Working Document: Enforcement Mechanisms for Responsible #AIforAll

NITI Aayog

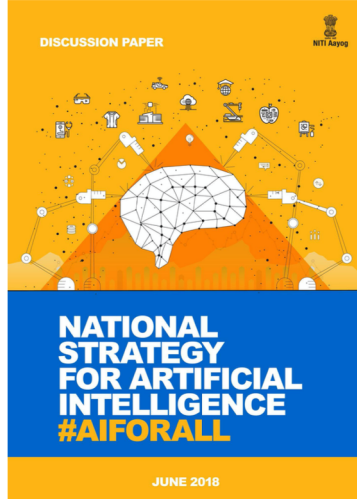DRAFT FOR DISCUSSION

———

Draft Document for discussion

The content of this draft document is solely for the purposes of discussion with stakeholders on the proposed subject.

The document was prepared based on expert consultations over the past year. The information contained herein is neither exhaustive nor final and is subject to change.

All stakeholders are requested to review the documents and provide comments on or before 15 December 2020, preferably on email at annaroy@nic.in

Towards Responsible AI

2018: National Strategy for Artificial Intelligence

2020: Towards Responsible AI for All (Part 1)

Advocated responsible use of AI to address ethical concerns

Proposes Principles for Responsible Management of AI in India

This document (Part 2 of the series on Responsible AI) proposes a framework for enforcement of responsible AI principles

AI - a general purpose technology showing rapid growth

Use cases for Artificial Intelligence have emerged across sectors and the technology has shown rapid growth over recent years
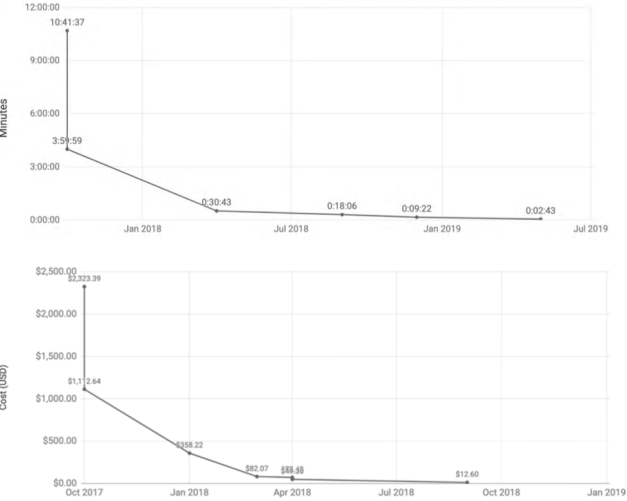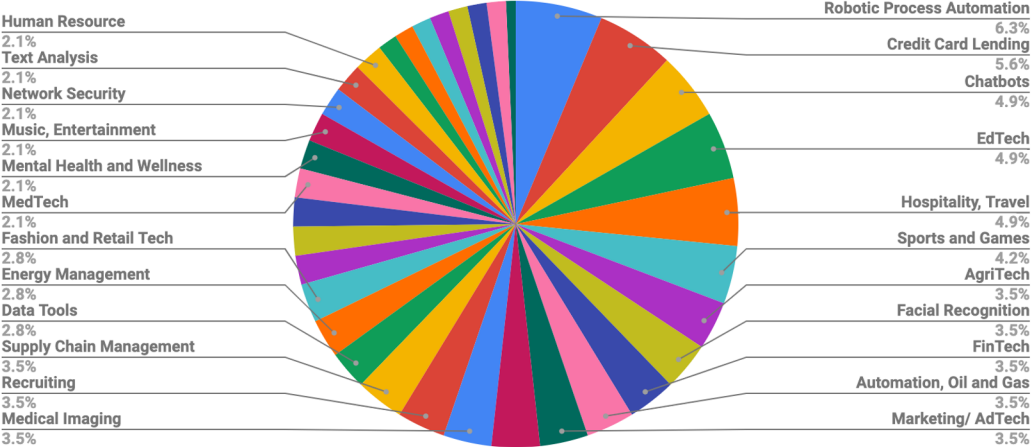
Startup investments in India for AI have happened across sectors

Time and cost to train ML system has come down drastically in just 3 years

Approach to manage risks cannot be isolated. Such approaches must be highly participatory and must keep pace with technology

Need for a context specific approach

Risk across use cases and contexts vary and also evolve over time. One-size-fits-all approach is not sustainable

Risk vary across use cases

Use case	Example Risk
Autonomous Vehicle	Safety
Credit lending	Discrimination
Fraud detection in healthcare	Inclusion

Risk depends on deployment context

Face Recognition	
Unlocking phone	Surveillance

Enforcement depends on regulatory environment

Sector	Regulators
Health	NeHA, National Medical Commission, Drug Controller General
Finance	SEBI, PFRDA, IRDAI

A flexible risk-based approach must be adopted. In this regard, the National Strategy for Artificial Intelligence proposes an Oversight Body

Role of the oversight body

The oversight body must play an enabling role under the following broad areas

1. Manage and update Principles for responsible AI in India

2. Research technical, legal, policy, societal issues of AI

3. Provide clarity on responsible behaviour through design structures, standards, guidelines, etc

4. Enable access to Responsible AI tools and techniques

5. Education and Awareness on Responsible AI

6. Coordinate with various sectoral AI regulators, identify gaps and harmonize policies across sectors

7. Represent India (and other emerging economies) in International AI dialogue on responsible AI

1. Manage and Update Principles for Responsible AI

The Principles should reflect the technology capabilities, risks, policy and legal environment and should adapt accordingly

Monitor and Update

Continuously monitor and update the Principles for responsible AI based on advances in use cases and technology

Mechanisms to translate principles to practice

Interface with various bodies, in designing specific mechanisms to translate principles into practice

2. Research into Responsible AI: Background

Research on responsible AI is vital to 'AI for Greater Good' and lags general AI research around the world

Number of papers with keywords relating to 'Classical', 'Trending' and 'Ethical' topics
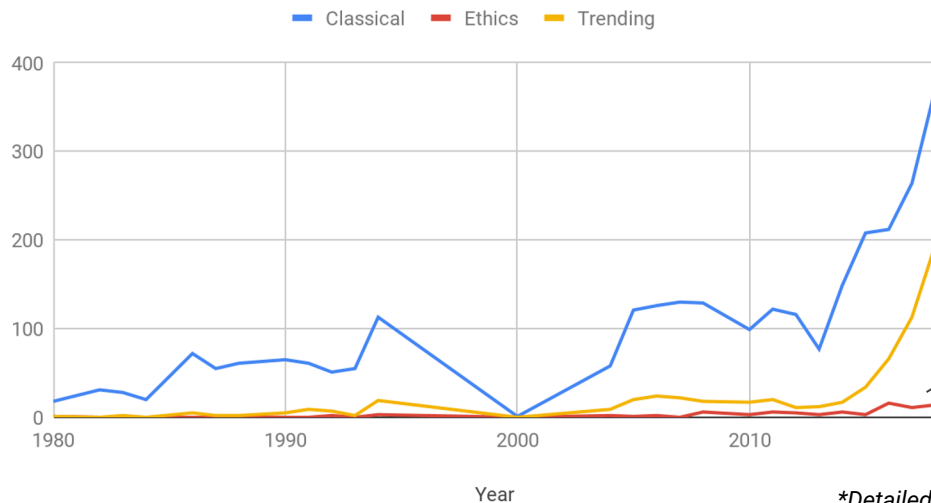
Research on ethics around the world has not kept pace with AI research on trending and classical topics*

*Detailed methodology on keywords used available in Prates (2018)

Research into ethics of AI is multi-disciplinary and must be aimed towards advancing the field, identifying issues, address concerns around AI and inform policy decisions and guidelines

2. Research into Responsible AI: Recommendations

Incentivise cross disciplinary research

The Government may support research on the impact of AI in Indian context and on fundamental research to advance Responsible AI by prioritising funding opportunities and fellowship programs.

International alliances may be leveraged to facilitate exchange of multi-disciplinary talent, data, and consolidation of research efforts, especially in areas of social good

Top conferences on ethics of AI may be incentivised to host in India so that challenges and approaches around the world can be studied and motivate indegenous research

Study and monitor impact on the ground

Engage with local communities, civil societies and other relevant organisations to study and monitor impact of various AI deployments on different communities and publish policy papers

3. Clarify responsible behaviour: Background (1/ 2)

Lack of clarity on responsible behaviour has inhibited the growth of AI in India

Example areas where guidance/ clarity will help

Healthcare*

- Doctor-Patient Confidentiality
- Informed Consent Process
- Explainability Standards
- Liability framework

Government Procurement

- Procurement mechanisms
- Monitoring mechanisms
- Liability framework

*Source: CIS report- AI in Healthcare

3. Clarify responsible behaviour: Background (2/ 2)

Standards and guidelines are being developed around the world on responsible ways of managing technologies under specific context and may be leveraged

ISO/IEC WD TS 4213
Information technology — Artificial Intelligence — Assessment of machine learning classification performance

ISO/IEC WD 5059
Software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Quality Model for AI-based systems

IEEE P7004 - Standard for Child and Student Data Governance

IEEE P7005 - Standard for Transparent Employer Data Governance

3. Clarify responsible behaviour: Recommendations

Oversight body may identify design standards, guidelines and acceptable benchmarks for priority use cases with sectoral regulators and experts. These may be made mandatory for public sector procurement

Standards	Leverage and ratify international standards when possible (in consultation with relevant Ministry/ sectoral regulator) Standards may also be created or augmented for local context when required in consultation with BIS and relevant sectoral regulators
Guidelines	Develop design guidelines, and frameworks for responsible AI through policy sandbox and controlled pilots Create guidelines for 'Model AI procurement' RFP for various priority use cases to guide responsible AI procurement in the public sector. Such documents may include risk assessment, best practices through the lifecycle, clarity on responsibility, liability and IP considerations.

4. Enable access to Responsible AI tools and techniques: Recommendations

Promote development and access to data and technology tools for responsible AI	
Support open technology projects	Hackathons, workshops, open challenge mechanisms may be used to identify and promote technology solutions for adherence to Principles Linguistic and NLP tools in local Indian languages may be promoted to facilitate access to benefits of AI across the country Promote projects for development of any tools and technologies to enable easy access to responsible AI practices
Enable data availability and sharing	Identify issues with data availability, sharing mechanisms and promote a) research into data generation, identifying proxies, b) creation and adoption of safe data sharing protocols (ex: through model protocols, data sharing agreements)

5. Awareness on Responsible AI: Background

Broad aim of awareness programs may be as follows,

Reduce Trust issues and apprehension of AI systems	Understand capabilities, strengths and weakness of AI systems	Learn about Responsible AI Management practices, tools and techniques

Reduce Information asymmetry	Develop skills to identify and think through ethical problems

Such programs may be entity specific (Public sector, Private sector, Academia, General Public, etc) and may be customized to the local context

5. Awareness on Responsible AI: Role based training

Training needs may depend on the role

Decision Maker	Procurer/ Influencer	Implementing Agency	User	Impacted Stakeholder
How AI/ML works Need for ethical thinking Best practices in procurement	How AI/ML works Identify and anticipate ethical problems Ability to reason on potential solutions Ability to communicate ways of addressing the problems	Standards, guidelines, best practices Tools and techniques for responsible AI Grievance redressal mechanisms, SOPs, etc	Capabilities of a specific AI technology Awareness of its limitations and safe usage protocols	Awareness of rights Awareness of role, capabilities, limitations of AI Awareness of grievance redressal mechanisms

*The topics mentioned are representative only. Actual needs may depend on individual context

5. Awareness on Responsible AI: Recommendations (1/2)

General Public	Local communities and regional social organisations may also be engaged to study the impact of AI, knowledge gaps and facilitate targeted awareness campaigns
Public Sector	Independent organisations may be leveraged for needs assessment, and developing targeted training curriculum for public sector officials. Academic Institutions, Private sector, and relevant experts may also be involved for training on use cases, and best practices. States, departments and bodies with experience in responsible deployment may host others and create sister-city agreements for knowledge transfer

5. Awareness on Responsible AI: Recommendations (2/2)

Private Sector	Private sector may be encouraged to create open knowledge resources on risks, case studies and best practices on responsible AI, in collaboration with academic institutes. Ethics-by-design standards for Responsible AI may mandate training for all stakeholders
Academic Institutes	Courses to be introduced at the earliest appropriate level to develop the skills to think through ethical issues early and learn to identify effective ways of addressing them Model curriculum may be created for Universities to leverage- (developed in collaboration with the Ministry of Education). In universities where multi-disciplinary faculty is not available, cross-university collaboration and guest lectures may be considered.

6. Coordinate with sectoral regulators

Multiple regulators across sectors are regulating data and AI. This requires coordination to prevent inconsistent policies and ambiguity, especially for cross sectoral AI use cases

Coordinate Approaches	Coordinate approaches across various regulators to avoid duplication of efforts and inconsistent policies
Identify risks	Assist regulators in identifying risks w.r.t AI use cases and design policies, benchmarks, or ratify standards as applicable
Monitor policies	Work with various civil societies, research institutions, industry bodies and other relevant agencies to monitor existing policies and regulations gaps, inconsistencies, and other issues and provide recommendations
	Publish policy papers and promote any such activities that contribute to realising benefits of AI while managing the risks

7. Represent India in International dialogue on AI

International collaboration on Responsible AI

Identify avenues for International collaboration on Responsible AI

Provide India's (and other emerging economies) perspective on responsible AI in International forum

Policies to enable International collaboration

Assist relevant ministries (MeitY, MEA) in development of cross border data sharing protocols to facilitate collaborative research

Assist in facilitating International University collaborations on Responsible AI

Design of Oversight Body

Global Practices

United Kingdom	Singapore

Centre for Data Ethics and Innovation	Advisory Council on Ethical Use of AI and Data

Under Department for Digital, Culture, Media & Sport	Under Infocomm Media Development Authority (IMDA)

Independent Board comprising expert and influential individuals from a range of fields relevant to its mandate

Eleven council members include international leaders in AI; advocates of social and consumer interests; and leaders of local companies

India's approach: Highly participatory advisory body is proposed

Considerations behind design of the Oversight mechanism

- Existing regulatory instruments are best placed to enforce rules, standards and guidelines. The oversight mechanism may serve in advisory capacity

- It must interface with existing regulators across sectors

- Have dedicated resources to drive each mandate

- Technology easily blends across other technologies and must not be viewed in silo. Ethics should be seen as not just limited to AI but also other emerging technologies such as ARVR, etc

Proposed composition of the advisory body

A *Council for Ethics and Technology* is proposed with a multi-disciplinary composition

For effective functioning, the body must include,

- Computer Science and AI experts,
- Legal experts,
- Relevant sectoral experts,
- Civil societies,
- Humanities and Social Science experts
- Industry representatives
- Representatives from Standard setting bodies
- Government support for interfacing across Ministries and Departments

Additional experts may be opted in by the body depending on the requirement

Procurement in Public Sector

Procurement of AI systems may include a review by an 'Ethical Committee'

Constitute Ethical Committee

An ethics committee may be constituted for the procurement, development, operations phase of AI systems and be made accountable for adherence to the Responsible AI principles

Composition depends on use case

Composition of the committee will depend on the use case. A model terms of reference and composition of such a committee is proposed in the following slides

Model Terms of Reference of Ethical Committees (1/2)

Ethical Committees are accountable for enforcement of principles in the AI system's lifecycle

- EC should assess the "potential of harm" and potential benefits, evaluate plan for mitigating risks and provide recommendations on whether the AI solution should be approved.

- Ethical Committees (EC) must ensure the AI system is developed, deployed, operated and maintained in accordance with the Principles

- EC should determine the extent of review needed for an AI system depending on inherent risks and benefits including but not limited to external audit.

- EC should ensure accessible and affordable grievance redressal mechanisms for decisions made by the AI system.

Model Terms of Reference of Ethical Committees (2/2)

- EC should ensure creation of structures within the entity for protection of 'whistleblowers' reporting unethical practices

- Every EC should have a documented Standard Operating Protocol (SOP) on functioning. The SOP may be reviewed and updated periodically to reflect changing requirements

- Every EC review must be documented, including the risks identified, mitigation strategy, and comments from the committee members

Model Composition of Ethical Committee (1/2)

Ethical Committees should have multi-disciplinary composition without Conflict of Interest

Member	Definition
Chairperson	● Nodal point of contact, accountable for independent and efficient functioning of the committee ● Must be able to ensure active participation of all members in discussions and deliberations ● Ratify minutes of EC meetings
Member Secretary	● Must be a member of the organization or institute and should be able to dedicate time for EC reviews ● Ensure effective procedures and protocols for EC review
Data Science and/or AI expert (one or more depending on requirement)	● Must be a qualified data scientist ● Must identify procedural or technical risks during development and deployment including, data collection, annotation, management, storage, processing, training, maintenance, and monitoring.

Model Composition of Ethical Committee (2/2)

Member	Definition
Sector expert	● Must have expertise in the sector and wide ranging deployment scenarios ● Must evaluate safety, reliability, access and affordability of grievance redressal mechanism
Legal expert	● Must have expertise in relevant rules and regulations relevant to the AI system ● Must evaluate legal considerations for the AI system
Social scientist/ ethicist (one or more depending on requirement)	● Must have background in social or behavioural science or relevant expertise. Must be sensitive to local cultural and moral values. ● Must assess impact on community, socio-cultural, religious, philosophical context
Representative of Stakeholder community (one or more, depending on requirement)	Must be a stakeholder of the AI solution. Serve as a representative of the user community

Private Sector (1/ 3)

Private sector may be encouraged to use ethics-by-design structures (defined by standards bodies) in the organisation. Adherence may further be incentivised through a carrot-and-stick approach

Encourage Self-Regulation in general

India currently lags in terms of private sector investment in AI and risks under local context are yet to be fully understood (*AI Index, 2019*).

In many countries around the world, public awareness and market forces have incentivised the private sector to self regulate

Voluntary self-regulation may be a good starting point for India as well. This may evolve as the risks become clear

Private Sector (2/ 3)

For high risk use cases*, adherence mechanisms may be mandated

Mandate adherence for high risk use cases through guidelines, standards and other instruments

High risk use cases include all such use cases that have the potential to cause significant harm to individuals

Such use cases, guidelines and adherence mechanisms may be defined by the '*Council for Ethics and Technology*' in consultation with sectoral regulators and experts. Adherence may be through self-declaration or through an independent third party audit, depending on the level of risk

International standards may not always to relevant, exhaustive or available for Indian context. Hence it is critical that the Government plays a role in ensuring the definition of 'acceptable behaviour' is clear

*We invite comments as a part of public consultation on a framework to identify high risk applications and practical means to ensure adherence

Private Sector (3/ 3)

Cost of compliance for ethical structures

Compliance to standards and guidelines has sometimes raised concerns in terms of creating a barrier to entry for the startups. However, startups around the world have found unique ways to manage the costs, some of them include,

a) assigning accountability for ethics to their existing leadership team;
b) leveraging online courses, workshops, open materials so the entire team is aware of the risks and develop the skill to ask the right questions;
c) leverage open tools and techniques;

Investment firms around the world are also being sensitized about the economic cost of non-adherence

Research Institutions

Existing Institute Review Board and Peer-review mechanism may be augmented

For research, a model ToR, composition and review mechanism for AI research may be developed by the '*Council for Ethics and Technology*' in collaboration with the *Ministry of Education*

The existing Institute Review Board and peer review mechanism may be augmented with necessary experts and cross-disciplinary skills. Cross-University collaboration may be considered in case the relevant skills are not available

Journal and Conference may be recommended to include of 'Statement of Ethical consideration' in all submissions

Government funding and fellowship opportunities on AI offered by various Ministries and Departments may mandate institutional adherence to responsible AI structures

Acknowledgement

We acknowledge the contribution of the following experts for reviewing this document and providing comments

Prof Mayank Vatsa (IIT Jodhpur)

Arghya Sengupta and Ameen Jauhar (Vidhi Centre For Legal Policy)

Google Team

John C Havens (IEEE)

Thank You