

STEERING AI RESPONSIBLY

A CASE FOR BUSINESS DECISION MAKERS IN INDIA TO CODE AI GOVERNANCE INTO THEIR ORGANISATIONAL DNA.

October 2024

CITIZEN DIGITAL
Foundation
Awareness. Accountability. Action.

TABLE OF CONTENTS

01 CONTEXT

02 WHAT IS AI GOVERNANCE ?

03 CONSEQUENCES OF UNGOVERNED AI

04 AI GOVERNANCE IN INDIA

05 INITIATIVES IN INDIA

06 CASE STUDIES

07 CDF AI GOVERNANCE ASSESSMENT RESOURCE

Compiled by: Nidhin Anil & Akash S S

© Fair Dealings Disclaimer
The materials used in this report, "AI Governance Resource Report" have been adapted and used for informational and educational purposes only. This transformation constitutes a 'fair dealing' of any such copyrighted material as provided for in Sections 52 of the Indian Copyright Act, 1957. All rights and credits go directly to the rightful owners of the resources.
No copyright infringement intended.



WORD FROM CDF CO-FOUNDER

Implementing AI without addressing its risks is akin to climbing a mountain without safety gear—you are courting a steep fall. It is no wonder AI governance leaves some business leaders uneasy. While many organisations articulate values such as privacy, fairness, and accountability, the challenge lies in converting those ideals into actionable steps. Too often, the breakdown occurs at the implementation stage. Leaders may speak at length about ethical AI, but unless they engage with the finer details, it remains mere rhetoric. The disconnect between values and business incentives frequently happens at this point. Sometimes, the groundwork is laid, but it proves either too technical or isolated. The key is ensuring that all teams—from HR to data science—are equally clear on their responsibilities in AI governance.

This report aims to contextualise the need for AI governance in India with a glimpse of ongoing initiatives, consequences to watch out for, case studies and an introduction to CDF's AI Governance Assessment Resource. Whether its businesses or the social sector, the sooner leaders begin assessing risks and preparing for governance, the more prepared they would be for compliances and contingencies.

So, where does one begin? What critical questions should your teams be asking? We have examined global AI frameworks and translated their criteria into a series of questions that can guide your teams through every stage of AI implementation—from planning to monitoring. Each team can complete this comprehensive assessment, receiving a score that reflects their current position. This scoring system highlights both strengths and areas in need of attention. By the end, you will possess a clear roadmap of action points, ensuring your AI initiatives align with your goals. You will also gain insight into potential risks and the means of preventing them. Moreover, we have included pertinent resources at various stages, allowing you to follow the pathway with ease.

Nidhi Sudhan
Co-founder
Citizen Digital Foundation

©Citizen Digital Foundation, 2024

01. CONTEXT

Artificial Intelligence (AI) has become an integral part of our lives, driving innovation across industries, automating decision-making processes and generating insights from vast datasets. This powerful, multi-purpose technology has the potential to transform both industrial and societal processes. Therefore utilising this potential requires meticulous governance, ethical deployment, transparency, and the mitigation of biases inherent in AI systems.



Photo credits: RationalQ

The rapid integration of AI into various sectors has given rise to significant governance challenges. Algorithms, while capable of emulating and automating intelligent human-like judgement, they are inherently imperfect models. They are susceptible to inaccurate interpretations and biased decisions, largely because they are only as good as the datasets they are trained on which may amplify the historical biases present in the society. Therefore delegating major social and commercial decisions to AI could lead to discrimination, particularly affecting vulnerable and marginalised communities.

As AI continues to play a major role in the society, it is important that governments, industries, and civil societies collaborate to establish robust frameworks for its development and deployment. While AI researchers, developers, and industry players can lay the groundwork for technical advancements, it is ultimately up to these broader stakeholders to determine the degree of regulations and legal oversight required. Thoughtful vigilance is essential across all sectors to ensure fairness and protect those whose lives are affected by AI.

> *"Robust AI governance is crucial for ensuring long-term business sustainability. By proactively investing in AI adoption that places a strong emphasis on user trust and safety, organisations can not only secure a competitive edge and boost their top-line growth, but also become more resilient to the fast-evolving regulatory landscape, minimising the risk of non-compliance and associated reputational and financial costs."*
>
> ***Raj Shekhar***
> ***Responsible AI Lead, NASSCOM***

Please note - all underlined texts have links that can be clicked on.

©Citizen Digital Foundation, 2024

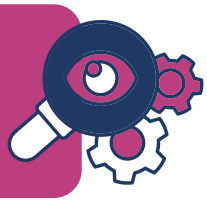
02. WHAT IS AI GOVERNANCE ?

Artificial Intelligence (AI) governance refers to the frameworks, rules, and practices that ensure the ethical and responsible development, deployment, and use of AI technologies. It encompasses a broad array of guidelines aimed at addressing the risks and challenges posed by AI, such as bias, privacy violations, and ethical dilemmas, while promoting innovation and development in AI systems.

03. CONSEQUENCES OF UNGOVERNED AI

01 Privacy invasion

Ungoverned AI can lead to mass surveillance, where personal data is collected, analysed, and exploited without consent.



02 Bias and discrimination

AI systems can perpetuate and exacerbate societal biases, influencing critical decisions in hiring and criminal justice.



03 Scalability issues

Lack of governance hinders the scalability of AI solutions, leading to inefficiencies, increased costs, and delayed AI-driven innovations.



04 Lack of accountability

The absence of governance makes it difficult to conduct audits, trace data lineage, and ensure accountability, which can bring down the confidence among stakeholders and limit AI adoption



05 Data quality issues

Ungoverned data can be incomplete, inaccurate, or inconsistent, leading to poor performance and unreliable outputs.



©Citizen Digital Foundation, 2024

04. AI GOVERNANCE IN INDIA

AI is set to play a pivotal role in driving India's digital economy, with projections indicating a contribution of USD 967 billion to the economy by 2035 and USD 450–500 billion to the GDP by 2025, making up 10% of the country's targeted USD 5 trillion GDP. Given this significant impact, effective governance of AI is crucial. Over the past few years, a series of advisories and regulatory measures have been introduced by the Indian government to address the challenges posed by AI-generated content, particularly in relation to misinformation, unlawful content, and the potential misuse of AI technologies.

Despite the existence of various legislative frameworks that address AI bias, enforcing these laws presents significant obstacles. For starters, the lack of a dedicated law addressing AI, creates a regulatory gap that makes it impossible to hold AI developers accountable for the biases in their products. Furthermore, there is a lack of experts and resources to adequately evaluate and monitor the biases present in AI systems. The lack of transparency in AI systems and their decision-making process makes identifying and correcting biases challenging. There is a crucial need to raise awareness about the need of mitigating AI bias among stakeholders such as legislators, legal professionals, and AI developers.

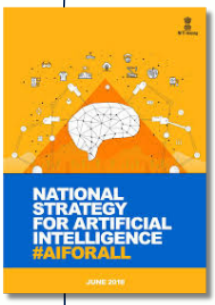
"AI governance is the cornerstone of responsible innovation, ensuring that artificial intelligence is developed and deployed with integrity, fairness, accountability, transparency, and ethical stewardship. It demands a proactive approach to safeguard against bias, protect privacy, and uphold human rights. Through thoughtful regulation and inclusive collaboration, we can shape AI systems that not only advance technology but also reinforce trust, equity, and the well-being of all people."

Dr. Anoop V S
Research Officer, School of Digital Sciences,
Digital University Kerala

Please note - all underlined texts have links that can be clicked on.

©Citizen Digital Foundation, 2024

05. INITIATIVES IN INDIA



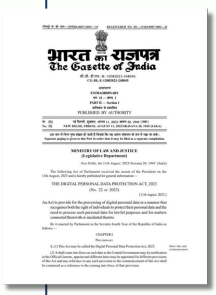
NATIONAL STRATEGY FOR ARTIFICIAL INTELLIGENCE

Released by NITI Aayog in 2018, this strategic document outlines a vision for AI development in India, focusing on sectors like healthcare, agriculture, and education. It aims to leverage AI for social good under the #AIForAll initiative, promoting inclusive growth and democratisation of technology



RESPONSIBLE AI

In 2021, NITI Aayog published guidelines emphasising ethical considerations in AI deployment. These guidelines address system and societal impacts, aiming to ensure accountability and inclusivity in AI applications



DIGITAL PERSONAL DATA PROTECTION ACT

Enacted in August 2023, this law marks India's first comprehensive privacy legislation, addressing concerns related to data protection in the context of AI



Photo credit: APAC News Network

NATIONAL CYBERSECURITY REFERENCE FRAMEWORK

Finalised in June 2023, the National Cybersecurity Reference Framework provides structured guidance for critical sectors, helping organisations improve their cybersecurity posture and comply with regulations while integrating AI technologies

GLOBAL PARTNERSHIP ON ARTIFICIAL INTELLIGENCE

India is a member of the Global Partnership on Artificial Intelligence (GPAI), which aims to promote responsible AI development and use. Built around a shared commitment to the OECD Recommendation on Artificial Intelligence, GPAI brings together engaged minds and expertise from science, industry, civil society, governments, international organisations and academia to foster international cooperation. India actively participates in GPAI initiatives, contributing to discussions on responsible AI governance and data ethics.

Please note - all underlined texts have links that can be clicked on.

©Citizen Digital Foundation, 2024

06. CASE STUDIES

> *"AI governance is essential for aligning technological progress with collective rights, democratising data access, and addressing historical injustices through a rights-based approach. This effort must redefine AI-related rights in areas like health, gender equality, work, and welfare to ensure individual and community well-being. Policies should check transnational corporations' power, develop open AI ecosystems, and address environmental impacts while ensuring strong accountability frameworks."*
>
> ***Malavika Rajkumar***
> ***Senior Associate – Research and Policy Engagement***
> ***IT for Change***

While AI has the potential to identify societal biases, the diversity of Indian society introduces certain challenges. While these biases may not be thoroughly documented in AI datasets, they are often evident in daily life.

CASE STUDY 1

MISIDENTIFICATIONS IN WELFARE SCHEMES



Photo credit: The Hindu

State governments in India are increasingly adopting algorithms to streamline the allocation of welfare schemes. However, their reliance on flawed models can lead to governance challenges and unintended consequences that harm the very people these schemes aim to help.

Please note - all underlined texts have links that can be clicked on.

©Citizen Digital Foundation, 2024

SAMAGRA VEDIKA - TELENGANA

In Telangana, Samagra Vedika, a machine learning system was deployed to assess eligibility criteria while processing welfare scheme applications. This initiative resulted in the cancellation of 100,000 ration cards among which 14,000 legitimate cards were wrongly identified as fraudulent and had to be later reinstated.



Photo credit: The Hindu

PARIVAR PEHCHAN PATRA - HARYANA

Another instance of a faulty AI system is Haryana's Old Age Samman Allowance scheme, which uses the Parivar Pehchan Patra (PPP) database to determine pension eligibility for citizens aged above 60. This algorithm has led to significant errors, one among which was declaring a 102-year-old man named Dhuli Chand as dead, halting his pension for several months despite his repeated attempts to correct this error.

This problem extends beyond a single case. Between 2020 and 2023, the PPP system halted the pensions of 277,115 elderly citizens and 52,479 widows, declaring them "dead." Upon review, it was found that 70 percent of those affected were still alive.

Please note - all underlined texts have links that can be clicked on.

©Citizen Digital Foundation, 2024

CASE STUDY 2

THE USE OF FACIAL RECOGNITION TECHNOLOGY IN INDIAN LAW ENFORCEMENT



Photo credit: Hindustan Times

WHAT IS FACIAL RECOGNITION TECHNOLOGY (FRT) ?

Facial recognition technology (FRT) is a technology that identifies or verifies individuals by analysing their facial features from digital images or video frames. The process begins with face detection, where the technology locates a face within an image and then further goes onto analyse key facial landmarks, such as the distance between the eyes and the shape of the nose, to create a unique mathematical representation known as a faceprint. This faceprint is then compared against a database of stored faceprints to confirm the individual's identity.

Indian law enforcement agencies have increasingly started to rely on facial recognition technology to monitor criminal activities nationwide. Currently, Indian authorities have installed 170 facial recognition technology systems (FRTs) throughout the country with an initial spending of INR 1,513.26 crore.

Please note - all underlined texts have links that can be clicked on.

©Citizen Digital Foundation, 2024

PUNJAB - PAIS

The Punjab Artificial Intelligence System (PAIS) is an application that uses an AI-assisted facial recognition algorithm to match photographs against a database of over 100,000 criminal records maintained by the state police.

Central to PAIS is a mobile app that features options such as Face Search, Text Search, and Gang Tree Search, enabling the police to digitise criminal records and automate searches. When officers encounter a suspect, they can capture a photo using their smartphones and compare it to a database compiled from images of convicts in jails across Punjab.

KARNATAKA - SAFE CITY PROJECT

As part of the Safe City Project, Bengaluru has setup a network of surveillance cameras equipped with FRT. The first phase, completed in March 2023, involved the installation of approximately 4,100 AI-enabled cameras across the city, with plans to expand this number to 7,500 in subsequent phases. These cameras are placed at key locations, including public spaces, transport hubs, and high-traffic areas, to facilitate real-time monitoring.

The FRT system allows law enforcement to capture images of individuals and compare them against a database of known offenders. The project also includes the establishment of command centers where video feeds from these cameras are monitored, providing police with immediate access to critical information. Additionally, 400 body-worn cameras are currently being used by police officials in the city.

The Facial Recognition System (FRS) should be able to match faces with images from its database/previous feeds, and track the past movements of these individuals. The system should generate alerts when a 'blacklisted person' is detected on the footage, or when someone is detected in an area where they are not permitted," the RTI response from the Bengaluru police read.

"Integrating AI governance within corporate structures is not just a regulatory necessity but a strategic imperative. Establishing dedicated AI Governance and Ethics Boards, which I have previously advocated for, can enhance decision-making by embedding principles of transparency, accountability, and ethical integration into the AI development pipeline. By doing so, organisations can mitigate risks like biased outcomes, both intentional and unintentional, and including issues such as algorithmic collusion. Such an internal framework will also ensure AI systems align with legal standards and the larger societal values while driving innovation responsibly"

Indranuj Pathak

Public Policy, Primus Partners India

Please note - all underlined texts have links that can be clicked on.

©Citizen Digital Foundation, 2024

TAMIL NADU - FACETAGR

The Tamil Nadu Police utilises the Facial Recognition Technology (FRT) application called FaceTagr as part of their efforts to enhance policing and public safety. The deployment of FaceTagr allows police personnel to capture images of individuals using their smartphones. These images are then matched against the database, which includes photographs of criminals and other relevant data. The system is reportedly used by over 46,000 personnel across Tamil Nadu, reflecting its integration into daily policing activities.

Despite its potential benefits, the use of FaceTagr raises significant concerns regarding privacy and the potential for misuse. Notably, there are no established criteria for identifying a person as a suspect; instead, the decision relies on the discretion and experience of the police officials. This lack of standardised procedures has led to fears of arbitrary use and potential abuse by rogue officers. Furthermore, officials have stated that they do not require a warrant to collect facial data, which exacerbates privacy concerns.



Photo credit: Coda Story

The security of the FRT system has come under scrutiny, particularly following a significant data breach in May 2024. A massive leak of 8 lakh data points in Tamil Nadu Police's FRT portal exposed the facial and personal data of about 50,000 accused or suspected persons, and compromising data related to FIRs, police officers, and other critical records. This incident highlighted vulnerabilities in the system and raised alarms about the security measures in place to protect sensitive data.

Please note - all underlined texts have links that can be clicked on.

©Citizen Digital Foundation, 2024

Facial recognition technology (FRT) is increasingly being used in India, despite significant concerns about its accuracy and the potential risks it poses to privacy. The technology relies on generating multiple possible matches with confidence scores, but the final identification is made by a police officer, introducing the risk of human error and bias. This process has been widely criticised for being prone to misidentification, as the officer's decision may be influenced by personal prejudices against certain races, religions, or communities.

India's law enforcement has a long history of institutional biases, particularly against marginalised groups, and relying on facial recognition technology may exacerbate these issues. The use of such technology can reinforce existing biases in police databases and reduce accountability for wrongful identifications and arrests.

Furthermore, the storage and regulation of centrally held facial biometric data are problematic. Not only does this practice inherently violate privacy and human rights, but it is also inadequately regulated under Indian law, leading to concerns about the security and misuse of this sensitive information. Despite these serious issues, the deployment of facial recognition technology continues to expand through various government initiatives.

S. No.	Region	Project Name	Sector it operates on	Status of Project	Commencement of the Project	Organization
1.	India	**National Automated Facial Recognition System (AFRS)**	Law enforcement- prevention and detection of crime and fast track document verification	Proposed project; Request for Proposal has been released to public[77]	Yet to commence	National Crime Records Bureau
2.	India	**Face Matching Technology**	Educational- identity authentication to access academic documents[78]	Implemented[79]	2020	Central Board for Secondary Education
3.	Telangana	**Degree Online Services Telangana**	Educational- identity authentication to access academic documents for Telangana State Board of Intermediate Education students.[80]	Implemented	2020	Telangana State Board of Intermediate Education
4.	Punjab	**PAIS**	Law enforcement- real-time identification of criminals, phonetic search and gang analysis	Implemented[81]	2018	Punjab Police
5.	Chennai and Madurai, Tamil Nadu	**FaceTagr**	Law enforcement- mobile application to match suspect photos against pre-existing crime records and CCTV footage.[82]	Implemented83	2017	Tamil Nadu Police
6.	Uttar Pradesh	**Trinetra**	Law enforcement- real-time identification of criminals	Implemented[84]	2018	Uttar Pradesh Police
7.	Maharashtra, Gujarat	**Indian Protocol Surveillance System**	Law enforcement- surveillance of individuals with the objective of ensuring women safety at railway stations[85]	Pilot project	2021	Indian Railways
8.	Telangana	**Face Recognition Application**	Authentication- additional tool in authenticating individuals in civic elections against the electoral rolls[86]	Pilot project	2020	Telangana State Election Commission
9.	India	**Authentication Based Facial Recognition**	Authentication: ease of authentication for Aadhar card[87]	Pilot project	2018	UIDAI
10.	India	**Authentication Based Facial Recognition**	Authentication- biometric authentication of Covid-19 vaccine recipients	Pilot Project launched in Jharkhand88	2021	UIDAI
11.	Telangana	**Real Time Digital Authentication of Identity**	Authentication- Authenticating the 'dead-or-alive' status of pension claimants under the T-folio app.[89]	Implemented as a voluntary program	2020	ITE&C Department, Telangana
12.	Prayagraj, Uttar Pradesh	**Pan Tilt and Zoom Surveillance Cameras**	Law enforcement- surveillance and maintenance of large crowds[90]	One-time functional project used in Kumbh Mela, 2021	2021	Uttar Pradesh Police
13.	New Delhi	**AI Vision**	Law enforcement- identifying criminals and detecting lost children	Functional Project[91]	2017	Delhi Police
14.	Pune, Maharashtra	**Selfie-App Based Face Recognition**	Monitoring- drones used to monitor quarantine adherence by COVID-19 positive patients[93]	Pilot Project	2020	Pune Municipal Corporation and Pune Police
15.	Telangana	**Darpan**	Identification- Used to match photos and identify missing children.[94]	Implemented[95]	2018	Telangana Police
16.	Surat, Gujarat	**Surat Safe City Project**	Surveillance: LFRT that integrates live video surveillance feeds[96] against a watchlist of suspected individuals.[97]	Implemented	2015	Gujarat Police

List of FRT systems deployed in India
Responsible AI - NITI Aayog

©Citizen Digital Foundation, 2024

CASE STUDY 3

UNDER-REPRESENTATION OF MUSLIM AND DALIT POPULATIONS IN AI DATASET



Photo credit: Al Jazeera

A recent poll conducted by Google revealed that the datasets used for developing AI systems in India do not adequately capture the experiences and perspectives of Muslim and Dalit communities. This underrepresentation is due to their lower levels of internet usage compared to other demographics. In India, internet access is unevenly distributed across various social, economic, and geographic groups. This lack of representation can lead to biased AI algorithms that can harm these marginalised groups. This case is not unique. Between 2020 and 2023, the PPP system halted the pensions of 277,115 elderly citizens and 52,479 widows, declaring them "dead." Upon review, it was found that 70 percent of those affected were still alive, highlighting the dangers of relying on faulty AI systems in governance.

CASE STUDY 4

UNDER-REPRESENTATION IN THE IMAGE-NET DATASET OF FACIAL RECOGNITION TECHNOLOGIES

The ImageNet dataset, a cornerstone in the development of facial recognition technologies, contains less than 3% of faces from Indian and Chinese populations. This stark under-representation raises significant concerns about bias in AI models trained on such data. Given that ImageNet comprises over 1.5 million images, the limited inclusion of diverse ethnicities can lead to facial recognition systems that perform inaccurately when applied to these groups. Consequently, this lack of diversity in training data can result in algorithms that not only fail to recognise individuals from these demographics effectively but may also perpetuate existing stereotypes and biases in automated decision-making processes.



Photo credit: Youth Ki Awaaz

Please note - all underlined texts have links that can be clicked on.

©Citizen Digital Foundation, 2024

07. CDF AI GOVERNANCE ASSESSMENT RESOURCE

At Citizen Digital Foundation, we recognise the growing complexity and importance of implementing AI responsibly. To address this, we have developed an AI Governance Assessment resource designed to simplify the process of evaluating AI readiness.

Our resource consolidates key elements from globally recognised AI governance frameworks, into an accessible, user-friendly Q&A format, providing organisations with a practical approach to assess their AI capabilities.

IDEAL FOR

- TECH COMPANIES
- GOVERNMENT AGENCIES
- NON PROFITS
- ANY ORGANISATIONS LEVERAGING AI



Take the assessment by clicking on the document here

The AI Governance Assessment Resource is envisioned as a foundational step for Indian businesses to navigate the complex landscape of AI developments. By drawing from multiple open AI governance frameworks, each linked within the toolkit, we offer a comprehensive starting point for organisations to take sufficient measures to prevent future risks and losses associated with AI deployment.

Please note - all underlined texts have links that can be clicked on.

©Citizen Digital Foundation, 2024

CDF's AI Governance Assessment resource incorporates the essential elements of the following open global AI governance frameworks and resources:

- NASSCOM RESPONSIBLE AI RESOURCE KIT
- AI BLINDSPOT FRAMEWORK
- NIST RISK MANAGEMENT FRAMEWORK (RMF)
- NIST RISK MANAGEMENT FRAMEWORK (RMF) PLAYBOOK
- KPMG GOVERNING AI RESPONSIBLY
- GERMANY FEDERAL OFFICE FOR INFORMATION SECURITY - GENERATIVE AI GUIDELINES
- GOOGLE - PERSPECTIVES ON ISSUES IN AI GOVERNANCE



The transformative power of AI brings with it significant responsibilities. While AI has the potential to revolutionise industries and society, its deployment must be carefully governed to prevent the perpetuation of biases, invasion of privacy, and other ethical concerns.

CDF's AI Governance Assessment acts as a crucial guide for organisations aiming to navigate the complexities of AI governance. By consolidating key elements from global frameworks, it provides a practical approach to ensuring that AI systems are developed and deployed with fairness, transparency, and accountability. As AI continues to shape our future, the need for comprehensive governance is clear: it is essential not only for safeguarding the rights of individuals and communities but also for fostering trust and ensuring that AI's benefits are equitably shared. The challenge of AI governance is ongoing, and it is a shared responsibility that requires vigilance, collaboration, and a commitment to ethical principles across all sectors.

Please note - all underlined texts have links that can be clicked on.

©Citizen Digital Foundation, 2024



Awareness for end-users



Accountability in tech businesses



Action in governance

Few of our partners & collaborators

Safe, responsible, equitable technologies

CITIZEN DIGITAL Foundation

Awareness. Accountability. Action.

Website: www.citizendigitalfoundation.org

Email: hello@citiendigitalfoundation.org

Phone: +91-81389-51979

Facebook | Instagram | Twitter | LinkedIn | YouTube