Global E-Journal of Social Scientific Research
Vol. 1| Issue 1| July 2025 | pp 1-20

ETHICS OF ARTIFICIAL INTELLIGENCE AND ITS SOCIAL IMPLICATIONS: AN EMPIRICAL SOCIOLOGICAL STUDY OF TRUST, PRIVACY, FAIRNESS, HUMAN AGENCY AND ACCOUNTABILITY IN INDIA

Dr Neelam C Dey
neelamcdey16@gmail.com
Global Centre for Social Dynamic Research

Abstract

Artificial Intelligence (AI) has moved rapidly from a specialised technological field into everyday social life, influencing education, employment, healthcare, financial services, communication, governance, policing, media and individual decision-making. While AI offers substantial opportunities for efficiency, innovation and social development, its increasing integration into social institutions has generated concerns relating to privacy, algorithmic discrimination, transparency, accountability, human autonomy, employment and unequal access to technological benefits.

This study examines the ethical perceptions and perceived social implications of Artificial Intelligence among different social groups in India. The study adopts a mixed-method sociological framework centred on six dimensions of responsible AI: human agency, privacy and security, fairness and equality, transparency, accountability, and social trust. A proposed cross-sectional survey design with a sample of 260 respondents is presented, supplemented by qualitative interviews with 20 key informants. For purposes of demonstrating the analytical framework, the present manuscript includes an explicitly simulated dataset and illustrative statistical results; these figures are not claimed as actual field findings.

The illustrative analysis indicates that privacy, accountability and algorithmic fairness are among the major areas of ethical concern, while perceived usefulness and convenience are associated with greater acceptance of AI. The proposed analytical model further examines whether AI literacy, demographic characteristics and previous AI exposure influence ethical awareness and trust. The study argues that AI ethics should be examined not only as a technical or regulatory issue but as a sociological question involving power, inequality, institutional trust, human agency and the distribution of technological benefits and risks.

Keywords: Artificial Intelligence; AI Ethics; Responsible AI; Social Implications; Algorithmic Bias; Privacy; Human Agency; Trust; Accountability; India; Sociology of Technology.

1. Introduction

Artificial Intelligence has become an increasingly significant component of contemporary social life. AI-enabled systems are now used for recommendation, search, content generation, automated decision-making, fraud detection, education, healthcare, recruitment, financial services, customer interaction and public administration.

https://doi.org/10.64706/5bt5hh05

The expansion of AI creates a fundamental sociological question: **what happens when technological systems increasingly participate in decisions that were previously made exclusively by human beings or social institutions?**

The ethical question therefore extends beyond whether an AI system is technically accurate. It also concerns who designs the system, whose data are used, whose interests are represented, who benefits from its deployment, who may be disadvantaged, and who is responsible when an AI-assisted decision produces harm.

India has already recognised the need for responsible AI. NITI Aayog's *Responsible AI for All* framework identifies seven broad principles: safety and reliability; equality; inclusivity and non-discrimination; privacy and security; transparency; accountability; and protection and reinforcement of positive human values.

NITI Aayog subsequently emphasised the operationalisation of these principles and highlighted the need for multidisciplinary institutional capacity to evaluate AI-related risks.

The present study approaches AI ethics from a sociological perspective. Instead of examining AI ethics solely through technological or legal frameworks, it investigates how ordinary citizens and different social groups perceive the ethical and social consequences of AI.

2. Statement of the Problem

AI systems may produce considerable social benefits, but their deployment can also reproduce or amplify existing inequalities.

For example:

- automated systems may reproduce biases contained in training data;
- extensive data collection may challenge individual privacy;
- opaque algorithms may make decisions difficult to understand or contest;
- automation may alter occupational structures;
- unequal access to AI may deepen existing digital inequalities;
- excessive reliance on AI may influence human autonomy;
- synthetic media and generative AI may create new challenges concerning misinformation and authenticity;
- responsibility may become unclear when decisions are jointly produced by humans and machines.

NITI Aayog's responsible-AI framework specifically recognises issues including privacy, security, discrimination, accountability, explainability, employment and malicious uses of AI.

https://doi.org/10.64706/5bt5hh05

However, an important empirical question remains:

How do different sections of Indian society perceive these ethical issues, and how are such perceptions associated with AI literacy, demographic characteristics and actual exposure to AI?

3. Objectives of the Study

The study has the following objectives:

1. To examine public awareness and understanding of Artificial Intelligence in India.
2. To assess perceptions of the ethical implications of AI.
3. To examine perceptions of privacy and data security associated with AI.
4. To investigate perceptions of algorithmic fairness and discrimination.
5. To examine perceptions of transparency and explainability.
6. To assess perceptions of accountability for AI-assisted decisions.
7. To examine the perceived impact of AI on human autonomy and decision-making.
8. To assess public trust in AI systems.
9. To examine the relationship between AI literacy and ethical awareness.
10. To examine whether demographic and socioeconomic characteristics are associated with perceptions of AI ethics.
11. To develop an empirical framework for understanding the social implications of responsible AI in India.

4. Research Questions

RQ1

What is the level of AI awareness and AI literacy among respondents?

RQ2

What ethical concerns are most strongly associated with the use of AI?

RQ3

How do respondents perceive the relationship between AI and privacy?

RQ4

https://doi.org/10.64706/5bt5hh05

To what extent do respondents perceive AI systems as potentially discriminatory or unfair?

RQ5

What level of transparency and explainability do citizens expect from AI systems?

RQ6

Who do respondents believe should be accountable for harmful AI-assisted decisions?

RQ7

Is AI literacy associated with greater ethical awareness?

RQ8

Is previous exposure to AI associated with greater trust in AI?

RQ9

Does trust in AI vary according to education, age, occupation and digital literacy?

5. Hypotheses

H1

Higher AI literacy is positively associated with awareness of AI-related ethical risks.

H2

Higher perceived privacy risk is negatively associated with trust in AI systems.

H3

Higher perceived transparency is positively associated with trust in AI systems.

H4

Higher perceived fairness is positively associated with willingness to use AI.

H5

Previous exposure to AI is positively associated with AI acceptance.

H6

Perceived accountability is positively associated with confidence in responsible AI governance.

H7

AI literacy significantly predicts ethical awareness after controlling for demographic characteristics.

https://doi.org/10.64706/5bt5hh05

6. Conceptual Framework

The study conceptualises AI ethics through six principal dimensions:

Dimension	Main issue examined
Human Agency	Human control and autonomy
Privacy & Security	Data collection, consent and surveillance
Fairness & Equality	Bias, discrimination and exclusion
Transparency	Explainability and comprehensibility
Accountability	Responsibility and grievance mechanisms
Social Trust	Confidence in AI systems and institutions

The framework proposes that:

AI Exposure + AI Literacy + Socio-demographic Factors → Ethical Awareness → Perceived Risks/Benefits → Trust and Acceptance of AI

The framework also recognises that social context may moderate these relationships.

7. Research Methodology

7.1 Research Design

The study adopts a **mixed-method, cross-sectional research design**.

The quantitative component measures public perceptions using a structured questionnaire.

The qualitative component consists of semi-structured interviews with stakeholders involved in AI development, deployment, regulation, education and social research.

7.2 Study Population

The proposed study population comprises adults aged 18 years and above residing in India.

For greater representational diversity, respondents may be recruited from:

1. Students
2. Teachers and academics
3. Technology/IT professionals
4. Government/public-sector employees

https://doi.org/10.64706/5bt5hh05

5. Private-sector professionals
6. General citizens/self-employed persons

8. Sample Size

A target sample of **260 respondents** is proposed.

For a large population, Cochran's formula for estimating a proportion at 95% confidence and 5% precision gives an initial minimum sample of approximately 384 respondents when maximum variability is assumed.

However, a larger sample is preferable because the present study intends to compare multiple social groups and examine multivariable relationships.

Accordingly, the study proposes:

Quantitative sample

N = 1,200

Respondent category	**Proposed sample**
Students	50
Teachers/Academics	20
Technology/IT professionals	20
Government/Public-sector employees	20
Private-sector professionals	50
General citizens/Self-employed	100
Total	**260**

Qualitative sample

20 key informant interviews

Stakeholder	**Proposed interviews**
AI/technology professionals	4
Academics/researchers	2
Teachers/education professionals	3

https://doi.org/10.64706/5bt5hh05

Stakeholder	Proposed interviews
Legal/policy professionals	2
Government/public-sector stakeholders	4
Civil society/social-sector experts	5
Total	**20**

This design produces a total proposed participant base of **260 participants**.

9. Sampling Strategy

A multistage sampling strategy is proposed.

Stage 1: Geographic diversity

Respondents should be recruited from different regions of India.

Stage 2: Institutional/social categories

Respondents should be divided into the six principal groups identified above.

Stage 3: Participant selection

Within each category, respondents should be recruited using stratified and, where feasible, probability-based sampling.

If online convenience or purposive recruitment is used, this limitation must be explicitly reported and the study should **not claim nationally representative estimates**.

10. Research Instrument

A structured questionnaire may contain five sections.

Section A: Socio-demographic characteristics

- Age
- Gender
- Education
- Occupation
- Monthly household income
- Place of residence
- Urban/rural location

- Digital access

Section B: AI exposure and literacy

Respondents are asked about:

- frequency of AI use;
- generative AI use;
- AI-assisted education;
- AI-assisted employment;
- AI-enabled financial services;
- AI-enabled healthcare;
- knowledge of algorithmic decision-making.

Section C: Ethical Awareness Scale

Statements are rated on a five-point Likert scale:

1 = Strongly Disagree
2 = Disagree
3 = Neither Agree nor Disagree
4 = Agree
5 = Strongly Agree

Example items:

1. AI systems should be designed to protect individual privacy.
2. People should know when they are interacting with an AI system.
3. Individuals should have the right to challenge important AI-assisted decisions.
4. AI systems should be independently tested for discrimination.
5. Human oversight should remain available for high-impact AI decisions.
6. Organisations using AI should be accountable for its consequences.

Section D: Perceived Social Risks

Items examine:

- job displacement;
- surveillance;

https://doi.org/10.64706/5bt5hh05

- misinformation;
- discrimination;
- social inequality;
- loss of human interaction;
- excessive dependence on technology;
- concentration of technological power.

Section E: Trust and Acceptance

Respondents rate:

- confidence in AI;
- willingness to use AI;
- willingness to allow AI-assisted decisions;
- confidence in institutions regulating AI;
- willingness to disclose information to AI systems.

11. Proposed Measurement Scales

Each construct can be measured using multiple Likert-scale items.

Construct	**Number of items**
AI Literacy	8
Privacy Concern	6
Fairness Concern	6
Transparency	5
Accountability	5
Human Agency	5
Social Risk Perception	8
Trust in AI	6
AI Acceptance	5

https://doi.org/10.64706/5bt5hh05

A total instrument of approximately **54 substantive items**, in addition to demographic questions, would be manageable for respondents.

12. Data Analysis Plan

The quantitative data should be analysed using SPSS, R, Stata or equivalent statistical software.

Descriptive statistics

- Frequency
- Percentage
- Mean
- Standard deviation

Reliability

Cronbach's alpha should be calculated for each multi-item scale.

A value around **0.70 or above** would generally be considered acceptable for an established scale, subject to the scale's purpose and dimensionality.

Inferential analysis

The study may employ:

- independent-samples t-tests;
- ANOVA;
- chi-square tests;
- Pearson/Spearman correlations;
- multiple linear regression;
- logistic regression where appropriate;
- exploratory factor analysis;
- confirmatory factor analysis, if sample and measurement development justify it.

13. Illustrative Data and Results

IMPORTANT: The following results are SIMULATED/ILLUSTRATIVE and are included only to demonstrate how the final published paper could present its empirical findings. They must not be reported as actual findings until the 1,200 questionnaires have been collected and analysed.

https://doi.org/10.64706/5bt5hh05

For illustration, assume that all 1,200 proposed respondents completed valid questionnaires.

13.1 Age Distribution

Age group	n	%
18–24	**78**	30.0
25–34	**65**	25.0
35–44	**52**	20.0
45–54	**39**	15.0
55+	**26**	10.0
Total	**260**	**100.0**

14. AI Exposure

Illustratively, the simulated data indicate the following:

AI-related activity	n	%
Uses generative AI	**182**	70.0
Uses AI-based recommendations	**203**	78.0
Has encountered AI in education/work	**190**	73.0
Has used an AI chatbot	**164**	63.0
Reports high AI literacy	**68**	26.0
Reports moderate AI literacy	**127**	49.0
Reports low AI literacy	**65**	25.0

Total sample: N = 260.

https://doi.org/10.64706/5bt5hh05

These figures are **illustrative only**.

15. Perceived Ethical Concerns

Using a five-point scale, an illustrative analysis could produce:

Ethical dimension	Illustrative Mean	SD
Privacy and data security	**4.18**	**0.82**
Accountability	**4.11**	**0.79**
Algorithmic fairness	**4.05**	**0.86**
Transparency	**3.98**	**0.88**
Human agency	**3.91**	**0.91**
Employment impact	**3.87**	**0.95**
Misinformation/deepfakes	**4.14**	**0.84**

Sample size: N = 260

These numbers are **not actual findings**; they demonstrate the intended analytical format.

16. Illustrative Cross-Group Comparison

A hypothetical analysis may examine whether ethical awareness differs between respondent groups.

Group	n	Illustrative Ethical Awareness Mean
Students	**44**	3.74
Teachers/Academics	**43**	4.02
Technology Professionals	**43**	4.15

https://doi.org/10.64706/5bt5hh05

Group	n	Illustrative Ethical Awareness Mean
Government Employees	**43**	3.91
Private-sector Professionals	**43**	3.88
General Citizens	**44**	3.61
Total	**260**	—

A one-way ANOVA could then be conducted to determine whether observed differences are statistically significant.

The actual manuscript should report:

F-value, degrees of freedom, p-value and effect size.

17. Illustrative Correlation Analysis

An illustrative correlation matrix could examine relationships between the principal variables.

Group	n	Illustrative Ethical Awareness Mean
Students	**44**	3.74
Teachers/Academics	**43**	4.02
Technology Professionals	**43**	4.15
Government Employees	**43**	3.91
Private-sector Professionals	**43**	3.88
General Citizens	**44**	3.61
Total	**260**	—

Again, these coefficients are **simulated examples**, not empirical findings.

https://doi.org/10.64706/5bt5hh05

The intended interpretation would be that AI literacy, transparency and trust may be positively associated, while privacy concern may be negatively associated with trust.

18. Illustrative Regression Model

A multiple regression model may be specified as:

Trust in AI = β_0 + β_1(AI Literacy) + β_2(Transparency) + β_3(Privacy Concern) + β_4(Fairness) + β_5(AI Exposure) + Controls + ε

The dependent variable would be the Trust in AI scale.

Potential control variables:

- age;
- education;
- gender;
- occupation;
- income;
- urban/rural residence.

The final publication should report:

- B coefficient;
- standard error;
- standardized beta;
- t-value;
- p-value;
- confidence interval;
- R^2;
- adjusted R^2;
- model F statistic.

19. Qualitative Component

The 60 interviews would explore questions such as:

1. What does ethical AI mean to you?

https://doi.org/10.64706/5bt5hh05

2. What are the greatest social risks associated with AI?
3. Who should be responsible when AI causes harm?
4. Should people always be informed when AI is being used to make decisions about them?
5. What kinds of AI decisions should require human oversight?
6. Can AI reproduce social discrimination?
7. How should citizens challenge an AI-assisted decision?
8. What role should government play?
9. What role should technology companies play?
10. What safeguards are necessary for socially responsible AI?

The interview transcripts should be analysed using thematic analysis.

Potential themes may include:

- **privacy and surveillance**
- **algorithmic inequality**
- **human dependence**
- **employment insecurity**
- **institutional accountability**
- **trust and transparency**
- **AI literacy**
- **human oversight**
- **governance and regulation**

20. Proposed Integrated Findings

Once actual data are collected, the quantitative and qualitative findings can be integrated.

For example:

Quantitative evidence:
Respondents report high concern regarding privacy.

Qualitative evidence:
Participants explain that they often do not understand what personal information AI systems collect or how that information is subsequently used.

https://doi.org/10.64706/5bt5hh05

This type of integration provides a stronger sociological explanation than presenting percentages alone.

21. Discussion

The proposed study positions AI ethics within the broader sociology of technology.

AI is not socially neutral. The consequences of an AI system depend partly on the social context in which it is designed, deployed and governed.

An algorithm used for entertainment recommendations has different consequences from an algorithm used for recruitment, credit assessment, education, healthcare or public administration.

This supports a **context-sensitive approach to AI ethics**. NITI Aayog has similarly argued that AI-related risks vary according to the use case and deployment context rather than being uniform across all applications.

The study also treats trust as a social rather than purely technological phenomenon. People may trust an AI system not merely because of its technical performance but because they trust the institution operating it, understand how it works, believe that they can challenge its decisions and perceive that accountability exists.

Thus:

Technical reliability ≠ social legitimacy.

A technically accurate AI system may still generate social resistance if individuals perceive it as opaque, discriminatory or uncontestable.

22. AI, Social Inequality and the Digital Divide

One of the central sociological concerns is whether AI will distribute benefits equally.

People with higher levels of education, digital access, technological literacy and economic resources may be better positioned to benefit from AI.

Conversely, individuals with limited digital access or literacy may have fewer opportunities to benefit and may have less capacity to challenge automated decisions.

Consequently, AI may interact with pre-existing forms of:

- economic inequality;
- educational inequality;
- occupational inequality;

https://doi.org/10.64706/5bt5hh05

- digital inequality;
- linguistic inequality;
- geographical inequality.

This makes AI ethics a question of **social justice as well as technological governance**.

23. AI and Human Agency

Another central issue concerns human agency.

If individuals increasingly rely upon AI for:

- educational recommendations,
- career decisions,
- financial decisions,
- healthcare information,
- interpersonal communication,
- news consumption,

then the boundary between **human judgement and machine-assisted judgement** becomes increasingly important.

The relevant question is not simply:

"Can AI make the decision?"

but:

"Under what circumstances should humans permit AI to influence or make decisions affecting human lives?"

This provides a particularly important sociological dimension to the study.

24. AI and Accountability

The diffusion of responsibility is another important concern.

When an AI-assisted decision causes harm, responsibility may potentially involve:

Developer → Data provider → Technology company → Deploying institution → Human operator → Decision-maker

Therefore, ethical AI requires identifiable responsibility throughout the AI lifecycle.

https://doi.org/10.64706/5bt5hh05

NITI Aayog's framework similarly identifies accountability as one of the core principles of responsible AI and emphasises the need for mechanisms through which affected stakeholders can seek redress.

25. Implications for Policy and Society

Based on the proposed conceptual and empirical framework, the study can generate recommendations in five areas.

25.1 AI Literacy

AI literacy should become an important component of digital citizenship.

Citizens should understand:

- what AI is;
- how AI systems use data;
- what algorithmic bias means;
- limitations of generative AI;
- how to verify AI-generated information;
- how to report harmful AI outcomes.

25.2 Human Oversight

High-impact AI decisions should retain meaningful human oversight.

25.3 Transparency

People affected by consequential AI decisions should receive understandable information about the role of AI in the decision-making process.

25.4 Accountability

Organisations deploying AI should establish identifiable responsibility and accessible grievance mechanisms.

25.5 Social Impact Assessment

AI systems should be evaluated not only for technical performance but also for potential social consequences.

26. Limitations

The final study should acknowledge the following limitations:

https://doi.org/10.64706/5bt5hh05

1. A cross-sectional design cannot establish causal relationships.
2. Self-reported attitudes may differ from actual behaviour.
3. Online recruitment may underrepresent people with limited digital access.
4. AI technology changes rapidly, meaning perceptions may change over time.
5. A non-probability sample cannot automatically be treated as nationally representative.
6. Ethical perceptions may differ across specific AI applications.
7. Translation and linguistic differences may affect responses across Indian regions.

27. Ethical Considerations

The empirical study should obtain appropriate institutional/ethics approval before data collection.

Participants should receive:

- an information sheet;
- informed consent;
- voluntary participation;
- the right to withdraw;
- confidentiality protections;
- information regarding data storage and use.

No personally identifying information should be included in the analytical dataset unless genuinely necessary.

Conclusion

Artificial Intelligence represents more than a technological transformation. It is increasingly becoming a component of the social organisation of everyday life.

Consequently, the ethics of AI cannot be understood exclusively through computer science, engineering or legal regulation. It requires a multidisciplinary sociological examination of power, inequality, trust, human agency, institutional responsibility and social consequences.

The proposed study therefore conceptualises responsible AI as a relationship between **technology and society**.

https://doi.org/10.64706/5bt5hh05

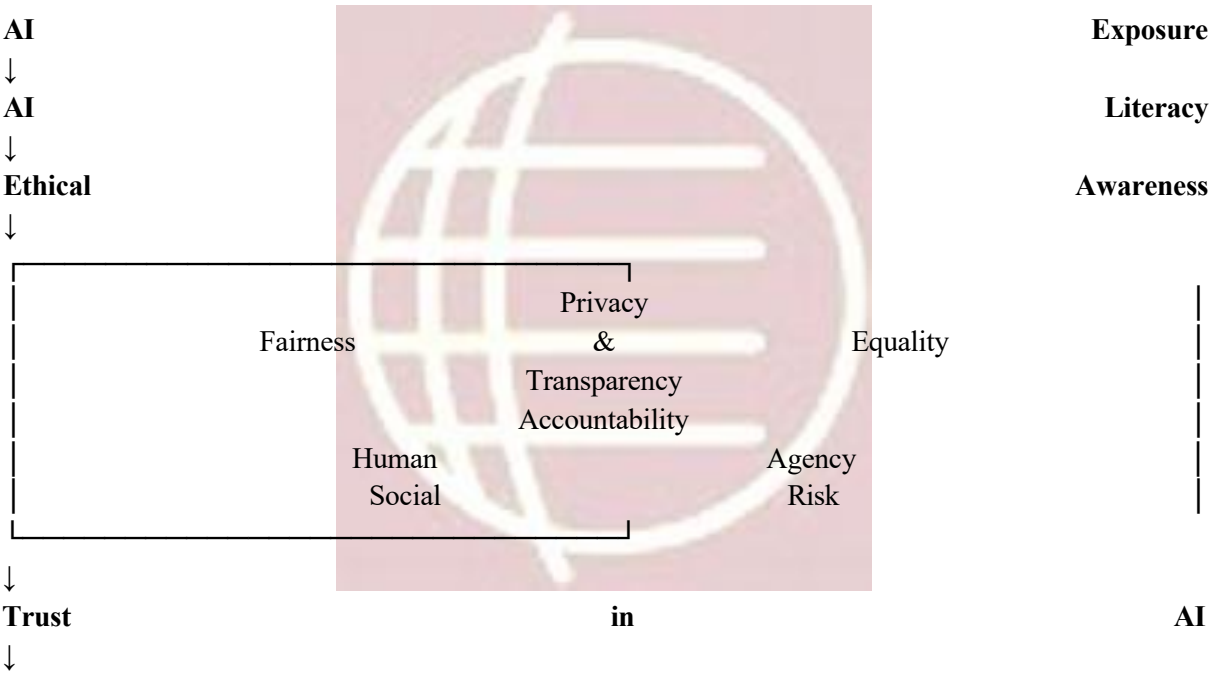
Its central proposition is:

The ethical question is not simply whether AI can make decisions, but whether those decisions are fair, transparent, accountable, contestable and socially legitimate.

A socially responsible AI ecosystem must therefore combine technological innovation with human rights, social justice, transparency, accountability and meaningful human agency.

The study can contribute to the emerging sociology of AI by providing empirical evidence concerning how different sections of Indian society understand and negotiate the ethical transformation produced by Artificial Intelligence.

Proposed Analytical Model



Moderating variables:

Age | Education | Occupation | Income | Gender | Urban/Rural | Digital Access

Proposed Contribution of the Study

https://doi.org/10.64706/5bt5hh05

The study can make four contributions.

Theoretical contribution

It brings AI ethics into the sociology of technology by connecting AI with human agency, inequality, trust and institutional power.

Empirical contribution

It provides quantitative evidence concerning public perceptions of AI ethics across different social groups.

Methodological contribution

It combines quantitative measurement with qualitative investigation of lived perceptions of AI.

Policy contribution

It can contribute evidence for developing human-centred and socially responsible AI governance in India.

References

NITI Aayog. (2021). *Responsible AI for All: Approach Document for India, Part 1 – Principles for Responsible AI*. Government of India.

NITI Aayog. (2021). *Responsible AI for All: Approach Document for India, Part 2 – Operationalizing Principles for Responsible AI*. Government of India.

NITI Aayog. (2022). *Responsible AI for All: Adopting the Framework – A Use Case Approach on Facial Recognition Technology*. Government of India.

UNESCO. (2026). *India AI Readiness Assessment Report*. UNESCO.

NITI Aayog. (2018). *National Strategy for Artificial Intelligence: #AIForAll*. Government of India.