The 17th Annual Social Informatics Research Symposium and the 3rd Annual Information Ethics and Policy Workshop: Sociotechnical Perspectives on Equity, Inclusion, and Justice (SIG-SI and SIG-IEP)

Critical Analysis of "Responsible AI #AIforAll: Approach Document for India"

Nga Than
City University of New York, USA
nthan@gradcenter.cuny.edu

Abhishek Gupta
Montreal AI Ethics Institute,
Canada
abhishek@montrealethics.ai

Ameen Jauhar
Vidhi Centre
Ameen.jauhar@vidhilegalpolicy.in

INTRODUCTION

Governments across the world have increasingly focused on creating national policy frameworks to take advantage of AI developments for their strategic national interests, as well as to adapt and adjust AI technologies that operate within their socio-cultural and political constraints (Schiff et al 2020). However, most empirical research has mainly utilized AI-related ethics documents produced by governments located in the Global North. In this study, we present a critical analysis of *Responsible AI #AIforAll: Approach Document for India (*thereafter, the Approach Paper*),* a national AI strategy document published by NITI Aayog, a premier public policy think-tank of the Government of India.[1] This document is one of the first of its kind in the Global South. Not only it would serve as an important public policy reference for creating and discussing responsible AI in India, but it also has potential to serve as an exemplary policy document for other developing countries.

We identify and discuss key missing elements in the document such as lack of Indian context, deterministic framing, epistemic incompleteness, and inaccuracies. We conclude with a list of recommendations for improving the process of generating a national strategy document on responsible AI.

LIMITATIONS OF THE APPROACH PAPER

The approach paper briefly describes how AI is being used in India in sectors such as agriculture, health, and medicine. However, most of the discussions in the paper are based on case studies in Global North and the Indian story is presented without much detail. Further, there is limited discussion of challenges in adopting AI in India. We view this as arguably the biggest oversight of this paper. The approach paper attempts to align "responsible AI" to constitutional morality- a concept that envisions the supremacy of the constitution and rights enshrined within it. However, in addition to philosophically "Indianizing" responsible AI, Indian case studies should have driven the identification of potential risks of AI, crafting of principles accordingly, and applying them to mitigate risks and harm in those incidents as a demonstration of the efficacy of the principles.

For instance, while discussing the contentiousness of the growing use of facial recognition technology in law enforcement, the paper delves deeper into Clearview's controversy.[2] This is in lieu of the ongoing integration of facial recognition technology by several state police forces in India (Smriti 2019), and more recently, the Indian government's proposal to create a pan-Indian system for all law enforcement agencies.[3] These efforts have raised alarm bells for the arbitrariness with which facial recognition is being used, and its tenuous legality. These are actual risks posed by AI within India, and a specimen of why local context and examples need to be the central point of such debates. The absence of such India-specific use cases, and parallelly the abundant reliance on Western conceptions, undermines the approach paper's stipulated objective- to evolve responsible AI for India, informed by Indians.

Another point that warrants more elaboration, is the current Indian legislative and policy landscape that may impact the design, development, and deployment of AI. The paper benchmarks its conversation on ethics and responsible usage around the Indian Constitution and rights enshrined therein. However, its engagement with Indian legislation like the Information Technology Act, 2000, is limited. It is devoid of discussions around intellectual property rights,

---

[1] https://niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf
[2] https://www.nytimes.com/interactive/2021/03/18/magazine/facial-recognition-clearview-ai.html
[3] https://www.medianama.com/wp-content/uploads/RFP_NAFRS.pdf

and domain specific legislation that may overlap with the use of AI (e.g., the potential inadequacy of the Criminal Procedure Code in governing the use of AI in policing and law enforcement). What is also lacking is clear actionable recommendations of how statutory and regulatory frameworks can be designed to facilitate responsible and ethical deployment of AI in India (approach paper, pp. 31-32). For instance, the issue of AI liability, which featured to some extent in the earlier national AI strategy document, does not find any mention in this approach paper. Instead, the paper has featured some legislative action in the United States (approach paper, pp. 30-31) and policy documents of the EU, which again makes a point that reads disjointed from the Indian context. Furthermore, it makes vaguely worded reductionist ideas around how AI specific legislation needs to keep pace with technology innovation. The fact that no clear and action-oriented ideas around legislation and regulation, stem from this paper, is indeed a significant lacuna.

A third key deficiency of the paper is what can be loosely termed as a reductionist discourse on what AI ethics entails and how India should prepare itself for the same. As a policy document, while it is imperative to articulate ideas and recommendations lucidly, it is equally crucial to not over-simplify, or sound inflexibly definitive of certain evolving thoughts. Within AI ethics, there is much to be unraveled and determined, yet the approach paper posits a particularly hard coded certainty on what AI ethics are and what they will entail. To that end, we advocate for a more flexible approach that starts with an acknowledgement that definitions and challenges in the space are ever evolving and will hence require constant investment and effort from the policymaking body and group of experts to stay up to date. A key problem that arises from a fixed mindset in this context is that it will lead us down the path determinism problem where the judicial framing of these issues might lock us out of being able to effectively address certain issues because of the choices made early on (as at the time of formulation and discussion of the current document).

The final element that is crucial from a policy standpoint, but fails to emerge from the approach document, is a coherent plan of action vis-a-vis stakeholder engagement, and the precise breakdown of action points at the central, state, and local levels, to responsibly integrate AI into public systems. E.g., the paper at several places acknowledges the need for grassroot level participation, to better the adoption of ethical practices in AI design and deployment. Participatory design principles have been shown to create better outcomes, e.g., in building better COVID-19 apps (Gupta & De Gasperis 2020), and similarly have benefits for more deliberately and meaningfully navigating complex, fast-changing, and uncertain landscapes like AI and AI ethics.

RECOMMENDATIONS

The approach paper gathers literature from many fields of studies such as Human Computer Interaction, Fairness in AI, Robotics, Law and Society. Even though the authors strived for

accessibility ("#AIforAll"), using layman's language, and stayed away from jargon, they should still strive for a high level of epistemological accuracy, i.e., not only the facts have to be correct, but the corresponding theoretical explanation should be plausible. This project is a public policy document that must be scientifically accurate. Without scientific accuracy, the document is removed from scientific discoveries, and could potentially result in harmful policy making. E.g., authors inaccurately state that Uber "stopped" their autonomous vehicle (AV) project after their AV killed a pedestrian, while the project was only paused. Such inaccurate statements can have public policy consequences. We propose that the org establish a review committee, composed of technical and social science experts, to proof-read the document before publishing.

The language must remain accessible to a wider population, but they should consult experts to ensure that both the facts and theoretical knowledge are recorded and explained adequately. Such a committee should be diverse. Its members shouldn't be restricted to "elite" circles nor only comprise of technical experts. Further, the selection process of its member should be transparent.

The approach paper also lacks case studies about India which defeats the purpose of crafting effective public policy proposals for India. Specifically, the document derives mostly from knowledge and case studies from the Global North which has a very different economic, social, and legal infrastructures from India. Consequently, many conclusions do not apply to India. We thus suggest that the authors focus on "indigenous knowledge production," i.e., focusing on case studies from India. This can be achieved by: (1) funding research that studies creation and deployment of AI technology in India, and (2) funding social science research that considers specificities of Indian society and culture.

Stimulating research on AI should be of national interest. Developments in AI are happening quickly and without proper funding and support, India's AI endeavors would not sustain. Recent developments often focus on the technical aspect of AI and AI ethics, while downplaying the contributions of social scientists in studying societal impacts of AI. A provision to stimulate "social science" research that examines impacts of AI in the Indian context is both essential

and necessary. This will require a shift in thinking of the committees and bodies that are created to debate, create, and deploy such policy documents, where traditional focal points continue to remain highly technical universities and traditional academic credentials overlooking the critical role played by social science scholarship and lived experiences of those on the ground who interact with the very people that this document is meant to serve.

Finally, to ensure the rigor of such a public policy document that would shape direct AI ethics discourse in India for years and decades to come, we propose the following steps to craft more effective public policies. India should adopt participatory policy making (Michels & De Graaf 2010), that argues for involving a diverse group of citizens for policy making. Since the tenet of the document is "AI for all," it concerns every citizen, not just policy makers, AI experts, and lawmakers. The common good here is the well-being of everyone who lives in India. Thus, having a transparent process that involves citizens, and different types of stakeholders should only benefit policies, and public support of India's AI discourse and directions in the long run.

We make the following recommendations for conducting participatory public policy making on AI. Firstly, transparency should be emphasized at every step including in the selection of experts,

documentation of their views, and how their views informed the final policy document. It is critical that such policy documents are not unilaterally formulated by the government, or its various think-tanks, but must be done through a consortium of entities to ensure that points which the government is uncomfortable with, or simply doesn't recognize, can still feature in the discussion. Secondly, the document should be impartial and should not become a government's AI marketing pitch. Both experts and ordinary citizens should be involved at every step rather than just a cursory consultation at the end of the formulation of the document at which stage drastic changes (if required) will face severe resistance, both because of the calcified structure and because of practical constraints like budget exhaustion at the end of the lifecycle. Finally, the authors should apply the suggestions in this paper to a case study from India. This would demonstrate how the policy document can be operationalized. Without proper guidance from principal writers, there will be misunderstanding of the suggestions, and polices. This would lead to more confusion, and harms than ensuring that AI is accessible and benefits everyone.

REFERENCES

Boddington, P. (2017). Towards a code of ethics for artificial intelligence. Cham: Springer.

Bostrom, N. (2020). Ethical issues in advanced artificial intelligence.

Cath, C. (2018). Governing artificial intelligence: ethical, legal, and technical opportunities and challenges.

Gupta, A., & De Gasperis, T. (2020). Participatory Design to build better contact-and proximity-tracing apps. arXiv preprint arXiv:2006.00432.

Kohli, P., & Chadha, A. (2019, March). Enabling pedestrian safety using computer vision techniques: A case study of the 2018 uber inc. self-driving car crash. In Future of Information and Communication Conference. Springer, Cham.

Michels, A., & De Graaf, L. (2010). Examining citizen participation: Local participatory policy making and democracy. Local Government Studies, 36(4).

Schiff, D., Biddle, J., Borenstein, J., & Laas, K. (2020, February). What's next for ai ethics, policy, and governance? a global overview. In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society

Smriti Parsheera 'Adoption and regulation of Facial Recognition Technologies in India: Why and why not?', (2019), Data governance network working paper No. 5, available online at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3525324.