

**THE REQUIREMENT OF AN INTELLIGIBLE DIFFERENTIA AS A CONSTRAINT
UPON MACHINE-LEARNED CLASSIFICATION BY THE STATE: ARTICLE 14
OF THE CONSTITUTION OF INDIA AND THE STATUTORY RESIDUE OF THE
CRIMINAL TRIBES ACT, 1871**

by

Sneha M

KLE Law College, Bengaluru

ABSTRACT

The literature upon algorithmic discrimination rests upon two foundations, neither of which is Indian. The first is the doctrine of disparate impact, an outcome-oriented enquiry directed at what a classifier has done to a protected group. The second is Article 22 of the General Data Protection Regulation, which the Digital Personal Data Protection Act, 2023 conspicuously declines to reproduce. The Indian equality guarantee is differently constructed. Since *State of West Bengal v. Anwar Ali Sarkar*, a classification effected by the State must rest upon an *intelligible* differentia bearing a rational nexus to its object; and the decision in *Shayara Bano v. Union of India* has added that a classification defensible by no determining principle is manifestly arbitrary and void. The adjective has never been required to perform any work. This article contends that intelligibility is a condition upon the legibility of the classification itself rather than a description of its outcome, and that a classifier whose decision boundary cannot be stated therefore fails the first limb of the classification test by reason of opacity alone, independently of proof of disparate effect. Three conclusions follow. The requirement is addressed simultaneously to the court, to the administrative authority and to the person classified, and post hoc explanation satisfies none of them. A defence founded upon demonstrated accuracy would convert Article 14 from a reason-giving norm into a consequentialist one. Finally, the records upon which Indian policing systems are trained carry the sediment of the Criminal Tribes

Act, 1871, so that what repeal removed from the statute book the training corpus has restored silently.

I. Introduction

A police officer who stops a man upon the ground that he belongs to a particular community may be asked to explain himself. The answer, once furnished, is susceptible of examination: the ground relied upon is either a permissible basis of classification or it is not, and Articles 14 and 15 of the Constitution supply the vocabulary in which that question is to be resolved. The enquiry presupposes that an answer exists. It presupposes, more precisely, that the officer possessed a rule, however inarticulate and however unstated, by which this man was distinguished from other men, and that the rule is capable of being reduced to words and thereafter scrutinised.¹

Alter a single element of that scenario. The officer stops the man because a system deployed by the State has placed his name upon a list, and the system generated that list by fitting a function across several hundred features drawn from a decade of police records. Enquire into the reason, and no answer of the kind which Article 14 is constructed to receive will be forthcoming. The difficulty is not that the State is withholding its reason. It is that the rule, in the sense in which the Constitution employs that term, was never formed. What exists in its place is a decision boundary situated within a space of many dimensions, produced by an optimisation procedure, which nobody is in a position to state as a proposition concerning persons: neither the vendor who supplied the system, nor the officer who acted upon its output, nor the officer who authorised its procurement.²

Indian constitutional scholarship has thus far addressed this development in a vocabulary borrowed from elsewhere. The prevailing frame is that of disparate impact: the system is said to be constitutionally suspect because its outputs fall unevenly upon Muslims, upon Dalits,

¹ The proposition that a classification must be capable of being stated before it can be tested is implicit in the structure of judicial review of administrative action generally. See H.W.R. WADE & C.F. FORSYTH, *ADMINISTRATIVE LAW* 293–312 (11th ed. 2014). The point is not that officers in fact articulate their reasons. It is that the legal system is organised upon the assumption that they are able to do so if required, and supplies remedies where they do not. See *Union of India v. M.L. Kapoor*, (1974) 1 S.C.C. 348, 364–65 (India).

² The distribution of ignorance is itself significant. In a conventional administrative hierarchy somebody knows the reason, even if the affected person does not. In the algorithmic case the knowledge is not withheld; it does not exist in any mind.

and upon the urban poor.³ That frame is accurate so far as it extends, and it extends some distance. It is nonetheless an imported frame, and it imports a burden which Indian law has no occasion to assume.⁴ A claim founded upon disparate impact obliges the claimant to establish the distribution of outcomes. Outcomes are precisely what a person subjected to a predictive policing system is unable to observe, for the operation of the system is not disclosed, its outputs are not published, and the population which it declined to flag is by construction invisible.⁵

There exists a second route, and it has lain unused within the oldest stratum of Indian equality doctrine. The classification test, as it has been restated over seven decades, does not merely enquire whether a differentia exists and whether that differentia is rationally connected to the statutory object. It enquires whether the differentia is *intelligible*.⁶ Every student of Indian constitutional law reproduces the adjective in an examination script. Remarkably few have paused to ask what function it discharges. It has not been required to discharge any, for until recently every classification which came before the Supreme Court was already stated, whether in a statute, a notification, a rule or an order, and the only questions genuinely in issue were whether the stated ground was rational and whether the nexus held. Legibility was a background condition so uniformly satisfied that it was never tested.⁷

Machine-learned classification tests it. The contention advanced here is that intelligibility is not ornamental; that it imposes an antecedent condition of stateability upon classification effected by the State; and that a learned classifier which cannot be reduced to a stateable rule fails that condition categorically. It fails not contingently, not pending the maturation of better techniques of explanation, and not only where disparate impact has been separately established. The constitutional infirmity consists in the impossibility of articulating the ground of differentiation, and it is complete at the moment of deployment.

³ See, e.g., Vidushi Marda & Shivangi Narayan, *Data in New Delhi's Predictive Policing System*, in PROCEEDINGS OF THE 2020 CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 317, 320–23 (2020); see also Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671 (2016).

⁴ It is not suggested that disparate impact analysis is unavailable in India. It is suggested that the evidentiary preconditions of such analysis are ordinarily unsatisfiable against an opaque system operated by the State.

⁵ The asymmetry is structural rather than contingent. A person who is not flagged is never informed that she was assessed, and a person who is flagged is ordinarily informed only of the consequence and not of the assessment.

⁶ *State of West Bengal v. Anwar Ali Sarkar*, A.I.R. 1952 S.C. 75, 91 (India) (Das, J.); *Budhan Choudhry v. State of Bihar*, A.I.R. 1955 S.C. 191, 193 (India).

⁷ A comparable dormancy attended the requirement in Article 311(2) of an inquiry, which acquired operative content only when the practice of summary dismissal was tested against it.

The argument is not that the State may never employ statistical instruments, nor that algorithmic governance is inherently objectionable. It is considerably narrower and, for that reason, considerably harder to evade. Where the State classifies persons for a coercive purpose, it must be in a position to state what the classification is. If it cannot state it, it may not classify. Whether the instrument is technically sophisticated is beside the point; whether it is empirically accurate is beside the point; whether its designers acted in good faith is beside the point. The Constitution requires a reason, and a reason is something capable of being stated.⁸

The question acquires a particular urgency in the Indian setting for a reason which has nothing to do with the sophistication of the technology. The records upon which Indian policing systems are necessarily trained are not neutral. They are the accumulated sediment of an administrative apparatus which, for eight decades, operated a statutory classification of persons by birth; and although that statute was repealed in 1952, its operative content survives within the files.⁹ A model fitted to such records recovers the classification as a statistical regularity. It does so without design and without anybody intending it, because the regularity is among the strongest signals which the data contains.

The argument proceeds in nine further Parts. Part II reconstructs the requirement of intelligibility within Article 14 and demonstrates that it survives, and is indeed strengthened by, the transition from the classification test to the doctrine of manifest arbitrariness. Part III examines what a learned classifier is, and why the object which it produces does not constitute a differentia in the constitutional sense. Part IV asks to whom the requirement of intelligibility is addressed, and contends that it is addressed to the court, to the person classified and to the administrative authority, for reasons which are not reducible to one another. Part V confronts the strongest reply available to the State, namely that demonstrated accuracy ought to be accepted in substitution for articulability. Parts VI and VII turn to the Indian material specifically: the former examines what occurs when opaque classification is trained upon records which encode a repealed colonial statute, and the latter identifies where

⁸ The proposition has an ancient lineage in the common law. See *Padfield v. Minister of Agriculture, Fisheries and Food*, [1968] A.C. 997, 1032–33 (H.L.) (Eng.) (a Minister who gives no reason may be found to have had none that would bear examination).

⁹ Criminal Tribes Act, No. 27 of 1871 (India), repealed by the Criminal Tribes Laws (Repeal) Act, No. 24 of 1952 (India). The formulation owes something to the historiography of colonial record-keeping, which has emphasised that the colonial state governed as much through its files as through its statutes. See BERNARD S. COHN, *COLONIALISM AND ITS FORMS OF KNOWLEDGE: THE BRITISH IN INDIA* 3–15 (1996).

the constitutional defect is properly to be located. Part VIII examines the principal objections. Part IX sketches the elements of a remedial framework, and Part X concludes.

II. The Structure of the Equality Guarantee and the Requirement of Intelligibility

A. The Origin and Content of the Classification Test

The two-limbed classification test entered Indian law in a form already substantially complete, and the sequence by which it did so repays attention, because the premise with which this article is concerned entered without argument at each stage.

In *State of West Bengal v. Anwar Ali Sarkar*, the Supreme Court struck down section 5(1) of the West Bengal Special Courts Act, 1950, which permitted the State Government to direct that offences, or classes of offences, or classes of cases, be tried by special courts under an abbreviated procedure, without stating any principle by reference to which the selection was to be made.¹⁰ What is ordinarily remembered of the decision is its outcome. What matters for present purposes is its reasoning. The Court did not hold that the classification produced unequal results; it could not have so held, for no classification had yet been effected. It held that the statute was void because it disclosed no basis upon which a classification might be effected at all. Fazl Ali, J. expressed the point with some force: the section conferred a power of selection without indicating the criterion upon which selection would proceed, and a power so framed is not a classification but the absence of one.¹¹ Sastri, C.J., dissenting, would have upheld the provision upon the footing that the preamble disclosed a sufficient policy;¹² but even the dissent proceeded upon the assumption that a stated principle was required, and differed only upon whether one had been supplied.¹³

That is a holding concerning legibility. It was arrived at before any outcome existed which might have been measured, and it did not depend upon one.

The decision in *Budhan Choudhry v. State of Bihar* thereafter supplied the canonical formulation. A classification is permissible if it is founded upon an intelligible differentia

¹⁰ *State of West Bengal v. Anwar Ali Sarkar*, A.I.R. 1952 S.C. 75, 80–81 (India). The Act was West Bengal Act X of 1950.

¹¹ *Id.* at 83–85 (Fazl Ali, J.).

¹² *Id.* at 78–80 (Sastri, C.J., dissenting).

¹³ The disagreement was thus located at the second stage of the enquiry rather than the first, which is itself evidence that the requirement of a stated principle was common ground.

distinguishing those grouped together from those left out, and if that differentia bears a rational relation to the object sought to be achieved by the statute in question.¹⁴ The decision in *Ram Krishna Dalmia v. Shri Justice S.R. Tendolkar* restated the formulation and elaborated the presumptions attending it,¹⁵ and the seven propositions enumerated in *In re Special Courts Bill, 1978* consolidated three decades of application.¹⁶ Throughout that development the adjective travels unexamined, treated as though it were a synonym for “real” or “non-arbitrary”.¹⁷

B. Intelligibility as an Epistemic Predicate

It is not a synonym. “Intelligible” is an epistemic predicate. It describes a relation subsisting between a classification and a mind capable of apprehending it. A differentia may be simultaneously real and unintelligible, for the world contains a great many true distinctions which nobody is able to state; and Indian doctrine, uniquely among the major traditions of equality jurisprudence, has written the requirement of graspability into the test itself.

American equal protection doctrine enquires into the fit between means and ends and into the suspectness of the trait employed; it does not, in terms, require that the trait be capable of expression.¹⁸ The German proportionality enquiry asks whether a measure is suitable, necessary and proportionate in the narrow sense, each of which presupposes an identified measure but none of which imposes a distinct requirement of articulability.¹⁹ Article 14 of the Constitution of India does impose such a requirement, and the reason it does so is structural rather than accidental. The two limbs of the classification test are sequential, and the second cannot be performed unless the first has been satisfied. A court is unable to assess whether a differentia bears a rational nexus to a statutory object unless it is first able to identify what the

¹⁴ *Budhan Choudhry v. State of Bihar*, A.I.R. 1955 S.C. 191, 193 (India).

¹⁵ *Ram Krishna Dalmia v. Shri Justice S.R. Tendolkar*, A.I.R. 1958 S.C. 538, 547–48 (India).

¹⁶ *In re Special Courts Bill, 1978*, (1979) 1 S.C.C. 380, 422–25 (India).

¹⁷ A survey of the Indian commentary bears this out. The standard texts reproduce the formulation and proceed directly to the second limb. See M.P. JAIN, *INDIAN CONSTITUTIONAL LAW* 861–79 (8th ed. 2018); H.M. SEERVAI, *CONSTITUTIONAL LAW OF INDIA* vol. 1, 436–55 (4th ed. 1991).

¹⁸ The American enquiry is organised around tiers of scrutiny keyed to the classification employed and the interest affected. See, e.g., *City of Cleburne v. Cleburne Living Center, Inc.*, 473 U.S. 432, 439–42 (1985). The closest American analogue is the requirement that a legislature’s purpose be identifiable for the purposes of rational basis review, but the doctrine permits the court to hypothesise a purpose which the legislature never entertained. See *Federal Communications Commission v. Beach Communications, Inc.*, 508 U.S. 307, 313–15 (1993). That permission is precisely what the Indian requirement of an intelligible differentia denies.

¹⁹ On the structure of the German enquiry, see Dieter Grimm, *Proportionality in Canadian and German Constitutional Jurisprudence*, 57 U. TORONTO L.J. 383 (2007); ROBERT ALEX, *A THEORY OF CONSTITUTIONAL RIGHTS* 66–69 (Julian Rivers trans., 2002).

differentia is. Review for nexus presupposes a stated ground. Withdraw the statement, and the second limb has nothing upon which to operate.

This is not a merely formal observation. It explains why the Supreme Court, in *Nagpur Improvement Trust v. Vithal Rao*, was able to examine whether a classification was under-inclusive or over-inclusive:²⁰ such an examination requires that the intended class be held in view alongside the actual class, and the intended class exists only in the form of a statement.²¹ It explains, similarly, the Court's method in *Deepak Sibal v. Punjab University*, where the distinction drawn between employees of government and semi-government institutions and employees of other institutions was tested against the object of an evening course of instruction in law.²² Remove the stated distinction and substitute an assertion that a system sorted applicants effectively, and nothing remains against which the object may be measured.

C. Manifest Arbitrariness and the Requirement of a Determining Principle

An obvious rejoinder presents itself. The classification test, it may be said, has been overtaken. The decision in *E.P. Royappa v. State of Tamil Nadu* recast equality as the antithesis of arbitrariness;²³ the decisions in *Maneka Gandhi v. Union of India* and *Ajay Hasia v. Khalid Mujib Sehravardi* entrenched that reading;²⁴ and the decision in *Shayara Bano v. Union of India* furnished the doctrine of manifest arbitrariness with its modern statement as a ground for invalidating legislation itself.²⁵ If the operative test is now arbitrariness rather than classification, does the requirement of intelligibility survive?

It survives, and it is sharpened. Nariman, J. defined manifest arbitrariness to include action taken capriciously or irrationally, and action taken “without adequate determining principle”.²⁶ That formulation deserves to be read slowly. It is not a test of outcome. A measure whose consequences are entirely benign remains manifestly arbitrary if no

²⁰ *Nagpur Improvement Trust v. Vithal Rao*, (1973) 1 S.C.C. 500, 512–13 (India).

²¹ This is why the Supreme Court has been able to strike down classifications as under-inclusive without any evidence concerning the persons excluded. The exclusion is apparent from the statement of the class itself.

²² *Deepak Sibal v. Punjab University*, (1989) 2 S.C.C. 145, 155–57 (India).

²³ *E.P. Royappa v. State of Tamil Nadu*, (1974) 4 S.C.C. 3, 38 (India) (“Equality is a dynamic concept with many aspects and dimensions and it cannot be imprisoned within traditional and doctrinaire limits.”).

²⁴ *Maneka Gandhi v. Union of India*, (1978) 1 S.C.C. 248, 283–84 (India); *Ajay Hasia v. Khalid Mujib Sehravardi*, (1981) 1 S.C.C. 722, 740–41 (India).

²⁵ *Shayara Bano v. Union of India*, (2017) 9 S.C.C. 1, 296–300 (India) (Nariman, J.).

²⁶ *Id.* at 298.

determining principle underlies it, and the reason is that Article 14 protects a relationship between the State and the governed, a relationship in which the exercise of coercive power is owed a reason, rather than protecting any particular distribution of goods. The vice identified in *Shayara Bano v. Union of India* was not that the practice of talaq-e-biddat produced adverse consequences on average. It was that the power which the practice conferred was untethered from any principle capable of statement and defence.²⁷

Both branches of Article 14 therefore converge upon the same demand, though they approach it from different directions. The classification test requires that the ground of differentiation be intelligible before nexus may be assessed. The doctrine of manifest arbitrariness requires that an adequate determining principle exist and be identifiable. Neither is satisfied by a State which says, in substance, that it is unable to disclose what the principle is, but gives its assurance that a principle exists and that it operates satisfactorily.

It should be acknowledged that the arbitrariness doctrine has attracted substantial criticism, both for its doctrinal indeterminacy and for the breadth of the discretion it confers upon the judiciary.²⁸ The argument advanced here does not depend upon the doctrine's more expansive applications. It requires only the narrow proposition, which even the doctrine's critics accept, that State action must be capable of being justified by reference to some principle which can be stated.²⁹

D. The Orientation of Indian Discrimination Law towards Grounds Rather than Effects

Indian equality law has never wholeheartedly embraced effects-based review. The words “only on grounds of” in Article 15(1) have been construed, with occasional and contested departures, as directed at the ground upon which the State acts rather than at the distribution which its action produces.³⁰ Scholars have long regarded this as a weakness, and in the

²⁷ *Id.* at 301–02. The same reasoning was applied in *Navtej Singh Johar v. Union of India*, (2018) 10 S.C.C. 1, 273–76 (India), and in *Joseph Shine v. Union of India*, (2019) 3 S.C.C. 39, 73–76 (India).

²⁸ See Shankar Narayanan, *Rethinking “Non-Arbitrariness”*, 4 J. NAT'L L. UNIV. DELHI 1 (2017); see also Tarunabh Khaitan, *Equality: Legislative Review under Article 14*, in THE OXFORD HANDBOOK OF THE INDIAN CONSTITUTION 699 (Sujit Choudhry, Madhav Khosla & Pratap Bhanu Mehta eds., 2016).

²⁹ Even the most restrictive account of the arbitrariness doctrine, which would confine it to executive action and exclude legislation, accepts this minimum. The argument in the text is concerned with executive deployment and therefore does not engage the contested extension.

³⁰ See *State of Bombay v. Bombay Education Society*, A.I.R. 1954 S.C. 561, 566 (India); *Nain Sukh Das v. State of Uttar Pradesh*, A.I.R. 1953 S.C. 384, 386 (India). For the argument that Indian discrimination law ought to be reoriented, see Tarunabh Khaitan, *Beyond Reasonableness: A Rigorous Standard of Review for Article 15 Infringement*, 50 J. INDIAN L. INST. 177 (2008).

ordinary run of discrimination cases it is one. Indirect discrimination is a real phenomenon, and a ground-facing doctrine apprehends it imperfectly. The Supreme Court's most sustained engagement with the problem, in *Lieutenant Colonel Nitisha v. Union of India*, recognised indirect discrimination in the evaluation criteria applied to women officers of the Army and adopted a framework drawn from comparative sources;³¹ but the recognition remains recent and its consolidation incomplete.³²

As against machine-learned classification, however, the orientation constitutes an advantage, and it is worth being precise concerning the reason. A ground-facing doctrine poses to the State a question which the State must answer in advance: upon what basis are persons being sorted? An effects-facing doctrine poses to the claimant a question which the claimant is ordinarily unable to answer at all: what has the sorting done, across the population, to persons situated as she is? In the algorithmic setting the second question approaches unanswerability by anybody outside the deploying agency. The claimant does not know that she was scored. She does not know which features were employed. She has no access to the counterfactual population which was not flagged, and no means of computing a base rate. Every empirical element of a claim founded upon disparate impact lies behind the wall which the opacity of the system itself erects.

The ground-facing question demands nothing of her. It requires the State to state its differentia, and treats the failure to do so as constituting the violation. That is a burden which the State is able to discharge if its classification is legible, and unable to discharge if it is not, which is precisely where the constitutional line ought properly to fall.

III. The Nature of Machine-Learned Classification

A. The Character of the Object the State Deploys

Constitutional lawyers are inclined to describe machine learning loosely, and the looseness conceals the argument. Precision is therefore necessary.

³¹ *Lieutenant Colonel Nitisha v. Union of India*, (2021) 15 S.C.C. 125, 172–80 (India) (adopting, inter alia, the framework articulated in *Fraser v. Canada (Attorney General)*, 2020 SCC 28 (Can.)).

³² The framework adopted in *Lieutenant Colonel Nitisha v. Union of India* requires the claimant to establish that a facially neutral criterion disproportionately affects a protected group, which reintroduces the evidentiary difficulty identified in the text.

A supervised classifier is the product of an optimisation procedure. The procedure is furnished with a body of labelled historical records, which in the policing context means records of who was arrested, charged, listed or found at the scene of an offence, together with a family of candidate functions; and it searches that family for the function which minimises a specified loss upon the historical data.³³ What emerges is a set of parameters. In a gradient-boosted ensemble of decision trees, the parameters may consist of several thousand branching conditions distributed across hundreds of trees whose outputs are summed. In a neural network they consist of matrices of weights. The resulting object possesses no propositional content. It is not a rule which was selected; it is a surface which was fitted.³⁴

Three properties of that surface are of constitutional significance, and none of them is an incidental defect susceptible of being engineered away.

B. Non-Articulability

The decision boundary is defined jointly over the whole set of input features, and it is characteristically non-linear and non-monotonic. A given feature may push a prediction upward across one region of the feature space and downward across another; the contribution of any feature is conditional upon the values of the others.³⁵ There exists in consequence no true sentence of the form “persons possessing property P are classified as X”, for the classifier does not operate upon properties in that sense. It operates upon positions. Asked to identify the differentia, the honest reply is the parameter file, and a parameter file is not a reason.

The difficulty is not one of disclosure. A State willing to publish its model in its entirety, together with its weights and its training code, would not thereby have stated a differentia. It would have supplied the means by which the classification may be reproduced, which is a

³³ For an accessible treatment addressed to lawyers, see Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 89–99 (2014); David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653 (2017). The choice of loss function is itself a normative choice with distributive consequences, for it determines how errors of different kinds are traded against one another. See Solon Barocas, Moritz Hardt & Arvind Narayanan, *FAIRNESS AND MACHINE LEARNING: LIMITATIONS AND OPPORTUNITIES* ch. 3 (2023).

³⁴ The distinction corresponds to Breiman’s contrast between the data modelling culture, which posits a stochastic model and interprets its parameters, and the algorithmic modelling culture, which treats the mechanism as unknown and evaluates only predictive accuracy. Breiman, *infra* note 38, at 199–201.

³⁵ Interaction effects of this kind are not incidental. They are frequently the source of the performance advantage which opaque models enjoy over additive ones, so that the property which defeats articulation is the same property which is offered as the justification for deployment.

different thing altogether.³⁶ The distinction between reproducing a classification and stating it is central to the argument of this article, and it is a distinction which the transparency literature has tended to elide.³⁷

C. Instability

Retrain the model upon data extended by a further quarter, adjust a hyperparameter, alter the random seed with which the optimisation is initialised, and the boundary shifts. Two models exhibiting indistinguishable aggregate accuracy may disagree upon a substantial proportion of individual cases, a phenomenon which the technical literature terms predictive multiplicity.³⁸ Under the classification test, a differentia is supposed to be that which explains why the group was constituted as it was. A boundary which might as readily have fallen elsewhere, with equivalent performance and no articulable difference, is not performing explanatory work.

The point may be sharpened by reference to *Kathi Raning Rawat v. State of Saurashtra*, in which the Supreme Court upheld a notification directing certain classes of offences to special courts because the preamble disclosed a policy of dealing with a particular species of disorder.³⁹ The differentia there was capable of statement, and having been stated it was capable of being assessed for nexus and, had the assessment failed, of being rejected. Predictive multiplicity denies the classification even that minimum stability. There is no proposition which the State could offer and later be held to.⁴⁰

Indian Journal of Contemporary Legal and Social Issues

³⁶ The distinction may be illustrated by analogy. Publishing the complete transcript of a jury's deliberations would enable the verdict to be reproduced; it would not thereby supply the ratio of a judgment.

³⁷ For the classic treatment of opacity in algorithmic governance, see FRANK PASQUALE, *THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION* 189–218 (2015); cf. Jenna Burrell, *How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms*, 3 *BIG DATA & SOC'Y* 1 (2016) (distinguishing opacity arising from corporate secrecy, from technical illiteracy, and from the intrinsic character of high-dimensional optimisation).

³⁸ Charles Marx, Flavio Calmon & Berk Ustun, *Predictive Multiplicity in Classification*, in *PROCEEDINGS OF THE 37TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING* 6765 (2020); see also Leo Breiman, *Statistical Modeling: The Two Cultures*, 16 *STAT. SCI.* 199, 206–08 (2001) (describing the multiplicity of good models as the "Rashomon effect").

³⁹ *Kathi Raning Rawat v. State of Saurashtra*, A.I.R. 1952 S.C. 123, 126–27 (India).

⁴⁰ Estoppel and consistency doctrines assume a stated position from which departure may be identified. Where the position is not stated, departure cannot be detected.

D. Non-Generality

Legal classification proceeds by general categories: persons convicted twice within five years, establishments employing more than twenty workers, vehicles exceeding a stated tonnage. The category is defined first and applied thereafter, and its generality is what enables a person to ascertain in advance upon which side of the line he stands. A learned classifier inverts this order. It produces a verdict in respect of each instance, and the “category” is merely the set of instances which happened to receive that verdict. There exists no antecedent rule which the subject might have consulted. The class exists only as an enumeration of its members.

This bears directly upon the rule of law values which Article 14 has been understood to protect. The Supreme Court in *Kartar Singh v. State of Punjab* recognised that a law which fails to give fair warning of the conduct it proscribes offends the constitutional guarantee, for a person of ordinary intelligence must be afforded a reasonable opportunity to know what is prohibited.⁴¹ A classification constituted only by enumeration gives no warning of any kind. It cannot, for there is nothing to warn of until the enumeration has been completed.⁴²

E. The Insufficiency of Post Hoc Explanation

The obvious rejoinder is that techniques of explanation close the gap. They do not, and the reason is conceptual rather than technical.

The technique known as LIME fits a simple and interpretable surrogate, ordinarily a sparse linear model, to the behaviour of the complex model within the immediate neighbourhood of a single input, and reports the coefficients of the surrogate.⁴³ The technique known as SHAP allocates the difference between a prediction and a chosen baseline across the input features by means of a solution concept drawn from cooperative game theory.⁴⁴ Both are legitimate

⁴¹ *Kartar Singh v. State of Punjab*, (1994) 3 S.C.C. 569, 643–44 (India); see also *Shreya Singhal v. Union of India*, (2015) 5 S.C.C. 1, 92–95 (India).

⁴² The point is distinct from, though related to, the vagueness doctrine. A vague law states a criterion whose boundaries are uncertain. An enumerative classification states no criterion at all.

⁴³ Marco Tulio Ribeiro, Sameer Singh & Carlos Guestrin, “Why Should I Trust You?”: Explaining the Predictions of Any Classifier, in PROCEEDINGS OF THE 22ND ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING 1135 (2016).

⁴⁴ Scott M. Lundberg & Su-In Lee, *A Unified Approach to Interpreting Model Predictions*, in ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 30 4765 (2017).

diagnostic instruments and both are valuable for the purposes for which they were designed.⁴⁵ Neither produces the ground upon which the classification rested.

The distinction merits being taken apart. An attribution generated by SHAP is a statement concerning a counterfactual decomposition: it asserts that, relative to a chosen baseline distribution and subject to assumptions concerning the independence of features which are almost never satisfied in practice, a given feature accounts for a given proportion of the deviation from that baseline. It does not assert that the model applied a criterion. The criterion is the fitted surface; the attribution is a summary computed afterwards by a second procedure, possessing its own assumptions, its own dependence upon the baseline selected, and its own well-documented instability.⁴⁶ Selbst and Barocas have captured the point precisely: methods of explanation inform us how a model behaves, and we mistake that for being informed why a decision was justified, because in human decision-making the two ordinarily coincide.⁴⁷ In the case of a fitted model they come apart, and the gap is not closing.

The argument advanced by Cynthia Rudin is stronger still and deserves to be received into constitutional law rather than left within computer science. Where decisions carry high stakes, she contends, the appropriate response is not to explain models which cannot be inspected but to decline to use them and to construct interpretable models instead, for a post hoc explanation of an uninspectable model is by definition a second model which may be mistaken concerning the first.⁴⁸ An explanation which may be mistaken concerning the model is not a differentia. It is a hypothesis concerning a differentia, advanced by the party bearing the burden of justification, and incapable of verification by the court.⁴⁹

There exists an analogue within Indian administrative law, and it is unexpectedly close. The Supreme Court has consistently declined to receive reasons supplied after the event in

⁴⁵ They were designed principally to assist practitioners in debugging models and detecting data leakage, which are engineering purposes and not forensic ones.

⁴⁶ I. Elizabeth Kumar, Suresh Venkatasubramanian, Carlos Scheidegger & Sorelle Friedler, *Problems with Shapley-Value-Based Explanations as Feature Importance Measures*, in PROCEEDINGS OF THE 37TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING 5491 (2020).

⁴⁷ Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 FORDHAM L. REV. 1085, 1109–18 (2018).

⁴⁸ Cynthia Rudin, *Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead*, 1 NATURE MACH. INTELLIGENCE 206, 206–07 (2019).

⁴⁹ The difficulty is compounded where the explanation is generated by the proponent using proprietary tooling which the opposing party cannot inspect. See Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343, 1368–92 (2018).

support of an order which recorded none. An order must stand or fall upon the grounds stated within it, and affidavits subsequently filed cannot cure the omission. Bose, J. stated the principle in *Commissioner of Police, Bombay v. Gordhandas Bhanji*, and Krishna Iyer, J. restated it in *Mohinder Singh Gill v. Chief Election Commissioner* in terms which have since been treated as canonical: when a statutory functionary makes an order based on certain grounds, its validity must be judged by the reasons so mentioned and cannot be supplemented by fresh reasons in the shape of an affidavit or otherwise.⁵⁰ The rule exists because a reason produced under the pressure of litigation is a rationalisation, and everybody understands that it is. An attribution generated by the deploying agency's own tooling, upon its own model, for the purpose of satisfying a court, stands in precisely that relation to the decision which it purports to explain.⁵¹

IV. The Addressee of the Requirement of Intelligibility

The decided cases have never been required to determine to whom the requirement of intelligibility is addressed. The question becomes live only when the three candidate addressees may be satisfied separately, a situation which machine classification creates for the first time. The candidates are considered in turn, and the conclusion is that the requirement runs to all three, for reasons which are not reducible to one another.

A. The Court and the Assessment of Rational Nexus

The judicial addressee is the least contentious. Review for nexus is impossible in the absence of a stated differentia, and this is a matter of the structure of the test rather than of judicial convenience. The Supreme Court in *Deepak Sibal v. Punjab University* examined whether a particular distinction bore a rational relation to the object of an evening course of instruction, and in order to do so was required to hold the distinction in view simultaneously with the

⁵⁰ *Commissioner of Police, Bombay v. Gordhandas Bhanji*, A.I.R. 1952 S.C. 16, 20–21 (India); *Mohinder Singh Gill v. Chief Election Commissioner*, (1978) 1 S.C.C. 405, 417–18 (India).

⁵¹ The analogy is not exact, since an attribution is generated mechanically rather than composed by counsel. The relevant similarity is that both are produced after the event, by the party bearing the burden, for the purpose of the litigation.

object.⁵² Substitute for the distinction an assertion that a model sorted applicants effectively, and nothing remains to be compared with anything.⁵³

Review under Article 14 further requires a court to test a classification for under-inclusion and over-inclusion, which is to say to enquire who has been swept in who ought not to have been, and who has been left out who ought to have been included.⁵⁴ That exercise is by definition comparative, and by definition requires a statement of the intended class. A model produces only the actual class. There exists no intended class against which the actual class might be measured, with the consequence that the enquiry into inclusiveness cannot be conducted at all. It is not conducted badly, nor conducted upon limited material; it is not conducted.⁵⁵

B. The Person Classified and the Capacity to Contest

The second addressee is the person classified, and the interest at stake is that of contestability. The requirements of natural justice in India have moved decisively towards a duty to record reasons, and the rationale which the Supreme Court has offered is instrumental as well as dignitary: reasons enable the affected party to know why a decision was taken against him and to challenge it.⁵⁶ In *Kranti Associates Private Limited v. Masood Ahmed Khan*, the Court collected the Indian and comparative authority and summarised the position in a series of propositions, observing that reasons constitute the live link between the material on record and the conclusion reached, and that their recording introduces clarity and excludes arbitrariness.⁵⁷

A person who is informed only that a system has placed him within a category is unable to mount a challenge of any description. He cannot contend that the criterion applied was bad in

⁵² *Deepak Sibal v. Punjab University*, (1989) 2 S.C.C. 145, 155–57 (India).

⁵³ A court might of course compare the system's outputs with the statutory object in an aggregate sense, asking whether the persons flagged were in fact more likely to offend. That is a validation enquiry and not a nexus enquiry, for the reasons developed in Part V.C.

⁵⁴ *Nagpur Improvement Trust v. Vithal Rao*, (1973) 1 S.C.C. 500, 512–13 (India); see also *Kedar Nath Bajoria v. State of West Bengal*, A.I.R. 1953 S.C. 404, 406–07 (India).

⁵⁵ This distinguishes the algorithmic case from the ordinary case of an inadequately reasoned order, where the court may remit for fresh consideration. There is nothing to remit.

⁵⁶ *A.K. Kraipak v. Union of India*, (1969) 2 S.C.C. 262, 272–73 (India); *S.N. Mukherjee v. Union of India*, (1990) 4 S.C.C. 594, 612–14 (India).

⁵⁷ *Kranti Associates Private Limited v. Masood Ahmed Khan*, (2010) 9 S.C.C. 496, 509–12 (India); see also *Assistant Commissioner, Commercial Tax Department v. Shukla and Brothers*, (2010) 4 S.C.C. 785, 791–93 (India).

law, for no criterion was disclosed to him. He cannot contend that the criterion was misapplied to his facts, for there exists no step of application to be attacked. He is left with the bare assertion that the output is mistaken, which is not an argument but an expression of dissatisfaction.⁵⁸

Where the classification is a policing classification, the consequence is more serious than inconvenience. The subject of a history sheet or of an entry in a surveillance register in India has, since the decisions in *Kharak Singh v. State of Uttar Pradesh* and *Malak Singh v. State of Punjab and Haryana*, enjoyed at best a thin set of protections, and the Court in *Malak Singh v. State of Punjab and Haryana* was explicit that such entries are made without notice to the person concerned, upon the footing that prior intimation would defeat the object of the surveillance.⁵⁹ To superimpose an unstateable classification upon a regime which already operates without notice is to produce a category of State action which is, for practical purposes, unreviewable from below. That precise configuration is presently before the Supreme Court in the writ petition challenging the deployment of facial recognition and associated biometric systems at sites of protest in Delhi.⁶⁰

C. The Administrative Authority and the Exercise of Statutory Discretion

The third addressee is the one most readily overlooked, and it carries with it the question of delegation. Indian administrative law requires that a statutory power be exercised by the authority upon whom it has been conferred, upon that authority's own satisfaction, formed upon relevant material.⁶¹ An officer who acts upon a machine output which he is unable to interrogate has not formed a satisfaction; he has adopted one.⁶² If he is unable to state the

⁵⁸ Compare the position of a person served with a show cause notice which does not disclose the material relied upon, which Indian courts have repeatedly held to be no notice at all.

⁵⁹ *Kharak Singh v. State of Uttar Pradesh*, A.I.R. 1963 S.C. 1295, 1302–04 (India); *Malak Singh v. State of Punjab and Haryana*, (1981) 1 S.C.C. 420, 425–27 (India). The Court in *Malak Singh v. State of Punjab and Haryana* was nonetheless careful to hold that the discretion was confined by the terms of the rule and that entries could be made only in respect of persons falling within the described categories, which is itself a requirement of legibility.

⁶⁰ The petition, taken up for examination by a Bench presided over by Surya Kant, C.J., challenges the deployment by the Delhi Police of facial recognition systems, vehicle-tracking systems, smart spectacles and mobile fingerprint-matching applications at sites of protest, and the retention of the resulting biometric data by private contractors.

⁶¹ *Barium Chemicals Limited v. Company Law Board*, A.I.R. 1967 S.C. 295, 322–24 (India); *Jaswant Sugar Mills Limited v. Lakshmi Chand*, A.I.R. 1963 S.C. 677, 684–85 (India).

⁶² The distinction between forming a satisfaction and adopting one is the same distinction which underlies the rule against acting under dictation. See *Purtabpore Company Limited v. Cane Commissioner of Bihar*, A.I.R. 1970 S.C. 1896, 1901–02 (India).

basis of the classification, he cannot have been satisfied of it in the sense which the statute requires, and the deployment accordingly raises a problem of delegation distinct from the problem of equality, namely the exercise of statutory power by an entity upon which the legislature conferred none.

The Conseil constitutionnel of France arrived at a closely comparable conclusion by an entirely different route, and its reasoning repays attention. Reviewing the legislation implementing the General Data Protection Regulation, it upheld the use of algorithms in individual administrative decisions only upon express conditions: that the administration be in a position to explain, in detail and in an intelligible form, the manner in which the processing operated in the particular case; and that algorithms which revise their own rules without human supervision not be employed for such decisions at all.⁶³ The second condition is the significant one. It is not a requirement of transparency but a prohibition, grounded upon the proposition that an administration which cannot state its own rule has not taken a decision which it is competent to take.⁶⁴

D. The Tripartite and Antecedent Character of the Requirement

Because the three addressees are differently grounded, in reviewability, in contestability and in the lawful exercise of conferred power respectively, satisfying one does not satisfy the others; and no method of explanation presently available satisfies any of them. A report of feature attributions does not furnish the court with a differentia capable of being tested for nexus. It does not furnish the subject with a criterion capable of being contested. It does not furnish the administrative authority with a basis upon which a satisfaction might be formed, for that authority is in no position to verify that the attribution is faithful to the model.

The condition of legibility is therefore tripartite, and it is antecedent. It must be satisfied before deployment, because each of the three functions which it serves operates at or before the moment of classification and not after litigation has commenced.

⁶³ Conseil constitutionnel [CC] [Constitutional Council] decision No. 2018-765DC, June 12, 2018, cons. 70–71 (Fr.).

⁶⁴ The French position is the more remarkable for having been reached under a constitutional text containing no equality provision comparable to Article 14, and in the course of reviewing legislation implementing a European instrument which itself does not require it.

V. Validation as a Proposed Substitute for Articulation

A. The Argument from Demonstrated Performance

The strongest reply available to the State is not that its model is explainable. It is that explanation is beside the point. The differentia, upon this view, is simply this: persons whom a validated system identifies as presenting elevated risk. That description is perfectly intelligible. The nexus to the statutory object, whether the prevention of crime, the efficient allocation of patrols or the targeting of investigative resources, is direct.⁶⁵ And the record of validation is offered in substitution for articulation: the State concedes that it cannot state what the model attends to, but undertakes to demonstrate that the model's outputs track the phenomenon with which the State is legitimately concerned.

The argument is serious and deserves better than dismissal. It is also, for three distinct reasons, mistaken; and the manner in which it is mistaken discloses what Article 14 exists to protect.

B. The Circularity of the Self-Referential Differentia

The first objection is one of circularity. The formula “persons identified by the system” names the output of the classification rather than its basis. Every classification whatsoever may be redescribed in that form: persons whom the officer selected, cases which the Government directed to special courts, establishments which the inspector identified.⁶⁶ The decision in *State of West Bengal v. Anwar Ali Sarkar* exists precisely in order to hold that such a redescription does not constitute a differentia.⁶⁷ Were the State able to satisfy Article 14 by naming the procedure of selection rather than the criterion of selection, the classification test would be empty, for every exercise of power names a procedure.

The point admits of a further refinement. What the State offers is not merely a redescription but a redescription which incorporates a warranty: persons identified by a system *which has been validated*. The warranty is doing the work, and it is a warranty concerning outcomes

⁶⁵ Indeed the nexus is tighter than in most classification cases, which is what gives the argument its initial appeal. The tighter the nexus appears, the less scrutiny the first limb is likely to receive.

⁶⁶ The vice was identified in exactly these terms in *State of West Bengal v. Anwar Ali Sarkar*, where the Government's power was to direct that “cases” or “classes of cases” be tried by special courts, a formulation which names the act of selection and nothing else.

⁶⁷ *State of West Bengal v. Anwar Ali Sarkar*, A.I.R. 1952 S.C. 75, 83–85 (India) (Fazl Ali, J.).

rather than grounds. Whether such a warranty may stand in place of a ground is the substance of the next two objections.⁶⁸

C. Contamination of the Target Variable and the Question of Nexus

The second objection is that accuracy cannot answer the question of nexus, because nexus concerns the relationship between the ground of differentiation and the statutory object, rather than the correlation between an output and an outcome.

Suppose a model predicts arrest with high fidelity. Arrest is not crime. It is the observable trace of police deployment, and police deployment is precisely what the model is being used to direct. The target variable is accordingly contaminated by the intervention, and a system trained upon it will be accurate concerning the phenomenon which it also causes. The technical literature terms this a runaway feedback loop, and has demonstrated formally that such a system will concentrate attention upon the areas in which it initially predicted crime, irrespective of the underlying rate of offending.⁶⁹ Metrics of validation cannot detect the difficulty, for the metric and the contamination share a common source.

There is a further and more elementary difficulty, which the literature on the construction of target variables has explored at length.⁷⁰ The decision as to what the model shall predict is a normative decision taken by the designer, and it is invisible in the output. A system which predicts arrest, a system which predicts conviction, and a system which predicts victimisation are three different systems embodying three different judgments about what matters, and each may be validated against its own target with equal success.⁷¹ A court enquiring into nexus is asking what the State is in truth sorting upon, and the record of validation is silent upon that question by construction.

Indian Journal of Contemporary
Legal and Social Issues

⁶⁸ It should be noted that the warranty is also, in practice, unverifiable by the court, since validation studies are conducted by or for the deploying agency and are rarely disclosed.

⁶⁹ Danielle Ensign, Sorelle A. Friedler, Scott Neville, Carlos Scheidegger & Suresh Venkatasubramanian, *Runaway Feedback Loops in Predictive Policing*, in PROCEEDINGS OF THE 1ST CONFERENCE ON FAIRNESS, ACCOUNTABILITY AND TRANSPARENCY 160 (2018); see also Kristian Lum & William Isaac, *To Predict and Serve?*, 13 SIGNIFICANCE 14 (2016). The formal demonstration proceeds by modelling the process as a Pólya urn, in which each observation increases the probability of further observation from the same source. *Id.* at 163–68.

⁷⁰ Barocas & Selbst, *supra* note 3, at 677–87.

⁷¹ The choice is rarely made deliberately. It is ordinarily dictated by which variable happens to be recorded, which means that the normative judgment is made by the historical record-keeping practices of the agency rather than by anybody exercising judgment.

D. Article 14 as a Reason-Giving Norm

The third objection is the most important. To accept the defence founded upon validation would be to convert Article 14 from a reason-giving norm into a consequentialist one, and that is not a modest doctrinal adjustment. It is an alteration in what the guarantee protects.

Article 14 is not a guarantee that State action will turn out well. Were it such a guarantee, the decisions in *E.P. Royappa v. State of Tamil Nadu* and *Shayara Bano v. Union of India* would be unintelligible, for a measure lacking a determining principle may be entirely benign in its effects and is void irrespective of them.⁷² What the guarantee protects is the standing of the governed to be given a reason and to answer it, an interest which Jeremy Waldron has argued distinguishes the rule of law from efficient administration, and which he locates in the procedural dimension of legality rather than in any particular substantive outcome.⁷³ An accurate but unstateable classification delivers good administration while withholding the reason. That is not a trade which the State is competent to make on the citizen's behalf.

A comparative datum sharpens the contrast. In *State v. Loomis*, the Supreme Court of Wisconsin permitted the use of a proprietary risk-assessment instrument at the stage of sentencing while acknowledging that neither the court nor the defendant was in a position to examine the manner in which the score was computed; the safeguard adopted was a written advisement to sentencing courts concerning the instrument's limitations.⁷⁴ The decision has been widely criticised, and the criticism has ordinarily fastened upon the requirements of due process. The deeper failure is that the court accepted demonstrated performance in place of a reason, and thereafter attempted to repair the substitution by means of a disclaimer.⁷⁵ Article 14, construed as this article suggests, does not permit the substitution in the first instance.

It is instructive that the European Union, legislating upon the subject afresh, has arrived at a distinction which is functionally the same. Article 5(1)(d) of the Artificial Intelligence Act

⁷² *Shayara Bano v. Union of India*, (2017) 9 S.C.C. 1, 298 (India).

⁷³ Jeremy Waldron, *The Rule of Law and the Importance of Procedure*, in NOMOS L: GETTING TO THE RULE OF LAW 3, 15–20 (James E. Fleming ed., 2011). The argument has an Indian counterpart in the insistence that Article 14 strikes at arbitrariness in order to ensure fairness and equality of treatment, and that the principle is one of the basic features of the Constitution. See *Ajay Hasia v. Khalid Mujib Sehravardi*, (1981) 1 S.C.C. 722, 740–41 (India).

⁷⁴ *State v. Loomis*, 881 N.W.2d 749, 760–64, 769–72 (Wis. 2016).

⁷⁵ The advisement required by the Supreme Court of Wisconsin directs sentencing courts to note, among other things, that the instrument is proprietary, that its accuracy has been questioned, and that it was developed for a purpose other than sentencing. The information is thus disclosed to the court while remaining unusable by it.

prohibits the placing on the market or use of systems which assess or predict the risk of a natural person committing a criminal offence where the assessment is based solely upon profiling or upon the assessment of personality traits; and it exempts from that prohibition systems which support a human assessment founded upon objective and verifiable facts directly linked to a criminal activity.⁷⁶ The operative distinction is between a conclusion derived from a profile, which cannot be stated as a proposition about conduct, and a conclusion derived from facts which can be stated and verified. That is the requirement of intelligibility, arrived at by a legislature rather than by a court, and expressed in the vocabulary of risk rather than of classification.⁷⁷

E. The Availability of Interpretable Alternatives

A practical corollary deserves statement, for it forecloses the contention that legibility is a luxury which effective administration cannot afford.

For the great majority of structured, tabular prediction problems of the kind which policing agencies actually confront, interpretable models, whether sparse scoring systems, short rule lists or shallow decision trees, perform within a narrow margin of opaque ensembles, and frequently indistinguishably.⁷⁸ The trade-off between accuracy and interpretability is real in the domains of vision and natural language. It is largely absent in risk scoring conducted over a few dozen administrative variables. A State which deploys an opaque model for such a task is therefore not making a painful sacrifice of transparency in the interests of effectiveness.⁷⁹ It is electing convenience, and Article 14 has never treated administrative convenience as a justification for dispensing with reasons.⁸⁰

Indian Journal of Contemporary Legal and Social Issues

⁷⁶ Regulation 2024/1689, art. 5(1)(d), 2024 O.J. (L, 2024/1689) 1 (EU).

⁷⁷ The Commission's guidelines on prohibited practices state that the elements of objective and verifiable fact must be real, substantial and meaningful if a system combining such elements with profiling is to escape the prohibition, which is a test of whether the stateable ground is doing genuine work.

⁷⁸ Rudin, *supra* note 48, at 207–08; Berk Ustun & Cynthia Rudin, *Supersparse Linear Integer Models for Optimized Medical Scoring Systems*, 102 MACHINE LEARNING 349 (2016).

⁷⁹ The claim is empirical and is contestable in particular applications. It is advanced as a defeasible presumption which the State should bear the burden of displacing in the case of the system it proposes to deploy.

⁸⁰ *Cf. Maneka Gandhi v. Union of India*, (1978) 1 S.C.C. 248, 284 (India) (administrative convenience cannot sustain a procedure which is not fair, just and reasonable).

VI. The Statutory Residue of the Criminal Tribes Act, 1871

A. The Colonial Classification

The Criminal Tribes Act, 1871 empowered the Local Government to declare that a tribe, gang or class of persons was addicted to the systematic commission of non-bailable offences, and, having made such a declaration, to register its members, to restrict their movement, to require them to report themselves at stated intervals, and to settle them in places designated for the purpose.⁸¹ The classification was hereditary in its operation and frequently hereditary in its terms. Its intellectual foundation lay in an ethnographic theory of inherited criminality which the colonial administration applied to communities whose itinerancy, occupational pattern or political inconvenience rendered them legible targets.⁸² Successive amendments extended and hardened the regime, culminating in the consolidating enactment of 1924, by which time the registered population numbered in the millions.⁸³

The historiography of the subject is now substantial and it is unanimous upon the essential point. The category was not descriptive of conduct. It was constitutive of status, and the status was ascribed at birth.⁸⁴ The scale of the population this classification touched was considerable and remains contested. The National Commission for Denotified, Nomadic and Semi-Nomadic Tribes, chaired by Balkrishna Renke, estimated in its 2008 report that persons belonging to the notified criminal tribes and their nomadic and semi-nomadic counterparts numbered approximately 10.74 crore on the basis of the 2001 Census; the subsequent Commission chaired by Bhiku Ramji Idate placed the figure, in its 2017 report, at closer to

Indian Journal of Contemporary
Legal and Social Issues

⁸¹ Criminal Tribes Act, No. 27 of 1871, §§ 2–3, 10–12 (India). The settlements were in several instances administered by missionary organisations under arrangements with the colonial government, a feature which has shaped the historiography of the subject.

⁸² MEENA RADHAKRISHNA, DISHONOUR BY HISTORY: “CRIMINAL TRIBES” AND BRITISH COLONIAL POLICY 1–20, 139–65 (2001); Sanjay Nigam, *Disciplining and Policing the “Criminals by Birth”, Part I: The Making of a Colonial Stereotype*, 27 INDIAN ECON. & SOC. HIST. REV. 131 (1990).

⁸³ Criminal Tribes Act, No. 3 of 1911 (India); Criminal Tribes Act, No. 6 of 1924 (India). Estimates vary and the figures in the colonial returns are unreliable. The order of magnitude is not in dispute.

⁸⁴ DILIP D’SOUZA, BRANDED BY LAW: LOOKING AT INDIA’S DENOTIFIED TRIBES 11–38 (2001); Rachel J. Tolen, *Colonizing and Transforming the Criminal Tribesman: The Salvation Army in British India*, 18 AM. ETHNOLOGIST 106 (1991). The point is not merely of historical interest. It determines the character of the trace which the regime left in the records, for a status ascribed at birth produces a pattern in the data which is stable across generations and therefore learnable.

fifteen crore. A classification of that magnitude could only have been administered, and can only now be inherited, through the standing registers described in Part VI.D below.⁸⁵

B. Repeal and Substitution

The Act was repealed with effect from 1952.⁸⁶ The communities which had been notified under it became, in the vocabulary which followed, denotified. Within a short period most States enacted habitual offender legislation upon a model circulated by the Union Government, and a variant of that model is presently in force in at least fourteen States and Union Territories.⁸⁷ The category continues to carry practical weight. The National Crime Records Bureau's Prison Statistics India 2022 recorded that 1.9 per cent of India's convict population was classified as a 'habitual offender', a proportion which rose to 21.5 per cent among convicts held in Delhi, a concentration difficult to explain except by reference to the disproportionate targeting which the Commissions discussed above have documented.⁸⁸

A substantial body of official and scholarly material records that the successor statutes were applied disproportionately to the same populations, in some instances by an administrative continuity so direct that the existing registers were simply relabelled. The National Commission for Denotified, Nomadic and Semi-Nomadic Tribes constituted under the chairmanship of Balkrishna Renke, which submitted its report in 2008, recommended repeal of the habitual offender legislation upon precisely that ground.⁸⁹ The subsequent Commission constituted under the chairmanship of Bhiku Ramji Idate, which reported in 2017, reiterated the concern.⁹⁰ Neither recommendation has been implemented.⁹¹

⁸⁵ National Commission for Denotified, Nomadic and Semi-Nomadic Tribes, Report (Ministry of Social Justice and Empowerment, Government of India, 30 June 2008) ('Renke Commission Report'); National Commission for Denotified, Nomadic and Semi-Nomadic Tribes, Report vol I ch 5 (Ministry of Social Justice and Empowerment, Government of India, 2017) ('Idate Commission Report').

⁸⁶ Criminal Tribes Laws (Repeal) Act, No. 24 of 1952 (India).

⁸⁷ See, e.g., Karnataka Habitual Offenders Act, No. 17 of 1961 (India); Madhya Pradesh Rajya Suraksha Adhiniyam, No. 21 of 1990 (India). The definitions vary between States, a circumstance which the Supreme Court in *Sukanya Shantha v. Union of India* found significant when directing that prison manuals conform to the definition applicable in the State concerned.

⁸⁸ 'India's Habitual Offender Laws: A Legacy of Discrimination' (Vajiram & Ravi, 21 March 2025) <<https://vajiramandravi.com/current-affairs/indias-habitual-offender-laws-a-legacy-of-discrimination>> accessed 23 September 2026, citing National Crime Records Bureau, Prison Statistics India 2022 (Ministry of Home Affairs, Government of India 2023).

⁸⁹ NATIONAL COMMISSION FOR DENOTIFIED, NOMADIC AND SEMI-NOMADIC TRIBES, REPORT (2008) (Renke Commission).

⁹⁰ NATIONAL COMMISSION FOR DENOTIFIED, NOMADIC AND SEMI-NOMADIC TRIBES, REPORT vol. I, ch. 5 (2017) (Idate Commission).

Repeal removed the statutory text. It did not remove the population from the records, from the reporting practices, from the surveillance registers, or from the operational intuitions of the police stations which had administered the regime for eight decades.

C. The Judicial Recognition in *Sukanya Shantha v. Union of India*

The Supreme Court said as much in 2024. In *Sukanya Shantha v. Union of India*, examining caste discrimination in the prison manuals of the States, the Court held the classification of “habitual offender” as deployed in the Model Prison Manual to be constitutionally suspect; observed that several State manuals equated habitual offenders with denotified or wandering tribes; and directed that references to the category conform to the definitions contained in the habitual offender legislation of the State concerned, with manuals to be revised where no such legislation exists.⁹² The Court further held that provisions of the manuals which effected caste-based segregation and caste-based allocation of labour were unconstitutional, and directed the Union and the States to amend their manuals within a stipulated period.⁹³

The judgment is important upon its own terms, and its significance for the present argument is evidentiary. The highest court in the country has now found, as a matter of record and after examination of the manuals of every State, that this particular administrative category carries a caste and tribal charge which the text of the category does not disclose. That finding is the hinge upon which the remainder of this Part turns.⁹⁴

D. The Composition of Indian Police Data

The predictive systems deployed by Indian police forces are trained upon, and operate over, police administrative records. The system deployed by the Delhi Police draws upon emergency call data and upon first information reports routed through the Crime and Criminal Tracking Network and Systems, and the documented operational practice includes the application of filters oriented towards what the officers interviewed described as ghettos,

⁹¹ The National Human Rights Commission and successive parliamentary committees have made recommendations to similar effect. The persistence of the statutes notwithstanding this weight of official opinion is itself evidence of the administrative entrenchment described in the text.

⁹² *Sukanya Shantha v. Union of India*, 2024 INSC 753, ¶¶ 175–90, 220 (India).

⁹³ *Id.* ¶¶ 200–20. The Court also directed that the caste column in prisoner registers be deleted and took suo motu cognisance of the subject to monitor compliance, which indicates that the Court regarded the practices as systemic rather than aberrational.

⁹⁴ The finding is, for present purposes, a finding of fact by the highest court, arrived at after notice to every State and upon material placed by them. It is therefore available to a subsequent petitioner without the need to prove the underlying history afresh.

migrant colonies and minority settlements.⁹⁵ The same ethnographic study records that the locations of emergency calls are frequently mis-plotted to police stations because the station address is the only reliable address database available to the call-taker; that the categorisation of calls depends upon the call-taker's own interpretation and is applied inconsistently; and that settlements are plotted as single points, so that the granularity of the record is itself a function of where the person resides.⁹⁶ These are not exotic failures. They are the ordinary texture of Indian police record-keeping, and they constitute the substrate.⁹⁷ Nor is the practice confined to Delhi. Comparable predictive systems drawing upon similar administrative and emergency-call data have been reported as under development or in use by police forces in Madhya Pradesh, Telangana, Himachal Pradesh and Jharkhand, which suggests that the operational features documented for the Delhi Police are unlikely to be an isolated case.⁹⁸

Superimposed upon that substrate are the surveillance registers and history sheets maintained under the police rules of the States, the subjects of which are selected by officers without notice and upon criteria which are, by design, not justiciable in the ordinary case.⁹⁹ Superimposed again is the Criminal Procedure (Identification) Act, 2022, which authorises the collection of an expansive set of “measurements” from persons arrested, convicted or detained, and provides for their retention in a central database for a period of seventy-five years.¹⁰⁰ The corpus available for training is therefore a seventy-five year accumulation of records generated by a policing apparatus whose practices of selection the Supreme Court has itself characterised as carrying a caste charge.

Nor is the deployment hypothetical. Facial recognition systems have been operated by the police forces of Delhi, Telangana, Tamil Nadu, Uttar Pradesh and Punjab, among others; the first legal challenge to such a deployment was brought before the High Court of Judicature at Hyderabad in *S.Q. Masood v. State of Telangana*, in which notice was issued in January

⁹⁵ Marda & Narayan, *supra* note 3, at 320–22.

⁹⁶ *Id.* at 321–23.

⁹⁷ The study is ethnographic and its findings are specific to the Delhi Police. Its value lies in documenting practices which there is no reason to suppose are peculiar to that force, and which are consistent with what is known of record-keeping elsewhere.

⁹⁸ Daksh, ‘Predictive Policing’ in Technology and Analytics for Law and Justice (Daksh 2023) <<https://www.dakshindia.org/Technology-and-Analytics-for-Law-and-Justice/part14.xhtml>> accessed 23 September 2026.

⁹⁹ *Malak Singh v. State of Punjab and Haryana*, (1981) 1 S.C.C. 420, 425–27 (India); Punjab Police Rules, 1934, rr. 23.4, 23.9 (India).

¹⁰⁰ Criminal Procedure (Identification) Act, No. 11 of 2022, §§ 2(1)(b), 3–4 (India).

2022;¹⁰¹ and the question of biometric surveillance at sites of protest is now before the Supreme Court.¹⁰² The category's practical reach was confirmed by the State itself in those very proceedings. In an affidavit filed before the Telangana High Court in January 2023, the Hyderabad City Police Commissioner described the Facial Recognition System as being used, among other purposes, to identify 'habitual offenders' and 'dreaded criminals', tying the deployment of the technology directly to the classification discussed in Parts VI.A and VI.B above.¹⁰³ The National Automated Facial Recognition System operated by the National Crime Records Bureau supplies the gallery against which such matches are computed.¹⁰⁴ Independent monitoring confirms that such deployment is neither confined to Delhi nor exceptional. Project Panoptic, an initiative of the Internet Freedom Foundation established to document facial recognition procurement across India, has recorded at least six distinct facial recognition projects operated by government authorities in Telangana alone, several of which retain facial data permanently and are used for policing purposes rather than the narrower administrative functions to which some other deployments are confined.¹⁰⁵

E. The Statistical Reconstruction of a Repealed Classification

The claim advanced here is stronger than the proposition that the data is biased, and the difference is material.

Bias, in the standard formulation, means that the training data reflects historical patterns of enforcement, with the consequence that the model reproduces them. That is true, and by now unremarkable. But the lineage of the Criminal Tribes Act, 1871 is not merely a pattern of enforcement. It was a *statutory classification*: a rule, enacted by a legislature, which sorted persons by community and attached legal consequences to the sorting. When the rule was

¹⁰¹ *S.Q. Masood v. State of Telangana*, W.P. No. 34771 of 2021 (Telangana H.C.) (India) (notice issued Jan. 2022).

¹⁰² See *supra* note 60.

¹⁰³ 'Facial Recognition System Does Not Trace Public Movement Through CCTVs, No Mass Surveillance: Hyderabad Police Tells High Court' LiveLaw (3 January 2023) <<https://www.livelaw.in/top-stories/telangana-high-court-hyderabad-police-facial-recognition-privacy-mass-surveillance-rti-act-cctv-217960>> accessed 23 September 2026.

¹⁰⁴ The system was procured pursuant to a request for proposals issued by the National Crime Records Bureau in 2019 and has been the subject of sustained criticism for operating without statutory authorisation. The gallery is populated in substantial part from the records of persons previously arrested, which means that the probability of being matched is conditioned upon prior contact with the police, and prior contact is precisely what the historical record over-represents for the communities described in Part VI.A.

¹⁰⁵ Mithun MK, 'Telangana Has the Largest Number of Facial Recognition Projects in India' The News Minute (Chennai) <<https://www.thenewsminute.com/telangana/telangana-has-largest-number-facial-recognition-projects-india-143837>> accessed 23 September 2026.

repealed, the classification did not disappear. It had been transcribed into administrative records which continued to be generated, maintained and inherited, and which remain in use.

A model fitted to such records recovers the classification as a statistical regularity. It does so without anybody designing it to, because the regularity is among the strongest signals which the data contains.¹⁰⁶ The model accordingly re-operationalises a repealed statute; and it does so in a form which is immune to the constitutional provision enacted to prevent exactly this.

Article 15(1) forbids discrimination “only on grounds of” religion, race, caste, sex or place of birth. The prohibition is engaged when the State acts *upon* such a ground.¹⁰⁷ A learned classifier never states a ground at all, for it states nothing. Caste enters not as a variable but as a reconstruction from correlated inputs: locality, surname where it is captured, prior contact with the police, the address of the settlement, the pattern of entries upon the history sheet. Each input is individually innocent; jointly they are sufficient.¹⁰⁸ The State is therefore able to say, and to say truthfully, that caste was not among the inputs. The classification is nonetheless the old one, restored.

This is the juncture at which the requirement of intelligibility performs work which no other doctrine is able to perform. A challenge founded upon impact under Article 15 requires proof that the outputs fall upon a protected group, which the claimant is unable to obtain. A challenge directed at the inputs fails, because caste is not an input. A challenge founded upon legibility under Article 14 requires neither. It asks the State to state its differentia. If the State cannot do so, the classification is void; and the reason the State cannot do so is, in this instance, precisely that stating it would expose what the classification has become.¹⁰⁹

F. Repeal without Decontamination

A final formulation may be offered, and it is advanced as this article’s principal contribution to the literature concerning denotified communities.

¹⁰⁶ An optimiser does not distinguish between regularities which are causally meaningful and regularities which are artefacts of the process that generated the data. It is indifferent to provenance by construction.

¹⁰⁷ India Const. art. 15(1); *State of Bombay v. Bombay Education Society*, A.I.R. 1954 S.C. 561, 566 (India).

¹⁰⁸ The phenomenon is known in the technical literature as redundant encoding, and it is the principal reason why removing a protected attribute from the feature set does not remove its influence from the model. See Barocas & Selbst, *supra* note 3, at 691–92, 720–22.

¹⁰⁹ It should be acknowledged that this is a rhetorical convergence rather than a logical entailment. The State’s inability to state the differentia follows from the architecture of the model, not from the content of the classification. The convergence is nonetheless what makes the remedy apposite.

Repeal is ordinarily understood as an event which extinguishes a law. In the case of a classificatory statute which operated through records, repeal is incomplete. The statute's operative content survives within the data which it caused to be created, and that data continues to circulate, to be copied into successive systems, and now to be employed as training material. The phenomenon may usefully be termed *statutory residue*. The residue is inert for so long as records are consulted individually by officers who must state a reason for acting upon them. It becomes active at the moment records are used to fit a function, for fitting is an operation which extracts precisely the regularities which the repealed statute inscribed.¹¹⁰

If that is correct, the constitutional duty attaching to repeal is larger than has been assumed. It is not discharged by deleting the text. Where a repealed classification has been inscribed within administrative records, a State which subsequently mines those records for classificatory purposes is exercising the repealed power by other means. The obligation is one of decontamination: an audit of provenance, the exclusion of the affected strata, and a burden upon the State to demonstrate that the residue has been removed before the corpus is employed. The Supreme Court in *Sukanya Shantha v. Union of India* directed that the category be cleansed from the manuals.¹¹¹ It did not reach the records, because the case as presented did not raise them. The next case will.¹¹²

VII. The Location of the Constitutional Defect

If the argument thus far is sound, something is constitutionally amiss in the deployment of an unstateable classifier for coercive purposes by the State. It matters where the defect is located, for the location determines the remedy.

A. Three Candidate Characterisations

The classification is void. Upon this characterisation the defect inheres in the classification itself: there exists no intelligible differentia, the first limb of the classification test therefore

¹¹⁰ This suggests a general principle of which the present case is an instance: the legal significance of a body of records changes when the mode of consulting them changes, and a repeal which was adequate to one mode of consultation may be inadequate to another.

¹¹¹ *Sukanya Shantha v. Union of India*, 2024 INSC 753, ¶ 220 (India).

¹¹² The writ petition presently before the Supreme Court concerning biometric surveillance at sites of protest may furnish the occasion, though the pleadings as reported are framed in terms of privacy and data protection rather than equality.

fails, and the enquiry is at an end. The remedy is the invalidation of the classification and of any action founded upon it. This is the cleanest characterisation and it follows the reasoning in *State of West Bengal v. Anwar Ali Sarkar* most directly.¹¹³

The delegation is void. Upon this characterisation the defect consists in the exercise of statutory power by a system upon which the legislature conferred none, the officer who adopts the output having formed no satisfaction of the kind the statute requires.¹¹⁴ The remedy is narrower: the deployment is ultra vires the enabling provision, which leaves open the possibility of express legislative authorisation.¹¹⁵

The failure to give a reason constitutes the violation. Upon this characterisation what is void is the particular act, whether a listing, a stop or a refusal, because it was taken without a determining principle capable of statement. The remedy is the quashing of individual acts, the system being left standing.

B. The Preferred Characterisation and its Consequences

The first characterisation is correct, and the second and third are consequences of it rather than alternatives to it.

The reason is that the defect is not contingent upon any particular application. A statute conferring a power of selection without stating a principle is void upon the day of its enactment, whether or not the power has yet been exercised; that is the holding in *State of West Bengal v. Anwar Ali Sarkar* and it has never been doubted.¹¹⁶ An unstateable classifier occupies the same position. Its incapacity to yield a differentia is a property of the object rather than of its outputs, and it is present before the first person has been classified. To locate the defect within the individual decision would be to make the constitutional question depend upon whether a particular claimant can be found and is able to litigate, which, given

¹¹³ It has the further advantage of being available at the instance of a petitioner who has not himself been classified, since the challenge is to the deployment rather than to its application.

¹¹⁴ *Barium Chemicals Limited v. Company Law Board*, A.I.R. 1967 S.C. 295, 322–24 (India).

¹¹⁵ Which is not a trivial qualification. A legislature which expressly authorised deployment would still confront the equality objection, but the delegation objection would fall away.

¹¹⁶ *State of West Bengal v. Anwar Ali Sarkar*, A.I.R. 1952 S.C. 75, 83–85 (India).

that the subjects of police classification are not notified, is a condition which almost nobody will satisfy.¹¹⁷

The delegation point survives as a second and independent ground, and it is worth pleading separately for a practical reason: it bites even where the State *is* able to state the differentia but the officer has not in fact applied his mind to it. The prohibition imposed by the Conseil constitutionnel upon self-revising algorithms in individual administrative decisions is best understood in these terms.¹¹⁸

C. Comparative Approaches and their Limited Application

There exists an instructive foreign instance of a court arriving at the correct result by what is, for Indian purposes, the wrong route.

The Federal Constitutional Court of Germany struck down the provisions of the Hesse and Hamburg police statutes which authorised automated analysis of lawfully held police data, holding that such analysis constitutes a fresh interference with the right to informational self-determination distinct from the original collection, and that the more far-reaching and opaque the analytical method, the stricter the threshold which must be satisfied. The judgment refers expressly to data mining and to self-learning systems, and observes that such methods may generate new information concerning the personality of the persons affected which would otherwise be inaccessible.¹¹⁹ The Court thus treated opacity as an aggravating factor operating within the proportionality enquiry, which escalates the burden of justification without ever rendering it impossible.¹²⁰

That is the natural analytical move in a jurisdiction whose equality guarantee contains no requirement of legibility. India need not make it. Here opacity is not a factor which raises the price of justification; it is a failure of the first limb of the classification test, and there exists no price at which it may be purchased. The distinction is not merely of theoretical interest. A

¹¹⁷ The problem is familiar from surveillance litigation generally, where the absence of notification defeats standing. See *Clapper v. Amnesty International USA*, 568 U.S. 398, 409–14 (2013), for the analogous American difficulty.

¹¹⁸ Conseil constitutionnel [CC] [Constitutional Council] decision No. 2018-765DC, June 12, 2018, cons. 71 (Fr.).

¹¹⁹ Bundesverfassungsgericht [BVerfG] [Federal Constitutional Court] Feb. 16, 2023, 1 BvR 1547/19, 1 BvR 2634/20, ¶¶ 65–73, 100–12 (Ger.).

¹²⁰ The German court did require that the statutes narrowly limit the categories of data and the methods of analysis permitted, which is a legibility requirement in substance; but it is imposed upon the legislature as a matter of statutory quality rather than upon the executive as a condition of deployment.

proportionality framework invites the State to argue that the gravity of the threat justifies the opacity of the means, an argument which will frequently succeed in the security context. A framework founded upon legibility admits no such argument, for the question is not how important the object is but whether the ground has been stated.

The Dutch litigation concerning the System Risk Indication is to similar effect and suffers from the same limitation. The District Court of The Hague held the enabling legislation incompatible with Article 8 of the European Convention on Human Rights, principally for want of transparency concerning the risk model and the indicators employed.¹²¹ Transparency, as Part III.B has argued, is not the same as legibility. A fully disclosed model whose decision boundary remains unstateable would satisfy the Dutch requirement and fail the Indian one.¹²²

VIII. Objections and Replies

A. The Opacity of Human Decision-Making

The most frequently advanced objection, and the most interesting, is that human decision-making is itself opaque. Officers act upon intuition. Judges rely upon impressions which they could not fully articulate. If unarticulable grounds vitiate a classification, the objection runs, the greater part of State action falls with it.

Two replies may be made. The first is that the law has never accepted this as an accurate description of what it licenses. It licenses discretion exercised for stated reasons, and where reasons are absent it supplies remedies: the duty to record reasons, review for non-application of mind, and the bar upon supplementing an order by affidavit.¹²³ That these remedies are imperfectly enforced is an argument for enforcing them rather than for extending the tolerance to a new class of decisions.¹²⁴

¹²¹ Rechtbank Den Haag [Rb.] [District Court of The Hague] 5 februari 2020, ECLI:NL:RBDHA:2020:1878, ¶¶ 6.49–6.95 (Nederlands Juristen Comité voor de Mensenrechten v. The Netherlands) (Neth.).

¹²² That is a claim about the reasoning rather than about the result. It is likely that the Dutch court, confronted with such a model, would have found other grounds for the same conclusion.

¹²³ *S.N. Mukherjee v. Union of India*, (1990) 4 S.C.C. 594, 612–14 (India); *Mohinder Singh Gill v. Chief Election Commissioner*, (1978) 1 S.C.C. 405, 417–18 (India).

¹²⁴ The objection, pressed to its conclusion, would establish that no reason-giving requirement can be justified anywhere, since human reasons are always incompletely articulated. That conclusion is not one the objector wishes to embrace.

The second reply is structural, and it is the reply which does the work. A human officer's reason is in principle capable of being stated and in fact capable of being contested. Pressed by a court, he must produce a ground; and once produced, the ground may be tested, for it is either a permissible basis or it is not, and if he has none the decision falls. The opacity is epistemic and it is remediable. The opacity of a fitted model is neither. No degree of judicial pressure will produce a differentia, because there exists none to produce. The difference between a reason which was not given and a reason which does not exist is the whole of the matter.

There is in addition an asymmetry of scale which ought not to be waved away. The unarticulated intuition of a single officer affects the persons whom that officer encounters. A deployed model applies the same unarticulated boundary across an entire jurisdiction, uniformly and simultaneously; and because officers characteristically defer to machine outputs, its errors are unlikely to be corrected by the humans nominally placed in the loop.¹²⁵

B. The Historical Under-Enforcement of the Requirement

It may be said that the requirement of an intelligible differentia has operated for seventy years as a formality which courts have not enforced, and that an argument resting upon it is therefore an argument resting upon a dead letter.

The objection is partly well founded and the concession should be made frankly. The classification test has frequently been applied with considerable deference, and the adjective has functioned as a recital rather than as a requirement. But a requirement which has not been enforced is not a requirement which has been abandoned, and the constitutional history of India furnishes numerous instances of dormant text becoming operative when conditions altered, the transformation of Article 21 following *Maneka Gandhi v. Union of India* being the most conspicuous.¹²⁶ The dormancy in the present instance has an identifiable cause: until now, every classification presented for review was already stated, so the limb concerned with legibility had nothing to do. The conditions have altered.

¹²⁵ Linda J. Skitka, Kathleen L. Mosier & Mark Burdick, *Does Automation Bias Decision-Making?*, 51 INT'L J. HUM.-COMPUTER STUD. 991 (1999); Raja Parasuraman & Victor Riley, *Humans and Automation: Use, Misuse, Disuse, Abuse*, 39 HUM. FACTORS 230 (1997). See also Ben Green, *The Flaws of Policies Requiring Human Oversight of Government Algorithms*, 45 COMPUTER L. & SEC. REV. 105681 (2022).

¹²⁶ *Maneka Gandhi v. Union of India*, (1978) 1 S.C.C. 248 (India). The doctrine of the basic structure furnishes a second instance, having lain implicit in the text until the conditions of 1973 required it. *Kesavananda Bharati v. State of Kerala*, (1973) 4 S.C.C. 225 (India).

C. Security, Confidentiality and Administrative Feasibility

The State will contend that disclosure of its criteria defeats the purpose of the exercise, in that a stated rule may be gamed by those whom it is designed to identify.

The reply is that legibility and publication are distinct requirements, and that the objection conflates them. What Article 14 demands is that a differentia exist and be capable of statement to the court and, in an appropriate form, to the person affected. Whether it must be published to the world at large is a separate question upon which procedures for inspection in camera, for redaction, and for disclosure to a special counsel all have a role to play. The State which says that it cannot publish its criteria is advancing an argument. The State which says that it has no criteria to publish is conceding the case.¹²⁷

D. The Anticipated Improvement of Explainability Techniques

Finally, it may be said that techniques of explanation will improve, and that an argument founded upon their present limitations will age badly.

The argument does not depend upon the present state of the art, and it is important to see why. The objection to post hoc explanation is not that the approximations are presently poor. It is that an approximation of a model's behaviour, however faithful, is a different thing from the ground upon which a decision was taken, and a court conducting review for nexus requires the second.¹²⁸ Improvements in fidelity narrow the gap between the model and its surrogate. They do not convert a surrogate into a reason. Should the research programme eventually yield models whose actual decision rules are capable of statement, those models will satisfy Article 14; and they will do so because they are interpretable, not because they are well explained.¹²⁹

¹²⁷ Procedures for disclosure to a court in camera are familiar in Indian practice, most recently in the context of sealed cover procedure, which has itself attracted criticism for the asymmetry it creates. The criticism is well founded and is not answered here, save to note that the choice is between imperfect disclosure and none.

¹²⁸ Selbst & Barocas, *supra* note 47, at 1109–18.

¹²⁹ Which is to say that the argument is not a prediction about the trajectory of a research programme. It is a claim about what would satisfy a constitutional requirement, and it would be satisfied by any model meeting the condition, whenever such models become available.

IX. Towards a Remedial Framework

A doctrinal argument which concludes without a remedy is incomplete. What follows is an outline rather than a draft.

A. Statability as a Condition Precedent to Deployment

Where the State deploys an automated system to classify persons for a coercive purpose, whether in policing, listing, surveillance, detention or the denial of a benefit, it must be in a position to state, before deployment, the criterion which the system applies, in terms which a court is able to test for nexus and which the person affected is able to contest. Inability to state the criterion is not a procedural shortfall capable of being cured by an affidavit; it constitutes the violation.¹³⁰

B. The Allocation of the Evidentiary Burden

The claimant cannot know that she was classified, cannot know which features were employed, and cannot compute the distribution of outcomes. Once she establishes that a classificatory system was applied to her, the burden ought to shift to the State to establish legibility and nexus. This is not novel in Indian law. The Supreme Court has shifted evidentiary burdens in equality litigation where the relevant material lies within the exclusive knowledge of the State,¹³¹ and the principle embodied in section 106 of the Indian Evidence Act, 1872, now section 109 of the Bharatiya Sakshya Adhiniyam, 2023, points in the same direction.¹³²

C. A Duty of Decontamination Attaching to Training Data

Where a classificatory statute has been repealed and its operation recorded, the State ought to bear a duty, before employing the resulting records to fit a classifier, to audit provenance and to exclude the strata generated under the repealed regime. Applied to the present subject, that means that records deriving from the registers maintained under the Criminal Tribes Act, 1871, and from the habitual offender and surveillance registers which inherited them, cannot

¹³⁰ The formulation deliberately tracks the language of condition precedent rather than of procedural requirement, in order to foreclose the argument that the defect is curable by subsequent compliance.

¹³¹ *Anuj Garg v. Hotel Association of India*, (2008) 3 S.C.C. 1, 14–16 (India); *Nikesh Tarachand Shah v. Union of India*, (2018) 11 S.C.C. 1, 45–50 (India).

¹³² Indian Evidence Act, No. 1 of 1872, § 106 (India); Bharatiya Sakshya Adhiniyam, No. 47 of 2023, § 109 (India).

be treated as neutral inputs. The direction issued in *Sukanya Shantha v. Union of India* is the beginning of this obligation and ought to be extended from manuals to data.¹³³

D. Interpretability by Design in Coercive Classification

Because the cost in accuracy of employing interpretable models upon structured administrative data is small or non-existent, the State ought to be required to employ such models wherever classification carries coercive consequences.¹³⁴ This is not a general regulation of artificial intelligence. It is a constitutional constraint upon a specific governmental function, and it follows from the premise that the exercise of coercive power requires a reason.¹³⁵

E. The Inadequacy of the Digital Personal Data Protection Act, 2023

It ought to be stated plainly that the Digital Personal Data Protection Act, 2023 does none of this. It contains no analogue to Article 22 of the General Data Protection Regulation, no right to human review, no right to an explanation, and no provision addressing profiling; and it empowers the Central Government to exempt instrumentalities of the State from its provisions upon notification in the interests of sovereignty, security, public order and cognate grounds.¹³⁶ The subordinate legislation since notified confirms rather than cures this deficit. The Digital Personal Data Protection Rules, 2025, notified by the Ministry of Electronics and Information Technology on 14 November 2025 after an extended public consultation, operationalise the Act's consent, breach-notification and data-retention provisions but introduce nothing resembling a right to an explanation, a right to human review, or a constraint upon profiling by State instrumentalities.¹³⁷ Whatever protection exists in India against algorithmic classification by the State is constitutional in character, or it does not exist at all. That is the reason for pressing Article 14 rather than awaiting legislation.

¹³³ *Sukanya Shantha v. Union of India*, 2024 INSC 753, ¶ 220 (India).

¹³⁴ Rudin, *supra* note 48, at 207–09.

¹³⁵ The limitation to coercive classification is important. Nothing in the argument prevents the State from using opaque models for purposes which do not involve the classification of persons for adverse treatment, such as the prediction of equipment failure or the optimisation of logistics.

¹³⁶ Digital Personal Data Protection Act, No. 22 of 2023, §§ 7(g)–(i), 17(2)(a) (India); *compare* Regulation 2016/679, art. 22, 2016 O.J. (L 119) 1, 46 (EU).

¹³⁷ ‘Digital Personal Data Protection Rules, 2025 Notified’ (Mondaq, November 2025) <<https://www.mondaq.com/india/data-protection/1708164/digital-personal-data-protection-rules-2025-notified>> accessed 23 September 2026.

X. Conclusion

The literature upon algorithmic accountability has largely assumed that India lacks the doctrinal equipment appropriate to the problem: no Article 22, no tradition of disparate impact, no administrative law duty to explain the outputs of machines. The assumption has produced a scholarship of deficit, within which the task is conceived as one of importing standards from elsewhere and arguing for their reception.

The equipment is already present. It is located within the first limb of the classification test, in an adjective which has been recited for seventy years without ever being required to do anything. Intelligibility is a demand that the ground of State differentiation be capable of statement. A learned classifier is unable to satisfy that demand; not because it is novel, nor because it is complicated, but because the object which it produces is a fitted surface rather than a rule, and a surface cannot be stated as a proposition concerning persons.

What renders the Indian case urgent rather than merely interesting is the character of the material to which those surfaces are being fitted. The records are not neutral. They are the sediment of a statute which classified persons by birth, which was repealed in 1952, and which was never removed from the files. To fit a model to such records is to perform an operation which recovers the classification and launders it, for the output carries no ground to which Article 15 might attach. The State is then able to say, truthfully, that it did not classify by caste. It did not. It classified by something which it cannot name, and which turned out to be caste.

Article 14 of the Constitution of India has an answer to that, and the answer requires the claimant to prove nothing whatever concerning outcomes. It requires the State to say what its differentia is. Where the State is unable to say, the classification fails, and it fails at the threshold. That is not a new doctrine. It is an old one, presented at last with the case for which it was written.