# Comments on the Working Document: Towards Responsible #AIforAll

By **Shweta Mohandas, Arindrajit Basu, and Ambika Tandon**

Edited by **Pranav M B**

**The Centre for Internet and Society, India**

# Relevance of the Document (High Level Comments)

1. The NITI Aayog Working Document on Responsible AI for All released on 21st July 2020 ("the report") serves as a significant statement of intent from NITI Aayog, acknowledging the need to ensure that any conception of "Responsible AI" must fulfill constitutional responsibilities, incorporated through workable principles. In doing so, it builds upon the National Strategy for Artificial Intelligence published in 2018, and is an important discussion starter. Since the report is merely a draft document for the purposes of discussion with stakeholders, it is important to highlight next steps, for research and policy maneuvers to build upon the report, and accordingly devise a constitutionally sound and actionable framework for governing AI in India. The multi-stakeholder approach identified by the report lays the foundation for robust policy engagement **(Slide 27)**.

2. India boasts a strong ethos of fundamental rights enshrined in the Constitution **(Slide 28)**, which has the potential to bolster its global reputation, and strengthen its position in global fora such as the Global Partnership on Artificial Intelligence (GPAI) **(mentioned on Slide 28).** This could also lead to a unique technological governance strategy that sets India apart from other countries that may seek to further authoritarian narratives around technological governance. The report is the first policy document of its kind that has referenced the Constitution and fundamental rights, while previous documents looking to regulate various aspects of AI or data have merely paid lip service to the more amorphous notions, like 'ethical AI.' By explicitly highlighting the need to derive responsible AI principles from constitutional fundamental rights, the report makes an important contribution. Future research must build on this and think through how this framework will play out with respect to specific use-cases.

3. Although the Report of the AI Task Force mentioned how AI could have a positive impact on accessibility, neither the National Strategy on AI, nor this report look at ways in which AI could help in making services more accessible. There are a number of access and accessibility issues that AI could help in. Some of these include voice interfaces, multilingual support, and easier avenues to accessing government aid. Highlighting the role of AI in accessibility is crucial to ensure the application of principles of inclusivity and non-discrimination **(Mentioned in Slide 29)**, within the goal of "AI for All."

# Sufficiency

1. Since the term Artificial Intelligence is an umbrella term with no universally accepted singular definition, and instead a number of definitions used contextually, it is necessary to set the parameters of what is considered to be 'AI' for the purposes of the report. During CIS's research on the adoption of AI in different sectors, we limited the scope and definition of AI to be a "dynamic learning system that can be used in decision making and actioning." This definition is important to ensure that any regulatory principles are tailored appropriately to the technology being regulated.

2. The envisioning of India as an AI Garage **(Slide 3)** is concerning, especially in the absence of a law or regulation on the use of AI. The essential concern here is that Indian citizens could be "...used as guinea pigs for AI-driven solutions, at the cost

of their fundamental rights." As we have previously stated in CIS's comments on the National Strategy for AI, the absence of a "strong ethical and regulatory framework, the use of new and possibly untested technologies in India could lead to unintended and possibly harmful outcomes." India should instead play a leadership role for emerging economies, in framing the norms that govern AI, and through these norms drive equitable and accessible solutions that benefit the vulnerable.

3. The distinction drawn in the report between direct impact and indirect impact **(Slide 5)** is unclear. "Overall deployment of AI", currently classified as indirect impact, could have a very real and tangible impact on lives. A better distinction to draw would be between direct discrimination where the technology is intentionally used to discriminate against an invidual or community, and disparate impact, where the discriminatory consequences of the use of a technology felt by a group or individual is unintended.

4. The principle of AI not causing harm **(Slide 29)** is a limited view principle. Although harm must be avoided at all costs, governing principles must also be framed from the perspectives of equity, empowerment, agency, openness, and upholding fundamental human rights. Relatedly, AI systems must be designed in a manner to enable fundamental rights rather than force citizens into trade offs, such as between the right to privacy and access to social protection. The report refers to more holistic framing in other parts **(Slide 28)** and this language should reflect in other principles as well.

5. While the introduction of "AI Ethics" into university curriculum is a positive step **(Slide 33)**, ethics is a limited lens to look at 'responsible AI' from. Instead, any curriculum on responsible AI must consider basic courses on constitutionality and empirical research methodology, and foster an understanding of international human rights norms in the AI space.

6. While setting up Ethical Committees is a good idea, there is no clarity on how Ethical Committees will fit into the present regulatory landscape in India, and which level of government they will form a part of **(Slides 36-39)**. Given the complexity and nuances in India's bureaucratic set up, this thinking needs to be done before the final policy sets up Ethics Committees.

7. The impact on jobs through automation **(Slide 24)** must be considered through a lens of inequality in skills and access to technology, as well as occupational segregation. Women workers are very poorly represented in the formal economy, with greater risk, in proportional terms, of job loss and disruption due to automation. Upskilling initiatives must also target such populations (taking into account socioeconomic factors such as caste, income, and geography) who are more likely to be excluded from reaping the benefits of the digital economy.

8. The report must consider the differential access to technology and skills to design and implement AI systems, which also includes the ability to meaningfully avail of grievance redressal systems **(Slide 34).** There are large gaps in digital access in India between rural and urban geographies, and by factors such as gender and age. Women and other marginalised groups are also less likely to be represented among developers of AI systems. These gaps must be addressed through (i) adequate representation of affected stakeholders in design and implementation

processes, and (ii) inclusive grievance redressal systems which do not require access to or knowledge of digital tools.

# Practicality

1. Any use of AI in administrative decisions by the State constitutionally requires the decision making process to also be explained **(Slide 16)**. Mere explainability however is an insufficient standard to govern the use of AI by the State. This is primarily due to the possibility of a power imbalance between the user of the AI (state official) and the impacted individuals and communities. Diverse regulatory spectrums however can potentially remedy some of these problems, especially when dealing with cases where a power asymmetry exists, like with AI used in predictive policing, and where the primary users of AI are also direct beneficiaries, like with AI driven agriculture, where farmers get direct alerts from AI systems. Decision making systems that involve factors such as these must not consider the decision of the AI as final, and instead these decisions must only aid an individual or a body responsible for making the final decision. In use cases where there is high potential for harm and present knowledge cannot predict the effects of that use case , the report must use the precautionary principle and consider a moratorium on AI deployment until its effects can be fully determined. This includes deployment in policing and judicial functions. Once again, this speaks to a broader focus on fundamental principles which not only avoid harm, but prioritise empowerment, agency, and civil liberties.

2. While accountability laws exist in governance systems in India ('The Consumer Protection Act'), a framework for AI accountability is specifically required **(Slide 16).** This conception of accountability however, must go beyond AI systems, and also apply to decisions made by the system, and to organisations using the AI to make decisions. The Algorithmic Accountability Act proposed in New York City is an example of such a legislation that holds individuals, organisations, and processes accountable, rather than the "AI system" itself.

3. While the report acknowledges the growing relevance of the private sector **(Slide 21)**, it does not devote any thought to holding the private sector accountable, especially in cases where AI is deployed through Public-Private partnerships (PPP). As per India's constitutional framework, whenever the State is involved in a 'public function,' it must be held accountable to all citizens for violation of fundamental rights, including in cases where a private sector partner is carrying out the function. With increased proliferation of PPP projects, policy makers need to think through how the contractual relationship is defined. There is a need for clearly drafted contracts with private sector developers which clarifies modes of liability, nature and frequency of audits, and a clarification that source code might need to be made public in case the algorithm is challenged in a court of law. This contractual certainty is also necessary to foster greater private participation in this space, so that the private sector partner has a holistic understanding of risks and obligations involved in the project before it is initiated.

4. There needs to be clarity on who qualifies as the "User of an AI system" **(Slide 9)**. There is a significant difference between the 'primary user' of an AI solution and the impacted stakeholders. For example, the primary user in the case of a predictive policing or facial recognition algorithm would be the police officers

using it but the impacted stakeholders will be individuals and communities impacted by the operation of the technology. The report does not make this nuanced distinction and appears to club stakeholders as groups developing the AI solution, and users as anyone using it, while disregarding the impact on those who are not primary users.

# Annex: References and Bibliography of Research on AI, undertaken by the Centre for Internet and Society

1. The Compendium of Artificial Intelligence in India: https://cis-india.org/internet-governance/blog/artificial-intelligence-in-india-a-compendium
    a. AI in Healthcare Case Study: https://cis-india.org/internet-governance/files/ai-and-healtchare-report
        i. Roundtable Report: https://cis-india.org/internet-governance/blog/ai-andhealthcare-in-india-looking-forward
    b. AI in Manufacturing and IT Services Case Study: https://cis-india.org/internet-governance/files/AIManufacturingandServices_Report_02.pdf/
        i. Roundtable Report: https://cis-india.org/internetgovernance/blog/ai-and-manufacturing-and-services-in-india-looking-forward
    c. AI in Banking and Finance Case Study: https://cis-india.org/internet-governance/files/ai-in-banking-and-finance
        i. Roundtable Report: https://cis-india.org/internet-governance/files/draft-roundtablereport-on-ai-and-banking/view
    d. AI in Governance Case Study: https://cis-india.org/internet-governance/ai-and-governance-case-study-pdf
        i. Roundtable Report: https://cis-india.org/internet-governance/blog/artificial-intelligence-in-governance-a-report-of-the-roundtable-held-in-new-delhi

2. Policy Recommendations:
    a. AI in India: A Policy Agenda: https://cis-india.org/internet-governance/files/ai-in-india-a-policy-agenda/
    b. CIS's response to the Niti Aayog Discussion Paper: https://cis-india.org/internet-governance/blog/niti-aayog-discussion-paper-an-aspirational-step-towards-india2019s-ai-policy
    c. CIS's response to the AI Task Force Report: https://cis-india.org/internet-governance/blog/the-ai-task-force-report-the-first-steps-towards-indias-ai-framework

3. Articles/Publications:
    a. People Driven and Tech Enabled – How AI and ML are Changing the Future of Cyber Security in India: https://cis-india.org/internet-governance/blog/people-driven-and-tech-enabled-2013-how-ai-and-ml-are-changing-the-future-of-cyber-security-in-india
    b. Discrimination in the Age of Artificial Intelligence: https://cis-india.org/internet-governance/blog/oxford-human-rights-hub-arindrajit-basu-october-23-2018-discrimination-in-the-age-of-artificial-intelligence
    c. The Srikrishna Committee Data Protection Bill and Artificial Intelligence in India: https://cis-india.org/internet-governance/blog/the-srikrishna-committee-data-protection-bill-and-artificial-intelligence-in-india
    d. What is the problem with Ethical AI?: An Indian Perspective: https://cis-india.org/internet-governance/blog/what-is-the-problem-with-2018ethical-ai2019-an-indian-perspective
    e. India needs a better AI Vision: https://cis-india.org/internet-governance/blog/fountain-ink-october-12-2019-arindrajit-basu-we-need-a-better-ai-vision
    f. Towards Algorithmic Transparency: https://cis-india.org/internet-governance/algorithmic-transparency-pdf
    g. Use of Algorithmic Techniques for Law Enforcement: https://www.epw.in/journal/2020/23/special-articles/use-algorithmic-techniques-law-enforcement.html
    h. Future of work India's IT/ITeS sector: https://cis-india.org/internet-governance/2018future-of-work2019-in-india2019s-it-it-es-sector-pdf
    i. Future of work in India's Automotive Sector: https://cis-india.org/internet-governance/pdf-automotive-case-study