Published by

British Journal of Contemporary Research

BEXFORD PUBLISHING

ARTIFICIAL INTELLIGENCE AND HUMAN RIGHTS: LEGAL CHALLENGES IN PUBLIC DECISION-MAKING

Original Research Article | Volume 1 | Issue 3 | 2026 | Article Number: 060

Accepted: 05 August 2026 | Published: 10 August 2026 | ISSN: 2979-8582 (Online)

Crossref : **https://doi.org/10.67693/BJCR-S2SEBJER**

Dr Pradnya Yadav

Assistant Professor, DES Shri Navalmal Firodia Law College, Pune Affiliated to Savitribai Phule Pune University, Pune, Maharashtra, India

Correspondence: Dr Pradnya Yadav, pradnya.magre@gmail.com

Abstract

Artificial intelligence is increasingly being incorporated into governmental decision-making in areas such as policing, criminal justice, welfare administration, taxation, healthcare, immigration, public employment, education and the delivery of essential services. AI-assisted governance may improve efficiency, consistency and the capacity of public institutions to process large quantities of information. Nevertheless, when algorithmic systems influence decisions affecting liberty, livelihood, privacy, equality or access to public benefits, their use raises serious constitutional and human-rights concerns. Biased datasets may reproduce historical discrimination; opaque models may prevent affected persons from understanding adverse decisions; automated systems may weaken procedural fairness; and fragmented responsibility among government agencies, technology vendors and individual officials may create an accountability deficit. This paper critically examines the human-rights implications of AI-based public decision-making, with particular reference to India and comparative international standards. It argues that ordinary data-protection and administrative-law principles, although important, are insufficient by themselves to address the distinctive risks of algorithmic governance. India requires a rights-based legal framework governing high-risk public-sector AI. Such a framework should require lawful authority, necessity, proportionality, algorithmic impact assessments, equality audits, meaningful human oversight, public registers, reasoned decisions, independent supervision and accessible remedies. AI should support public authorities but should not become an unreviewable substitute for constitutionally accountable human judgment.

Keywords: Artificial Intelligence, Human Rights, Automated Decision-Making, Algorithmic Bias, Public Administration, Transparency, Accountability, Privacy, Administrative Law, India

1. INTRODUCTION

Artificial intelligence has moved from experimental technological application to an increasingly important instrument of public administration. Governments and public bodies use, or are considering using, algorithmic systems to identify welfare fraud, assess eligibility for benefits, predict crime, conduct facial recognition, prioritise tax investigations, evaluate visa applications, allocate public housing, support medical diagnosis, shortlist employment candidates and determine the level of governmental scrutiny to which individuals should be subjected.

The appeal of AI in public governance is understandable. Public agencies frequently operate under financial, administrative and human-resource constraints. AI systems can process large datasets more quickly than individual officers, identify patterns that may otherwise remain invisible and potentially reduce inconsistency in routine decision-making. The OECD recognises that governmental use of AI may improve public-sector productivity and responsiveness, but it also identifies surveillance, data-protection concerns and skewed outcomes as major risks.

The legal problem is not that technology participates in administration. Public authorities have long used statistical models, databases and computerised decision-support systems. The fundamental concern is that contemporary AI systems may influence decisions through complex correlations, probabilistic classifications and predictive inferences that are difficult for officials, affected individuals and courts to understand. When these systems operate in areas involving fundamental rights, administrative convenience cannot displace constitutional accountability.

An adverse algorithmic decision is not merely a technical output. It may mean denial of food assistance, increased police surveillance, exclusion from public employment, refusal of medical treatment, cancellation of a pension or prolonged investigation. The consequences are experienced by human beings, often by those who are already socially, economically or politically vulnerable.

This paper argues that AI used in public decision-making must be regulated as an exercise of public power. The state cannot avoid constitutional responsibility by delegating technical functions to a private software provider or by describing an algorithm as merely "advisory." Whenever an automated recommendation materially influences an administrative decision, the entire decision-making chain must comply with constitutional rights, human-rights obligations and principles of administrative fairness.

2. RESEARCH PROBLEM

The central research problem is the growing mismatch between the expanding use of AI in public administration and the limited legal safeguards available to individuals affected by algorithmic decisions.

Most existing legal systems regulate particular components of algorithmic governance through constitutional rights, administrative law, anti-discrimination law, data-protection legislation, procurement rules and sector-specific regulations. These protections are fragmented. They may not clearly establish:

1. when public authorities may lawfully use AI;
2. which uses should be prohibited;
3. when an individual must be informed that AI was used;
4. what explanation must accompany an adverse decision;
5. whether source code or model logic must be disclosed;
6. who bears responsibility for discriminatory or inaccurate outcomes;
7. what standard courts should apply when reviewing algorithmic decisions; and
8. what remedies should be available to affected persons.

The absence of clear rules is particularly concerning in jurisdictions where administrative decision-making affects large populations but institutional capacity for technological auditing and rights-based oversight remains limited.

3. RESEARCH OBJECTIVES

This paper seeks:

1. to examine the use of AI in public decision-making and its implications for human rights;
2. to identify the constitutional and administrative-law principles applicable to algorithmic governance;
3. to analyse the risks of discrimination, opacity, privacy violations and diluted accountability;
4. to evaluate the emerging Indian and international regulatory frameworks;
5. to assess whether existing Indian law adequately protects individuals affected by AI-assisted governmental decisions; and
6. to propose a rights-based legal framework for responsible public-sector AI.

4. RESEARCH QUESTIONS

The paper addresses the following questions:

1. How does AI-assisted public decision-making affect fundamental and human rights?
2. Can existing constitutional, administrative and data-protection laws adequately regulate algorithmic governmental power?
3. Who should bear responsibility when an AI-assisted public decision causes unlawful discrimination or other harm?
4. What procedural protections should be guaranteed to individuals affected by automated or AI-assisted decisions?
5. What lessons can India draw from international AI-governance frameworks?

5. HYPOTHESIS

The paper proceeds on the hypothesis that the use of AI in public decision-making, without a legally enforceable framework of transparency, equality, human oversight and accountability, creates a substantial risk of violating fundamental and human rights. Existing Indian constitutional and data-protection principles provide an important foundation, but they do not constitute a complete regulatory framework for high-risk public-sector AI.

6. RESEARCH METHODOLOGY

The study adopts a doctrinal, analytical and comparative research methodology.

The doctrinal component examines constitutional principles, legislation, judicial decisions and administrative-law doctrines relevant to privacy, equality, natural justice, proportionality and accountability. The analytical component evaluates how the technical characteristics of AI interact with legal standards. The comparative component considers the European Union Artificial Intelligence Act, the Council of Europe Framework Convention, UNESCO's Recommendation on the Ethics of Artificial Intelligence, OECD principles and selected foreign judicial developments.

The study primarily relies on secondary and publicly available legal sources. It does not undertake empirical testing of a specific algorithm. Its conclusions therefore concern the legal design of public-sector AI governance rather than the technical performance of an individual system.

7. UNDERSTANDING AI-BASED PUBLIC DECISION-MAKING

For the purposes of this paper, AI-based public decision-making refers to the use of computational systems by governmental or public authorities to make, recommend, rank, predict, classify or materially influence decisions affecting individuals or groups.

Such systems may perform different functions:

- **Automated decision-making:** The system produces a decision with little or no meaningful human intervention.
- **Decision support:** The system generates a score, recommendation or risk classification that an official considers.
- **Predictive administration:** The system predicts conduct, risk, eligibility, fraud or future need.
- **Biometric identification:** Facial, voice, gait or other biometric information is used to identify or verify individuals.
- **Resource allocation:** Algorithms prioritise inspections, medical services, welfare assistance or other public resources.
- **Profiling:** Individuals are categorised based on personal characteristics, behaviour or inferred attributes.

The distinction between fully automated and assisted decision-making is legally important but can be misleading. A nominally human decision may still be effectively automated when the official lacks time, expertise, authority or information to depart from the algorithmic recommendation. This phenomenon is sometimes described as automation bias: the tendency to place excessive confidence in computer-generated outputs.

Meaningful human oversight therefore requires more than a human officer clicking an approval button. The official must understand the system's purpose and limitations, have access to relevant information, possess authority to reject its recommendation and provide independent reasons for the final decision.

8. POTENTIAL BENEFITS OF AI IN PUBLIC ADMINISTRATION

A balanced legal analysis must acknowledge the legitimate benefits of AI. Properly designed systems may:

1. improve consistency in repetitive administrative processes;
2. reduce processing delays;
3. detect errors, duplication and fraud;
4. assist in distributing limited public resources;
5. improve accessibility through translation and assistive technologies;
6. identify previously neglected patterns of inequality;
7. help officials analyse large and complex datasets; and
8. enable earlier governmental intervention in healthcare, disaster management and social protection.

AI may also reduce certain forms of individual bias if it is trained on reliable data, tested for discriminatory outcomes and embedded in an accountable decision-making process. Technology is therefore not inherently hostile to human rights. The legal challenge is to ensure that efficiency remains subordinate to legality, dignity, fairness and democratic control.

9. HUMAN-RIGHTS CHALLENGES

9.1 Equality and Non-Discrimination

Algorithmic discrimination is among the most serious risks associated with AI-based public decisions. Machine-learning systems identify patterns from historical data. When historical data reflect social

inequality, discriminatory policing, unequal access to education or gendered employment practices, the model may reproduce these inequalities.

Discrimination may arise through:

- unrepresentative training data;
- inaccurate or incomplete government records;
- proxy variables correlated with protected characteristics;
- biased labelling of past cases;
- unequal error rates among demographic groups;
- deployment in communities already subjected to disproportionate surveillance; or
- system design that ignores disability, language or socioeconomic disadvantage.

An algorithm need not directly use caste, race, religion, sex or disability to produce discriminatory outcomes. Postal codes, family structures, education history, consumption behaviour or prior contact with public authorities may function as indirect proxies.

In India, Article 14 of the Constitution guarantees equality before the law and equal protection of the laws. Article 15 prohibits discrimination on specified grounds, while Article 16 protects equality of opportunity in public employment. An algorithmic classification adopted by the state must therefore have a lawful objective, rest on a rational basis and avoid manifest arbitrariness.

The difficulty is evidentiary. An affected individual may know that a benefit or opportunity was denied but may not know which variables were used, how the model was trained or whether similarly situated persons received different treatment. Without transparency and access to evidence, constitutional equality becomes difficult to enforce.

Equality audits should consequently test not only whether a system expressly uses protected attributes, but also whether its outcomes disproportionately burden protected or vulnerable groups. Public authorities should be required to evaluate false-positive and false-negative rates across relevant populations.

9.2 Privacy and Data Protection

AI systems frequently depend on extensive collection, linkage and analysis of personal data. Public authorities may combine information obtained for different purposes, including welfare records, tax information, biometric identifiers, health data, education records, location information and law-enforcement databases.

This creates risks of:

- excessive data collection;
- function creep;
- indefinite retention;
- inaccurate data matching;
- unauthorised sharing;
- inferential profiling;
- re-identification of supposedly anonymous data; and
- mass surveillance.

In *Justice K.S. Puttaswamy (Retd.) v Union of India*, the Supreme Court of India recognised privacy as a constitutionally protected right. The Court linked privacy with dignity, liberty and individual autonomy under the constitutional framework. Any state interference with privacy must satisfy legality, legitimate purpose and proportionality.

The Digital Personal Data Protection Act, 2023 establishes a statutory framework for processing digital personal data and expressly recognises both the individual's interest in protecting personal data and the need for lawful processing.However, data-protection compliance does not by itself ensure that an AI-generated decision is substantively fair, explainable or non-discriminatory. A system may process data lawfully yet produce an arbitrary or unjust outcome.

AI regulation must therefore supplement rather than merely replicate data-protection law.

9.3 Transparency and Explainability

Administrative law ordinarily requires public power to be exercised according to known legal standards. Individuals affected by adverse decisions must generally be able to understand the case against them and the reasons for the decision.

AI complicates this requirement. Some models are technically complex, proprietary or dynamically updated. Government departments may purchase systems from private vendors that claim trade-secret protection over model architecture, source code or training data.

Opacity can exist at several levels:

1. **Institutional opacity:** The public does not know that an AI system is being used.
2. **Data opacity:** Individuals do not know what data were collected or inferred.
3. **Model opacity:** The logic or operation of the system is difficult to understand.
4. **Decision opacity:** The connection between the system's output and the final decision is unclear.
5. **Responsibility opacity:** It is uncertain whether the agency, official, developer or vendor is accountable.

A meaningful explanation need not always disclose every line of code. It should, however, identify the nature and purpose of the system, the significant factors affecting the decision, the relevant data, the limitations of the model, the role of human judgment and the procedure for challenging the outcome.

The United Nations Special Rapporteur has emphasised that persons affected by data processing and AI should receive timely, comprehensive, accessible and understandable information. Algorithmic transparency in the public sector is connected to the democratic right to know how governmental systems affect individuals and communities.

9.4 Natural Justice and Procedural Fairness

The principles of natural justice require, among other things, absence of bias and a meaningful opportunity to be heard. AI systems create new forms of procedural unfairness.

An individual may be unable to challenge an adverse decision because:

- the authority does not disclose that AI was used;
- the system relies on incorrect or outdated data;
- the reasons are expressed only as a risk score;
- the official cannot explain the model;
- the vendor refuses disclosure;
- there is no opportunity to submit contextual information; or
- the appeal is reviewed through the same automated system.

Procedural fairness requires more than notification of a final result. Before an adverse high-impact decision is implemented, the affected person should ordinarily have an opportunity to correct data, contest assumptions and provide relevant contextual evidence.

The principle *audi alteram partem* becomes ineffective when the individual does not know the basis of the case to be answered. A black-box decision can therefore violate natural justice even where the underlying prediction is statistically accurate.

9.5 Human Dignity and Autonomy

Human dignity requires that individuals be treated as persons rather than merely as data points or risk profiles. AI classification can reduce complex human circumstances to numerical indicators. A person may be labelled "high risk," "likely fraudulent," "low employability" or "undeserving" without adequate consideration of personal circumstances.

Public decision-making necessarily uses categories, but AI can intensify categorisation by making it continuous, predictive and difficult to contest. Dignity is threatened when individuals are permanently judged by past data, inferred behaviour or group-based probabilities.

The UNESCO Recommendation on the Ethics of Artificial Intelligence places human dignity and human rights at the centre of AI governance. It states that human rights and fundamental freedoms should be respected, protected and promoted throughout the AI lifecycle.UNESCO also stresses proportionality, prevention of harm, fairness, privacy, transparency and human oversight.

A dignity-based framework requires public authorities to preserve space for explanation, rehabilitation, change and individual circumstances.

9.6 Freedom of Expression and Association

AI-enabled surveillance may affect freedom of expression, protest and association. Facial-recognition systems deployed at demonstrations, public meetings or religious gatherings can identify participants and create records of political or associational activity.

Even where no person is arrested, the perception of continuous surveillance may produce a chilling effect. Individuals may avoid lawful protests, controversial speech or association with minority groups because they fear governmental profiling.

Restrictions on such freedoms must satisfy legality, necessity and proportionality. Indiscriminate or continuous biometric surveillance of public spaces should therefore be prohibited or subjected to extremely narrow statutory conditions and independent authorisation.

9.7 Right to an Effective Remedy

Human rights are meaningful only when violations can be challenged. Algorithmic systems create difficulties concerning standing, evidence, causation, jurisdiction and remedies.

An individual may not be able to prove that the AI system caused the harm because the final decision is formally attributed to a public official. Conversely, the official may state that the result was produced by the software. The agency may blame inaccurate data, while the vendor may rely on contractual limitations.

The law should prevent such responsibility shifting. The state must remain answerable for public power exercised through technology. Individuals should have access to:

- notice that AI materially influenced the decision;
- an understandable explanation;
- correction of inaccurate data;
- human reconsideration;
- administrative appeal;
- independent complaint mechanisms;

- judicial review;
- compensation where appropriate; and
- suspension of systems presenting serious or systemic risks.

10. ACCOUNTABILITY FOR ALGORITHMIC DECISIONS

A central challenge is identifying the legally responsible actor. The AI decision chain may include:

- the government department that procures the system;
- the public authority that deploys it;
- officials who rely on its recommendations;
- private developers;
- cloud-service providers;
- data suppliers; and
- consultants responsible for implementation.

Responsibility should be allocated according to control, knowledge and legal duty, but the public authority must retain primary responsibility toward affected individuals. Constitutional obligations cannot be outsourced.

Government contracts should contain mandatory provisions concerning audit access, data quality, security, discrimination testing, documentation, explanation, incident reporting, model changes and cooperation with courts or regulators. Trade secrecy should not be permitted to defeat due process.

Vendors may bear contractual, statutory or tortious responsibility for defective systems, misleading claims or negligent design. However, an affected individual should not be required to identify the exact technical failure before obtaining administrative relief from the government.

11. JUDICIAL REVIEW OF AI-ASSISTED DECISIONS

Judicial review traditionally examines legality, procedural fairness, relevance of considerations, rationality, proportionality and constitutional compliance. These principles remain applicable to AI-assisted decisions, but courts may require new forms of evidence and expertise.

Judicial review should ask:

1. Did the authority possess lawful power to use the system?
2. Was the system used for a legitimate and defined purpose?
3. Was its use necessary and proportionate?
4. Were the data relevant, accurate and lawfully obtained?
5. Was the system tested for discriminatory outcomes?
6. Did the official exercise independent judgment?
7. Were understandable reasons provided?
8. Could the affected person challenge the data and inference?
9. Was there an effective appeal?
10. Has the authority maintained sufficient records for review?

Courts should be cautious about accepting claims that technical complexity prevents disclosure. A public authority that cannot explain the basis of its decision may be unable to demonstrate that it acted lawfully.

At the same time, judicial review should not require every judge to become a software engineer. Independent technical experts, specialised regulatory bodies and standardised documentation can assist courts in assessing algorithmic systems.

12. INTERNATIONAL REGULATORY DEVELOPMENTS

12.1 European Union Artificial Intelligence Act

The European Union's Artificial Intelligence Act, Regulation (EU) 2024/1689, creates a risk-based regulatory structure governing AI systems.The Regulation recognises that AI can create risks to public interests and fundamental rights.

The framework distinguishes between prohibited practices, high-risk systems and lower-risk applications. High-risk systems include several applications affecting essential services, employment, education, law enforcement, migration and access to public benefits.

Important safeguards include:

- risk-management systems;
- data-governance requirements;
- technical documentation;
- record keeping;
- transparency;
- human oversight;
- accuracy, robustness and cybersecurity;
- post-market monitoring; and
- fundamental-rights impact assessments in specified contexts.

The EU model is significant because it transforms general ethical principles into legally enforceable duties. Nevertheless, its effectiveness will depend on implementation, institutional capacity, enforcement and interpretation of exemptions.

12.2 Council of Europe Framework Convention

The Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law was adopted in May 2024 and opened for signature on 5 September 2024. It is the first international legally binding treaty directed specifically at ensuring consistency between AI lifecycle activities and human rights, democracy and the rule of law.

The Convention adopts a risk-based approach based on the severity and likelihood of adverse impacts. It emphasises dignity, equality, privacy, transparency, accountability, reliability and procedural safeguards.

Its importance lies in locating AI governance within established public-law and human-rights principles rather than treating AI merely as a matter of industrial innovation or voluntary ethics.

12.3 UNESCO Recommendation

UNESCO's Recommendation on the Ethics of Artificial Intelligence was adopted by 193 member states in 2021.It advances a human-rights-centred approach based on dignity, proportionality, safety, privacy, fairness, transparency, accountability and human oversight.

UNESCO has also developed an Ethical Impact Assessment mechanism intended to help governments evaluate AI systems before deployment.The Recommendation opposes social scoring and mass surveillance and emphasises that final accountability must remain with human actors rather than being assigned to AI systems.

Although the Recommendation is not binding in the manner of legislation, it provides an influential normative framework for domestic regulation.

12.4 OECD AI Principles

The OECD AI Principles promote trustworthy AI that respects human rights and democratic values. Originally adopted in 2019, the principles were updated in 2024 to reflect technological and policy developments.

They emphasise:

- inclusive growth and well-being;
- human rights, fairness and privacy;
- transparency and explainability;
- robustness and safety; and
- accountability.

These principles are useful for policy design, but voluntary principles must be translated into enforceable duties when AI affects fundamental rights.

13. INDIAN CONSTITUTIONAL AND REGULATORY POSITION

13.1 Constitutional Framework

India does not require an entirely new constitutional philosophy to regulate AI. Existing constitutional principles provide a strong foundation.

Article 14 prohibits arbitrary state action and requires equal protection.

Article 19 protects freedoms including speech, movement, association and occupation, subject to constitutionally permissible restrictions.

Article 21 protects life and personal liberty and has been interpreted to include dignity, privacy, autonomy and procedural fairness.

Government use of AI must therefore satisfy legality, non-arbitrariness, reasonableness and proportionality. When AI affects a protected right, the state should demonstrate a lawful basis, legitimate purpose, rational connection, necessity and adequate safeguards.

The constitutional problem is that these principles are generally applied after an individual has suffered harm and approached a court. A dedicated regulatory framework is needed to prevent harm before deployment.

13.2 Digital Personal Data Protection Act, 2023

The Digital Personal Data Protection Act, 2023 regulates the processing of digital personal data and imposes obligations on data fiduciaries. It is relevant to AI because data are central to model development, profiling and decision-making.

However, the Act is not a comprehensive AI law. It does not independently regulate all forms of algorithmic bias, explainability, model validation, human oversight or the substantive fairness of automated governmental decisions.

The law of data protection answers questions about lawful processing, but algorithmic administrative law must additionally answer whether a public decision was fair, rational, proportionate and reviewable.

13.3 NITI Aayog Responsible AI Principles

NITI Aayog's *Principles for Responsible AI* presents a roadmap for responsible AI adoption in India and places "AI for All" at its centre. Its principles include safety, equality, inclusivity, privacy, transparency,

accountability and the protection of positive human values.

A subsequent document addressed operationalisation and recognised the importance of incorporating responsible AI principles into governance and research.NITI Aayog has also examined facial-recognition technology as a use case for responsible AI governance.

These documents are valuable policy foundations, but soft-law principles cannot substitute for enforceable rights, independent oversight and remedies.

13.4 India AI Governance Guidelines

India's AI Governance Guidelines released in 2025 seek to support safe and responsible AI adoption and include recommendations concerning government capacity, training and institutional governance.

The guidelines represent an important evolution from general ethical principles toward governance mechanisms. Nevertheless, the central legal question remains whether affected persons can enforce transparency, obtain explanations and challenge harmful algorithmic decisions.

A public-sector AI regime should therefore combine innovation-oriented guidelines with statutory duties and constitutional remedies.

14. GAPS IN THE INDIAN FRAMEWORK

India's present framework contains important constitutional, statutory and policy elements, but several gaps remain.

14.1 Absence of a Comprehensive Classification of High-Risk AI

Indian law does not yet provide a single binding classification of high-risk public-sector AI. Systems affecting liberty, policing, welfare, healthcare, education, employment and essential services should be subject to stricter obligations than low-risk administrative tools.

14.2 Limited Right to Explanation

Affected persons need a clear legal right to know when AI materially influenced a decision and to receive understandable reasons. General administrative-law duties may not adequately address probabilistic or technically complex systems.

14.3 Lack of Mandatory Algorithmic Impact Assessments

Public authorities should assess rights, equality, privacy and security risks before procuring or deploying high-risk AI. At present, such assessment is not uniformly required across governmental sectors.

14.4 Weak Public Transparency

Citizens often cannot determine which agencies use AI, for what purpose, with which vendors and subject to what safeguards. A national public register of governmental algorithmic systems is necessary.

14.5 Procurement-Related Opacity

Public procurement may create dependence on proprietary systems. Contracts may restrict access to technical information, obstruct audits or allow vendors to modify systems without meaningful public supervision.

14.6 Fragmented Accountability

Responsibility may be divided among ministries, state governments, vendors and officials. India requires a clear rule that the deploying public authority remains responsible for legality and rights compliance.

14.7 Unequal Ability to Challenge Decisions

Persons most affected by welfare, policing and public-service algorithms may lack technical knowledge, legal assistance or financial resources. Remedies must therefore be simple, accessible and available in regional languages.

15. PROPOSED RIGHTS-BASED FRAMEWORK FOR INDIA

15.1 Statutory Authority

A public body should not deploy high-risk AI merely on the basis of administrative convenience. The use must have a clear legal foundation defining its purpose, scope, data sources, responsible authority and safeguards.

15.2 Prohibited Uses

Certain practices should be prohibited or presumptively unlawful, including:

- generalised social scoring by public authorities;
- indiscriminate real-time biometric surveillance;
- systems designed to manipulate vulnerable persons;
- decisions based solely on highly sensitive or constitutionally impermissible classifications;
- predictive policing based primarily on group characteristics or location; and
- fully automated final decisions involving detention, criminal liability or deprivation of essential subsistence.

15.3 Risk Classification

Public-sector AI should be classified according to the seriousness and likelihood of harm. High-risk applications should include those affecting:

- criminal justice and policing;
- immigration and citizenship;
- welfare and essential public benefits;
- healthcare;
- education;
- public employment;
- biometric identification;
- child protection;
- housing; and
- access to justice.

15.4 Algorithmic Impact Assessments

Before deployment, the responsible authority should conduct an Algorithmic and Human Rights Impact Assessment addressing:

1. the legitimate public purpose;
2. the necessity of using AI;
3. less intrusive alternatives;
4. data quality and representativeness;
5. equality and discrimination risks;
6. privacy and surveillance implications;
7. accuracy and error rates;

8. effects on vulnerable groups;
9. human-oversight arrangements;
10. cybersecurity;
11. environmental costs;
12. complaint and remedy mechanisms; and
13. conditions for suspension or withdrawal.

A non-confidential summary should be publicly available.

15.5 Public Algorithm Register

Each public authority should disclose:

1. the name and purpose of the system;
2. the responsible department;
3. the legal basis;
4. the vendor;
5. the categories of data used;
6. the population affected;
7. whether the system makes or recommends decisions;
8. the date of deployment;
9. audit results;
10. known limitations; and
11. contact details for complaints.

Security-sensitive details may be protected, but secrecy should be narrowly justified.

15.6 Meaningful Human Oversight

High-risk AI should not operate without officials who:

- understand the system's function and limitations;
- are trained to identify errors and bias;
- have access to relevant evidence;
- can depart from the recommendation;
- record reasons for acceptance or rejection; and
- remain individually and institutionally accountable.

15.7 Notice and Explanation

An individual affected by a materially AI-influenced decision should receive:

- notice that AI was used;
- the purpose of the system;
- the principal factors affecting the result;
- the source and categories of data;
- the role of the human decision-maker;
- information on correcting inaccurate data; and
- instructions for obtaining reconsideration and appeal.

15.8 Equality Audits

Independent audits should measure differential error rates and adverse impacts across relevant groups, including caste, tribe, religion, sex, disability, age, language, region and socioeconomic status, subject to

lawful and privacy-protective methods.

A system showing persistent discriminatory outcomes should be modified, suspended or withdrawn.

15.9 Procurement Safeguards

Government contracts should require:

- audit access;
- technical documentation;
- data provenance records;
- performance testing;
- notice of model changes;
- incident reporting;
- retention of decision logs;
- cooperation with regulators and courts;
- security standards; and
- termination rights for rights violations.

15.10 Independent Regulatory Oversight

India should establish or designate an institution with technical, legal and human-rights expertise to supervise high-risk public-sector AI. Its functions should include:

- maintaining the algorithm register;
- reviewing impact assessments;
- issuing standards;
- conducting inspections;
- investigating complaints;
- ordering corrective action;
- imposing penalties; and
- recommending suspension of dangerous systems.

15.11 Effective Remedies

Affected individuals should have access to prompt human review, correction of data, reconsideration, reasoned appeal and judicial review. Public-interest litigation and representative complaints should be available where algorithmic harm is systemic.

Where a public authority cannot produce adequate records explaining a decision, courts should be permitted to draw an adverse inference regarding legality and fairness.

16. DISCUSSION

AI does not eliminate discretion; it redistributes and conceals it. Human choices determine what problem the system addresses, which data it uses, how outcomes are labelled, what level of error is acceptable and how officials respond to its recommendations.

The language of technological neutrality can therefore obscure normative judgments. A risk score reflects institutional choices concerning whose behaviour is considered risky, which harms matter and how uncertainty is allocated.

The burden of algorithmic error is also unequal. A false-positive fraud classification may be a minor statistical error for a government database but a devastating event for a family dependent on welfare. A facial-recognition mismatch may be a tolerable performance metric for a vendor but a serious threat to

liberty for the misidentified person.

Human-rights law requires the legal system to evaluate technology from the perspective of the affected individual, not merely from the aggregate efficiency of the institution.

Regulation should not assume that innovation and rights are opposing interests. Clear safeguards can improve public trust, data quality and administrative legitimacy. Systems that cannot survive transparency, independent testing or meaningful appeal should not be used to exercise coercive public power.

17. FINDINGS

The study arrives at the following findings:

1. AI can improve administrative efficiency but also magnify existing structural inequalities.
2. Algorithmic bias may arise even where protected characteristics are not expressly used.
3. Data-protection law does not fully regulate the substantive fairness of AI-based decisions.
4. Human involvement is not meaningful unless the official can understand, question and reject the system's output.
5. Trade-secret claims should not override natural justice or judicial review.
6. Public authorities must remain responsible for decisions made through privately developed systems.
7. Indian constitutional principles provide a strong basis for regulating AI but need statutory and institutional operationalisation.
8. Soft-law principles are valuable but insufficient for high-risk governmental applications.
9. Impact assessments, equality audits, public registers and accessible remedies should be mandatory.
10. The legality of AI should be assessed throughout its lifecycle rather than only after individual harm occurs.

18. CONCLUSION

Artificial intelligence is transforming the nature of public administration. It can assist governments in addressing complex social problems, improving public services and using resources more effectively. Yet the same technologies can create opaque, discriminatory and unaccountable forms of public power.

The rule of law requires that governmental decisions remain attributable to legally responsible institutions, based on lawful and relevant considerations, accompanied by reasons and open to challenge. These requirements do not disappear when a decision is generated or influenced by an algorithm.

India's Constitution, particularly Articles 14, 19 and 21, offers a principled foundation for rights-respecting AI governance. The recognition of privacy, dignity, equality, proportionality and procedural fairness provides strong standards against arbitrary algorithmic power. The Digital Personal Data Protection Act, NITI Aayog's responsible AI principles and India's emerging AI-governance guidelines represent meaningful progress. Nevertheless, important regulatory gaps remain.

India should adopt a dedicated legal framework for high-risk public-sector AI. Such a framework should prohibit unacceptable uses, classify high-risk systems, require human-rights impact assessments, ensure meaningful human oversight, mandate public transparency, protect the right to explanation and establish independent supervision.

The appropriate principle is not that human beings must always make every decision without technological assistance. Rather, no person should lose liberty, livelihood, dignity or access to essential public services because of a system that neither the government nor the affected individual can adequately understand, question or challenge.

AI may assist the state, but it must not become an invisible and unaccountable sovereign.

References

Constitution of India, arts 14, 15, 16, 19 and 21.

Digital Personal Data Protection Act 2023.

Justice K.S. Puttaswamy (Retd.) v Union of India, Supreme Court of India, 2017.

European Parliament and Council, Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence.

Council of Europe, *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, CETS No 225, 2024.

UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, 2021.

OECD, *Principles on Artificial Intelligence*, adopted 2019 and updated 2024.

NITI Aayog, *Responsible AI: Principles for Responsible AI*, 2021.

NITI Aayog, *Operationalizing Principles for Responsible AI*, 2021.

NITI Aayog, *Responsible AI for All: Adopting the Framework—A Use Case Approach on Facial Recognition Technology*, 2022.

Office of the United Nations High Commissioner for Human Rights, reports concerning privacy, profiling and automated decision-making.

Government of India, *India AI Governance Guidelines*, 2025.