Jus Corpus Law Journal

Open Access Law Journal – Copyright © 2026 – ISSN 2582-7820

Editor-in-Chief – Prof. (Dr.) Rhishikesh Dave; Publisher – Ayush Pandey

Algorithmic Caste Enumeration: Constitutional Validity of Digital Census 2027's GIS-Based Self Count Algorithm

Midhila Jayan[a]

[a]National Law School of India University, Bangalore, India

Received 20 May 2026; *Accepted* 20 June 2026; *Published* 24 June 2026

India's 2027 census marks the first fully digital caste-based enumeration exercise since 1931. The census of 2027 incorporates a new inclusion of Geographic Information System (GIS)-based self-count algorithm to carry out the survey for millions of households through mobile app-based data collection, geographic information system geotagging, and machine learning validation checks under the Census Act, 1948. The use of such a technological system raises various constitutional concerns with regard to the intersection of rights, such as the right to privacy, the right to equality and the governance of algorithms. In this paper, the constitutionality of such an algorithmic architecture is tested against Article 14(equality) and Article 21(privacy) in light of the Supreme Court's decision in K.S. Puttaswamy v Union of India (2017). By employing the methodology of doctrinal legal analysis, the paper analyses whether the caste enumeration on the basis of GIS could meet the triple test of legality, necessity and proportionality as enunciated by the Puttaswamy judgment. Moreover, the analysis is extended to the extent of whether the location-based caste map is in violation of Article 14's prohibition against arbitrary classification and Article 15(1)'s prohibitory rule against caste-based discrimination. The paper also provides for a comparative analysis of frameworks such as the Digital Personal Data Protection Act 2023 and the EU General Data Protection Regulation, to identify the lacunas in India's digital census system. The paper also identifies certain critical constitutional challenges, such as GIS geotagging violating Article 14 through location-based discrimination, and algorithmic self-count mechanisms infringing the informational privacy principle. The paper concludes with certain legislative recommendations, which include establishing the National Digital Census Tribunal, obligatory algorithmic audits, and opt-out GIS provisions, therefore, allowing the continuation of this system based on algorithmic constitutionalism while addressing its challenges in India's first digital caste census constitutional framework.

Keywords: *census 2027, algorithmic cate enumeration, article 21, article 14, data protection.*

INTRODUCTION

The 1931 census of India was the last census in which a comprehensive caste-based enumeration occurred, which took place under the British colonial administration.[1] Ever since independence, Indian censuses have only documented the Scheduled Castes (SC) and the Scheduled Tribes (ST), without including the Other Backwards Classes (OBC) and the general caste.[2] It was the Union Cabinet approval of Census 2027 on December 12, 2025, which heralded the dawn of India's first completely digital caste-based census after 96 years, which is aimed at enumerating around 267 million households with mobile app-based data collection, GIS-based geotagging of data and a self-count algorithm for the disclosure of castes.[3]

The 2027 census comprises a three-tiered technical framework. The first level involves a mobile app-based data collection system for households to disclose caste information using a government-encrypted mobile application, which is equipped with Aadhaar-linked biometrics and real-time geolocations. The second tier involves GIS-based geo-tagging, which provides coordinates for each household's caste disclosures, building a 'caste map' at the village or ward level. The third tier involves algorithm processing of the self-counted caste disclosures and flags inconsistencies with the help of machine learning models trained on 2011 census data to ensure that anomalies in caste disclosures are manually verified.

This algorithmic caste-based enumeration system raises serious constitutional questions. The triple test established under Article 21 in the K.S. Puttaswamy v Union of India case of 2017 requires that the collection of personal data should be lawful, necessary and proportional.[4] At the same time, Article 14 forbids arbitrary state action and provides for substantive

[1] 'The last caste census was in 1931. A look back at its findings' (*Vision IAS*, 01 May 2025) <https://visionias.in/current-affairs/upsc-daily-news-summary/article/2025-05-01/the-indian-express/polity-and-governance/the-last-caste-census-was-in-1931-a-look-back-at-its-findings> accessed 08 May 2026

[2] 'Census 2027 UPSC: India's First Digital Census and Caste Data Debate' (*Textbook*) <https://testbook.com/ias-preparation/census-2027-digital-caste-data-self-enumeration> accessed 08 May 2026

[3] 'Census 2027: India's First Digital Enumeration Exercise' (*Press Information Bureau*, 25 April 2026) <https://www.pib.gov.in/PressReleasePage.aspx?PRID=2255461&reg=48&lang=2> accessed 08 May 2026

[4] *Justice K S Puttaswamy (Retd) & Anr v Union of India & Ors* (2017) 10 SCC 1

equality, and Article 15 specifically protects against discrimination on the grounds of caste.[5] The algorithmic caste enumeration system violates the above mentioned constitutional principles due to mandated disclosures of private caste identities with limited privacy safeguards, the provision of local caste data that would enable localised caste-based discrimination, and the possibilities of biases within the automatic-validation algorithm, which could disproportionately affect the marginalised castes.

DOCTRINAL FRAMEWORK: CENSUS ACT 1948 AND CONSTITUTIONAL PRINCIPLE OF PRIVACY

The Census Act 1948 provides legislative support for the 2027 Census. Section 3 of the act allows the central government to carry out the census operations, and Section 4 requires the appointment of census officers.[6] The confidential rights provided in section 6(1) state that 'No individual census record shall be disclosed except for census purposes.'[7] However, Section 6(1) was framed before the advent of the Digital Age and does not address encryption standards of digital data, algorithmic processing of data based on caste, right of an individual for data correction and deletion, thereby creating an ample scope for violation of privacy rights under Article 21.

The nine-judge constitutional bench in K.S. Puttaswamy v Union of India (2017) held that the 'right to privacy is a fundamental right' which is protected under Article 21 of the Constitution as part of the right to life and personal liberty.[8] A three-tier test of legality, necessity and proportionality in all state actions that impinge upon privacy interests was proposed in this judgment. Firstly, 'legality' demands that there must be 'existing law' in place and not a decision by the executive for a state to infringe on an individual's privacy. Secondly, 'necessity' demands that there should be a 'legitimate state aim' in infringing privacy, that it should be 'necessary in a democratic society'. Thirdly, 'proportionality' demands that there should be 'proportionality' in the invasion of privacy and that the state should use the least restrictive method possible to fulfil its aim. These three principles must be satisfied by the Census 2027 for it to be constitutional.

[5] Constitution of India 1950, arts 14 and 15
[6] Census Act 1948, ss 3 and 4
[7] Census Act 1948, s 6(1)
[8] *Justice K S Puttaswamy (Retd) & Anr v Union of India & Ors* (2017) 10 SCC 1

The first tier of the test is 'legality'. The Census Act of 1948 allows the census to be conducted, and thus it qualifies as an 'existing law', fulfilling the legality test.[9] However, since section 6(1) only mentions confidentiality rights within the context of the pre-digital era, these rights may not qualify as 'law' when processing digital data, as mentioned in the K.S. Puttaswamy ruling. The court ruled that 'legislation that infringes upon privacy has to satisfy specific requirements, that is, a legal sanction, that the objective for which it seeks to infringe must be a legitimate one, and there must be proportionality in such infringement.' The Census 2027 has no specified legal authority to perform GIS geotagging of caste data, machine learning validation checks on data, cross reference data with the Aadhaar database.

The second tier of the test is 'necessity'. The Census 2027 aims to formulate an 'evidence-based affirmative action policy' by collecting data on caste. However, in order for the use of GIS geotagging for this purpose to be constitutional, it has to be proven that it is 'strictly necessary' for achieving this purpose. There are various alternatives for caste data collection, such as collecting caste data that is not geotagged, obtaining aggregate statistics of caste at the district level and anonymising the caste data for policy-making purposes, etc. The third tier of the test, 'proportionality', requires a balance between the 'individual privacy rights' and the 'state's interest'. However, geotagging caste data may lead to an infringement of individual privacy rights by enabling profiling and discriminatory action by private actors and data breaches that would reveal the caste status of millions of households to potential harassment based on caste.

ARTICLE 14: EQUALITY: ANALYSIS OF GIS GEOTAGGING AND LOCATION-BASED DISCRIMINATION

Article 14 of the Indian Constitution deals with ‘equality before law’ and ‘equal protection of laws’.[10] The Supreme Court of India, in construing Article 14, has taken it to mean the prohibition of 'arbitrary state action' and the upholding of 'substantive equality'.[11] In E. P Royappa v State of Tamil Nadu (1974), it was argued that 'equality is a dynamic concept with many aspects and dimensions' that cannot be confined within the ‘traditional doctrinal classifications’. Applying the 'arbitrariness doctrine' in Ajay Hasia v Khalid Mujib Sehravardi

[9] Census Act 1948
[10] The Constitution of India 1950, art 14
[11] *E P Royappa v State of Tamil Nadu & Anr* AIR 1974 SC 555

(1981), the Apex Court has banned 'arbitrary, irrational or manifestly arbitrary' state action.[12] Geotagging of caste using GIS may be a violation of Article 14 since it leads to an arbitrary categorisation of persons by placing them at particular locations.

By geotagging caste data, it is possible to classify populations based on location (village/ward boundaries, neighbourhoods, spatial arrangements of caste groups). This allows the classification of people based on their neighbourhood location and has resulted in 'caste hotspots,' where specific caste groups are concentrated. The constitutional issue with the use of these maps lies with 'discriminatory targeting', that is, when these maps, which contain location and caste data, can be accessed by individuals and private organisations like employers, landlords and colleges, they can be used to discriminate against people based on caste. For example, the landlord could find out whether the people belonging to certain castes are residing in a specific area, then they could restrict membership in that particular area to only people belonging to some castes.

Discrimination on grounds only of religion, race, caste, sex, place of birth or any of them is prohibited under Article 15 of the Indian Constitution.[13] In State of Madras v Champakam Dorairajan (1951), it was held that Article 15(1) creates an 'absolute prohibition' on any kind of caste-based discrimination.[14] The issue with the GIS geotagging of caste data lies in the 'instrumentalisation of caste data'. The state has collected the information and made caste maps easily accessible, and in doing so has allowed for and facilitated private caste discrimination, which Article 15(1) seeks to prohibit. This is a violation of the state's affirmative duty under Article 15(3) to protect against caste discrimination.

About privacy protections, the EU General Data Protection Regulation (GDPR) imposes stringent measures with regard to location data. Article 9 states that processing of 'special category data' that reveals their 'racial or ethnic origin' is prohibited.[15] Recital 51 explicitly states that 'geolocation data' falls under sensitive personal data requiring 'explicit consent', which means that explicit consent has to be obtained before this data can be processed.[16] This mandatory processing of location data under the Census 2027 without the required explicit

[12] *Ajay Hasia Etc v Khalid Mujib Sehravardi & Ors Etc* (1981) 1 SCC 722
[13] Constitution of India 1950, art 15(1)
[14] *State of Madras v Srimathi Champakam Dorairajan & Ors* AIR 1951 SC 226
[15] General Data Protection Regulation (EU) 2016, art 9
[16] General Data Protection Regulation (EU) 2016, recital 51

consent is a violation of equivalent protections under Article 14's principle of integration with international law.[17]

DPDP ACT 2023 CONFLICT: CENSUS CONFIDENTIALITY V ERASURE RIGHTS

The Digital Personal Data Protection Act 2023 lays down data protection rights. Section 8(5)(a) of the act gives data principals the right to erasure, where it states that the data principal shall have the right to seek erasure of personal data when it is no longer necessary for the purpose for which it was collected.[18] Section 4 of the act lays down 'data fiduciary obligations' for processing personal data, which includes the obligation to implement 'reasonable security safeguards' for personal data, and section 10 provides for the 'Data Protection Board of India' for data protection grievance redressal.[19]

On the other hand, the Census Act Section 6(1) requires data to be retained for a '100-year duration' and that the data cannot be revealed for any purpose other than 'census purposes'.[20] This provision is in direct conflict with the right of erasure in Section 8(5)(a) of the DPDP Act. This conflict requires 'statutory harmonisation' and can be resolved by amendment to the Census Act for it to be in conformity with DPDP Act provisions, or a DPDP Act overriding provision for census data.

Data principals' sensitive personal data is defined in DPDP Act Section 2(o) as data revealing 'racial or ethnic origin, religious beliefs or caste'.[21] Caste data, hence, falls under sensitive personal data and requires further protection. DPDP Act Section 5(1) requires 'free, informed, unconditional, and specific' consent to be obtained to process sensitive personal data.[22] The self-count algorithm for the 2027 Census is compulsory, with no 'opt-out', meaning explicit consent has not been obtained. DPDP Act Section 5(2) provides that personal data should be processed for 'specified purposes' only.[23] This means that GIS geotagging for 'policy targeting' may not be in the scope of 'census enumeration', which violates this provision.[24]

---

[17] *Justice K S Puttaswamy (Retd) & Anr v Union of India & Ors* (2017) 10 SCC 1
[18] Digital Personal Data Protection Act 2023, s 8(5)(a)
[19] Digital Personal Data Protection Act 2023, ss 4 & 10
[20] Census Act 1948, s 6(1)
[21] Digital Personal Data Protection Act 2023, s 2(o)
[22] Digital Personal Data Protection Act 2023, s 5(1)
[23] Digital Personal Data Protection Act 2023, s 5(2)
[24] Digital Personal Data Protection Act 2023, s 6

ALGORITHMIC BIAS AND MACHINE LEARNING VALIDATION CHALLENGES

The self-count algorithm for Census 2027 uses machine learning (ML) models trained on 2011 census data to cross-check self-disclosed castes. 'Algorithmic bias' is thus established in the algorithm through historical bias, representational bias and feedback loops. Historical bias stems from the 2011 data having undercounted disadvantaged castes because of societal stigma. The ML models, trained on 2011 data, end up replicating and reinforcing these historical undercounts. Representational bias happens because models trained on data with lower proportions of young castes, urban Dalits, and inter-caste households will systematically fail to detect caste disclosures from such populations as valid, therefore, leading them to be disproportionately affected.

The 'arbitrariness doctrine' makes the exercise of 'unfettered discretionary power' illegal where no 'guiding principles' are present.[25] The Census 2027 algorithm has used 'black box' machine learning models that lack a clear standard for validating data. The algorithm, thus, violates Article 14 in terms of lack of transparency, lack of accountability and lack of remedy. A household has no understanding as to why its disclosure is marked 'suspicious', and no provision is made to appeal against any error.

On the other hand, under the EU AI Act 2024, 'AI systems used for law enforcement, border control or employment' are declared 'high risk' and, thus, need to meet risk assessment requirements, human oversight and transparency.[26] Mandate to conduct AI impact assessment before deploying systems requires human review of algorithmic decisions that impact fundamental rights, and requires explanation of the rationale behind the algorithmic output. The algorithm under Census 2027 does not meet these safeguards, therefore violating the due process of law.

REFORM RECOMMENDATIONS: CONSTITUTIONAL COMPLIANCE FRAMEWORK

Census 2027 should enforce Algorithmic Caste Impact Assessment (ACIA), which must open-source the machine learning (ML) validation code, have it audited by expert technologists, and certified by a newly established Caste Algorithm Ethics Board, before deployment. This addresses the arbitrariness under Article 14 by replacing non-transparent

[25] *Ajay Hasia Etc v Khalid Mujib Sehravardi & Ors Etc* (1981) 1 SCC 722
[26] Artificial Intelligence Act 2024

'black box' algorithms with community-empowered and auditable algorithmic governance, allowing for independent scrutiny of biases.

In place of compulsory GIS geotagging, Census 2027 should use opt-in consent registered on an immutable blockchain ledger; smart contracts should automatically delete data after 10 years and provide cryptographic proof of deletion. Households must have the right to revoke access, view data accessors, and receive notification of data usage by the state, complying with the explicit consent mandated in Section 5(1) DPDP Act and mitigating risks under Article 14 concerning location-based discrimination arising from caste hotspots.

A 3-tiered temporal anonymisation regime using quantum encryption is required: Tier 1 (0-5 years) with complete quantum encryption, Tier 2 (5-10 years) with district-level anonymisation, and Tier 3 (10+ years) with cryptographic deletion. This resolves the 100-year retention under the Census Act and the storage limitations of the DPDP Act, while also mitigating risks from future quantum computing.

Caste Data Trusts at the district-level, with 50% representatives from affected castes, must have veto powers over any government request for access to caste data, shifting ownership from state property to community property to fulfil Article 15(3) affirmative duties and rectifying procedural arbitrariness under Article 14.

Post-Census, individual caste data shall be transferred to Digital Caste Archives with tiered access, where individuals retain ownership of past data, researchers will need ethical clearance, and corporations will be barred from access, while the government will only have access upon trust approval. This approach balances privacy rights under Article 21 with academic interests under Article 19(1)(a).

CONCLUSION

Census 2027's framework for algorithmic caste enumeration is a constitutional inflexion for digital rights laws in India. Doctrinal analysis of this paper suggests that the GIS-based self-count algorithms would violate privacy under Article 21 (as under Puttaswamy's triple test), equality under Article 14 (based on locational discrimination), and the DPDP Act 2023's erasure rights. Lack of mechanisms for algorithmic audit and dedicated tribunals creates

accountability deficits contrary to due process rights guaranteed by the Constitution. The recommended reforms provide actionable pathways for constitutional compliance.