Collecting Data, Constructing the Criminal:

Connecting Contemporary Datafication Practices To The Casteist and Colonial History of Biometric Datafication in 19th and 20th Century India

NIKITA SONAVANE (CPA PROJECT),
DR. KERRY MCINERNEY (UNIVERSITY OF CAMBRIDGE), AND
NIKHIL DHARMARAJ (UNIVERSITY OF CAMBRIDGE)

aapti institute
DataPrivacyBR Research
PARADIGM INITIATIVE

DATAFICATION AND DEMOCRACY FUND

Final Narrative Report

Project title	Collecting Data, Constructing the Criminal: Connecting Contemporary Datafication Practices To The Casteist and Colonial History of Biometric Datafication in 19th and 20th Century India
Grant duration	from April 15, 2024 to October 15, 2024
Reporting period	from June 30, 2024, to October 15, 2024
Prepared by	Nikita Sonavane (CPA Project), Dr. Kerry McInerney (University of Cambridge), and Nikhil Dharmaraj (University of Cambridge)
Contact email	nikitasonavane5@gmail.com

In 2009, India launched Aadhar, the world's largest biometric identification system. The system was praised by the World Bank and the International Monetary Fund, despite notable concerns over Aadhar's privacy and security issues, and the possibility that the system would be used as a mechanism for state surveillance and control. The Indian state now exports Aadhar around the world, most notably to Kenya as well as other parts of East Africa. Indeed, over the last decade, organizations such as ID4Africa, the World Bank, and more have consistently looked to India as an exemplar for replicating national biometric platforms in Sub-Saharan Africa. Aadhar was closely followed with the launch of the Crime and Criminal Tracking Network & Systems (CCTNS) which aimed to integrate all the data and records of crime by the police. With CCTNS, the State Police departments across the country found an opportunity to digitise existing data, link it to a common database accessible across the state, and use this dataset to trial new forms of 'predictive policing'. A decade later, an upgrade on the CCTNS was envisioned in the form of the Interoperable Criminal Justice System (ICJS), which will integrate existing centralised databases the CCTNS, e-prisons, and e-courts. In doing so, it promises "seamless exchange of live data" among these branches and to inaugurate a new age of "SMART policing".

The expansion of digitalised police powers is especially concerning given the recent passing of the Criminal Procedure Identification Act (CPIA) in 2022. The CPIA replaced an earlier piece of colonial legislation called the Identification of Prisoners Act 1920 (IPA), which allowed for the collection of biometric data through methods such as photographing, fingerprinting, and other forms of bodily measurement. The CPIA empowers the police to collect biometric data including DNA from accused, witnesses or any other person and store and share this data at a federal level, for 75 years. While existing critiques of the CPIA primarily focus on how the Act may threaten individual privacy rights, this research project aimed to situate the CPIA (and associated technological projects such as Aadhar, the CCTNS, and the ICJS) within a much longer trajectory of datafication practices that began under British colonial rule and, entwined with casteist power relations, extend into the present day. We intended to show how the supposedly neutral category of the 'habitual offender' that lies at the heart of both the IPA and the CPIA is merely a re-invention of the concept of the 'hereditary criminal', which emerged as a key figure in British colonial policing and was codified into Indian criminal law. This figure of the 'hereditary criminal' was, from the start, underpinned by colonial and casteist forms of discrimination.

Therefore, through the research, we sought to explore the following two questions:

1. How did biometric data collection evolve as a means of criminal identification in early 20th century India? In particular, how was biometric datafication used to create criminal categories?

2. How does this history of early biometric datafication influence emerging models of technological surveillance in India today?

METHODS

In order to trace the casteist and colonial lineage of the Identification of Prisoners Act (1920) and understand how it laid the groundwork for contemporary forms of datafication in India, we undertook extensive archival research in Hyderabad, Cambridge, and London. Through the funding provided by the DDF, two of our researchers (Mallika Dhamaraj and Nikita Sonavane) were able to travel to the United Kingdom to access vital historical documents currently sequestered in British archives relating to the colonial governance of India through early biometric practices. We focused on documents from the time period 1890-1950, or the first half of the 20th century. After scanning 70-80 documents, we created a database of 34 files. This database will then be categorised according to; 1) regional focus; 2) decade/time period; 3) file type; 4) author; 5) core themes with the help of our research assistant.

In our archival research, we also worked from the position that historical research is inherently political. Rather than treating the archive as a source of objective truth and the historical process as one of ascertaining neutral facts, we see history as something that is constantly produced and imbued with specific political ideologies and power relations. Historical fact is always entwined with narrative, and through our archival research, we aim to tell a different story about datafication, caste, coloniality and power. To do this, we drew on Umesh Bagade's anti-casteist and non-brahminical approach to historical research. Bagade demonstrates how brahminism shaped the work of British colonial historians, and shows how anti-casteist historians developed new tools and concepts for writing about caste's social history. We follow in this rich tradition of the "non-Brahmin method of history

writing" in how we approach and analyse archival documents throughout this project (Bagade 2012: 4). In addition to Bagade's insights, we also drew on the scholarship of academics such as Saidiya Hartman and Ann Stoler who explore the tensions and the possibilities of reading both with and against the historical archive. Finally, we also took into account our own positionalities and how they might affect our reading of the historical documents we sourced. We discuss some of the initial archival findings below.

DISCUSSION

The existing literature on colonial policing in British India explores how caste shaped the colonial production of the 'criminal' (specifically through the figure of the 'criminal tribe', as established through the Criminal Tribes Act (1871)). Scholars such as Mark Brown, Sagnik Bhattacharya and Radhika Singha trace how British colonial policing practices were legally codified in India through the Thuggee and Dacoity Acts Suppression Acts (1836-1848). These laws were followed in 1871 by the Criminal Tribes Act (CTA), which constituted entire communities (such as caste-oppressed groups, nomadic and indigenous groups, and hijras or third-gender people) as 'habitual criminals'. Tribes and groups labelled as 'criminal' under the CTA were deemed to be 'addicted' to committing non-bailable crimes and labelled 'criminals by birth'. The Act granted the colonial administration sweeping powers to monitor, surveil, and resettle so-called 'criminal tribes'. At this particular time, policing could be broadly conceived as serving two primary functions – repression/response (of/to dissident political activity), and surveillance/prevention (of the so-called habitual criminal's future offense).

These practices of biometric identification laid the groundwork for contemporary datafication practices today; Shivangi Narayan traces how contemporary predictive policing in India is built upon the 18th century colonial criminal registries and their extension through the CTA and later the Code of Criminal Procedure (1973). Casteist discourses about the so-called "habitual criminal" coalesced with a period of great colonial advances in technological innovation to produce what Erin Margaret Giuliani terms a 'surveillance order'. Early forms of datafication and biometric surveillance played a key role in this process. Analog instruments including the fingerprint (as well as its associated schema of indexing and classification formulae), cartography, and more lent a 'scientific' veneer to the policing project of the British Raj that continued on into the twentieth century. For example, Mrinal

Satish foregrounds how criminalised tribes were surveilled by the colonial administration, from being forced to stay in specific settlements, being compelled to carry ID cards, and taking part in thrice-daily roll calls. In 1911, the law was amended to include more extensive regimes of biometric surveillance (such as fingerprinting) and more methodically registering these criminalised communities.

The expansion of biometric data collection was both caused by and reinforced the pseudoscientific belief that criminality was framed as a heritable property that could be identified and located in the face and the body. The regulatory regime that unfolded under British imperial rule in India was built upon the positivist school of criminology that had been enacted in the United Kingdom (as evidenced by Habitual Criminals Act of 1869). Building on the physiognomical and pseudoscientific scholarship of Italian criminologist Cesare Lombroso, the positivist school believed that people were genetically predisposed to criminal behaviour. As Narayan writes, police would record "the facial features and other peculiarities of the persons such as 'appearance, gait, speech etc' so that the supposed criminals may be readily distinguished by their biological features" (Narayan 2021: 121).

Datafication practices (like fingerprinting, photographing, and other biometric data collection techniques) thus played a key role in establishing caste-based hierarchies through the concept of hereditary criminality, and also in linking false physical and biological markers of 'criminality' to individual bodies. The control of individual bodies, especially through biometric systems, marked a departure from previous forms of social control. Unlike earlier mechanisms that governed populations based on racial or caste categories, biometric systems emphasised the direct management of the physical body, abstracting individuals from their subjective identities and reducing them to biological data. This shift reflects Michel Foucault's theory of biopolitics, where control moves away from population-level management to direct control over individuals. In the contemporary era, biometric systems, like India's Aadhaar, have expanded this individual control, linking personal data with services and resources. This form of governance is described as moving from a disciplinary society, which imposed rules and moulded individuals into groups, to a control society where codes and data determine an individual's access to spaces, services, and rights. This digitalization of control fragments the individual into various data points, shifting focus from personal identities to their 'objective' electronic indices.

Our archival findings both corroborated and extended many of these secondary findings. Over one month of visits to the British Library, the National Archives, and the UCL Reading Room, we viewed and recorded a litany of case studies, police memos, correspondences, crime reports, legal statues, and more that had to do with the CTA and its biometric implementation across various parts of British India (Bengal, Punjab, Madras, etc.). Various themes emerged throughout these documents. Many of them reflected the aforementioned interplay between the group and the individual in constructing criminality: for instance, The Criminal Tribes Act of 1924 and the Criminal Tribes (Amendment) Act of 1925 [IOR/L/PJ/6/1873, File 846] explicitly amended the CTA framework to use the word "who" instead of "which" when referring to criminalised groups. While it may appear simply like a minor semantic, this change in fact exactly reflects the Foucauldian turn to the individual body as a site of data to be harvested and controlled, as discussed above.

Manuals guiding police officers on the most 'precise' ways to exact anthropometric and phrenological measurements only further testify to this trend. One manual from 1893 provides a detailed 13-step process, complete with sketched diagrams, on how to take footprint casts of suspected criminals (IOR/L/PJ/6/342, File 598), while another from 1890 explicitly instructs officers to not just record height/width/fingerprint measurements but correlate them with a subject's caste and document it legibly (IOR/L/PJ/6/282, File 1282). These memos evince an immense level of faith in the scientific policing project, one that seemed verifiable and foolproof to top British administrators of the day: as the last page of this manual notes, "[biometric] identity is the safeguard of the innocent, dread of the guilty." The collection of other police manuals we viewed, as well as workbooks directly from Francis Galton's personal library, provide ample material to understand the way in which colonial India was used as a testing ground for many pseudo-scientific biometric technologies of criminal identification, grounded on the pre-existing axis of caste. We systematically captured images of everything ranging from fingerprints taken from colonial Bengal to diagrams from police manuals to case studies of criminal groups and ethnographies aiming to crack their intra-community languages. In toto, this selection offers rich material for a historical-political analysis of casteist biometric identification in India, including much material hereto unexamined in much of the leading secondary literature in this field.

NEXT STEPS

Legal and Policy Advocacy

The findings of this study will be used in an ongoing constitutional litigation by the CPA Project at the Supreme Court of India challenging the federal Criminal Procedure Identification Act (CPIA, 2022). We hope to use the findings of the research as a way to demonstrate to the Court the colonial origins of the CPIA and the implications of the expanding powers of the police for criminal identification in India.

Educational Materials

Additionally, we will also use the findings of this research to drive forward community-led policy advocacy around biometric surveillance by creating audio-visual materials in English and Hindi for dissemination among criminalised communities. These resources may be used for further research and advocacy by the following stakeholders:

1. Lawyers, public policy experts and think tanks who can utilise this data to challenge and highlight the criminalisation in Courts, and with government agencies.

2. Community organisers can utilise this data to delve further into modes and forms of caste based criminalisation.

Transnational Coalition Building

Finally, it is our hope that the results of this study will also serve to create more thoughtful, nuanced, and resilient transnational coalitions fighting biometric surveillance globally. The mutual flow of surveillance technologies between India and other parts of the Global South in fact did not start with Aadhaar: rather, even as early as the nineteenth century, the fingerprint, acts like the CTA, and techniques of anthropometric photography that were first inaugurated in colonial India boomeranged quickly to other regions of the British empire. Indeed, a plethora of secondary sources document how the brahminical and colonial ideology behind these forms of casteist surveillance mutated and merged with local hierarchies of anti-Blackness and indenturehood in colonies such as Mauritius and Tanganyika throughout the 1800 and 1900s.

With this history in mind, we hope our work will contribute to ongoing efforts to connect anti-surveillance movements across the Global South: if the violence of biometric surveillance has been transnational from its ugly colonial beginnings, so too must be our resistance. In our personal capacity as collaborators with organisations such as Logic(s) Magazine, the Ida B. Wells Just Data Lab, and more, we appreciate the fertile ground offered to build with grassroots coalitions of criminalised, caste-oppressed, racialized, and trans groups fighting biometric datafication and surveillance in various contexts, and hope this work might be the seed for both online and in-person gatherings with scholar-activists in India, the Horn of Africa, and beyond.