Policy Brief

Ashank Desai Centre for Policy Studies
Indian Institute of Technology Bombay
Insight • Dialogue • Impact

Public AI Systems in India

Addressing Governance Gaps for a Rigorous and Democratic Oversight

Anupam Guha & Aayushi Vishnoi

www.cps.iitb.ac.in/pbs-2026-010

PBS № 2026-010

Introduction

Since the development of a broader AI vision for India around 2018 as articulated by the AI Task Force report under the Ministry of Commerce, and the various NITI Aayog vision documents, Indian policymakers at both the central and state levels have propelled a developmentalist and market driven view of AI with the hope that it will increase prosperity and access to public services (Marda, 2018). Our research indicated that there is a significant ecosystem of what can be called 'Public AI Systems', i.e., AI system primarily imagined, researched, designed, deployed, or funded by the public sector, in the public domain, interacting with the citizenry.

Formally the ownership structures of these systems differ, they may be sourced from private, profitmaking or non-profit organisations. For example, the Ministry of Civil Aviation has created the "Digi Yatra" app incorporating facial recognition for airports around India. Law enforcement across various state jurisdictions has also been active in the furtherance of AI deployments, primarily via the use of facial recognition technology (FRT), as well as other uses of machine learning. As of the writing of this brief, there are at least 170 public FRT systems, with at least 12 centrally funded ones.

There is a wide gap in the work done in India on impact of public AI, let alone the policymakers' understanding around these systems. While scholarship and gray literature exists on the impact of such systems on individual rights like privacy, the impacton structural violence, equity, etc is lacking. From a public policy point of view the proliferation of AI systems in the public domain in the absence of an active oversight, scrutiny, and feedback mechanism from the Indian citizenry presents a retreat of democratic decision making. From a technical and a public policy lens this is of concern, the normalization of ubiquitous AI ignores some fundamental technical and social issues inherent in machine learning systems, like their data hunger (Paullada et al., 2021), stochasticity, bias (Schwartz, 2022), epistemic issues, and opacity which makes their use inappropriate in areas like governance, law, and judiciary where human arbitrariness might have some democratic recourse but machine arbitrariness does not.

As the publicly deployed AI systems proliferate, they alter the relations of Indian citizens with the State and with public systems and resources, representing a kind of unconscious yet active policymaking via machine learning.

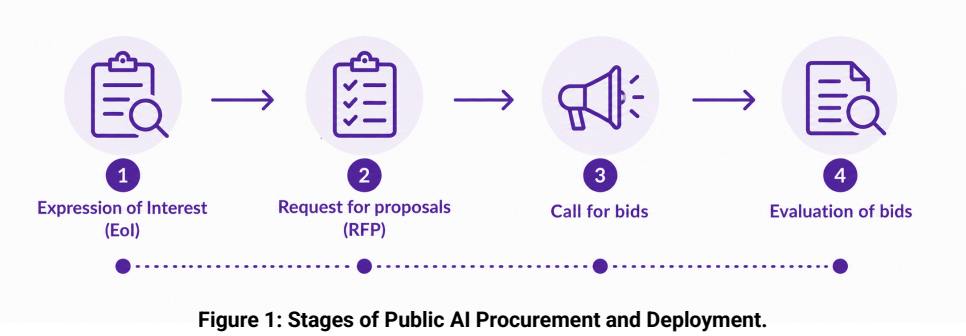


Figure 1: Stages of Public AI Procurement and Deployment.

This content is available under CC BY-NC-SA 4.0 (Attribution-NonCommercial-ShareAlike) License. For more information, see: https://creativecommons.org/licenses/by-nc-sa/4.0

India's Public Procurement Framework

While there is currently no AI-specific legislation, such regulatory gaps underscore the potential for misuse and unaccountable deployment of AI systems. In order to round off our understanding of the Indian State's vision on public AI, we also went through 10 representative policy documents from ministries and state affiliated think tanks from 2018 to 2022. Our study of these documents reveals that currently the vision of public AI being thought by the Indian state lacks a concrete understanding of the specific impacts of the technology, and so even when there is an attempt to build frameworks to evaluate use cases, these are abstract and often vague normative lines. As of now there is a significant gap in understanding the instrumental and the political economic aspects of AI use cases.

Our research first looks into the procurement ecosystem of these systems. It maps who is providing the goods and services for public AI use. For this we tracked the government ministries, departments and state-owned enterprises which purchased AI related goods and services from the private sector through the public procurement process. India is one of the few countries that do not have a specific public procurement law. The public procurement bill has been pending since 2012, it was revamped in 2015 but failed to pass in the parliament, and in the absence of a comprehensive legislation, there is a lack of uniformity in the public procurement process. Figure 1 illustrates the key stages through which public AI systems move from initial interest to deployment and monitoring. The General Financial Rules (GFR), 2017, were brought to address this, but they apply only to Central government and Central Public sector enterprises. State governments, on the other hand, may have altogether different norms, some states like Tamil Nadu and Karnataka have state specific procurement laws. Similarly, Defence procurement is separate, and follows the Defence Acquisition Procedure, 2020. To increase openness and transparency in public procurement, the Government of India launched GeM, its online procurement portal, in 2017. It is now widely used; however, state governments, defence, military and railways continue to have their own separate portals. Sometimes, the EoI and RFP steps can be merged into one, or either of them could be missing.

However, the Call for bids and evaluation of bids remains consistent across the board. There could also be pre-bid meetings to address questions from bidders and to discuss technical specifications, especially at the RFP stage. The minutes of such meetings are circulated among the prospective bidders, and sometimes also uploaded on public portals. GFR 2017 recommends a two-stage bidding process- a technical bid and a financial bid. Entities qualifying the criteria set in technical specifications are then evaluated on the financial bids and selection is done through the L1 method, which selects for the least amount quoted in bids.

A review of 1334 government tenders issued between 1 January 2022 and 11 May 2024 identified 353 public AI tenders, including 142 GeM tenders, following manual screening and cross-verification of procurement records.

Table: Year-wise Distribution of AI-related Tenders and Public AI Procurement (2022–May 2024)

Year	Total AI-related Tenders	Share of Total AI-related Tenders (%)	Public AI Tenders	GeM AI Tenders	Public AI as % of AI-related Tenders	GeM as % of Public AI Tenders
2024*	305	22.9%	88	41	28.9%	46.6%
2023	586	43.9%	167	81	28.5%	48.5%
2022	443	33.2%	98	20	22.1%	20.4%
Total	**1,334**	**100%**	**353**	**142**	**26.5%**	**40.2%**

**2024 data covers January–May 2024.*

All the documents were manually read and analysed, and the following information categories were extracted: Location, State, issuing authority, Description, Area, Function, Scope of work, Tender amount (if provided), EMD amount, the step where a specific tender is in the tendering process (if accessible), and results (if available). For accessing the results, we used the search functionalities in GeM portal, state procurement websites and department websites.

GeM portal hosts information on where the tender is in the process - with results displayed if the process has been completed. It occasionally also has the contract document uploaded with the result. While the other websites also have similar functionalities, acessing some of the results of the tender process is an issue since they are only made available if one has registered as a vendor. One common problem that was encountered in building this dataset was that often the documents did not fully state the purpose of the AI system.

Instead, generic terms like analytics, predictive analytics, and automation were used. In general, there is a huge dearth of information of what an AI deployment is needed for and is it actually an improvement from the *status quo*. We further analysed the documents falling into the category of public AI to construct a typology of usage. Document analysis identified eight recurring categories of public-sector AI procurement, spanning pre-bidding partnerships, AI infrastructure, Centres of Excellence, training, software licensing, vendor empanelment, and public AI applications. A key finding was limited transparency in tender documents, where broad terms such as analytics and automation often replaced clear descriptions of the AI system's purpose or intended use as shown in figure 2 below.

Major Policy Issues in Public Deployments of AI Systems

1. Function Creep: Function creep (Koops 2021) refers to the common problem observed and rigorously documented in machine learning systems of both the datasets and the models being used in contexts they were not meant for and causing harm. In the Indian context, we observe this problem is often the result of ad hoc attempts at pushing through machine learning deployments without understanding the scope needs to be well defined. At times, this is also connected to an optimistic but false assumption that processes will improve by using AI on them, or in some cases the assumption is that it is human intervention or discretion which causes errors and bias, thus AI has to be introduced at all costs, resulting in use of models and data which a particular use case was not meant for and causes harm.

2. Technological Determinism: A commonality was discovered in documents as early as the 2018 AI Task Force report, the 2018 NITI Aayog documents, and the more recent policy documents which have accompanied the public deployment of AI systems in India. These frame AI deployments as a 'Sector + AI' trope which envisages AI as a sort of panacea in an otherwise resource constrained environment. This tilts the policy discourse towards thinking of a digital 'solution' rather than any attempt to fix the resource constraint, and incentivises band aid 'solutions' which do not address the fundamental problems. This, alongside the imagination of regulation of sectoral AI rather than horizontal regulations based on first principles indicate technological determinism (Smith et al., 1994) being the guiding principles of deployments, which is a significant policy gap.

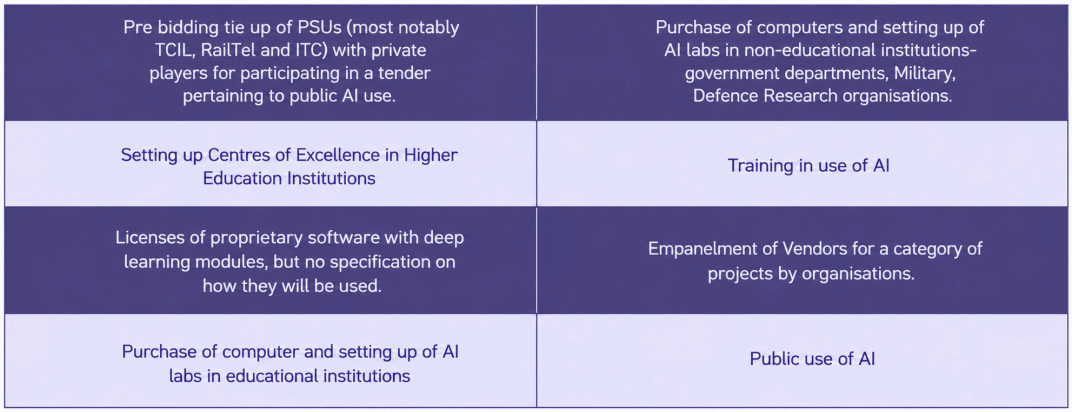


Figure 2: Classification of Tenders

3. Consultocracy: This leads to the common but incorrect idea that it is the AI technologies which are evolving by themselves and leading social change instead of considering its policy by government(s), decisions by the private sector impelled by the profit principle, and academic processes which determine how the technology evolves. At worst, the policy space gets abandoned by policymakers *in lieu* of trusting consultants and industry voices who have the incentive to frame deployments and data usage in terms of inevitability, as well as to downright falsify the capability and relevance of the technology to frame it as a solution to complex policy problems.

In our study of the documents, we discover several questionable deployments which reveals an overconfidence in the technology compounded by tech illiteracy, a belief that all regulation is inverse of innovation, that the industry can and ought to self-regulate, and that fundamentally human discretion and human oversight is undesirable. While these are all questionable priors which are not backed by evidence, these are common in policy and consultancy circles which leads to an echo chamber of ideas governing deployments, with very little academic or critical input in the process.

4. Pseudoscientific Priors Abound in these Deployments. One egregious example is emotion tech (Marda and Ahmed, 2021) which has been repeatedly characterised by scholars as fraudulent, and yet find space in diluted reframed versions in public AI systems. For example, we discovered the various safe city projects have ambitions of detecting distress among women, dangerous behaviours etc in public spaces, which are irresponsible use cases. Amongst AI scholars, fraudulent uses of AI are called 'AI snake oil' (Narayanan and Kapoor, 2024) and present a grave threat to public trust in policymaking as well as individual, social, and economic rights.

NITI Aayog's Responsible AI Framework

The most mature articulation of responsible AI deployments in India from a government adjacent organisation is found in NITI Aayog's Responsible AI framework.

Our work thus analyses this framework by studying the NITI Aayog documents and released case studies using both instrumental and normative approaches, finds gaps, and suggests additions and alterations to make it robust. The primary technical gap in the Responsible AI framework as given in the NITI Aayog documents is the mischaracterisation of explainable AI.

The framework recognises that sophisticated machine learning systems are generally opaque so when they go wrong, as all statistical systems do, it is impossible to tell why they gave a particular result. The framework thus gives a solution that AI deployments should use explainable AI as much as possible. The framework mentions that explainable machine learning systems exist as well as standards for explainability have been developed, they refer the National Institute of Standards and Technology under the US Department of Commerce.

For example, the NITI Aayog framework refers "the four principles of explainability from the Biometrics and Facial Forensic Algorithms from the National Institute of Standards and Technology" paper (Phillips and Przybocki, 2020). The creators of those four principles in their own paper acknowledge in that very paper that "the authors failed to find a significant effort in the face recognition and biometrics literature that addresses one of the principles, let alone all four principles".

The creators of these principles recognise that the industry does not have the technology and has no regulatory or economic incentive right now to develop it. While techniques for explainability like heat maps exist (Jung et al., 2021), 'explainability' itself means very different things for developers who see it as a tool for refining models, and policy experts who think it can be used to actually decipher the actions on an AI system (Minh, Dang, et al., 2022). The latter meaning is currently incorrect. There are other technical gaps in this framework like the complete absence of accounting for stochasticity which is a quality of all AI systems.

Their pseudo-random behaviour implies that there will always be errors and that is not a consequence of bad code and thus cannot be audited away but rather the errors and thus a degree of arbitrariness is a fundamental quality of AI systems, and thus responsible AI deployment and usage implies not using machine learning in any public facing system where arbitrary outcomes can cause harm, especially of the legal and social rights of a person or their economic wellbeing.

The primary policy gap in this framework is that it identifies correctly the dominant use cases but does not single out the tech involved, namely FRT or the domain it is used in, namely law enforcement. From a policy perspective, any analysis of FRT needs to consider the Puttaswamy judgement (https://indiankanoon.org/doc/127517806/) which concluded that the right to life is accompanied with a right to dignity which assures privacy as a right for citizens, this was a significant alteration from previous judgements on privacy. Any mass deployment of FRT in the public space, being at its core data hungry (Paullada et al., 2021), is against the core of this judgement revealing a significant gap in policy thinking and a governance challenge, making FRT for law enforcement fundamentally irresponsible. Scholars have also pointed out that expanding ubiquitous FRT is against the right to autonomy and freezes the public space, which public policy needs to guard against. Merely a data protection regime and establishing data processing standards does not solve this gap as the deployment itself violates collective dignity, there is no recourse to consent based solutions in such an environment.

Future Work and Policy Recommendations

1. Our work reveals both non-transparency and ad hocness in the procurement systems for public AI at all levels. This prompts further research into the complex procurement system and mapping the actors and processes involved. Such work will reveal the specificities of AI procurement and can help bolster a future draft of a public procurement bill. Such work also needs to investigate the dataset sourcing and management used in public AI deployments. A general public procurement act is vital to ensure oversight into this complex system.

2. Public AI systems need to be guided by the precautionary principle rather than techno-optimism, as harm is hard to contain and rectify post facto. "Getting on the AI bus" thinking should be avoided. Technological determinism is a popular but harmful idea which can cause damage to the national economy and policymaking around automation. AI is not a monolithic technology, rather it is an umbrella term for a wide range of disparate use cases using machine learning, and is a field rife with fraud while containing very little scope for self-regulation. Thus, the policy space must not be abandoned to industry voices and Conflict of Interest (CoI) must be kept in mind while designing consultancies. AI is also a data hungry field which makes any mass deployment of AI systems in the public domain prone to violating constitutional priors and Supreme Court judgements.

3. Any framework for responsible AI deployments must consider extant fraudulent/snake-oil use case detection mechanism (we will provide a future policy brief on this matter).

4. Frameworks for responsible AI deployments cannot depend on non-existent but theoretically possible requirements like 'explainable AI' because not only are these requirements impossible to ensure compliance on, they are meaningless post facto for a policymaker, bureaucrat, or a legal officer. Rather, frameworks should mandate use of simpler AI models which are transparent as is, rather than sophisticated deep learning models which are opaque. Also, post model uses of generative AI etc to explain why a decision was taken (for example the 'chain of thought' model) cannot be taken as demonstration of explainability as it is completely opaque to the actual processes happening inside the model.

5. Related to the last recommendation, anthropomorphisation of AI models should be avoided to prevent inflation in claims of what a model can do. Claims of 'human like' understanding, 'reasoning', 'empathy', 'affective intelligence' should be treated with scepticism by policymakers till proven. This prompts future research into a typology of AI hype claims which can harm public AI deployments.

Acknowledgment

This document is made possible by funding from the TRUST LAB at IIT Bombay. The views expressed herein are those of the authors and do not reflect the official position of the TRUST LAB, IIT Bombay.

References

Jung, Yeon-Jee, Seung-Ho Han, and Ho-Jin Choi (2021). "Explaining CNN and RNN using selective layer-wise relevance propagation." *IEEE Access* 9: 18670-18681.

Koops, Bert-Jaap (2021). "The Concept of Function Creep." *Law, Innovation and Technology* 13 (1): 29–56 doi:10.1080/17579961.2021.1898299.

Marda, Vidushi (2018). "Artificial intelligence policy in India: a framework for engaging the limits of data-driven decision-making." *Philosophical Transactions of the Royal Society A*: 376: 20180087.

Marda, Vidushi, and Shazeda Ahmed (2021). "Emotional Entanglement: China's emotion recognition market and its implications for human rights." https://www.article19.org/wp-content/uploads/2021/01/ER-Tech-China-Report.pdf

Minh, Dang, et al (2022). "Explainable artificial intelligence: a comprehensive review." *Artificial Intelligence Review* 55.5: 3503-3568.

Narayanan, Arvind, and Sayash Kapoor (2024). "AI snake oil: What artificial intelligence can do, what it can't, and how to tell the difference." AI Snake Oil. Princeton University Press.

Paullada, Amandalynne, Inioluwa Deborah Raji, Emily M. Bender, Emily Denton, and Alex Hanna (2021). "Data and its (dis) contents: A survey of dataset development and use in machine learning research." *Patterns* 2, no. 11.

Phillips, P. Jonathon, and Mark Przybocki (2020). "Four principles of explainable AI as applied to biometrics and facial forensic algorithms." arXiv preprint arXiv:2002.01014.

Schwartz, Reva, Reva Schwartz, Apostol Vassilev, Kristen Greene, Lori Perine, Andrew Burt, and Patrick Hall (2022). Towards a standard for identifying and managing bias in artificial intelligence. Vol. 3. US Department of Commerce, National Institute of Standards and Technology.

Smith, Merritt Roe, and Leo Marx, eds (1994). Does technology drive history?: The dilemma of technological determinism. MIT Press.

Previously Published Policy Briefs

PBS-2025-009 Revamping Urban Planning Practice for Resilience.

PBS-2025-008 The Hidden Circuits of Data Governance Voids in India's Personal Data Economy.

PBS-2026-007 Data Brokering Practices: Conduct and Credibility.

PBS-2025-006 Towards Sustainable Menstrual Waste Management in India.

PBS-2025-005 Policy Lessons From India's Privatized Waste-to-Energy Models.

PBS-2025-004 The Impact of Interface Design on Aadhaar Authentication for 'Non-Dominant' Name Patterns: Lessons from SARAL policy of Maharashtra.

PBS-2025-003 Transitioning Towards Digital and Data-Driven Decision Support System The Case of Jal Jeevan Mission (JJM) in Maharashtra.

PBS-2025-002 Jal Jeevan Mission (JJM) Database: From Monitoring to Decision Support Tool.

PBS-2025-001 Digital Federalism and Service Delivery Governance Transformations in India's Rural Water Sector.

Design and Layout: Stuti Upadhyay, Sarath K. and Jaisal E. K.