WHEN ALGORITHMS REFUSE: THE RIGHT TO A REASONED GOVERNMENT DECISION IN THE AGE OF ARTIFICIAL INTELLIGENCE

Sideque N, B.S. Abdur Rahman Crescent School of Law

ABSTRACT

Governments increasingly use artificial intelligence to support decisions on welfare, taxation, licensing, immigration and public security. These systems may improve administrative efficiency, but they can also make it difficult for an affected person to understand why a decision was taken. This paper examines whether an AI-assisted government decision can satisfy the legal duty to give reasons. It analyses the black-box problem, the implications of opaque automated decision-making for equality, privacy, fair procedure and effective remedies, and the allocation of responsibility between public authorities and private technology providers. Drawing on international human-rights instruments and comparative case law from the Netherlands, the United Kingdom, the United States and the European Union, it argues that a legally adequate reason must be case-specific, intelligible and capable of challenge. The paper proposes mandatory human review for high-stakes decisions, case-specific explanations, independent audits, discrimination testing, proper records, effective appeals and clear institutional accountability.[1]

Keywords: artificial intelligence; automated decision-making; right to reasons; algorithmic transparency; human rights; administrative law.

1 See U.N. Secretary-General, Governing AI for Humanity, U.N. Doc. A/79/2 (2024).

1. Introduction

Artificial intelligence (AI) is a general term for computer systems that can perform tasks which normally require human judgement, such as recognising patterns, making predictions, or ranking options. In the last decade, governments have started to use AI to assist, and sometimes to replace, human officials in making decisions that affect people's lives. Tax authorities use algorithms to select taxpayers for audit. Welfare agencies use risk-scoring tools to decide which claimants should be investigated for fraud. Police forces use facial-recognition software to identify persons of interest in public places. Immigration authorities use automated systems to sort visa applications. Courts and parole boards use risk-assessment software to help decide on bail, sentencing, or release.

These tools can make government more efficient. They can process large numbers of cases quickly and can, in principle, apply the same rules consistently to everyone. But they also create a serious problem for a basic legal guarantee: the right to know why a government decision was made against you. When a human official refuses a licence or a benefit, the affected person can normally ask the official to explain the reasoning, and a court can later review that reasoning. When the decision is made, or heavily influenced, by an algorithm, the reasoning may be locked inside a statistical model with thousands of variables that no single person, including the official who signed the decision, can fully explain. This is often called the "black box" problem.

The central research problem of this paper is therefore straightforward to state but difficult to resolve: can an AI-assisted government decision ever be a "reasoned decision" in the legal sense, and if so, under what conditions? This question matters because the right to reasons is not a technical formality. It is closely tied to the right to a fair hearing, the right to challenge a decision, and, ultimately, the right to an effective remedy. Without an adequate reason, a person cannot know whether a decision was lawful, whether it was based on accurate facts, or whether it treated them unfairly compared with others.[2]

This paper is guided by four research questions. First, can an AI-assisted government decision satisfy the legal requirement to give reasons? Second, what human-rights problems arise when governments use AI to make decisions? Third, what level of transparency and human review

2 See Cary Coglianese & David Lehr, Regulating by Robot: Administrative Decision Making in the Machine-Learning Era, 105 Geo. L.J. 1147, 1155–61 (2017).

should be legally required? Fourth, who should be responsible when an AI-assisted government decision violates a person's rights? The paper pursues these questions using doctrinal legal research, meaning that it studies international treaties, regional human-rights instruments, domestic legislation, and court decisions, supported by comparative analysis across several legal systems. The paper does not attempt to describe every national AI law in the world. Instead, it focuses on a representative set of real judgments and legal instruments that illustrate the recurring legal problems, so that the analysis remains grounded and precise rather than speculative.[3]

The scope of the paper is limited to government, or "public sector", decision-making, as distinct from private commercial uses of AI, although private companies that build systems for government use are discussed where relevant to the question of responsibility. The paper connects throughout to the JILPAC 2027 theme, "Rights, Platforms, and Algorithms: Human Rights in the Digital Age", because the right to a reasoned decision sits exactly at the meeting point of individual rights, the platforms and algorithms that governments now rely on, and the broader digital transformation of public administration.

2. AI and Government Decision-Making

"AI-assisted government decision-making" describes any process in which a public authority uses a computer system, ranging from a simple scoring rule to a complex machine-learning model, to help decide, recommend, or directly determine the outcome of a case involving a member of the public. It is useful to separate three levels of automation. At the first level, AI simply organises information for a human decision-maker, for example by flagging a file for review. At the second level, AI produces a recommendation or a risk score that a human official is expected to consider, but is formally free to reject. At the third level, the system produces the final decision with no meaningful human involvement. The legal risk grows as one moves from the first level to the third, but even at the second level, research on automation bias shows that officials tend to follow a machine's recommendation far more often than they reject it, so that a supposedly advisory tool can function, in practice, as a decisive one.[4]

Government AI is used in welfare and public benefits, immigration, taxation, policing,

3 INDIA CONST. art. 21; Maneka Gandhi v. Union of India, (1978) 1 SCC 248, 281–82 (India).
4 See Danielle Keats Citron & Frank Pasquale, The Scored Society: Due Process for Automated Predictions, 89 Wash. L. Rev. 1, 3–10 (2014).

licensing, allocation of public services, and criminal-justice risk assessment. The legal significance varies with the stakes and degree of automation. A useful illustration is a welfare claimant whose payment is suspended because an algorithm assigns a high-risk score: a statement that the system flagged the case does not tell the claimant which facts mattered, whether those facts were accurate, or how the result can be challenged. The problem of reasons is therefore practical rather than abstract.[5]

3. The Right to Know Why a Government Decision Was Made

The idea that a public authority must explain its decisions is one of the oldest principles of administrative law. It flows from the doctrine of natural justice, which requires fair procedure before a decision that affects a person's rights or interests. Natural justice has two traditional pillars: the right to be heard before a decision is made, and the right to an unbiased decision-maker. Modern administrative law has added a third element that is closely related to both: the duty to give reasons. A decision-maker who must explain a decision is less likely to decide arbitrarily, and a person who receives reasons is better able to accept an unfavourable outcome, or to challenge it if the reasons appear unlawful, irrational, or based on a mistake of fact.[6]

The European Court of Human Rights has developed this principle under Article 6(1) of the European Convention on Human Rights, which guarantees the right to a fair hearing. In Van de Hurk v the Netherlands, the Court explained that fair-hearing rights include an obligation on courts to give reasons for their judgments. It clarified this duty further in the twin cases of Ruiz Torija v Spain and Hiro Balani v Spain, both decided on the same day in 1994. The Court held that Article 6(1) obliges courts to give reasons for their judgments, although it does not require a detailed answer to every argument raised; the extent of the duty depends on the nature of the decision and the circumstances of the case. Crucially, the Court found a violation where a domestic court failed to address a submission that was decisive to the outcome. That principle transfers naturally to the question of AI-assisted decisions: if an algorithm relies on a factor that could change the outcome, such as an inaccurate data point, and the affected person cannot find out about that factor, the reasoning has failed in the same way that the Spanish courts failed in Ruiz Torija and Hiro Balani.[7]

5See Virginia Eubanks, Automating Inequality 11–18 (2018).
6S.N. Mukherjee v. Union of India, (1990) 4 SCC 594, 603–04 (India).
7Van de Hurk v. Netherlands, 160 Eur. Ct. H.R. (ser. A) (1994).

The right to reasons is also protected in international human-rights law more broadly. Article 8 of the Universal Declaration of Human Rights guarantees an effective remedy for acts violating fundamental rights, and Article 10 guarantees a fair hearing; both are difficult to exercise meaningfully without knowing the basis of a decision. Article 14 of the International Covenant on Civil and Political Rights (ICCPR) protects the right to a fair hearing, and Article 2(3) requires states to provide an effective remedy to anyone whose rights have been violated. An effective remedy is, by definition, one that a person can actually use, which in turn depends on the person understanding the reasons that led to the decision being challenged. The right to reasons is therefore not a stand-alone technical rule; it is the practical condition for exercising several other rights, including the right to challenge a decision and the right to equal treatment.[8]

4. The "Black Box" Problem

A "black-box" algorithm is a system whose internal decision process cannot easily be understood, either because its design is kept secret for commercial or security reasons, or because its internal structure, such as a deep neural network with millions of parameters, is too complex for a human being to trace step by step, even with full access to the code. Both forms of opacity create the same practical result for the person affected: no clear, human-readable account of why the decision came out the way it did.

Commercial secrecy is a recurring theme in the case law. In the United States case of State v Loomis, a Wisconsin court used a risk-assessment tool called COMPAS, made by a private company, to help decide the defendant's sentence. Because the tool's internal formula was a trade secret, Loomis could not examine exactly how his score had been calculated, and he argued this violated his right to due process. The Wisconsin Supreme Court held that a court may consider a COMPAS score at sentencing without violating due process, provided the score is not the determinative factor in the sentence and the sentencing report includes written warnings about the tool's limitations, including that it is based on group data and had not been validated for the local population. The decision shows a court trying to reconcile the practical reality of proprietary software with the individual's right to a fair process, but it also shows the limits of that compromise: Loomis still could not test the accuracy of the specific score applied

8 Universal Declaration of Human Rights arts. 8, 10, G.A. Res. 217 (III), U.N. Doc. A/RES/217(III) (Dec. 10, 1948).

to him, only the general cautions that accompanied it.[9]

Technical opacity was central to the Dutch SyRI case discussed in detail in Section 12. The government could not tell the court exactly what risk indicators the algorithm used, because the underlying risk model itself had not been made public even to the court. The court held that this lack of transparency, on its own, meant the system did not strike a fair balance between the public interest in detecting fraud and the individual's right to privacy.[10]

The black-box problem is compounded when a government cannot explain its own system's output. A machine-learning model may identify statistical correlations between input data and past outcomes without anyone, including its designers, being able to state in ordinary language why a particular case received a particular score. In such cases, an explanation of "how the system generally works" is not the same as an explanation of "why this decision was made about this person". A description of an algorithm's architecture, its training data, or its accuracy rate on a test set is useful for oversight and audit, but it does not answer the question a refused applicant actually asks: what, about my case, led to this result, and was it correct? Purely technical explanations, without a case-specific account, therefore fail to satisfy the legal purpose of the duty to give reasons, which is to let the individual understand and, if necessary, challenge the decision made about them.[11]

5. Human Rights Problems

AI-assisted government decisions can affect a wide range of human rights, often at the same time.

Equality and non-discrimination

Machine-learning systems learn patterns from historical data. If that data reflects past discrimination, for example because a particular neighbourhood was historically over-policed or over-investigated for fraud, the algorithm can reproduce and even amplify that pattern under a veneer of mathematical neutrality. This was central to the SyRI case, where the risk-detection system was used only in low-income neighbourhoods with a high proportion of residents from immigrant backgrounds, prompting concern from the United Nations Special Rapporteur on

9 International Covenant on Civil and Political Rights art. 14, Dec. 16, 1966, 999 U.N.T.S. 171.
10 See Frank Pasquale, The Black Box Society 3–10 (Harvard Univ. Press 2015).
11 Case C-634/21, SCHUFA Holding AG (Scoring), ECLI:EU:C:2023:957, ¶¶ 42–60 (Dec. 7, 2023).

extreme poverty and human rights that the system risked stigmatising and disproportionately targeting the poor. In the United Kingdom, the Court of Appeal in Bridges (discussed in Section 12) held that South Wales Police had failed to take reasonable steps to satisfy itself that its facial-recognition software did not have an unacceptable bias on grounds of race or sex, in breach of the public sector equality duty.

Privacy and data protection

AI-assisted decisions typically depend on combining large volumes of personal data drawn from multiple government databases. The SyRI case is again instructive: the Hague court found that linking data from separate government agencies for algorithmic risk-profiling interfered with the right to respect for private life under Article 8 of the European Convention on Human Rights, and that this interference was not adequately justified given the lack of transparency about how the data was used.

Fair procedure, dignity, and access to justice

Being reduced to a risk score, without an opportunity to explain one's individual circumstances, can be experienced as a denial of dignity as well as a procedural defect. Where a person cannot understand or challenge an automated decision, their access to justice is also undermined, because a court asked to review the decision faces the same opacity that the individual does. Article 47 of the Charter of Fundamental Rights of the European Union guarantees the right to an effective remedy and to a fair trial, and this right is hollowed out where neither the applicant nor, in some cases, the reviewing body can reconstruct the reasoning behind an automated decision.

Right to an effective remedy

Article 13 of the European Convention on Human Rights and Article 2(3) of the ICCPR require an effective domestic remedy for rights violations. An effective remedy must be capable of examining the substance of a complaint, not merely its form. Where the substance of a decision is a statistical score that no one can fully explain, an oversight body can review the legal framework surrounding the tool, as the Dutch court did in SyRI, but it becomes much harder to review whether the specific decision made about a specific person was correct. This gap between institutional oversight of a system and individual review of a decision is one of the

most difficult human-rights problems raised by government use of AI.

6. Government Responsibility

When an AI-assisted government decision is wrong, several actors could, in principle, bear responsibility: the individual official who acted on the system's output, the government department that deployed the system, and the private technology company that built it. In practice, responsibility is often unclear, and this unclear allocation is itself a human-rights problem, because a person seeking a remedy needs to know whom to hold accountable.

Government officers who rely on an automated recommendation without applying independent judgement risk what has been called "automation bias", treating the machine's output as if it were self-evidently correct. Administrative law generally requires that discretion be exercised by the person or body legally empowered to exercise it; an official who simply rubber-stamps a computer-generated result may not be lawfully exercising discretion at all. Government departments carry institutional responsibility for procuring, testing, and deploying a system, and for ensuring adequate safeguards exist before the system is used on real cases; the SyRI judgment shows a court holding the state responsible for failing to build sufficient transparency and safeguards into the legal framework surrounding a fraud-detection tool, even though the tool was operated through a separate implementing body.

Technology companies that design and sell decision-support systems to governments occupy an intermediate position. They are usually private actors and therefore not directly bound by human-rights obligations in the same way as the state, yet their design choices, including what data the model is trained on, what accuracy the model must achieve before deployment, and what limitations must be disclosed, substantially determine whether a resulting government decision can be explained at all. The Loomis case shows the tension this creates: the private company's insistence on trade-secret protection for its algorithm directly limited what the court, and the defendant, could learn about the reasoning behind a decision that affected his liberty.

Three institutional requirements follow from this analysis. First, human supervision must be genuine and documented, not a formality; an official must be able to show, in the individual case, what independent consideration was given to the automated output. Second, governments need proper records of how a system reached a particular result in a particular case, not only records of the system's general design, so that an individual decision can later be reconstructed

and reviewed. Third, accountability structures should make clear, in advance of deployment, which public body is legally responsible for an AI-assisted decision and how a complaint against that decision will be handled, including where a private contractor built or operates the underlying system.

7. Right to Human Review

A recurring proposal in this field is that a person affected by an AI-assisted government decision should have a right to ask for human review: a right to challenge the decision, a right to receive meaningful reasons, a right to submit evidence, and a right to appeal or seek a remedy. Article 22 of the EU General Data Protection Regulation (GDPR) already provides a version of this right in the data-protection context: a person generally has the right not to be subject to a decision based solely on automated processing that produces legal effects concerning them or similarly significantly affects them, subject to certain exceptions, and where those exceptions apply, the person is entitled to at least obtain human intervention, to express their point of view, and to contest the decision.[12]

The scope of this right was clarified by the Court of Justice of the European Union in its 2023 judgment in the SCHUFA case. A German credit-reference agency generated automated probability scores predicting whether an individual would repay a loan; a lender then relied heavily on that score to refuse credit. The Court held that generating the score itself amounted to automated individual decision-making within the meaning of Article 22, because the third party relying on the score "drew strongly" on it, even though the agency itself did not formally make the final lending decision. This is directly relevant to government AI, because public authorities frequently rely on scores or recommendations generated by another body, whether an internal analytics unit or an external contractor, without that body being the final decision-maker on paper. The Court's approach suggests that the source of the score, not merely the name on the final decision letter, should be treated as part of the automated decision-making chain for the purposes of individual rights. The same judgment, however, also confirmed that the agency was not required to disclose the precise mathematical formula behind the score, illustrating the limits of the transparency obligations even where a right to explanation exists in principle.[13]

12 Rechtbank Den Haag [District Court of The Hague], Feb. 5, 2020, ECLI:NL:RBDHA:2020:865 (SyRI).
13 R (Bridges) v. Chief Constable of South Wales Police, [2020] EWCA Civ 1058, [2020] 1 W.L.R. 5037.

Arguments in favour of a strong right to human review

A strong right to human review protects dignity, because it ensures a person is not treated purely as a data point. It corrects errors that a rigid algorithm cannot anticipate, because human decision-makers can consider context that was never captured in the training data. It also supports the rule of law, because a human reviewer, unlike an opaque model, can be asked to explain the reasoning behind a decision in ordinary language that a court can then assess.

Arguments against an unqualified right to human review

An unqualified right to human review for every automated decision, however minor, risks overwhelming administrations that adopted AI specifically to manage high case volumes, undermining the efficiency gains that justified using the technology in the first place. It may also produce a false sense of security if the "human reviewer" simply re-approves the algorithm's recommendation without genuine independent scrutiny, which brings the analysis back to the automation-bias concern discussed in Section 6. A workable rule should therefore scale the right to human review to the seriousness of the decision: routine, low-stakes, easily reversible decisions may tolerate a lighter-touch review mechanism, while decisions affecting liberty, livelihood, immigration status, or essential welfare should require a documented, substantive human review before the decision takes effect or immediately upon request.

8. International Legal Framework

The Universal Declaration of Human Rights (UDHR), adopted in 1948, is not binding as a treaty but expresses foundational principles that later treaties made binding. Article 8 guarantees an effective remedy for violations of fundamental rights, and Article 10 guarantees a fair and public hearing by an independent tribunal. Applied to AI-assisted decisions, these articles support the argument that a remedy against an automated decision must be capable of actually testing the reasoning behind it, not merely confirming that a process was followed.

The International Covenant on Civil and Political Rights (ICCPR) makes several of these guarantees binding on states parties. Article 14 protects the right to a fair hearing, which the UN Human Rights Committee has interpreted to include, in relevant proceedings, a reasoned judgment. Article 17 protects against arbitrary or unlawful interference with privacy, which is engaged whenever a government links or profiles personal data through an algorithmic system.

Article 26 guarantees equality before the law and protection against discrimination, which is directly engaged by the bias risks discussed in Section 5. Article 2(3) requires states to ensure an effective remedy, developed by competent authorities, for anyone whose rights are violated, and to ensure that any remedy granted is enforced.

Within the European human-rights system, Article 6 of the European Convention on Human Rights (ECHR) protects the right to a fair hearing and has been interpreted, in the case law discussed in Section 3, as including a duty to give reasons. Article 8 protects private and family life and was the basis for striking down the SyRI legislation. Article 13 requires an effective remedy before a national authority. Article 14 prohibits discrimination in the enjoyment of Convention rights, reinforcing Article 26 ICCPR in this context.

European Union law adds further, more technologically specific layers. The Charter of Fundamental Rights of the European Union protects the right to respect for private life (Article 7), the right to protection of personal data (Article 8), the right to good administration, which includes an obligation on the administration to give reasons for its decisions (Article 41), and the right to an effective remedy and a fair trial (Article 47). The GDPR gives these rights operational content: Articles 13 to 15 require controllers to provide meaningful information about the existence and logic of automated decision-making, and Article 22, as clarified in the SCHUFA judgment discussed in Section 7, restricts purely automated decisions with legal or similarly significant effects and guarantees a right to human intervention where such decisions are permitted. Most recently, the EU Artificial Intelligence Act, Regulation (EU) 2024/1689, adopted in 2024, classifies many government uses of AI, including systems used for eligibility for public benefits, for law enforcement risk assessment, and for migration and border control, as "high-risk", triggering obligations of risk management, data governance, technical documentation, human oversight, transparency to affected persons, and accuracy and robustness testing before and after deployment.

These instruments do not merely describe rights in the abstract; they interact directly with AI-assisted government decisions. The duty to give reasons under Article 41 of the EU Charter, for example, cannot be satisfied by a general statement that "the algorithm flagged this case", because that statement does not identify which facts about the individual, correctly or incorrectly recorded, produced the outcome. Similarly, the GDPR's requirement of "meaningful information about the logic involved" has been interpreted by data-protection

authorities and commentators to require more than disclosure of source code, since source code alone is rarely meaningful to a lay person; it requires an intelligible, case-specific account.

9. Case Law

NJCM and Others v The State of the Netherlands (the SyRI case)

In this 2020 judgment, the District Court of The Hague considered a legal challenge brought by a coalition of civil-society organisations and two individual citizens against the Dutch government's System Risk Indication (SyRI) legislation, which allowed public authorities to combine data from separate government databases to build a hidden risk model predicting who was likely to commit welfare, tax, or labour-law fraud. The system was used only in selected neighbourhoods with a high proportion of low-income and immigrant residents. The court held that the SyRI legislation violated Article 8 ECHR because it did not strike a fair balance between the government's legitimate aim of preventing fraud and the individual's right to privacy. Central to the court's reasoning was the lack of transparency: the government could not tell the court what risk indicators the model used or how it weighted them, and the court held it could not assess whether the interference with privacy was proportionate without that information. The court also noted, without deciding the point, that it could not rule out that operating SyRI in "problem neighbourhoods" risked stigmatising and discriminating against residents of those areas on the basis of their socio-economic status. The judgment did not resolve every legal question that AI-assisted government decisions raise, but it establishes an important precedent: a government cannot rely on an algorithmic system to justify an interference with rights if it cannot explain, at least to the reviewing court, how the system actually works.

R (Bridges) v Chief Constable of South Wales Police[14]

In 2020, the England and Wales Court of Appeal considered a challenge brought by a privacy campaigner, Edward Bridges, against South Wales Police's use of live automated facial-recognition technology in public places. The Court of Appeal held, unanimously, that the use of the technology was not "in accordance with the law" as required by Article 8(2) ECHR, because the legal framework left too much discretion to individual police officers as to who could be placed on a watchlist and where the technology could be deployed, without adequate

14 State v. Loomis, 881 N.W.2d 749, 761–68 (Wis. 2016).

published criteria. The Court also held that the police had failed to comply with the Data Protection Impact Assessment requirements under domestic data-protection law, and, significantly for the discussion of equality in Section 5, that the force had not taken reasonable steps to satisfy itself, as required by the public sector equality duty, that the software did not have an unacceptable bias on the grounds of race or sex. The case demonstrates that the duty to explain and justify an automated government tool applies before deployment, at the level of the legal framework and equality assessment, and not only after an individual complains about a specific decision.

State v Loomis

The 2016 decision of the Wisconsin Supreme Court, discussed in Section 4, considered whether a sentencing court's reliance on a proprietary risk-assessment tool, COMPAS, violated the defendant's right to due process, given that the tool's methodology was a trade secret and that it used group-based statistical data rather than an individualised assessment. The court held that such reliance does not violate due process, provided the score is not the determinative factor in sentencing and the court is given, and considers, written warnings about the tool's limitations. The United States Supreme Court later declined to review the case. The decision remains controversial because it permits continued reliance on an unexplainable score, subject only to procedural safeguards around its use, rather than requiring that the score itself be explicable.

OQ v Land Hessen (the SCHUFA case)

The Court of Justice of the European Union's 2023 ruling, discussed in Section 7, interpreted Article 22 GDPR for the first time and held that the automated generation of a credit-probability score by a private agency constitutes automated individual decision-making where a third party draws strongly on that score to make a decision with legal effects. Although the case arose in a private commercial context, its reasoning is directly transferable to government use of AI, because many public authorities rely on risk scores generated by a separate unit, agency, or contractor, rather than generating the score themselves, and the judgment confirms that this separation does not remove the individual's rights under Article 22.

Read together, SyRI, Bridges, Loomis and SCHUFA show a recurring concern with transparency and explicability across welfare fraud detection, policing, criminal sentencing and credit scoring. Their contexts differ, but the underlying legal difficulty is similar: a court and

the affected person must be able to understand enough of the decision-making process to assess whether the decision was lawful and fair.

10. The Main Legal Problem

For high-stakes decisions, a government should not deploy a system whose output cannot be explained in a case-specific way. At minimum, an affected person should receive notice that AI was used, the principal factors influencing the outcome, the factual data relied upon, and information on human review or appeal. Where liberty, immigration status, essential welfare, or another fundamental interest is affected, substantive human review should be mandatory or immediately available.

The public authority that makes or adopts the decision must remain legally responsible for explaining it. Contracts with private vendors should guarantee sufficient access to the system and its records to enable the authority to meet that duty. Commercial confidentiality may protect source code from public disclosure, but it should not prevent disclosure of the case-specific basis of a decision to the affected person or a reviewing court.

An AI-assisted government decision can qualify as a "reasoned decision" only where the reasons are case-specific, intelligible to the affected person, and capable of meaningful challenge. A general description of an algorithm, its architecture, or its average accuracy is not enough.

11. Proposed Solutions

Building on the analysis above, this paper proposes the following legal safeguards for AI-assisted government decision-making.

First, mandatory human review for high-stakes decisions, meaning any decision materially affecting liberty, immigration status, essential welfare, licensing needed for a livelihood, or a criminal-justice outcome, with the review documented and reasoned in its own right, not a mere confirmation of the algorithm's output. Second, clear, case-specific reasons, requiring authorities to disclose the main factors that influenced a decision and the underlying facts relied upon, in language an ordinary person can understand, distinct from, and in addition to, general system documentation. Third, transparency about the existence of automation, requiring that any person subject to an AI-assisted decision be told, at the time of the decision, that AI was

used and what role it played.[15]

Fourth, regular independent AI audits, including pre-deployment and periodic post-deployment testing for accuracy and for disparate impact on protected groups, with published summaries, building on the equality-duty reasoning in Bridges. Fifth, systematic checking for discrimination, requiring authorities to test training data and outputs for patterns correlated with protected characteristics such as race, ethnicity, sex, and socio-economic status, and to justify, or discontinue, any system that cannot pass such testing, informed by the concerns raised in SyRI about geographically targeted profiling. Sixth, proper records, requiring that the specific inputs and outputs relied upon in each individual case be retained and made available on request, so that a decision can be reconstructed and reviewed after the fact, addressing the evidentiary gap that hampered the court in SyRI.

Seventh, effective notice to affected persons, given at a time and in a form that allows a realistic opportunity to respond before serious consequences occur, rather than only after the fact. Eighth, a genuine right to challenge, allowing an affected person to contest both the facts used and the fairness of the outcome, not merely to request a repeat automated assessment. Ninth, a clear appeal pathway to an independent body or court, with the reviewing body given full access to the case-specific reasoning regardless of any commercial confidentiality claim by a vendor. Tenth, effective legal remedies, including the possibility of suspending a decision, correcting inaccurate data, and awarding compensation where an unexplainable or discriminatory system has caused harm. Eleventh, clear government accountability, requiring that, before any AI-assisted system is deployed, the responsible public authority be identified by law, together with the standard of review its decisions will be subject to and the minimum contractual guarantees of explainability and access that it must secure from any private vendor.[16]

RESEARCH GAP AND CONTRIBUTION

Existing scholarship on automated public decision-making frequently addresses transparency, bias, data protection, or administrative efficiency as separate concerns. This paper connects these concerns through the legally enforceable duty to give reasons. Its central contribution is to distinguish between a technical explanation of an algorithm and a legally sufficient reason

15A.K. Kraipak v. Union of India, (1969) 2 SCC 262, 272–73 (India).
16Regulation (EU) 2024/1689, arts. 5, 6, 9, 13, 14, 2024 O.J. (L) 1.

for an individual governmental decision. A system may be explainable in technical terms yet still fail the affected person if it does not identify the facts relied upon, the decisive considerations, the applicable legal standard, and the route for challenge. The paper therefore treats reason-giving as a bridge between algorithmic governance, procedural fairness, equality, accountability, and effective remedies.

ANALYTICAL FRAMEWORK: WHEN IS AN AI-ASSISTED DECISION REASONED?

For the purposes of this paper, an AI-assisted governmental decision should be treated as reasoned only when it satisfies five connected requirements. First, the legal authority and purpose for using the system must be identifiable. Second, the affected person must be told the material facts and data relied upon. Third, the authority must disclose the principal factors that materially influenced the outcome in a form that is intelligible to the affected person. Fourth, the decision must remain open to meaningful human review, correction, and appeal. Fifth, an adequate record must exist so that a court, tribunal, auditor, or other reviewing body can examine whether the system was lawfully used and whether the individual decision was justified. These requirements do not demand disclosure of source code in every case. They demand legally relevant, case-specific reasons sufficient to make the exercise of public power contestable.

INDIAN ADMINISTRATIVE-LAW PERSPECTIVE

The Indian constitutional framework provides a strong basis for applying this approach to automated public decision-making. Article 14 is relevant where automated systems produce arbitrary, irrational, discriminatory, or unexplained classifications. Article 21 is relevant where automated decisions affect liberty, livelihood, reputation, privacy, access to welfare, or other legally protected interests. The principles of natural justice also require attention where a decision has serious civil consequences. In this setting, the right to reasons is not merely an administrative courtesy: it supports fairness, judicial review, institutional accountability, and the affected person's ability to challenge an adverse outcome.[17]

The Indian position should nevertheless be developed with doctrinal precision. Not every administrative action requires the same level of disclosure, and the content of reasons may vary

17Regulation (EU) 2016/679, arts. 12, 13, 14, 15, 22, 2016 O.J. (L 119) 1.

according to the nature of the power, the urgency of the decision, statutory confidentiality, national security, privacy, and the rights of third parties. The appropriate principle is proportionality: the more seriously an automated decision affects rights or interests, the stronger the justification, human oversight, record-keeping, and review safeguards should be.

12. Conclusion

This paper has argued that an AI-assisted government decision can satisfy the duty to give reasons only when the explanation connects the operation of the system to the specific facts of the individual case. The right to reasons is closely connected to fair procedure, equality, privacy, the right to challenge a decision and the right to an effective remedy. SyRI, Bridges, Loomis and SCHUFA demonstrate, in different legal settings, the difficulty of protecting those interests when decision-making becomes opaque.

The appropriate response is not to reject AI in public administration, but to impose legal conditions on its use. High-stakes decisions should receive genuine human review; affected persons should receive case-specific reasons and notice of automation; authorities should conduct independent audits and discrimination testing; relevant inputs and outputs should be recorded; and effective appeal and remedy mechanisms should be available. Above all, the public authority must remain accountable for a decision even when a private system produced or influenced the underlying recommendation. These safeguards preserve the basic rule that a person affected by government action should be able to know why it happened.[18]

18 INDIA CONST. art. 14; E.P. Royappa v. State of Tamil Nadu, (1974) 4 SCC 3, 38–39 (India).