanizing people by reducing them to data points. So making sure that risky high stakes decisions are accompanied by meaningful human intervention can reduce such risks and preserve agency of individuals.

Finally, principles of ethical AI design such as fairness, accountability, transparency and inclusivity should be built into the lifecycle of AI systems. This requires interdisciplinary cooperation between technologists, legal experts and ethicists, in order to reconcile technological innovation with human rights values.

8. SUGGESTIONS AND POLICY RECOMMENDATIONS

To make above framework operational, a number of specific policy interventions are proposed.

First there is a need to use hybrid protected ground frameworks to transcend traditional categories like race and gender to something like algorithmic groups. These are data driven classifications, which may not have any correspondences with legally recognized classifications but which can still yield factors and effects that descend into discrimination. Understanding such groups in terms of the legal frameworks can help fill in the gaps already existing in doctrine.

Second, Algorithmic Impact Assessments should be mandated for all high-risk AI systems before they are deployed. These assessments must be standardized, be available to the public and undergo regulatory review.

Third, transparency requirements should be supplemented with clear and enforceable requirements for explainability. This includes not only technical explanations, but user-friendly disclosure that will enable the affected people to understand decision-making processes.

[53] OECD, Principles on Artificial Intelligence (2019).
[54] European Commission, Ethics Guidelines for Trustworthy AI (2019)

Fourth, governments should establish independent regulatory bodies that are focused on AI oversight. These authorities should have a multidisciplinary expertise and power to carry out audits and ensure compliance and settle disputes.

Fifth, access to justice must be enhanced through the simplification of challenging process around algorithmic decisions. This could include collective redress mechanisms, legal aid and changing the burden of proof in certain cases to redress the asymmetries of evidence.

Finally, there is a dire need to promote international cooperation in AI governance. Given the transnational nature of AI systems, harmonisation of regulatory standards across jurisdictions can help to prevent regulatory arbitrage and ensure consistent protection of human rights. Initiatives [led by] organizations such as the Organisation for Economic Cooperation and Development (OECD) and United Nations can help to develop global norms and best practices.

9. CONCLUSION

The accelerated introduction of artificial intelligence into decision making processes in public and private institutions is a revolution in the way in which rights, opportunities and resources are distributed in modern societies. As has been argued in this paper, AI systems are no longer marginal instruments but rather central players in governance structures affecting employment, creditworthiness, welfare distribution, policing and access to justice. While such systems are promising when it comes to increased efficiency, objectivity and scalability, there are also immense threats to the security of fundamental human rights. In particular, the emergence of algorithmic discrimination challenges the core principles of equality, non-discrimination, privacy, due process and human dignity on which modern legal systems are based. A crucial insight about this study is that algorithmic systems are not neutral in and of themselves. Instead they too reproduce and amplify pre-existing social biases that are embedded into historical data and institutional practices. This is leading to discriminatory results which can be indirect, opaque and hard to challenge.

The analysis further indicates that the existing legal frameworks are, in particular against the European context, not well prepared to adequately address these challenges. Although instruments such as anti-discrimination law or data protection regimes most notably the GDPR offer important safeguards these remain limited by doctrinal and practical limitations. The reliance on predefined protected grounds does not reflect the novel forms of classification produced by AI systems which in many cases are referred to as "algorithmic groups." Similarly, the concept of indirect discrimination is more complicated when it comes to multi-layered and data-driven decision-making processes. Enforcement gaps, information asymmetries and the limited scope of rights like the "right to explanation" add further to the ineffectiveness of such frameworks.

In the light of these shortcomings, this paper makes a case for a comprehensive and integrated regulatory approach which is above the fragmentation of existing legal regimes. Such an approach requires the harmonization of the anti-discrimination law, data protection principles and sector-specific regulations in order to find a coherent approach that can address the multifaceted nature of algorithmic harms. At the centre of this framework is the acceptance of human dignity as a basic principle of social norm. A dignity-based approach moves away from a focus on formal equality, and emphasizes the importance instead of substantive fairness in decision-making, by emphasizing the importance of ensuring that individual autonomy is respected, that dehumanization is avoided, and that human oversight plays a central role in decision-making processes that involve automation.

In order to make this vision practical, execution of strong safeguards are a must. Algorithmic audits and impact assessments should become a requirement, particularly in high-risk sectors, to ensure that biases are able to be proactively recognised and mitigated before becoming deployed. Transparency obligations should be enhanced so that people have access to meaningful explanations of automated decisions about them. Additionally, independent oversight bodies with technical expertise to monitor compliance and ensure accountability should be established. Sector-specific regulations based on areas such as employment, credit, welfare and law enforcement may further deal with risks in relation to the context and ensure that safeguards are proportionate. Ultimately the challenge is not to regulate AI as a technological phenomenon, but to channel the development of AI with reference to the values of the Constitution and human rights. This includes a forward looking and adaptive legal framework that is sensitive to technological change and is

founded on fundamental principles of justice and equality. As AI continues to alter the contours of decision-making, the responsibility of the legal systems is that it does not come at the expense of fundamental rights. By making human rights central to AI control, policymakers can develop a digital ecosystem that empowers people and leads not just to efficiency and progress, but also to fairness, accountability and respect for human dignity in an increasingly automated world.

Acknowledgments

The authors express sincere gratitude to Prof. (Dr) Santosh Kumar Tripathy, Vice-Chancellor, Fakir Mohan University, Balasore; Prof. (Dr) Amulya Kumar Acharya, Coordinator, PG. Department of Law, Fakir Mohan University, Balasore; Prof. (Dr) Binay Kumar Das, Dean, PG. Department of Law, Fakir Mohan University, Balasore; the esteemed Judges, senior advocates and junior fellow advocates of Orissa High Court for their invaluable support.

10. REFERRENCES

[1] Case C-634/21, SCHUFA Holding AG (Scoring), ECLI:EU:C:2023:957 (Ct. Just. Dec. 7, 2023).
[2] Case C-83/14, CHEZ Razpredelenie Bulgaria AD v. Komisia za zashtita ot diskriminatsia, ECLI:EU:C:2015:480 (Ct. Just. July 16, 2015).
[3] Cathy O'Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy (Crown 2016). DOI: Not verified for the published book edition.
[4] Charter of Fundamental Rights of the European Union arts. 20–21, 2012 O.J. (C 326) 391.
[5] Council Directive 2000/43/EC, of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin, 2000 O.J. (L 180) 22.
[6] European Commission, Ethics Guidelines for Trustworthy AI (2019). DOI: 10.2759/346720.
[7] European Commission, Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, COM (2021) 206 final (Apr. 21, 2021).
[8] European Parliamentary Research Service, Civil Liability Regime for Artificial Intelligence (2020). DOI: 10.2861/737677.
[9] European Union Agency for Fundamental Rights, Bias in Algorithms: Artificial Intelligence and Discrimination (2022). DOI: 10.2811/25847.
[10] Frank Pasquale, The Black Box Society: The Secret Algorithms That Control Money and Information (Harvard Univ. Press 2015). DOI: 10.4159/harvard.9780674736061.
[11] Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson & Harlan Yu, Accountable Algorithms, 165 U. Pa. L. Rev. 633 (2017).
[12] Karen Yeung, Algorithmic Regulation: A Critical Interrogation, 12 Regulation & Governance 505 (2018). DOI: 10.1111/rego.12158.
[13] Lilian Edwards & Michael Veale, Slave to the Algorithm? Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking For, 16 Duke L. & Tech. Rev. 18 (2017).
[14] Luciano Floridi, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke & Effy Vayena, AI4People - An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations, 28 Minds & Machines 689 (2018). DOI: 10.1007/s11023-018-9482-5.
[15] Mireille Hildebrandt, Smart Technologies and the End(s) of Law (Edward Elgar 2015). DOI: 10.4337/9781849808774.
[16] NJCM c.s./De Staat der Nederlanden (SyRI), ECLI:NL:RBDHA:2020:865 (Rb. Den Haag Feb. 5, 2020).
[17] OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449 (May 22, 2019).
[18] Paul De Hert & Vagelis Papakonstantinou, The New General Data Protection Regulation: Still a Sound System for the Protection of Individuals? 32 Comput. L. & Sec. Rev. 179 (2016). DOI: 10.1016/j.clsr.2016.02.006.

[19] 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), 2016 O.J. (L 119) 1.
[20] Sandra Wachter & Brent Mittelstadt, A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI, 2019 Colum. Bus. L. Rev. 494 (2019). DOI: 10.7916/cblr.v2019i2.3424.
[21] Sandra Wachter, Brent Mittelstadt & Luciano Floridi, Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation, 7 Int'l Data Priv. L. 76 (2017). DOI: 10.1093/idpl/ipx005.
[22] Sandra Wachter, Brent Mittelstadt & Chris Russell, Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR, 31 Harv. J.L. & Tech. 841 (2018). DOI: 10.2139/ssrn.3063289.
[23] Solon Barocas & Andrew D. Selbst, Big Data's Disparate Impact, 104 Calif. L. Rev. 671 (2016). DOI: 10.15779/Z38BG31.
[24] Tal Z. Zarsky, Transparent Predictions, 2013 U. Ill. L. Rev. 1503. DOI: Not verified for the published version.
[25] Tarunabh Khaitan, A Theory of Discrimination Law (Oxford Univ. Press 2015). DOI: 10.1093/acprof:oso/9780199656967.001.0001.
[26] U.N. High Comm'r for Human Rights, The Right to Privacy in the Digital Age, U.N. Doc. A/HRC/39/29 (Aug. 3, 2018).
[27] Victoria Hendrickx, Rethinking the Judicial Duty to State Reasons in the Age of Automation? 1 Cambridge Forum on AI: L. & Governance e26 (2025). DOI: 10.1017/cfl.2025.11.