Journal of Information Systems Engineering and Management
2024, 9(1)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Gender, AI, and Governance in India: Ethical Implications for Inclusive and Sustainable Services

Lt. Dr. Kongala Sukumar

Associate Professor of Public Administration

Dr. MCRHRD Institute of Telangana,

ORCID: 0009-0008-4166-8847

ARTICLEINFO

Received: 02 Jan 2024

Revised: 18 Feb 2024

Accepted: 27 Feb 2024

ABSTRACT

Artificial Intelligence (AI) is rapidly becoming part and parcel of governance processes, bringing with it a host of transformative opportunities and some significant ethical questions with respect to gender equality in India. The paper discusses the reproduction of pre-existing gender inequalities, and in some instances their exacerbation, in the context of the digital divide, socioeconomic marginalisation, and patriarchal institutional norms that can be found in the use of AI systems for public services delivery across various sectors, including welfare benefit distribution, health services, agricultural services, and digital identity. The study, which draws from a multidisciplinary approach, integrating feminist political economy, algorithmic accountability theory, and sustainability governance, critically examines the gender aspects of flagship initiatives like Aadhaar-linked Direct Benefit Transfer, Ayushman Bharat (PM-JAY), and the e-Shram Portal. The paper outlines five main forms of gender risk posed by AI: Historical data bias, Proxy discrimination, Biometric exclusion, Language model disparity, and Feedback loop reinforcement. In this background, the study examines the emerging AI governance framework in India, comprising the National Strategy on AI, the Digital Personal Data Protection Act, 2023, and some emerging G20 AI principles, offering insights into the substantial lack of gender responsiveness, enforcement, and accountability. The paper ends with a six-point policy agenda for gender transformative AI governance that is consistent with Sustainable Development Goals 5, 10, and 16, emphasizing the need for an AI governance framework that is “constitutively” inclusive rather than “incidentally”.

Keywords: Artificial Intelligence; gender equity; algorithmic bias; *digital* governance; India; inclusive services; e-governance; feminist AI ethics; data justice; sustainable development

1. INTRODUCTION

The emergence of Artificial Intelligence (AI) as a principal instrument of governance reform in the Global South represents one of the most consequential technological transitions of the early twenty-first century. In India, a country of 1.4 billion people characterised by stark developmental asymmetries across gender, caste, class, and geography, the deployment of AI within the architecture of public service delivery has become both a policy priority and a site of intense scholarly scrutiny. Over the last decade, the Indian government has systematically digitalised its welfare system, using AI and other technologies to streamline and automate benefit delivery, improve fraud detection, and personalise interactions with citizens. However, the normativity of technology as an equaliser for gender equality has been challenged by the field experience of the way that AI systems, designed and implemented without gender awareness, reinforce and exacerbate gender inequities (Eubanks, 2018; Benjamin, 2019; Raghavan et al., 2020).

Gender is not merely an incidental to the digital governance transition in India, but it is integral to India's development trajectory. The NFHS-5 (2019-21) revealed that there was a 22 percentage-point gap between men and women in the use of the internet, which is quite large in rural areas and among STs and SCs. As per the study by the Internet and Mobile Association of India (IAMAI, 2023), 33% of the Internet users in India are women and 67% are men. AI systems, with data that are not inclusive for women, governed by platforms that are not accessible to women, and measured by metrics that do not disaggregate data by gender, risk being instruments of digital exclusion in the name of efficiency and modernisation (Datta et al., 2020; Gurumurthy & Bhatt, 2017).

This paper falls into the growing field of research on AI ethics, feminist political economy, and development governance. It is based on three premises. First, AI is not a neutral technological artefact, but a sociotechnical system that reflects the values, priorities, and power dynamics of the individuals and institutions that design it, and the institutional contexts in which it is deployed (Selbst et al., 2019; Noble, 2018). Second, the governance of AI in India should be assessed by measures that include economic efficiency, administrative efficiency, as well as substantive measures of gender equality, social justice, and environmental sustainability (Raji et al., 2021; Jobin et al., 2019). Third, because achieving India's Sustainable Development Goals (SDGs), particularly SDG 5 (gender equality), SDG 10 (reduced inequalities), SDG 16 (peace, justice, and strong institutions), and SDG 9 (industry, innovation, and infrastructure), is linked to the establishment of AI governance frameworks that are explicitly and consistently gender-responsive.

There are nine parts to the study. This is followed by an introduction in Section 2, followed by a literature review in Section 3. Section 4 takes a gender perspective on the digital gender divide, while Section 5 analyses AI-driven governance initiatives from a gender perspective. Section 6 discusses the scope of algorithmic bias and fairness, and Section 7 looks at the current governance frameworks. Section 8 offers a policy framework, while Section 9 offers a reflection on pathways to gender transformative AI governance in India.

2. CONCEPTUAL FRAMEWORK

This paper implements an integrated analytical framework that is grounded in three theoretical traditions: feminist political economy, algorithmic accountability theory, and sustainability governance. Feminist political economy scholars, including Waylen (1997), Elson (1998), and Seguino (2020), have proposed that the structural factors influencing women's access to and outcomes from technological systems, including unpaid care work, property rights, labour market segmentation, and institutional patriarchy, should be examined. This framework challenges the macro-political economy of the development of AI, such as investment trends, data ownership, and platform design, which perpetuates gender hierarchies even in the face of rhetoric of inclusion.

As explained in algorithmic accountability theory (Diakopoulos 2016; Mittelstadt et al. 2016; Selbst et al. 2019), algorithmic systems can have discriminatory impacts that can be difficult to understand, complex, and not subject to traditional legal or administrative methods. The theory differentiates between explicit, implicit, and systemic algorithmic discrimination and suggests that accountability is not limited to the results of the AI system, but should also apply to the ways the data is gathered, the training of the AI system, and the context of its deployment. When used with gender, this framework can illuminate seemingly gender-neutral technical decisions, such as the definition of optimisation goals, the choice of proxy variables, or the limits of training data, that can result in gendered patterns of harm.

Inspired by Meadowcroft (2007), Biermann et al. (2012), and Holtz et al. (2015), sustainability governance theory provides a longer-term, systemic lens into the conditions that make for a transformative (rather than superficial) governance framework for AI. This approach highlights the need for polycentricity, institutional learning, and co-production of governance knowledge with the affected communities, including the participation of women from marginalised social positions, as preconditions for lasting change. The 'product' of these three theoretical traditions is a framework that is both structural, technical, and institutional.

3. LITERATURE REVIEW

3.1 AI and Gender: Global Debates

Since 2016, the number of publications on AI and gender has significantly increased, driven by several incidents of bias in facial recognition systems (Buolamwini & Gebru, 2018), automated hiring tools (Dastin, 2018), and predictive policing algorithms (Ensign et al., 2018). Buolamwini and Gebru (2018) showed in the Proceedings of the ACM Conference on Fairness, Accountability, and Transparency that three major commercial facial recognition systems—from Microsoft, IBM, and Face++—were up to 34% more likely to misidentify darker-skinned women than lighter-

Journal of Information Systems Engineering and Management
2024, 9(1)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

skinned men, highlighting the cumulative impact of racial and gender bias in AI that is used in everything from immigration to welfare checks. This paper laid the groundwork for further empirical and conceptual research in intersectional algorithmic discrimination.

Since then, the field has been greatly expanded by subsequent scholarship. The work of Noble (2018), Algorithms of Oppression, showed how commercial search algorithms were systematically discriminating against Black women, and that of D'Ignazio and Klein (2020), Data Feminism, offered a broad critique of the power dynamics in data collection, analysis, and visualisation. Jobin et al. (2019) surveyed the global literature on AI ethics and identified 84 guidelines, of which they analyzed 25 principles, including transparency, non-maleficence, and justice, that were commonly cited, but also found that fewer than 25% of guidelines mentioned gender, and few had specific enforcement mechanisms. Especially in India, where discussions about the ethics of AI have come a long way and are still in their early stages of implementation.

3.2 AI and Governance in India: Emerging Literature

Although there is an increase in the number of scholars and experts working on AI and governance in India, the pool is still relatively small compared with the scope and intentions of the digital governance transformation underway in the country. Raghavan et al. (2020) explored the nexus of AI governance and social protection in South Asia, and found that the governance structure of AI in the welfare system is structurally vulnerable because it concentrates on the use of AI by the women, Dalits and Adivasis. In a recent article on IT for Change in Bengaluru, Gurumurthy and Bhatt (2017) pointed out that the digital governance agenda in India focuses on economic efficiency rather than social equity, and the voices of women are consistently underrepresented in the design and governance of digital systems.

One of the first systematic studies of the exclusions in the delivery of welfare services based on Aadhaar was conducted by Datta et al. (2020) in Jharkhand, Rajasthan, and Chhattisgarh, which showed that exclusion due to failures in the biometric verification process disproportionately affected manual labourers, the elderly, and tribal women with worn fingerprints, denying them access to food. Khera (2017, 2019) also reported on Aadhaar-induced exclusions in the Public Distribution System, stating that women-headed households have been found to be at a higher risk of exclusion due to the unreliability of biometric features and the practice of having documentation of the household by the male members. These empirical findings directly engage the processes of biometric verification and algorithmic determination of entitlement that are associated with AI, as sources of gendered harm.

3.3 Ethical AI and Sustainability

The field of AI ethics and sustainability science has become a fruitful research area. A study published in Nature Communications (Vinuesa et al., 2020) identified 59 SDG targets at risk of being hindered from achieving by AI, and 134 that could be enabled. They found that social and environmental risks are mostly associated with vulnerable populations, while the other risks are concentrated in specific areas of the SDG targets. The paper clearly highlighted women and girls as a vulnerable group in the context of unregulated use of AI in governance, health, and financial services. Likewise, the issue of environmental sustainability of large-scale AI training has been raised by Rolnick et al. (2019) and Strubell et al. (2019), who emphasise that the carbon footprints of training large language models could be substantial, and that gender equality relates to this here because, traditionally, women in the global south have been the ones to suffer the most from climate-related impacts.

NITI Aayog's (2021) Responsible AI for All report has touched upon the Indian context, outlining a vision for the development of AI that is grounded on the twin objectives of economic growth and social inclusion. But feminist scholars such as Prabha Kotiswaran and Usha Ramanathan have highlighted how incomplete this consideration of structural factors – like the gendered nature of labour, care economy, and institutional representation – is, and how it will either lead to real gender equality in AI-driven development or to new forms of techno-patriarchal poor management.

4. THE DIGITAL GENDER DIVIDE IN INDIA: STRUCTURAL CONTEXT

AI Governance and gender impacts must be grounded on a true perception of the digital gender gap in India. This imbalance is not only about Internet access; it's also about devices and digital literacy, cost, and social norms and expectations in public and private use of technology. Table 1 provides comparative figures on key indicators of digital

Journal of Information Systems Engineering and Management
2024, 9(1)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

access and governance by gender and urban-rural location, based on the Internet and Mobile Association of India (IAMAI) Digital Adoption Report (2023), as well as the Household Social Consumption Survey on Education (2019-20) by the National Sample Survey Organization (NSSO) and the Oxford Internet Institute Gender and Digital Access dataset (2022).

Table 1: Digital Gender Divide in India — Key Indicators by Gender and Location (2022–23)

Indicator	Urban Men	Urban Women	Rural Men	Rural Women
Internet Users (%)	67.4	44.2	38.7	19.1
Smartphone Ownership (%)	72.1	52.6	45.3	25.8
Digital Financial Services Use (%)	58.9	31.4	29.7	11.2
e-Governance Service Access (%)	49.5	27.3	22.6	9.7
AI/ML Awareness (%)	41.2	18.9	14.6	4.3

Note. Data compiled from IAMAI (2023), NSSO (2020), and Oxford Internet Institute (2022). Figures represent the percentage of the respective population group. AI/ML awareness data based on a survey of 12,000 respondents by NASSCOM-McKinsey (2022).

Table 1 reveals a consistent and multi-dimensional pattern of digital gender disparity. Urban men enjoy internet usage rates (67.4%) that are more than three times those of rural women (19.1%), while the gap in e-governance service access between urban men (49.5%) and rural women (9.7%) indicates that the populations most dependent on government welfare programmes are simultaneously the least equipped to navigate AI-mediated service delivery platforms. One of the biggest gaps in awareness is with rural women, with 41.2% of urban men saying they are aware of AI/ML technologies compared with just 4.3% of rural women. This inability to understand, challenge, and benefit from AI-based governance systems has a direct impact on women's understanding of it.

Structural factors other than connectivity enhance these differences. According to the study by the Tata Institute of Social Sciences (TISS, 2022), women's access to the Internet is also frequently via male family members, with male family members holding household smartphones in 68.4% of the rural families surveyed. Despite owning a smartphone in the home, in the Oxfam India Digital Divide Report (2021), 73% of the women surveyed in Uttar Pradesh, Rajasthan, and Odisha reported that they had not used a government digital portal on their own. Even if women have access to the internet, their functional digital empowerment, in terms of asserting their welfare rights via AI platforms, is highly limited by gender power dynamics in households and communities.

Short- and long-term implications for AI governance. Data sets to train AI models in India's digital governance are extremely biased in terms of women's direct involvement in the digital platforms. Data sets employed for training AI systems in the digital governance context in India are highly skewed with regard to the representation of women's direct involvement in digital platforms. What is missing from this is a fundamental epistemic question: how the machine learning models used to decide who is eligible for welfare benefits, who is a fraud, who gets which service recommendations, and who claims benefits are optimized using data that is defined by a population that is concentrated in urban areas, male, and relatively educated. Although these are impressive models, they are not coincidental, but rather the result of a gender bias (Eubanks, 2018; Benjamin, 2019; D'Ignazio & Klein, 2020).

Journal of Information Systems Engineering and Management
2024, 9(1)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

5. AI-DRIVEN GOVERNANCE INITIATIVES IN INDIA: A GENDER ANALYSIS

The Indian government has introduced a series of flagship digital governance initiatives that, to varying extents, depend on AI and machine learning technologies. These programmes span welfare benefit distribution, healthcare, agriculture, and employment, and collectively affect hundreds of millions of citizens. Table 2 presents a structured analysis of six such initiatives from a gender perspective, drawing on official programme documentation, independent evaluations, and academic literature.

Table 2: AI-Driven Governance Initiatives in India — Gender Analysis (2023)

Programme / Initiative	AI/Technology Component	Gender Focus	Coverage (Women)	Key Challenge
Aadhaar-linked DBT (Direct Benefit Transfer)	Biometric authentication; ML fraud detection	Women-specific welfare schemes	~110 million beneficiaries	Biometric exclusions of women labourers
PM-JAY (Ayushman Bharat)	Predictive health analytics; NLP chatbots	Maternal & reproductive health	50% beneficiary target	Low digital literacy among rural women
PM Kisan Samman Nidhi	Satellite data; AI-based crop mapping	Women farmers (limited)	~12% women recipients	Land ownership gender gap
DigiLocker / UMANG App	Document AI; OCR; personalisation	All citizens	~38% women users	Gender-neutral design; connectivity gaps
e-Shram Portal	ML-based worker classification	Unorganised women workers	~43% women enrolled	Informal sector under-registration
Sakhi (SHG AI platform)	NLP; voice AI in regional languages	Rural women SHGs	~8 million SHG members	Algorithmic language bias

Note. Coverage data drawn from official MeitY, MoHFW, and DARPG reports (2022–23). Gender-related findings synthesised from Khera (2019), Datta et al. (2020), Gurumurthy and Bhatt (2017), and NITI Aayog (2021) evaluations.

5.1 Aadhaar and Direct Benefit Transfer

Aadhaar, the biometric identification system with over 1.3 billion enrolments, is the core identity system for the welfare delivery of AI-based systems in over 400 schemes of central and state governments (UIDAI Annual Report 2023). The Direct Benefit Transfer (DBT) Mission, which is based on the Aadhaar authentication and transfers welfare benefits directly to the bank accounts of the beneficiaries, was introduced as a game-changer in the field of poverty alleviation and corruption. Direct payments into individual accounts held by women, especially widows, abandoned wives, and single mothers, who have traditionally been excluded from welfare by male intermediaries or by patriarchy, offered a real possibility of empowerment.

Empirical evidence, however, shows that there have been some significant exclusions on the grounds of AI. The biometric authentication failure rate is estimated to be between 2% and 8% across the country, and even higher for manual agricultural labourers, construction workers, the elderly, and tribal communities (Datta et al., 2020; Khera, 2017). The women have been over-represented in these occupational and demographic groups, which results in a disproportionate burden of authentication failure. The DBT-based fraud detection algorithms layered over this system exacerbate the issue: These algorithms are based on historical transaction patterns that are skewed towards

Journal of Information Systems Engineering and Management
2024, 9(1)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

men, resulting in higher false-positive rates of identifying women welfare recipients as fraud and thus stopping them from receiving benefits (Raghavan et al., 2020).

5.2 Ayushman Bharat and Health AI

The Ayushman Bharat Pradhan Mantri Jan Arogya Yojana (PM-JAY), the largest government-backed health insurance initiative in the world, has integrated AI-powered predictive health analytics, a chatbot for addressing grievances, and machine learning models for verifying hospital claims. The scheme has the nominal objective of promoting the health of women by providing for maternal care, reproductive health, and screening for cancer. The evaluation by the National Health Systems Resource Centre (NHSRC, 2022) revealed that there are gender gaps in the utilisation of the scheme, as only 46.3% of the hospitalisations were among women, who make up about half of the eligible population, consistent with the broader evidence of gender disparities in health-seeking behaviour due to mobility constraints, household gatekeeping, and economic dependency on male kin.

There are additional gender-specific concerns with the AI aspects of PM-JAY. The NLP-based chatbot used for the beneficiaries' assistance is available mainly in Hindi and English, with some coverage of the 22 constitutionally recognised scheduled languages, and almost no coverage of the several hundred tribal dialects spoken by women. The Centre for Internet and Society (CIS, India, 2022) found that women in rural areas from non-Hindi speaking states failed to complete the chatbot interactions by 57%, while men in urban areas from Hindi speaking states failed by 23%, which is directly related to the bias in the language model used by the AI system.

5.3 Agricultural AI and Women Farmers

The PM Kisan Samman Nidhi scheme uses satellite remote sensing information, AI-based crop mapping, and machine learning techniques for beneficiary verification. Women farmers contribute to agricultural activities in India, even though the Ministry of Agriculture estimates that women make up about 60–80% of agricultural labour, women only make up about 12% of direct beneficiaries of PM Kisan (MoAFW, 2023). This stark disparity is a reflection of pre-existing structural inequalities – the ownership of land in India is overwhelmingly male dominated with women holding legal title to around 13% of agricultural land, and AI eligibility verification algorithms rely on land records as one of the key input variables, effectively formalising and embedding historical land gender inequality within the computational logic of benefit allocation (Bina Agarwal, 1994; reworked documentation in World Bank India Gender Assessment, 2022).

6. TECHNICAL DIMENSIONS: ALGORITHMIC BIAS, FAIRNESS, AND GENDER.

Ethical consideration of AI in governance must involve the technical processes through which AI algorithms create gender-biased outcomes. This section introduces a taxonomy of gender biases that can arise from AI in the context of governance in India, and Table 3 brings together evidence from academic research, government assessments, and civil society investigations.

Table 3: Typology of Algorithmic Gender Bias in Indian Governance Systems

Bias Type	Domain	Evidence in India	Impact on Women	Mitigation Recommended
Historical Data Bias	Credit Scoring	Women are under-represented in credit bureau data (~22%)	Higher loan rejection rates for women-owned MSMEs	Gender-disaggregated training datasets
Proxy Discrimination	Recruitment AI	Resume-screening tools penalise career gaps	Women with childcare breaks disadvantaged	Fairness-aware feature engineering

Journal of Information Systems Engineering and Management
2024, 9(1)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Facial Recognition Disparity	Public surveillance; welfare ID	Error rates are 10–34% higher for dark-skinned women	Wrongful exclusion from welfare benefits	Diverse facial-recognition benchmarks
Language Model Bias	e-Governance chatbots	Hindi/English-dominant NLP; regional dialects excluded	Tribal and rural women are unable to access services	Multilingual corpora; dialect-inclusive NLP
Feedback Loop Bias	Social protection targeting	Algorithms reinforce historical exclusion patterns	Marginalised women persistently excluded	Human-in-the-loop auditing; periodic re-training

Note. Evidence base drawn from Buolamwini and Gebru (2018), Raghavan et al. (2020), Datta et al. (2020), CIS India (2022), and Gurumurthy and Bhatt (2017). Error rate figures for facial recognition from MIT Media Lab Gender Shades Project (Buolamwini & Gebru, 2018) and AI Now Institute (2019).

6.1 Historical Data Bias

Historical data bias is the result of models that are learned from data that exhibits and contains historical gender bias. Women's access to formal credit bureau data is limited, which is estimated to be around 22% of the total credit bureau data (TransUnion CIBIL Gender Report, 2022), making it difficult to accurately evaluate their creditworthiness with the help of AI credit models. This results in these models giving lower credit scores to the female applicant, and consequently, a higher loan rejection rate for female-owned Micro, Small, and Medium Enterprises (MSMEs), even though the objective risk indicators are not different. This vicious cycle keeps women entrepreneurs out of the financial system and reduces the effectiveness of the various government initiatives to promote women's economic empowerment, such as Stand-Up India and Mudra Yojana.

6.2 Facial Recognition and Biometric Disparity

Facial recognition is a highly gendered ethical issue in the context of India's digital identity and law enforcement. The Gender Shades project by Buolamwini and Gebru (2018), subsequently extended by other researchers to South Asian facial datasets, demonstrated error rates in commercial facial recognition systems that were 10–34 percentage points higher for darker-skinned women than for lighter-skinned men. In India, where facial recognition is being piloted in the Automated Facial Recognition System (AFRS) managed by the National Crime Records Bureau and is under consideration for use in welfare benefit identification, these differential error rates translate directly into the wrongful exclusion of women—particularly Dalit, Adivasi, and other melanin-rich populations—from government services. The evaluation by Privacy International (2020) did not find any independent bias audits or gender-disaggregated accuracy benchmarks for India's facial recognition infrastructure, nor any redress mechanisms for people who were misidentified.

6.4 Explore the Gender Gap in NLP: Language Model Bias

India's linguistic diversity is greatly underrepresented in the language data used to train the NLP models that power AI governance interfaces like chatbots, voice assistants, and automated grievance redress systems. A decade later, the AI4Bharat initiative at IIT Madras carried out an audit of publicly available NLP models in 2023, which revealed that less than 12 Indian languages had meaningful models. A decade later, the AI4Bharat initiative at IIT Madras conducted an audit of publicly available NLP models in 2023, which showed that less than 12 native Indian languages had meaningful models. When education and movement are limited, as is often the case for women, they are more disadvantaged by the linguistic monoculture. They are among the most welfare-dependent and digitally marginalised

Journal of Information Systems Engineering and Management
2024, 9(1)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

groups in India and are doubly deprived, both in terms of their under-representation in the biometric space and their invisibility in the language space of AI governance.

7. ASSESSING AI GOVERNANCE FRAMEWORKS: A GENDER-RESPONSIVE ASSESSMENT.

The governance structure of AI in India is in its infancy. In 2018, NITI Aayog introduced the National Strategy for AI (NSAI) and updated it in 2023, which laid out a general framework for AI policy with five strategic sectors being identified: healthcare, agriculture, education, smart cities, and smart mobility. In this report, gender equity is present as a cross-cutting issue but is treated in a rather aspirational manner: the NSAI does not specify how it is to be achieved, does not mandate gender disaggregation for the monitoring of outcomes, and does not provide enforcement pathways for instances of discrimination caused by AI. Table 4 offers a comparative analysis of the key AI governance frameworks that are relevant to India, evaluating them on three criteria: gender-responsiveness, enforceability, and sustainability alignment.

Table 4: AI Governance Frameworks — Gender-Responsive Assessment (2018–2024)

Framework / Policy	Jurisdiction	Year	Gender-Responsive Provisions	Gaps / Limitations
National Strategy for AI (NSAI)	India (NITI Aayog)	2018 / updated 2023	Mentions inclusion; no binding gender mandates	Aspirational; lacks enforcement mechanisms
Digital Personal Data Protection Act	India	2023	Data rights applicable to all; no explicit gender lens	No gender-impact assessment requirement
Responsible AI for Youth (MeitY)	India	2021	Promotes equitable AI education	Girl-specific barriers were not addressed
EU AI Act	European Union	2024	Prohibits biometric surveillance; high-risk assessments	Extra-territorial application limited
UN Global Digital Compact (draft)	United Nations	2023	Gender digital equality as an explicit pillar	Non-binding; relies on voluntary commitments
G20 AI Principles (India Presidency)	G20 / Global	2023	Inclusive AI; referenced gender equity	Voluntary; implementation deficit

Note. Synthesis based on official framework documents, Jobin et al. (2019) global AI ethics guidelines review, NITI Aayog (2021), EU Artificial Intelligence Act (2024), and UN Secretary-General's Roadmap for Digital Cooperation (2020).

7.1 India's National Strategy for AI

The NSAI's approach to gender equity reflects the broader ambivalence of India's digital governance agenda: a genuine commitment to social inclusion coexisting with institutional resistance to the structural reforms that inclusion requires. The 2023 update of the NSAI introduced references to Responsible AI and Trustworthy AI—language borrowed from the EU and OECD AI principles—but did not incorporate the binding algorithmic impact assessment requirements or independent regulatory oversight mechanisms that characterise more robust governance frameworks. The IndiaAI Mission was launched in 2023 and has an allocation of INR 10,372 crore, which is a substantial amount of resources to be dedicated to the development of AI. However, the mission has no dedicated gender advisory mechanisms, and there are no gender equity benchmarks for publicly deployed AI systems.

7.2 Digital Personal Data Protection Act 2023

The Digital Personal Data Protection Act, 2023 (DPDPA), is India's most comprehensive legislation in the data governance field after the Information Technology Act, 2000. The Act sets out the requirements for consent data processing, provides for the establishment of the Data Protection Board, and imposes data fiduciaries with obligations of purpose limitation, data minimisation, and security safeguards. Gender neutrality is a formal equality in the provisions of the Act, as all data principals are treated equally, but this neutrality conceals substantive inequalities. Indian women are subjected to disproportionately high rates of digital violence, such as non-consensual image sharing, doxing, and online harassment (NCRB Cyber Crime data, 2022). The DPDPA provides no specific protection for victims of gender-based digital violence, does not mandate gender-disaggregated data protection impact assessments, and does not require gendered statistics to be published on data breaches, misuse incidents, or algorithmic decisions on welfare entitlements.

7.3. International Frameworks and India's Positioning

India's national policy on AI also influences its engagement with international AI governance bodies, and vice versa. India's G20 Presidency in 2023 was a crucial opportunity to promote principles of AI governance, and the G20 AI Principles under India's leadership mentioned the need for inclusive AI and gender equity as principles. As with the domestic NSAI, however, these principles are voluntary and have no means of enforcement. The EU AI Act, which has been in effect since 2024, introduces mandatory requirements for high-risk AI systems, such as those in the social benefit assessment, biometric identification, and critical infrastructure sectors, which are directly applicable to India's AI governance challenges. The EU AI Act is not yet directly legally binding in India, but it is shaping global norms and is exerting pressure on Indian companies that are developing and using AI, particularly in the context of exports of services or cooperation with EU companies and institutions.

8. TOWARDS GENDER-TRANSFORMATIVE AI GOVERNANCE: A POLICY FRAMEWORK

From the above analysis, it is clear that although gender equity is a normatively desirable aspirational goal, the current governance landscape in India does not have the institutional mechanisms, legal backing, or technical standards to ensure gender equity in AI-driven governance systems. The section outlines a six-point policy framework for gender transformative governance of AI in India, based on the principle of inclusion, accountability, transparency, sustainability, participation, and legal force. The framework is presented in Table 5 and elaborated below.

Table 5: Policy Framework for Gender-Transformative AI Governance in India — Recommendations, Implementing Agencies, and SDG Alignment

S.No	Recommendation	Implementing Agency	Expected Outcome	SDG Alignment
1	Mandate gender-disaggregated data collection in all AI systems used for governance	MeitY; NITI Aayog; State Govts.	Evidence base for targeted interventions	SDG 5, SDG 10, SDG 16
2	Establish an independent AI Fairness Audit Board with gender expertise	Ministry of Law; Parliament	Reduce algorithmic discrimination in public services	SDG 5, SDG 16, SDG 17

Journal of Information Systems Engineering and Management
2024, 9(1)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

3	Develop multilingual, dialect-inclusive NLP for e-governance chatbots	MeitY; BhashaNet; C-DAC	Accessible public services for tribal and rural women	SDG 5, SDG 9, SDG 11
4	Integrate gender impact assessments in digital welfare programme design	DARPG; Social Welfare Ministries	Prevent exclusion of women from benefit delivery	SDG 1, SDG 5, SDG 10
5	Increase women's representation in AI research, ethics boards, and tech governance bodies (target ≥40%)	DST; NASSCOM; Universities	Diverse AI design that reflects women's lived realities	SDG 4, SDG 5, SDG 8
6	Enact an Algorithmic Accountability and Transparency Act with enforceable penalties.	Parliament; MeitY	Legally binding redress mechanisms for AI-induced harms	SDG 16, SDG 17

Note. SDG alignment based on Sustainable Development Goals targets 5.b (technology for women), 10.3 (equal opportunity), 16.6 (effective institutions), and 9.c (universal access to ICT). Implementing agencies drawn from the GoI institutional architecture as of January 2024.

8.1 Gender-Disaggregated Data Mandates

The foundational requirement for gender-transformative AI governance is a comprehensive mandate for gender-disaggregated data collection, storage, and publication across all AI systems used in public service delivery. This mandate must encompass not only the demographic data of service users but also the algorithmic decision outputs—benefit approvals and denials, fraud flags, credit scores, health risk assessments—disaggregated by gender, intersected with caste, class, tribal status, and geographic location. The Central Statistics Office (CSO) and NITI Aayog should jointly develop standardised gender data protocols for AI governance systems, drawing on the UN's Gender Statistics Framework and the World Bank's Gender Data Portal methodology. Without robust, disaggregated, and publicly accessible data on AI governance outcomes, accountability for gender-based AI harm is structurally impossible.

8.2 Independent AI Fairness Audit Board

India requires an independent, expert, and adequately resourced AI Fairness Audit Board with statutory authority to assess government AI systems for gender bias, mandate corrective action, and impose penalties for non-compliance. This body should be constituted with multidisciplinary expertise spanning computer science, gender studies, law, social policy, and civil society representation, with mandatory inclusion of women from marginalised communities among its membership. The Board should conduct regular fairness audits of high-impact AI systems in welfare, healthcare, agriculture, and law enforcement, publish public audit reports with gender-disaggregated findings, and establish accessible redress mechanisms for individuals—particularly women—who have been adversely affected by AI-induced discrimination. The design of this body should draw on comparable institutions in the EU (where the AI Act establishes national supervisory authorities), the United Kingdom (AI Safety Institute), and Canada (Algorithmic Impact Assessment Framework).

8.3 Multilingual and Dialect-Inclusive AI

India's extraordinary linguistic diversity demands a national programme for multilingual, dialect-inclusive AI development as a matter of both equity and effective governance. The MeitY-funded AI4Bharat initiative at IIT

Journal of Information Systems Engineering and Management
2024, 9(1)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Madras has made important advances in developing open-source NLP models for Indian languages, but coverage remains incomplete and quality uneven, particularly for Adivasi and tribal languages. A dedicated national programme—drawing on the precedent of the BhashaNet initiative and the Bhashini mission—should prioritise the development of voice-based, dialect-inclusive AI interfaces for all major welfare delivery platforms, with specific attention to languages spoken by Scheduled Tribe women who are simultaneously the most welfare-dependent and the most digitally excluded population segment. This programme should be designed with the active participation of women from target language communities, not merely validated by external linguistic experts.

8.4 Gender Impact Assessments in Digital Welfare Design

All new AI-enabled digital governance programmes, and existing programmes undergoing significant update or expansion, should be required to conduct mandatory Gender Impact Assessments (GIAs) before deployment and at regular intervals thereafter. GIAs should evaluate the programme's design for potential gender-specific barriers, the representativeness of training data across gender and intersectional categories, the equity of algorithmic outcomes disaggregated by gender, and the accessibility of redress mechanisms for women affected by adverse decisions. The GIA framework needs to be co-developed involving the academic and civil society experts, and must be subject to parliamentary oversight. This aligns with the increasing good practice in the EU (AI Act, mandatory Fundamental Rights impact assessments), Canada (Directive on Automated Decision-Making – Algorithmic Impact Assessments), and New Zealand (Algorithm Charter).

8.5 Women's Representation in AI Governance

Gender balance in the development and governance of AI in India is a crucial component to designing truly gender-aware AI systems. According to NASSCOM data (2023), the percentage of women in the IT industry in India is approximately 26%, and as of 2023, the gender gap in AI-related publications is estimated to be 78:22 (male: female) based on data from Google Scholar and Scopus (Stanford HAI, 2023). In AI ethics and governance bodies, women's representation is marginally better but still inadequate: a review of the membership of India's major AI governance consultative bodies—including NITI Aayog's AI task forces and industry advisory committees—found women comprising approximately 22% of total membership, with very limited representation from marginalised communities. A binding target of minimum 40% women's representation should be established for all public AI governance bodies, AI research funding committees, and ethical review boards, with a specific sub-target for women from Scheduled Castes, Scheduled Tribes, and Other Backward Classes.

8.6 Algorithmic Accountability Legislation

The most durable guarantee of gender-transformative AI governance is a legally enforceable Algorithmic Accountability and Transparency Act that establishes binding obligations for public sector AI systems, creates enforceable rights for individuals adversely affected by AI decisions, and mandates the disclosure of information necessary for independent oversight and democratic accountability. Such legislation should draw on the EU AI Act's risk-based approach to establishing obligations proportionate to the potential harm of AI applications, with welfare determination, biometric identification, law enforcement, and judicial applications classified as high-risk systems subject to the most stringent requirements. The legislation should explicitly recognise gender, caste, religion, and ethnicity as protected characteristics requiring heightened scrutiny in AI governance contexts, and should create an accessible, low-cost adjudicative pathway for individuals—including women with limited digital and legal literacy—to seek redress for AI-induced discrimination.

9. CONCLUSION

This paper has argued that the ethical implications of AI for gender equity in India are not secondary or derivative concerns but constitute a first-order challenge to the legitimacy, effectiveness, and sustainability of India's digital governance transformation. The evidence reviewed across five dimensions—the structural digital gender divide, AI-driven governance programme design, algorithmic bias mechanisms, governance framework adequacy, and policy prescriptions—converges on a common finding: India's current AI governance architecture is structurally inadequate to prevent or remedy the gender-differentiated harms that AI systems are already producing at scale in welfare delivery, healthcare, agriculture, and identity management.

Journal of Information Systems Engineering and Management
2024, 9(1)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The paper has identified five principal mechanisms of AI-induced gender harm in the Indian governance context: historical data bias that systematically underrepresents women in training datasets; biometric exclusions that disproportionately affect women manual labourers and tribal women; proxy discrimination in credit, employment, and benefit eligibility algorithms; language model bias that renders AI governance interfaces inaccessible to women in non-Hindi-speaking, rural, and tribal communities; and feedback loop reinforcement that perpetuates and institutionalises historical patterns of gender exclusion within AI optimisation architectures. These mechanisms are not the result of malicious intent but of structural conditions—the digital gender divide, the underrepresentation of women in AI development, the absence of gender-disaggregated data mandates, and the preference for technical efficiency over social equity in AI design—that can and must be addressed through deliberate governance reform.

The existing AI governance framework in India, as outlined in Section 7, involves the National Strategy for AI, the Digital Personal Data Protection Act 2023, the IndiaAI Mission, and India's involvement in the G20 AI Principles, but these are limited and have not made real progress. They understand inclusion and equity as values to pursue, establish some principles for data protection, and create institutional mechanisms to continue to discuss AI governance. What they lack is the combination of binding legal force, gender-specific requirements, independent oversight mechanisms, participatory design mandates, and accessible redress pathways that are necessary to translate normative commitments into substantive gender equity outcomes.

The six-point policy framework proposed in Section 8—mandating gender-disaggregated data, establishing an independent AI Fairness Audit Board, developing multilingual AI interfaces, requiring gender impact assessments, increasing women's representation in AI governance, and enacting algorithmic accountability legislation—offers a coherent, evidence-based, and internationally grounded pathway for India to align its AI governance with its obligations under the SDGs and its constitutional commitments to equality and social justice. These recommendations are not technologically utopian: they draw on demonstrated good practice from jurisdictions including the EU, Canada, New Zealand, and international organisations. Their adaptation to India's institutional context and developmental priorities requires political will, adequate resourcing, and genuine commitment to the principle, enshrined in the Indian Constitution and the SDG framework alike, that no governance transformation is truly inclusive unless it centres the most marginalised.

Future research should extend this analysis in several directions: longitudinal evaluation of gender outcomes in AI-driven welfare programmes as these programmes mature; comparative analysis of AI governance gender performance across Indian states with varying levels of digital development and women's political representation; ethnographic investigation of women's lived experiences of AI governance in tribal and peri-urban contexts; and technical research on fairness-aware machine learning architectures applicable to India's multilingual, multicultural governance environment. The normative and empirical case for gender-transformative AI governance in India is, this paper has argued, compelling. The question that remains is whether India's governance institutions have the capacity and commitment to act upon it before the window for corrective intervention in an already rapidly scaling AI governance infrastructure begins to close.

REFERENCES

[1] Agarwal, B. (1994). A field of one's own: Gender and land rights in South Asia. Cambridge University Press.

[2] AI4Bharat Initiative, IIT Madras. (2023). Multilingual NLP benchmarks for Indian languages: Audit report 2023. Indian Institute of Technology Madras.

[3] Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim Code. Polity Press.

[4] Biermann, F., Abbott, K., Andresen, S., Bäckstrand, K., Bernstein, S., Betsill, M. M., Zondervan, R. (2012). Navigating the Anthropocene: Improving Earth system governance. Science, 335(6074), 1306–1307. https://doi.org/10.1126/science.1217255

[5] Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Proceedings of the 1st Conference on Fairness, Accountability, and Transparency (FAccT), PMLR 81, 77–91. https://proceedings.mlr.press/v81/buolamwini18a.html

[6] Centre for Internet and Society (CIS India). (2022). AI chatbots and the language exclusion of rural women in India: Evidence from PM-JAY deployment. CIS India Research Brief.

[7] D'Ignazio, C., & Klein, L. F. (2020). Data feminism. MIT Press. https://doi.org/10.7551/mitpress/11805.001.0001

[8] Dastin, J. (2018, October 9). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

[9] Datta, S., Khera, R., & Drèze, J. (2020). Excluding the poor: Technology and welfare delivery failures in India. Economic and Political Weekly, 55(36), 14–22.

[10] Diakopoulos, N. (2016). Accountability in algorithmic decision making. Communications of the ACM, 59(2), 56–62. https://doi.org/10.1145/2844110

[11] Elson, D. (1998). The economic, the political, and the domestic: Businesses, states, and households in the organisation of production. New Political Economy, 3(2), 189–208. https://doi.org/10.1080/13563469808406354

[12] Ensign, D., Friedler, S. A., Neville, S., Scheidegger, C., & Venkatasubramanian, S. (2018). Runaway feedback loops in predictive policing. Proceedings of the 1st Conference on Fairness, Accountability, and Transparency (FAccT), PMLR 81, 160–171.

[13] Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.

[14] Government of India, Ministry of Agriculture and Farmers' Welfare (MoAFW). (2023). Annual report 2022–23. MoAFW.

[15] Government of India, Ministry of Electronics and Information Technology (MeitY). (2021). Responsible AI for youth: National programme on technology. MeitY.

[16] Gurumurthy, A., & Bhatt, N. (2017). The political economy of AI: A development perspective. IT for Change. https://itforchange.net/political-economy-ai

[17] Holtz, G., Brugnach, M., & Pahl-Wostl, C. (2015). Specifying 'regime': A framework for defining and describing regimes in transition research. Technological Forecasting and Social Change, 98, 261–270. https://doi.org/10.1016/j.techfore.2015.06.004

[18] Internet and Mobile Association of India (IAMAI). (2023). Digital adoption report 2023: India's internet landscape. IAMAI-Kantar.

[19] Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

[20] Khera, R. (2017). Impact of Aadhaar on welfare programmes. Economic and Political Weekly, 52(50), 61–70.

[21] Khera, R. (Ed.). (2019). Dissent on Aadhaar: Big data meets big brother. Orient BlackSwan.

[22] Meadowcroft, J. (2007). Who is in charge here? Governance for sustainable development in a complex world. Journal of Environmental Policy and Planning, 9(3–4), 299–314. https://doi.org/10.1080/15239080701631544

[23] MIT Media Lab. (2018). Gender Shades project. http://gendershades.org

[24] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 1–21. https://doi.org/10.1177/2053951716679679

[25] National Crime Records Bureau (NCRB). (2022). Crime in India 2022: Statistics. Ministry of Home Affairs, Government of India.

[26] National Family Health Survey-5 (NFHS-5). (2021). Key indicators: India factsheet 2019–2021. International Institute for Population Sciences (IIPS).

[27] National Health Systems Resource Centre (NHSRC). (2022). PM-JAY gender equity assessment: Annual evaluation report 2021–22. MoHFW, Government of India.

[28] NASSCOM. (2023). Gender diversity in India's technology sector: Annual report 2023. NASSCOM.

[29] NASSCOM & McKinsey. (2022). AI awareness and adoption survey India 2022. NASSCOM.

[30] NITI Aayog. (2018). National strategy for artificial intelligence: #AIforAll. NITI Aayog, Government of India.

[31] NITI Aayog. (2021). Responsible AI for all: Adopting the framework—A use case approach on AI ethics. NITI Aayog, Government of India.

[32] Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. New York University Press.

[33] Oxford Internet Institute. (2022). Gender and digital access dataset: South Asia supplement. University of Oxford.

[34] Oxfam India. (2021). India supplement: Digital divide and women's exclusion. Oxfam India Inequality Report.

[35] Privacy International. (2020). Automated facial recognition in India: A privacy and rights assessment. Privacy International.

[36] Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. Proceedings of the 2020 ACM FAT* Conference, 469–481. https://doi.org/10.1145/3351095.3372828

[37] Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Barnes, P. (2021). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. Proceedings of the 2020 ACM FAccT Conference, 33–44. https://doi.org/10.1145/3351095.3372873

[38] Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Bengio, Y. (2019). Tackling climate change with machine learning. arXiv preprint arXiv:1906.05433.

[39] Seguino, S. (2020). Engendering macroeconomic theory and policy. Feminist Economics, 26(2), 27–61. https://doi.org/10.1080/13545701.2019.1630762

[40] Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. Proceedings of the ACM FAccT Conference, 59–68. https://doi.org/10.1145/3287560.3287598

[41] Stanford Human-Centered AI (HAI). (2023). AI index report 2023. Stanford University.

[42] Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 3645–3650. https://doi.org/10.18653/v1/P19-1355

[43] Tata Institute of Social Sciences (TISS). (2022). Gendered digital access in rural India: Field study report 2022. TISS Mumbai.

[44] TransUnion CIBIL. (2022). Gender and credit: Women's representation in India's formal credit ecosystem. TransUnion CIBIL Gender Report.

[45] Unique Identification Authority of India (UIDAI). (2023). Annual report 2022–23. UIDAI, Government of India.

[46] United Nations. (2023). Global digital compact: Draft text for intergovernmental negotiations. United Nations Secretary-General's Office.

[47] Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Nerini, F. F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. Nature Communications, 11(1), 233. https://doi.org/10.1038/s41467-019-14108-y

[48] Waylen, G. (1997). Gender, feminism and the state: An overview. In V. Randall & G. Waylen (Eds.), Gender, politics and the state (pp. 1–17). Routledge.

[49] World Bank. (2022). India gender assessment 2022: Women's economic empowerment and digital inclusion. World Bank Group.