

BEYOND THE VEIL OF IGNORANCE: RAWLSIAN JUSTICE AND THE JURISPRUDENCE OF ALGORITHMIC GOVERNANCE IN INDIA

by

Gregory Koshy Thomas

Assistant Professor of Law

School of Law, PRIST Deemed to be University

Tamil Nadu, India

&

M.S. Amrithavarshini

B.A., LL.B. Student

School of Law, PRIST Deemed to be University

Tamil Nadu, India

ABSTRACT

Algorithmic systems increasingly mediate welfare delivery, policing, taxation, public services, and judicial administration in India. Their attraction lies in scale, speed, and the promise of consistent decision-making. Their constitutional difficulty lies in the fact that a computational output can distribute burdens and benefits without making its reasons intelligible to the person affected. This article develops a Rawlsian jurisprudence for Indian algorithmic governance. It argues that the original position and the veil of ignorance provide a disciplined test for institutional design: decision-makers should choose rules without knowing whether they will be digitally included or excluded, data-rich or data-poor, socially dominant or historically marginalised, or the beneficiary of an automated classification or its adverse subject. Rawls's equal basic liberties, fair equality of opportunity, and difference principle are translated into constitutional requirements of legality, non-arbitrariness, proportionality, explainability, contestability, and protection against wrongful exclusion. The article situates this normative framework within Indian debates on Articles 14 and 21, privacy, Aadhaar-linked welfare, algorithmic bias, and administrative accountability. It proposes a rights-preserving architecture

consisting of a public algorithm register, mandatory pre-deployment impact assessments, meaningful reasons, human review, independent audits, accessible remedies, and a non-waivable analogue route for essential entitlements. Rawlsian reasoning does not require India to reject automation. It requires the State to ensure that automation remains publicly justifiable to those who bear its risks.

Keywords: Rawls; veil of ignorance; algorithmic governance; India; Article 14; Article 21; privacy; administrative law; welfare technology; artificial intelligence.

1. Introduction

Algorithmic governance is no longer confined to private technological innovation. In India, automated and data-intensive systems are increasingly associated with welfare targeting, policing, taxation, municipal services, facial recognition, and judicial administration. The contemporary legal question is therefore not whether algorithms exist in public administration, but when their use becomes an exercise of public power requiring constitutional justification. Recent scholarship describes this transition as a movement from technological assistance toward the algorithmic mediation of state action, with corresponding concerns about opacity, discrimination, accountability, and due process.¹

The constitutional stakes are particularly high where an automated classification determines whether an individual receives food, a subsidy, a public service, or police attention. A welfare system may promise efficiency while producing wrongful exclusion through data mismatch, opaque backend processing, or the treatment of an automated classification as final.² An Aadhaar-enabled public distribution system, moreover, is not simply a fixed computational

¹ A. Surana, A. Pandey, R. K. Dubey, R. Sharma, S. Tiwari, and S. Sharma, “Algorithmic Governance and the Rule of Law in India: Regulating Artificial Intelligence Beyond Automation,” *International Journal of Drug Delivery Technology*, 2026, doi: 10.25258/ijddt.16.10s.43.

² W. R. Khan, “Victims Of Algorithmic Harm In India’S Welfare System: Articles 14 & 21 Remedies,” *LawFoyer International Journal of Doctrinal Legal Research.*, 2026, doi: 10.70183/lijdlr.2025.v03.216.

sequence: its practical operation is enacted through people, databases, machines, documents, and shifting institutional settings.³ The legal subject encounters not an abstract model but a socio-technical arrangement that can fail at several points.

This article asks how Rawlsian political philosophy can guide the jurisprudence of such arrangements in India. The argument is not that John Rawls supplied a technical blueprint for artificial intelligence. Rather, his method supplies a public test for institutional choice. Behind the veil of ignorance, no constitutional designer knows whether they will possess a stable digital identity, reliable connectivity, legible documents, biometric compatibility, technical literacy, or social privilege. A just design must therefore be acceptable from the position of the least advantaged and must protect the basic liberties and opportunities of every person subject to it.

The article makes three claims. First, algorithmic governance should be evaluated as institutional rule-making, not merely as software procurement. Second, Indian constitutional doctrines of equality, dignity, privacy, and non-arbitrariness can operationalise Rawlsian concerns without importing Rawls as a substitute for Indian constitutional interpretation. Third, a rights-centred framework must regulate the entire life cycle of public algorithms: authorisation, design, data collection, procurement, deployment, explanation, review, audit, and remedy.

2. Method and Conceptual Scope

This is a doctrinal and normative study. It synthesises scholarship on Rawlsian approaches to governmental machine learning, Indian algorithmic governance, welfare automation, privacy, equality, institutional oversight, and administrative accountability. The method is interpretive rather than empirical: it does not claim to measure the performance of a particular Indian

³ B. Chaudhuri, “Programmed welfare: An ethnographic account of algorithmic practices in the public distribution system in India,” *New Media & Society*, 2022, doi: 10.1177/14614448221079034.

algorithm or establish the incidence of exclusion across schemes. Its purpose is to develop a constitutional framework that can be applied to high-impact public-sector systems.

The analysis treats “algorithmic governance” broadly. It includes automated decision systems, predictive models, biometric authentication, risk scoring, eligibility engines, data-linkage systems, and human decisions materially shaped by algorithmic recommendations. The relevant object is not only artificial intelligence in the narrow sense. A simple rule-based database can exercise public power if its output determines a person’s legal or material position. Conversely, not every use of automation requires the same scrutiny. The intensity of review should increase with the system’s effect on basic liberties, essential services, livelihood, bodily integrity, or exposure to coercive state action.

3. Rawlsian Justice Beyond the Metaphor

3.1 The original position as an institutional test

The original position asks what principles free and equal persons would choose under fair conditions of choice. The veil of ignorance prevents the parties from knowing facts about themselves that could produce a self-serving constitutional bargain. Applied to algorithmic governance, the veil removes the assumption that the designer will be a digitally privileged citizen. The designer may instead be a person whose fingerprints do not authenticate, whose address is unstable, whose data profile is incomplete, whose language is poorly represented in a model, or whose social history is encoded through discriminatory proxies.

This application has a practical consequence. An administrative system cannot be treated as just merely because it is formally uniform or statistically efficient. A rule that rejects every failed

authentication equally may still impose unequal burdens if the failures are concentrated among people who lack alternative documents, transport, money, or institutional access. Rawlsian reasoning therefore directs attention from the average user to the position of the person who is worst placed by the system. Scholarship on Rawls and machine learning similarly identifies the original position and the difference principle as resources for thinking about distributive justice through algorithmic systems, while insisting that socially responsible data practices must exceed purely technical considerations.⁴

The veil of ignorance also disciplines institutional optimism. A policymaker who does not know whether they will be classified as low risk or high risk would demand more than a general assurance that the model is accurate. They would ask who defines the objective, which errors count, what data are used, how a decision can be challenged, and whether a human institution remains responsible. The question is not “Does the model work?” in the abstract. It is “Would a person who might bear its most serious error accept the legal conditions under which it operates?”

3.2 Equal basic liberties and public reason

Rawls’s first principle gives priority to an equal scheme of basic liberties. In algorithmic governance, this principle supports limits on systems that affect privacy, bodily integrity, expression, association, movement, or access to justice. A predictive policing model, facial-recognition deployment, or extensive data-linkage system cannot be justified solely by administrative convenience. Its legal authority, purpose, necessity, safeguards, and review must be publicly defensible.

⁴ H. Ashrafian, “Engineering a social contract: Rawlsian distributive justice through algorithmic game theory and artificial intelligence,” *AI and Ethics*, 2022, doi: 10.1007/s43681-022-00253-6.

Public reason adds a requirement of intelligibility. Citizens need not receive every line of source code, but they must receive reasons that permit them to understand the legal and factual basis of an adverse decision and to challenge it. A black box is not only an engineering problem. It can become a constitutional defect when the affected person cannot know what was decided, why it was decided, what evidence was used, or how correction is possible. Scholarship on automated public decision-making identifies the “legal black box” and “technical black box” as threats to due process, equal protection, and transparency.⁵

Public reason also rejects the idea that technical expertise alone can settle questions of justice. Whether a welfare criterion is acceptable, whether a surveillance purpose is proportionate, or whether a proxy for risk is discriminatory are normative and constitutional questions. Engineers, administrators, judges, affected communities, and elected institutions may contribute different forms of knowledge. The ultimate standard must remain a public legal standard rather than a proprietary model score.

3.3 Fair equality of opportunity and the difference principle

Fair equality of opportunity requires more than formally equal access. If an automated employment, education, credit, or welfare system relies on historical data shaped by caste, gender, disability, regional, linguistic, or economic disadvantage, it may reproduce unequal opportunity while appearing neutral. Indian scholarship on algorithmic bias emphasises that machine-learning systems trained on socially produced data can reproduce histories of caste discrimination, gender inequality, and economic exclusion.⁶ The constitutional concern is therefore not confined to intentional discrimination. It includes proxy variables, skewed data, opacity, and institutional practices that make unequal treatment difficult to detect.

⁵ H. Liu, C.-F. Lin, and Y.-J. Chen, “Beyond State v Loomis: artificial intelligence, government algorithmization and accountability,” *International Journal of Law and Information Technology*, 2019, doi: 10.1093/ijlit/eaz001.

⁶ S. K. A. Rauf, S. M. Yaseen, S. Aqeeb, S. Rifath, and Prof. G. Jain, “Algorithmic Bias and the Right to Equality,” *Zenodo*, 2026, doi: 10.5281/zenodo.20642190.

The difference principle provides a second test: inequalities are acceptable only when they improve the position of the least advantaged. Applied carefully, this does not mean that every algorithm must produce identical outcomes. It means that an efficiency gain cannot justify a distributional design that shifts error, delay, or proof burdens onto those with the fewest resources. A digital welfare system may be permissible where it improves access for marginalised communities, but its legitimacy depends on whether it includes safeguards against exclusion and a real alternative when the automated pathway fails.

This is the central distinction between automation as assistance and automation as final authority. An algorithm that helps an official locate applications for support may expand administrative capacity. An algorithm that conclusively denies a benefit without reasons or review transforms a probabilistic output into a legal status. The former may be compatible with Rawlsian justice; the latter requires heightened scrutiny because it makes the least advantaged bear the cost of uncertainty.

4. Indian Constitutional Translation

4.1 Article 14 and the problem of computational arbitrariness

Article 14 supplies a direct constitutional language for evaluating algorithmic classification. A system may be facially neutral yet arbitrary in operation if its categories lack a rational connection to a legitimate public purpose, encode unexplained proxies, or make materially different decisions without reasons. Recent Indian scholarship argues that algorithmic opacity and skewed training data can make computational outcomes vulnerable to scrutiny under Article 14, including the doctrine of manifest arbitrariness.

The Rawlsian contribution is to sharpen what equality review should ask. Courts and regulators should inquire whether the system's design was chosen under conditions that took the vulnerable person seriously, whether the relevant classification is publicly defensible, whether error rates and burdens are distributed fairly, and whether the State has tested foreseeable effects on disadvantaged groups. Equality review should not stop at the model's aggregate accuracy. An average accuracy rate can conceal concentrated errors against a minority, a linguistic group, persons with disabilities, or people whose documentation is unstable.

A legally relevant audit should therefore examine at least four dimensions: the quality and provenance of training or reference data; the choice of variables and proxies; the distribution of false positives and false negatives; and the consequences of error. The fourth dimension matters because equal treatment in computation does not guarantee equal treatment in life. A false negative in a low-stakes recommendation is not equivalent to a false negative that cuts off food, shelter, healthcare, or livelihood.

4.2 Article 21, privacy, dignity, and proportionality

The privacy jurisprudence associated with Justice K.S. Puttaswamy (Retd.) v. Union of India provides an important constitutional structure for data-intensive governance. The scholarship surveyed here describes privacy as a fundamental right under Article 21 and identifies legality, legitimate purpose, necessity, and proportionality as the relevant sequence for reviewing state measures.⁷ Analysis of the post-Puttaswamy surveillance landscape likewise treats

⁷ B. S and N. Sree, "REGULATING ARTIFICIAL INTELLIGENCE IN INDIA: CONSTITUTIONAL CHALLENGES TO PRIVACY, EQUALITY, AND DUE PROCESS," INDIAN JOURNAL OF LEGAL REVIEW, 2026, doi: 10.65393/ijlrvi6i718.

proportionality and judicial oversight as tools for counterbalancing executive power over personal information.⁸

For algorithmic governance, proportionality must be applied to the whole system rather than only to the initial collection of data. The reviewing institution should ask: Is there a valid legal basis? Is the objective legitimate? Is the data practice rationally connected to that objective? Is the scale and intrusiveness necessary, or is a less intrusive method available? Do safeguards, retention limits, access controls, audit trails, and remedies prevent excessive harm? Does the benefit justify the burden imposed on privacy and dignity?

The Aadhaar experience illustrates why dignity and recognition matter alongside abstract data security. Scholarship on biometric identification argues that welfare-dependent populations may bear a disproportionate burden in proving identity and may experience authentication failure as a threat to dignified living.⁹ A Rawlsian system must therefore treat privacy not as an elite preference opposed to welfare, but as a condition of equal citizenship for those who depend most on public systems.

4.3 Procedural fairness and the right to contest

A person affected by an automated decision needs more than a notice that a computer generated the result. Procedural fairness requires a meaningful opportunity to know the case against them, correct inaccurate information, present relevant circumstances, and obtain review by a competent human authority. The explanation need not disclose trade secrets indiscriminately, but it must be specific enough to make contest possible.

⁸ V. Bhandari and K. Lahiri, “The Surveillance State: Privacy and Criminal Investigation in India: Possible Futures in a Post-Puttaswamy World,” SSRN Electronic Journal, 2020.

⁹ P. Singh, “Aadhaar and data privacy: biometric identification and anxieties of recognition in India,” Information, Communication & Society, 2019, doi: 10.1080/1369118X.2019.1668459.

This requirement is particularly urgent in welfare administration. Research on Indian automated welfare systems identifies wrongful exclusions arising from data mismatches, opaque backend processing, and automated classifications treated as final. An appeal that exists only in formal terms is not meaningful if it requires repeated travel, expensive documentation, digital literacy, or access to the same institution that produced the error. Contestability must be designed around the claimant's practical position.

The rule should be simple: where an automated or substantially automated output adversely affects an individual's access to an essential entitlement or imposes a coercive burden, the State must provide notice, reasons, a correction channel, human reconsideration, and a time-bound remedy. The burden should not be placed entirely on the citizen to prove that an opaque system is wrong. Public bodies should preserve logs, disclose the relevant data categories, and explain the decisive factors.

5. From Principles to a Jurisprudence of Algorithmic Governance

5.1 Legality and delegated public power

The first question is whether the State has lawful authority to deploy the system and to allow its output to affect rights or entitlements. Procurement contracts, executive instructions, or internal software policies should not silently create new grounds for exclusion, surveillance, or coercion. The legal rule should identify the decision-maker, the purpose, the classes of affected persons, the permissible data, the safeguards, and the remedy.

This is a constitutional response to what recent scholarship calls the risk that technical efficiency will produce “mechanized legality,” where a formally functioning system displaces democratic accountability and human dignity.¹⁰ Delegation to a vendor does not delegate constitutional responsibility. Nor does the involvement of a human official cure an unlawful system if the official merely rubber-stamps an unexplained score.

5.2 Transparency without technological reductionism

Transparency should be layered. At the institutional level, the public should know which agencies use algorithmic systems, for what purposes, with which categories of data, and with what consequences. At the individual level, an affected person should receive a comprehensible statement of the reasons for their particular outcome. At the technical level, authorised auditors and courts should be able to inspect documentation, validation, performance, logs, and relevant model behaviour.

A public algorithm register would make these obligations administratively visible. Each entry should identify the responsible authority, purpose, legal basis, data sources, affected population, risk classification, procurement or development history, audit status, retention policy, human-review process, and complaint route. Publicity would allow civil society and affected communities to identify systems before harm becomes irreversible.

5.3 Mandatory algorithmic impact assessments

India should require an impact assessment before deployment of a high-risk public algorithm and after material changes to its data, objective, or operating environment. The assessment should address equality, privacy, dignity, accessibility, federal and linguistic effects, error distribution,

¹⁰ U. S. & V. Chaudhary, “Algorithmic Governance And The Crisis Of Legal Morality: Revisiting The Hart–Fuller Debate In Automated Decision-Making,” Zenodo, 2026, doi: 10.5281/zenodo.20261091.

cybersecurity, environmental and financial costs, and available non-automated alternatives. It should include consultation with affected communities rather than treating technical documentation as public participation.

The case for such assessments is supported by Indian scholarship that finds no binding pre-deployment review requirement in force and identifies fragmented implementation across ministries.¹¹ The proposed assessment should not be a paper exercise. It should be linked to a power to delay, modify, suspend, or prohibit deployment where the system cannot satisfy constitutional requirements.

A Rawlsian assessment would include a “veil-of-ignorance test”: would the design be acceptable if the evaluator did not know whether they would be a direct beneficiary, a false positive, a false negative, a person with poor connectivity, or a member of a historically excluded group? This is not a substitute for statistical testing. It is a normative screen that ensures that statistical performance is interpreted through the standpoint of equal citizenship.

5.4 Independent oversight and institutional capacity

Oversight must be independent, technically capable, and accessible to ordinary claimants. Comparative work on India’s emerging oversight architecture identifies statutory independence, technical capacity, redress accessibility, and cross-sectoral coordination as core variables, while

¹¹ Lt. Dr. K. Sukumar, “Auditing the Automated State: Impact Assessments and Institutional Capacity for Ethical AI in Indian Governance,” *International journal of innovative research and creative technology*, 2023, doi: 10.62970/ijirct.v9.i4.2608014.

describing current arrangements as fragmented across ministries, sectoral regulators, and voluntary codes.¹²

India need not copy a foreign institutional model. It should, however, establish clear responsibility. A central coordinating authority could maintain the public register, issue minimum standards, accredit auditors, publish guidance, and receive systemic complaints. Sectoral regulators and state-level bodies should retain authority over context-specific risks. Courts should remain available for constitutional remedies. Civil society organisations, legal services institutions, and local administrative officers should be integrated into redress pathways because many affected citizens will not be able to litigate technically complex claims.

5.5 Human oversight and the analogue guarantee

Human oversight must be substantive rather than ceremonial. The reviewer should have authority to depart from the model, access the relevant record, understand the basis of the output, and provide independent reasons. A human signature added after an automated denial is not review if the institution has no time, information, or incentive to reconsider.

For essential benefits, India should retain an analogue guarantee: no person should lose food, emergency healthcare, shelter, or another basic entitlement solely because an automated identity, eligibility, or risk process failed. The alternative route may require verification, but it must not make digital compliance the hidden condition of citizenship. This proposal follows the Rawlsian priority of basic liberties and the difference principle's insistence that institutional innovation must not worsen the position of those with the least resources.

¹² Lt. Dr. K. Sukumar, "Governing the Governors: Comparative Institutional Approaches to Ethical Algorithmic Oversight in India," *International Journal For Multidisciplinary Research*, 2024, doi: 10.36948/ijfmr.2024.v06i01.85929.

6. Applications

6.1 Welfare distribution

Welfare automation presents the clearest test of the proposed framework. Digital systems can reduce delay, detect duplication, and improve targeting. They can also convert an uncertain data signal into a denial that is difficult to correct. The practical operation of Aadhaar-linked welfare demonstrates that the system's effects arise from an assemblage of databases, documents, machines, officials, and local conditions rather than from code alone.

A just welfare architecture should therefore separate authentication from entitlement wherever possible, provide multiple modes of verification, record and publish exclusion data, and ensure immediate provisional support while disputes are resolved. The system should distinguish fraud prevention from proof of deservingness. A failed biometric match is evidence of a failed authentication attempt, not conclusive evidence that a person is ineligible.

6.2 Predictive policing and surveillance

Predictive policing and facial recognition raise more severe concerns because false positives may expose people to coercion, stigma, or surveillance. The State's use of data to prevent crime must satisfy legality, legitimate purpose, necessity, proportionality, and independent oversight. The same infrastructure that can improve service delivery can also enable predictive policing and

biometric tracking; legitimacy depends on the accountability architecture surrounding the technology.¹³

A Rawlsian approach is sceptical of systems that classify citizens without their knowledge and provide no effective route to challenge the classification. The burden should be especially high where a model relies on historical policing data that may reflect prior patterns of over-policing. A prediction about risk cannot itself become proof of wrongdoing.

6.3 Administrative discretion and public services

Automation is often defended as a remedy for arbitrary human discretion. That defence is incomplete. It may replace visible discretion with hidden discretion in the selection of objectives, data, thresholds, labels, and error tolerances. Public law protects dignity through rationality, transparency, and responsiveness, and these principles remain relevant when officials use deterministic or predictive systems.¹⁴

The proper response is not to romanticise human decision-making. Human officials can be biased, inconsistent, and corrupt. The response is to subject both human and automated power to reasons, records, review, and equality constraints. An algorithm may assist judgment, but the institution must remain answerable for the legal decision.

7. Objections and Limits

¹³ Lt. Dr. K. Sukumar, “Between Surveillance and Service: Reconciling AI-Driven Governance with Democratic Legitimacy in the Indian State,” *Journal of Advances in Developmental Research*, 2024, doi: 10.71097/ijaidr.v15.i1.2110.

¹⁴ W. Bateman, “Automating Discretionary Decision-Making in the Public Sector: Legal Dimensions,” *SSRN Electronic Journal*, 2019, doi: 10.2139/ssrn.3493433.

One objection is that Rawlsian theory is too abstract for complex technical systems. The objection has force if the veil of ignorance is treated as a complete audit methodology. It is not. It is a constitutional orientation that must be joined to empirical bias testing, usability research, security assessment, domain expertise, and affected-community participation.

A second objection is that explainability may reduce accuracy or reveal proprietary information. The answer is not an absolute right to source code in every case. It is a right to reasons proportionate to the decision's impact, supported by confidential inspection where necessary. Trade secrecy cannot become a blanket immunity from constitutional review when a private vendor's system exercises delegated public power.

A third objection is that strict safeguards will slow welfare delivery and innovation. That concern should be taken seriously. But speed cannot be measured only by processing time. A system that issues rapid wrongful exclusions may shift costs to citizens, local officials, courts, and families. Institutional design should measure both administrative efficiency and the time required to repair errors. A fair system may be slower at the point of decision because it is faster and more humane at the point of correction.

The article also has limits. It does not offer a quantitative assessment of any individual Indian deployment, and the available scholarship varies in method and evidentiary depth. Some sources are doctrinal or conceptual, while others are qualitative or based on secondary material. The proposed framework should therefore be tested through case studies, administrative data, interviews with affected communities, and audits of actual systems. It should also be adapted to India's federal structure, linguistic diversity, and variation in state capacity.

8. Conclusion

The jurisprudence of algorithmic governance should begin with a modest but demanding question: what rules would citizens choose if they did not know whether the system would classify them as eligible or suspicious, documented or deficient, safe or risky, digitally visible or administratively invisible? Rawls's veil of ignorance makes the answer difficult to evade. A just system must preserve equal basic liberties, protect fair opportunity, and ensure that institutional inequalities work to the advantage of those who are least well placed.

For India, these ideas can be translated through Articles 14 and 21, privacy and proportionality, procedural fairness, non-arbitrariness, and meaningful remedies. The result is neither technological prohibition nor uncritical automation. It is a jurisprudence of accountable assistance. Algorithms may support public administration, but they cannot displace constitutional responsibility. The State must disclose when consequential systems are used, test their distributional effects, provide reasons, preserve human review, retain analogue routes for essential entitlements, and establish independent oversight with real corrective power.

Beyond the veil of ignorance lies not a demand for perfect prediction, but a demand for institutions that no citizen would fear simply because they might occupy the wrong side of a data classification. In a constitutional democracy, the legitimacy of algorithmic governance depends on whether those subject to its power can understand it, contest it, and live with its consequences as equal members of the political community.

ISSN: 2583-2794