



INTERNATIONAL JOURNAL OF HUMAN RIGHTS LAW REVIEW

An International Open Access Double Blind Peer Reviewed, Referred Journal

Volume 5 | Issue 3 | 2026

Art. 9

**Explainability as a Constitutional Requirement:
Can “Black Box AI” ever be Judicially
Legitimate?**

Reclaiming Reasoned Justice in the age of Algorithms

Arkapriya Ghosh

Law Student, 5th Year, BA.LL.B.,

Jogesh Chandra Chaudhuri Law College, Calcutta University

Recommended Citation

Arkapriya Ghosh, *Explainability as a Constitutional Requirement: Can “Black Box AI” ever be Judicially Legitimate?*, 5 IJHRLR 116-127 (2026).

Available at www.humanrightlawreview.in/current-issues/.

This Article is brought to you for free and open access by the International Journal of Human Rights Law Review by an authorized Lex Assisto & Co. administrator.

For more information,
please contact humanrightlawreview@gmail.com

Explainability as a Constitutional Requirement: Can “Black Box AI” ever be Judicially Legitimate?

ABSTRACT

Artificial Intelligence is no longer a speculative presence in judicial systems. Artificial Intelligence is used for case management software to algorithmic risk assessments used for bail and sentencing decisions across jurisdictions. Courts are increasingly engaging and investing themselves in automated tools which promise efficiency, consistency and speed. Yet many of these systems still operate as “Black Boxes”- producing outcomes without intelligible explanations. The inclusion of Artificial Intelligence in judicial administration challenges the foundational constitutional question - Can judicial legitimacy survive when reasons are replaced by results? Through an analysis of precedents set by the Supreme Court, on reasoned orders and non-arbitrariness, the article demonstrates that opacity in decision making technologies often undermines equality before laws and erodes procedural dignity. Comparative insights from the European Union’s regulatory approach and the experimentation of the United States with algorithmic risk assessment tools further reveal a global judicial reluctance in application of unexplainable systems in adjudication. Through this article, we cross-examine whether the use of non-explainable Artificial Intelligence in judicial decision making is compatible with the principles set by the Indian Constitution. This article argues that explainability is not a mere desirable feature of algorithmic systems but a constitutional requirement in accordance with the provisions of Article 14 and Article 21 of the Indian Constitution. The duty to give reasons, the guarantee of a fair and just procedure and prohibition of arbitrariness collectively demand that any tool influencing judicial outcomes must be capable of meaningful explanation.

KEYWORDS

Artificial Intelligence, Explainability, Black Box Algorithms, Judicial Reasoning, Article 14, Article 21, Natural Justice, Fair Trial, Constitutional Adjudication.

INTRODUCTION

The legitimacy of judicial power in a constitutional democracy is not merely based on outcomes but on the reasons which form the basis of such outcomes. Courts do not command obedience because they are

efficient or accurate but because they justify their decision making through intelligible reasoning. This obligation to give reasoning has been recognised as an integral component of the rule of law, theory of natural justice and constitutional accountability¹. Recent developments in Artificial Intelligence have challenged this foundational premise.

Judicial systems of both intra territorial jurisdictions and extra territorial jurisdictions are experimenting with algorithmic tools for case management, case allocation, legal research and decision making in complex matters like bail and sentencing². While these technologies promise efficient administration, many of them operate through complex models whose internal logic is obscure and incomprehensible even to their own designers³. Such systems which generate conclusions without offering reasons are incapable of being questioned, examined or challenged.

This opacity, however, is contradictory to the principles of the Constitution. Judges are bound by Article 14 and Article 21 of the Indian Constitution to act non - arbitrarily and to follow fair, just and reasonable procedures⁴. Adjudication requires transparency in the decision-making process, ability of affected parties to understand the basis of adverse outcomes and the possibility of a judicial review. When judicial reasoning cannot explain itself and is influenced - directly or indirectly by an algorithmic system, the constitutional commitments get challenged.

UNDERSTANDING “BLACK BOX” AI IN THE JUDICIAL CONTEXT

The term “Black Box”⁵ Artificial Intelligence refers to decision making systems whose internal reasoning processes are not accessible or intelligible to users, affected parties, or in most of the cases their own technical creators. While such systems may often generate output with impressive speed or statistical accuracy, they do so without providing proven judicial explanations which can be meaningfully understood,

¹ S.N. Mukherjee v. Union of India, (1990) 4 SCC 594,35-36 (India) available at <https://indiankanoon.org/doc/400596/> (Last visited January 17, 2025).

² Sonja B. Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, 66 Stan. L. Rev. 803, 809-15 (2014).

³ Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, 3-8 (Harvard Univ. Press 2015).

⁴ *Maneka Gandhi v. Union of India*, (1978) 1 SCC 248, 281-84 (India) available at <https://indiankanoon.org/doc/1766147/> (Last visited January 17, 2025)

⁵ Jenna Burrell, *How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms*, 3 Big Data & Soc’y 1, 3-5 (2016).

tested and contested against.

Different types of algorithmic systems are used in decision making. Rule based systems offer through pre-defined instructions - *if certain conditions are met, a specific outcome follows*. This logic is transparent and explainable. By contrast, most “black box” systems rely on machine learning models trained on large data sets. These models identify patterns and co relations autonomously adjusting internal parameters in ways that are not interpretable in human terms. The system has been trained to predict an outcome, but cannot articulate and portray the factual basis on which that prediction tests⁶.

Opacity in these systems often arise due to multiple reasons;

- Firstly, machine learning models often rely on co-relations which defer explanations. Even when both input and outputs are visible, the link between them remains obscured.
- Secondly, many algorithms used in judicial or quasi-judicial contexts are developed by private technologists who can claim proprietorship⁷ over their models.

This restricts disclosure not only to litigants but sometimes even to institutions who are deploying the technology. A commonly cited legal example is the use of algorithmic risk assessment tools in bail or sentencing decisions across various jurisdictions⁸. The systems generate probable scores via statistical analysis predicting likelihood of re-offending or absconding based on variables derived from past precedents and data. This score, while influencing a judicial decision, cannot provide sufficient explanation to the individuals subject to this discretion. The reasoning is embedded in code, inaccessible to scrutiny or challenge.

Courts do not merely pronounce outcomes, they justify them. The adjudication shouldn't be exhausted by accuracy or efficiency and instead should be fulfilled by reasoned explanation. In legal contexts, the absence of explainable reasoning is not a peripheral concern - it goes to the heart of adjudicatory legitimacy. Courts deal with reasons, not

⁶ Finale Doshi-Velez & Been Kim, *Towards a Rigorous Science of Interpretable Machine Learning*, arXiv:1702.08608, at 2–4 (2017).

⁷ Andrew D. Selbst & Julia Powles, *Meaningful Information and the Right to Explanation*, 7 Int'l Data Privacy L. 233, 239–42 (2017).

⁸ *State v. Loomis*, 881 N.W.2d 749, 761–63 (Wis. 2016) available at <https://courts.ca.gov/sites/default/files/courts/default/2024-12/btb24-21-3.pdf> (Last visited January 18, 2025).

results⁹.

THE CONSTITUTIONAL DUTY TO GIVE REASONS

A. Reasoned Orders as a Pillar of Natural Justice

The obligation to give reasons is not a procedural formality but a substantive component of natural justice. It works in accordance with the principle of the “Rule of Law”. Indian constitutional jurisprudence has consistently treated reasoned decision-making as an indispensable and essential safeguard against arbitrariness. In *Siemens Engineering v. Union of India*¹⁰, the Supreme Court held that the requirement of reasons ensures fairness and excludes the possibility of caprice.

This ruling was reaffirmed and elaborately portrayed and acted as a precedent in *Kranti Associates v. Masood Ahmed Khan*¹¹, where the court recognised reasoned orders as essential to transparency, accountability and public confidence in adjudication.

Reasons and explanations perform multiple constitutional functions. They operate as a check upon arbitrary exercise of power by compelling the decision maker to rationally justify conclusions. They serve as a tool of accountability by exposing judicial reasoning to public and appellate scrutiny. Most importantly, reasons form the basis of the right to appeal enabling higher courts to assess legality and validity of decisions¹².

Without proper explanations, adjudication degenerates into an authority with justification- an outcome fundamentally inconsistent with constitutional governance.

B. Explainability as a Constitutional Extension

If judges are constitutionally bound to provide reasons, any tool influencing judicial reasoning should itself be capable of giving an

⁹ *Mohinder Singh Gill v. Chief Election Commissioner*, (1978) 1 SCC 405, 417 (India) available at <https://indiankanoon.org/doc/1831036/> (Last visited January 18, 2025).

¹⁰ *Siemens Engineering and Manufacturing Company of India Limited v. Union of India*, (1976) 2 Supreme Court Cases 981, 986 (India) available at <https://ipleader.in/wp-content/uploads/2024/06/siemens-engineering-vs-uo-1976.pdf> (Last visited January 18, 2025).

¹¹ *M/s. Kranti Associates Pvt. Ltd. v. Masood Ahmed Khan*, (2010) 9 SCC 496, 47-51 (India), available at <https://indiankanoon.org/doc/1304475/> (Last visited January 18, 2025).

¹² *Union of India v. Mohan Lal Capoor*, (1973) 2 Supreme Court Cases 836, 843-44 (India) available at <https://indiankanoon.org/doc/714743/> (Last visited January 18, 2025).

explanation. This is not an additional burden imposed upon by technology but a logical extension of existing constitutional principles. Judicial reasoning cannot be replaced by opaque computational outputs without undermining the very duty of arbitrariness which legitimizes judicial power.

When an algorithm shapes a judicial decision, its opacity is constitutionally significant. As a matter of constitutional logic, a judge cannot outsource reasoning to a system which cannot reason publicly and provide proper explanations¹³. Explainability is not a technical preference - it's a constitutional necessity inherent in the duty to give reasons¹⁴.

ARTICLE 14 AND THE PROBLEM OF ALGORITHMIC ARBITRARINESS

Article 14 of the Indian Constitution embodies a guarantee against arbitrariness as much as it makes a promise of equality. Since *E.P. Royappa v. State of Tamil Nadu*¹⁵ arbitrariness has been recognised as the antithesis of equality, rendering any state action that is unreasonable or irrational is constitutionally void. This principle was further strengthened in *Maneka Gandhi v. Union of India*, where the court held that fairness and non-arbitrariness are inseparable from constitutional validity¹⁶.

Black Box AI systems pose a distinct challenge to this framework. By this very design, such systems generate outcomes which are neither visible nor intelligible to those affected. In doing so, they create hidden classifications - grouping and risk profiles which are not articulated rationally. Individuals may be treated differently based on patterns detected in data - without any disclosure of the basis of criteria¹⁷ that produced such differentiation. This absence of transparency makes it impossible to assess whether classification bears a rational nexus to legitimate objective¹⁸, a core requirement under Article 14.

¹³ *Mohinder Singh Gill v. Chief Election Commissioner*, (1978) 1 Supreme Court Cases 405, 417 (India) available at <https://indiankanoon.org/doc/1831036/> (Last visited January 18, 2025).

¹⁴ *S.N. Mukherjee v. Union of India*, (1990) 4 Supreme Court Cases 594, 613-14 (India) available at <https://indiankanoon.org/doc/400596/> (Last visited January 17, 2025)

¹⁵ *E.P. Royappa v. State of Tamil Nadu*, (1974) 4 Supreme Court Cases 3, 38 (India) available at <https://indiankanoon.org/doc/1327287/> (Last visited January 18, 2025).

¹⁶ *Maneka Gandhi v. Union of India*, (1978) 1 Supreme Court Cases 248, 281-84 (India) available at <https://indiankanoon.org/doc/1766147/> (Last visited January 17, 2025)

¹⁷ Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 Calif. L. Rev. 671, 677-80 (2016)

¹⁸ *State of West Bengal v. Anwar Ali Sarkar*, 1952 Supreme Court Reports 284, 297 (India)

Moreover algorithmic systems are generally trained on precedents which often replicate existing structural biases embedded in society¹⁹. When these biases are absorbed in opaque models, discrimination becomes harder to detect and easier to accept. What appears as neutral may often perpetuate unequal treatment under the disguise of objectivity. Opacity thus aggravates arbitrariness. When neither party can scrutinise the basis of differential treatment, constitutional review is often biased which goes against the principles of Right to Equality under Article 14.

Black Box AI introduces a new form of arbitrariness - one that is automated, concealed and therefore constitutionally suspect. Where classification cannot be explained, equality before law is reduced to an illusion.

ARTICLE 21 - FAIR TRIAL AND PROCEDURAL JUSTICE

Article 21 guarantees not merely the protection of life and personal liberty, but also provides the assurance that any deprivation of the same will follow a procedure which is fair, reasonable and just²⁰. The right to a fair trial is a central component of this guarantee²¹. At its core lies the basic right of an individual to know a basis of a decision affecting their rights has been taken and to meaningfully challenge any adverse adjudication²² by the judiciary.

The use of Black Box in AI in quasi-judicial or judicial context undermines these procedural safeguards. When an algorithm influences certain adjudication without providing any intelligible reasons, the affected individual is denied an effective hearing. A party cannot respond to, rebut or contextualise reasoning that remains concealed within computational processes. The absence of specified reasons often frustrate meaningful scrutiny by higher courts²³. These concerns are more severe and acute where judicial decisions directly affect the

available at <https://blog.iplers.in/state-of-west-bengal-vs-anwar-ali-sarkar-1952/> (Last visited January 18, 2025).

¹⁹ Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* 11–15 (St. Martin’s Press 2018)

²⁰ *Maneka Gandhi v. Union of India*, (1978) 1 Supreme Court Cases 248, 281–84 (India) available at <https://indiankanoon.org/doc/1766147/> (Last visited January 17, 2025)

²¹ *Zahira Habibullah Sheikh v. State of Gujarat*, (2004) 4 Supreme Court Cases 158, 181–82 (India) available at <https://indiankanoon.org/doc/105430/> (Last visited January 17, 2025).

²² *Union of India v. Mohan Lal Capoor*, (1973) 2 Supreme Court Cases 836, 843–44 (India) available at <https://indiankanoon.org/doc/714743/> (Last visited January 18, 2025).

²³ *State of Punjab v. Jagdev Singh Talwandi*, (1984) 1 Supreme Court Cases 596, 603 (India) available at <https://indiankanoon.org/doc/1158281/> (Last visited January 18, 2025).

fundamental rights of an individual - right to liberty and dignity²⁴ in matters of criminal proceedings, bail determinations, sentencing and even matrimonial disputes involving custody and maintenance.

Procedural justice under Article 21 is not satisfied by efficiency alone. It requires participation, explanation and the right to appeal. When judicial reasoning is displaced and obscured by opaque algorithmic outputs, the individual is effectively excluded from the reasoning process.

The concern is fundamental: a person is reduced to a data point where reasons are replaced by risk scores. Such action is incompatible with the constitutional promise that justice is to be provided with human reasoning and empathy and not inscrutable computation.

COMPARATIVE JUDICIAL AND REGULATORY APPROACHES

Comparative experiences often provide insight into how various jurisdictions and their legal systems respond to the threat posed by the algorithmic opacity in adjudication. While approaches differ, a common point is the judicial unease with unexplainable and non - scrutiny of decision-making processes.

The European Union has adopted a precautionary regulatory stance through proposed EU Artificial Intelligence Act. The framework recognizes certain AI systems such as “high risk”, particularly those used in law enforcements and judicial contexts²⁵. It mandates explainability, transparency and human reasoning, ensuring that automated systems affecting human rights must remain accountable²⁶ and subordinate to human judgement. The regulatory emphasis is not on prohibiting AI, but on embedding safeguards that preserve legal responsibility and reasoned decision-making.

The United States presents a more fragmented picture. Algorithmic risk assessment tools like COMPAS has been used in bail and sentencing decisions, prompting significant judicial controversy²⁷. Courts have often expressed dissatisfaction with reliance on proprietary algorithms whose

²⁴ *Mohd. Arif @ Ashfaq v. Registrar, Supreme Court of India*, (2014) 9 Supreme Court Cases 737, 751–52 (India) available at <https://blog.ipleaders.in/mohd-arif-ashfaq-vs-the-reg-supreme-court-of-india-2014/> (Last visited January 17, 2025).

²⁵ Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), arts. 6, 9–14, COM (2021) 206 final (Apr. 21, 2021)

²⁶ European Union Agency for Fundamental Rights, Getting the Future Right – *Artificial Intelligence and Fundamental Rights* 41–45 (2020)

²⁷ *State v. Loomis*, 881 N.W.2d 749, 761–66 (Wis. 2016) available at <https://courts.ca.gov/sites/default/files/courts/default/2024-12/btb24-21-3.pdf> (Last visited January 18, 2025).

workings are shielded from disclosure, raising questions about equal protection and due process. While some judicial opinions permitted limited use, they have acknowledged the constitutional risks posed by opacity and difficulty of maintaining adversarial fairness in reasoning.

India, by contrast lacks proper comprehensive regulatory framework governing use of AI in decision making. Existing usage of Artificial Intelligence is largely administrative and formal adjudication still relies on human discretion. In the absence of statutory guidance, constitutional principles still remain a primary safeguard²⁸. Articles 14 and 21 along with the established principle of Natural Justice and Rule of Law provide a framework against which any integration of AI must be assessed.

As jurisdictions grapple with technological innovations, constitutional values serve as a limiting principle against uncritical adoption. For India, the absence of regulation does not create a vacuum but an obligation that India's constitutional morality demands higher safeguards and not blind adoption.

CAN BLACK BOX AI EVER BE JUDICIALLY LEGITIMATE?

The legitimacy of any tool used for adjudication must always be measured in accordance with constitutional standards²⁹ and not technological capability. From this perspective, Black Box AI can only be accommodated in judicial systems in a limited manner. The answer to the question, *Can Black Box Ai Ever Be Judicially Legitimate?* is therefore a qualified NO.

Black Box AI may legitimately assist courts in administrative functions which do not play a role or shape judicial reasoning³⁰. Applications such as case management, digitalization of records and even precedents retrieval operate at level of efficiency rather than adjudication. These usages do not influence the quality of judicial decisions and do not displace the judge's inherent duty to reason, explain and justify outcomes.

However, the constitutional difficulty arises when the opaque algorithmic tools move from administrative assistance to decision

²⁸ *Justice K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 Supreme Court Cases 1, 247–48 (India) available at <https://indiankanoon.org/doc/91938676/> (Last visited January 18, 2025).

²⁹ *State of Madras v. V.G. Row*, 1952 Supreme Court Reports 597, 607–08 (India) available at <https://indiankanoon.org/doc/554839/> (Last visited January 18, 2025).

³⁰ *Central Board of Secondary Education v. Aditya Bandopadhyay*, (2011) 8 Supreme Court Cases 497, 523–24 (India) available at <https://indiankanoon.org/doc/1519371/> (Last visited January 18, 2025).

making³¹. Tools that assess risk, predict outcomes and recommend bail actions, sentencing or custody matters often directly influence judicial discretion. When such influence is exerted through certain algorithmic processes which cannot be explained or scrutinized, constitutional safeguards get questioned regarding their reliability. Judicial reasoning becomes immune to review and accountability.

The distinction between assistance and adjudication is therefore critical. While precedent retrieval tools merely perform administrative tasks like organising for human evaluation, risk assessment systems introduce new judgments without giving any reasons for basis. It alters the reasoning process itself under the guise of advisory.

Efficiency is valuable but is not a constitutional justification for opacity³². The legitimacy of judicial decisions flows from reasoned justification and not computational speed and predictive accuracy. Efficiency can aid justice, but cannot replace reasoning. Where explanation is absent, judicial legitimacy cannot co-exist.

SAFEGUARDS: DESIGNING CONSTITUTIONALLY COMPLIANT AI

If Artificial Intelligence is to be integrated into judicial systems, it should be done without eroding constitutional values. Its design and deployment must be governed by clear safeguards. The object is not technological restraint, but constitutional alignment.

Various safeguards for designing constitutionally compliant AI are :-

- First, AI Systems to be used within judicial processes must meet mandatory explainability standards. Where an algorithm influences decision making, it must be capable of providing intelligible reasons which can be understood, evaluated and contested. Outputs without explanations cannot satisfy constitutional requirements of fairness and non-arbitrariness³³.

³¹ *State v. Loomis*, 881 N.W.2d 749, 761–66 (Wis. 2016) available at <https://courts.ca.gov/sites/default/files/courts/default/2024-12/btb24-21-3.pdf> (Last visited January 18, 2025).

³² *Maneka Gandhi v. Union of India*, (1978) 1 Supreme Court Cases 248, 281–84 (India) available at <https://indiankanoon.org/doc/1766147/> (Last visited January 17, 2025)

³³ Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* 155–60 (Harvard University Press 2015)

- Secondly, systems must be subject to regular audit including independent assessments for bias, error and disproportionate impact³⁴.
- Thirdly, there must be a prohibition on some reliance on AI generated outputs in adjudication. No decision affecting rights or liberties should rest exclusively on algorithmic assessments, particularly in criminal or family law contexts.
- Fourthly, meaningful human override must be preserved. Judges must retain complete discretion to accept, modify or reject algorithmic inputs without pressure or deference. Responsibilities for decision must be human and judicial³⁵.
- Finally, disclosure to affected parties is essential and mandatory. Litigants must be informed when AI tools are used in their case and provided a basis on which such tools operate³⁶.

The safeguards reflect a principled framework - *AI AS A JUDICIAL AID, NOT A JUDICIAL ACTOR*.

CONCLUSION

The integration of artificial intelligence in judicial systems presents both promise and peril. While technological tools may enhance efficiency and administrative capacity, they also test constitutional foundations upon which adjudication rests. Explainability in judicially deployed AI is not a matter of technological practice but a constitutional requirement³⁷. While Artificial Intelligence can legitimately aid judges in administrative and research functions, the use of “Black Box” systems in decision shaping roles is incompatible with the constitutional duty to give reasons³⁸. Explainability is not an optional feature of judicial decision making in the age of algorithms but a constitutional requirement

³⁴ European Commission, *Ethics Guidelines for Trustworthy Artificial Intelligence*, 18–20 (Apr. 8, 2019)

³⁵ Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), arts. 14–16, COM (2021) 206 final (Apr. 21, 2021)

³⁶ Andrew D. Selbst & Julia Powles, *Meaningful Information and the Right to Explanation*, 7 Int'l Data Privacy L. 233, 239–42 (2017)

³⁷ *S.N. Mukherjee v. Union of India*, (1990) 4 Supreme Court Cases 594, 613–14 (India) available at <https://indiankanoon.org/doc/400596/> (Last visited January 17, 2025)

³⁸ *Mohinder Singh Gill v. Chief Election Commissioner*, (1978) 1 Supreme Court Cases 405, 417 (India) available at <https://indiankanoon.org/doc/1831036> (Last visited January 18, 2025).

grounded in Article 14 and Article 21 of the Indian Constitution.

This article does not reject or stand against the usage and inclusion of AI in judiciary per se but tries to advance a calibrated framework which may attempt to distinguish between administrative assistance and decision shaping influence. While AI may aid in judicial efficiency, it cannot constitutionally substitute or obscure judicial reasoning.

As India cautiously navigates integration of AI in judicial processes, constitutional principles must act as a compass. Technology must adapt to Constitutional values and not the reverse. Any future use of AI in adjudication must reinforce and not erode the culture of reasoned justice that defines constitutional courts.

In a constitutional democracy, justice must be based on reasons - not in code.