NITI Aayog Discussion Paper: An aspirational step towards India's AI policy

Authors: Sunil Abraham, Elonnai Hickok, Amber Sinha, Swaraj Barooah, Shweta Mohandas, Pranav M Bidare, Swagam Dasgupta, Vishnu Ramachandran and Senthil Kumar

Introduction

The National Strategy for Artificial Intelligence — a discussion paper on India's path forward in AI, is a welcome step towards a comprehensive document that reflects the government's AI ambitions.[1] The 115-page discussion paper attempts to be an all encompassing document looking at a host of AI related issues including privacy, security, ethics, fairness, transparency and accountability. The paper identifies five focus areas where AI could have a positive impact in India.[2] It also focuses on reskilling as a response to the potential problem of job loss due the future large-scale adoption of AI in the job market.[3] This blog is a follow up to the comments made by CIS on Twitter[4] on the paper and seeks to reflect on the National Strategy as a well researched AI roadmap for India. In doing so, it identifies areas that can be strengthened and built upon.

Identified Focus Areas for AI Intervention

The paper identifies five focus areas—Healthcare, Agriculture, Education, Smart Cities and Infrastructure, Smart Mobility and Transportation, which Niti Aayog believes will benefit most from the use of AI in bringing about social welfare for the people of India.[5] Although these sectors are essential in the development of a nation, the failure to include manufacturing and services sectors is an oversight. Focussing on manufacturing is fundamental not only in terms of economic development and user base, but also regarding questions of safety and the impact of AI on jobs and economic security. The same holds true for the service sector particularly since AI products are being made for the use of consumers, not just businesses. Use of AI in the services sector also raises critical questions about user privacy and ethics. Another sector the paper fails to include is defense, this is worrying since India is chairing the Group of Governmental Experts on Lethal Autonomous Weapons Systems (LAWS) in 2018.[6] Across sectors, the report fails to look at how AI could be utilised to ensure accessibility and inclusion for the disabled. This is surprising, as aid for the differently abled and accessibility technology was one of the 10 domains identified in the Task Force Report

[1] National Strategy for Artificial Intelligence, http://niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf

[2] ibid pg. 24.

[3] National Strategy for Artificial Intelligence, pg.8.

[4] https://twitter.com/cis_india/status/1003986361261514752

[5] National Strategy for Artificial Intelligence, pg.7.

[6] https://www.unog.ch/80256EE600585943/(httpPages)/7C335E71DFCB29D1C1258243003E8724

on AI published earlier this year. [7]This should have been a focus point in the paper as it aims to identify applications with maximum social impact and inclusion.[8]

In its vision for the use of AI in smart cities, the paper suggests the adoption of a sophisticated surveillance system as well as the use of social media intelligence platforms to check and monitor people's movement both online and offline to maintain public safety.[9] This is at variance with constitutional standards of due process and criminal law principles of reasonable ground and reasonable suspicion. Further, use of such methods will pose issues of judicial inscrutability. From a rights perspective, state surveillance can directly interfere with fundamental rights including privacy, freedom of expression, and freedom of assembly. Privacy organizations around the world have raised concerns regarding the increased public surveillance through the use of AI.[10] Though the paper recognized the impact on privacy that such uses would have, it failed to set a strong and forward looking position on the issue - such as advocating that such surveillance must be lawful and inline with international human rights norms.[11]

Harnessing the Power of AI and Accelerating Research

One of the ways suggested for the proliferation of AI in India was to increase research, both core and applied, to bring about innovation that can be commercialised.[12] In order to attain this goal the paper proposes a two-tier integrated approach: the establishment of CORE s (Centres of Research Excellence in Artificial Intelligence) and ICTAI (International Centre for Transformational Artificial Intelligence).[13] However the roadmap to increase research in AI fails to acknowledge the principles of public funded research such as free and open source software (FOSS), open standards and open data. The report also blames the current Indian Intellectual Property regime for being "unattractive" and averse to incentivising research and adoption of AI.[14] Section 3(k) of Patents Act exempts algorithms from being patented, and the Computer Related Inventions (CRI) Guidelines have faced much controversy over the patentability of mere software without a novel hardware component.[15] The paper provides no concrete answers to the question of whether it should be permissible to patent algorithms, and if yes, to to what extent. Furthermore, there needs to be a standard either in the CRI Guidelines or the Patent Act, that distinguishes between AI algorithms and non-AI

[7] Report of the AI Task Force, pg 21.
[8] National Strategy for Artificial Intelligence, pg.5.
[9] Ibid pg 40.
[10]The National Strategy For AI http://niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf, p. 40
[11] CIS's submission to the stated that any surveillance must comply with the International Principles on the Application of Human Rights to Communications, such as Legality,Legitimate Aim,Necessity,Adequacy and Proportionality, etc. https://cis-india.org/internet-governance/files/data-protection-submission
[12] The National Strategy For AI, pg. 50.
[13] *Centre of Research Excellence (CORE) will* focus on developing greater understanding of existing core research and pushing technology frontiers through creation of new knowledge while the *International Centers of Transformational AI (ICTAI)* with a mandate of developing and deploying application-based research. Private sector collaboration is envisioned to be a key aspect of ICTAIs.
[14] The National Strategy For AI, pg. 46
[15]https://spicyip.com/2017/07/patent-office-reboots-cri-guidelines-yet-again-removes-novel-hardware-requirement.html

algorithms. Additionally, given that there is no historical precedence on the requirement of patent rights to incentivise creation of AI, innovative investment protection mechanisms that have lesser negative externalities, such as compensatory liability regimes[16] would be more desirable. The report further failed to look at the issue holistically and recognize that facilitating rampant patenting can form a barrier to smaller companies from using or developing AI. This is important to be cognizant of given the central role of startups to the AI ecosystem in India and because it can work against the larger goal of inclusion articulated by the report.

Ethics, Privacy, Security and Safety

In a positive step forward, the paper addresses a broader range of ethical issues concerning AI including transparency, fairness, privacy and security and safety in more detail when compared to the earlier report of the Task Force.[17] Yet despite a dedicated section covering these issues, a number of concerns still remain unanswered.

Transparency

The section on transparency and opening the Black Box has several lacunae.[18] First, AI that is used by the government, to an acceptable extent, must be available in the public domain for audit, if not under Free and Open Source Software (FOSS). This should hold true in particular for uses that impinge on fundamental rights. Second, if the AI is utilised in the private sector, there currently exists a right to reverse engineer within the Indian Copyright Act,[19] which is not accounted for in the paper. Furthermore, if the AI was involved both in the commission of a crime or the violation of human rights, or in the investigations of such transgressions, questions with regard to judicial scrutability of the AI remain. In addition to explainability, the source code must be made circumstantially available, since explainable AI[20] alone cannot solve all the problems of transparency. In addition to availability of source code and explainability, a greater discussion is needed about the tradeoff between a complex and potentially more accurate AI system (with more layers and nodes) vs. an AI system which is potentially not as accurate but is able to provide a human readable explanation.[21] It is interesting to note that transparency within human-AI interaction is absent in the paper. Key questions on transparency, such as whether an AI should disclose its identity to a human have not been answered.

Fairness

With regards to fairness, the paper mentions how AI can amplify bias in data and create unfair outcomes.[22] However, the paper neither suggests detailed or satisfactory solutions nor

[16]This model has been suggested in microbial research stating "The economic logic underlying this model is that the providers of microbial materials would presumably obtain more potential reciprocity benefits from the vast upstream research opportunities generated by the semicommons than would accrue from operating in isolation." https://www.ncbi.nlm.nih.gov/books/NBK92720/

[17]Report of the AI Task Force http://dipp.nic.in/whats-new/report-task-force-artificial-intelligence

[18] The National Strategy For AI, pg. 63

[19] Section 107 of the Indian Copyrights Act

[20] Ibid pg. 86

[21] Similarity Cracks the Code Of Explainable AI https://simmachines.com/similarity-cracks-code-explainable-ai/

[22] Ibid pg. 85

does it deal with biased historical data in an Indian context. More specifically, there seems to be no mention of regulatory tools to tackle the problem of fairness, such as:

- Self-certification
- Certification by a self-regulatory body
- Discrimination impact assessments
- Investigations by the privacy regulator

Such tools will proactively need to ensure inclusion, diversity, and equity in composition and decisions.[23]

Additionally, with reference to correcting bias in AI, it should be noted that the technocratic view that as an AI solution continues to be trained on larger amounts of data , systems will self correct, does not fully recognize the importance of data quality and data curation, and is inconsistent with fundamental rights. Policy objectives of AI innovation must be technologically nuanced and cannot be at the cost of intermediary denial of rights and services.

Further, the paper does not deal with issues of multiple definitions and principles of fairness, and that building definitions into AI systems may often involve choosing one definition over the other. For instance, it can be argued that the set of AI ethical principles articulated by Google[24] are more consequentialist in nature involving a a cost-benefit analysis, whereas a human rights approach may be more deontological in nature. In this regard, there is a need for interdisciplinary research involving computer scientists, statisticians, ethicists and lawyers.

Privacy

Though the paper underscores the importance of privacy and the need for a privacy legislation in India - the paper limits the potential privacy concerns arising from AI to collection, inappropriate use of data, personal discrimination, unfair gain from insights derived from consumer data (the solution being to explain to consumers about the value they as consumers gain from this), and unfair competitive advantage by collecting mass amounts of data (which is not directly related to privacy).[25] In this way the paper fails to discuss the full implications on privacy that AI might have and fails to address the data rights necessary to enable the right to privacy in a society where AI is pervasive. The paper fails to engage with emerging principles from data protection such as right to explanation and right to opt-out of automated processing, which directly relate to AI. Further, there is no

[23] The Toronto Declaration notes: "intentional and inadvertent discriminatory inputs throughout the design, development and, use of machine learning systems create serious risks for human rights; systems are for the most part developed, applied and reviewed by actors which are largely based in particular countries and regions, with limited input from diverse groups in terms of race, culture, gender, and socio-economic backgrounds. This can produce discriminatory results." See: https://www.accessnow.org/the-toronto-declaration-protecting-the-rights-to-equality-and-non-discrimination-in-machine-learning-systems/

[24] AI at Google: our principles, https://blog.google/topics/ai/ai-principles/

[25] The National Strategy For AI, pg. 87 http://www.niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf

discussion on the issues such as data minimisation and purpose limitation which some big data and AI proponents argue against. To that extent, there is a lack of appreciation of the difficult policy questions concerning privacy and AI. The paper is also completely silent on redress and remedy. Further the paper endorses the seven data protection principles postulated by the Justice Srikrishna Committee.[26] However CIS has pointed out that these principles are generic and not specific to data protection.[27] Moreover, the law chapter of IEEE's '*Global Initiative on Ethics of Autonomous and Intelligent Systems*'[28] has been ignored in favor of the chapter on '*Personal Data and Individual Access Control in Ethically Aligned Design*'[29] as the recommended international standard.[30] Ideally, both chapters should be recommended for a holistic approach to the issue of ethics and privacy with respect to AI.

AI Regulation and Sectoral Standards

The discussion paper's approach towards sectoral regulation advocates collaboration with industry to formulate regulatory frameworks for each sector. However, the paper is silent on the possibility of reviewing existing sectoral regulation to understand if they require amending. We believe that this is an important solution to consider since amending existing regulation and standards often takes less time than formulating and implementing new regulatory frameworks.[31] Furthermore, although the emphasis on awareness in the paper is welcome, it must complement regulation and be driven by all stakeholders, especially given India's limited regulatory budget. The over reliance on industry self-regulation, by itself, is not advisable, as there is an absence of robust industry governance bodies in India and self-regulation raises questions about the strength and enforceability of such practices. The privacy debate in India has recognized this and reports, like the Report of the Group of Experts on Privacy, recommend a co-regulatory framework with industry developing binding standards that are inline with the national privacy law and that are approved and enforced by the Privacy Commissioner.[32] That said, the UN Guiding Principles on Business and Human Rights and its "protect, respect, and remedy" framework should guide any self regulatory action.[33]

Security and Safety of AI Systems

In terms of security and safety of AI systems the paper seeks to shift the discussion of accountability being primarily about liability, to that of one about the explainability of AI.[34] Furthermore, there is no recommendation of immunities or incentives for whistleblowers or researchers to report on privacy breaches and vulnerabilities. The report also does not recognize certain uses of AI as being more critical than others because of their potential harm to the human. This would include uses in healthcare and autonomous transportation. A

---

[26] Submission to the Committee of Experts on a Data Protection Framework for India https://cis-india.org/internet-governance/files/data-protection-submission

[27] ibid.

[28] http://standards.ieee.org/news/2017/ieee_global_initiative.html

[29] https://standards.ieee.org/develop/indconn/ec/ead_personal_data_v2.pdf

[30] The National Strategy For AI, pg. 88

[31] Ibid pg 87

[32] See pg. 58 http://planningcommission.nic.in/reports/genrep/rep_privacy.pdf

[33] Guiding Principles on Business and Human Rights http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf

[34] The National Strategy For AI, pg. 85

key component of accountability in these sectors will be the evolution of appropriate testing and quality assurance standards. Only then, should safe harbours be discussed as an extension of the negligence test for damages caused by AI software. Additionally, the paper fails to recommend kill switches, which should be mandatory for all kinetic AI systems.[35] Finally, there is no mention of mandatory human-in-the-loop in all systems where there are significant risks to safety and human rights. Autonomous AI is only viewed as an economic boost, but its potential risks have not been explored sufficiently. A welcome recommendation would be for all autonomous AI to go through human rights impact assessments.

Research and Education

Being a government think-tank, the NITI Aayog could have dealt in detail with the AI policies of the government and looked at how different arms of the government are aiming to leverage AI and tackle the problems arising out of the use of AI. Instead of tabulating the government's role in each area and especially research, the report could have also listed out the various areas where each department could play a role in the AI ecosystem through regulation, education, funding research etc. In terms of the recommendations for introducing AI curriculums in schools, and colleges,[36] the government could also ensure that ethics and rights are part of the curriculum - especially in technical institutions. A possible course of action could include corporations paying for a pan-Indian AI education campaign.This would also require the government to formulate the required academic curriculum that is updated to include rights and ethics.

Data Standards and Data Sharing

Based on the amount of data the Government of India collects through its numerous schemes, it has the potential to be the largest aggregator of data specific to India. However the paper does not consider the use of this data with enough gravity. For example, the paper recommends Corporate Data Sharing for "social good" and making government datasets from the social sector available publicly.[37] Yet this section does not mention privacy enhancing technologies/standards such as pseudonymization, anonymization standards, differential privacy etc. Additionally there should be provisions that allow the government to prevent the formation of monopolies by regulating companies from hoarding user data. The open data standards could also be applicable to the private companies, so that they can also share their data in compliance with the privacy enhancing technologies mentioned above. The paper also acknowledges that AI Marketplaces require monitoring and maintenance of quality. It recognises the need for "continuous scrutiny of products, sellers and buyers"[38], and proposes that the government enable these regulations in a manner that private players could set up the marketplace. This is a welcome suggestion, but the legal and ethical framework of the AI Marketplace requires further discussion and clarification.

An AI Garage for Emerging Economies

---

[35] Google DeepMind Researchers Developing A.I. Kill Switch, Just In Case https://www.forbes.com/sites/curtissilver/2016/06/07/google-deepmind-researchers-developing-a-i-kill-switch-just-in-case/#6f96077f952b

[36] The National Strategy For AI, pg. 92

[37] Ibid. Pg. 112

[38] The National Strategy For AI, pg. 82

The discussion paper also qualifies India as an "ideal test-bed"[39] for trying out AI related solutions. This is problematic since questions of regulation in India with respect to AI have yet to be legally clarified and defined and India does not have a comprehensive privacy law. Without a strong ethical and regulatory framework, the use of new and possibly untested technologies in India could lead to unintended and possibly harmful outcomes.The government's ambition to position India as a leader amongst developing countries on AI related issues should not be achieved by using Indians as test subjects for technologies whose effects are unknown.

Conclusion

In conclusion, NITI Aayog's discussion paper represents a welcome step towards a comprehensive AI strategy for India. However, the trend of inconspicuously releasing reports (this and the AI Task Force) as well as the lack of a call for public comments, seems to be the wrong way to foster discussion on emerging technologies that will be as pervasive as AI. The blanket recommendations were provided without looking at its viability in each sector.[40] Furthermore, the discussion paper does not sufficiently explore or, at times, completely omits key areas. It barely touched upon societal, cultural and sectoral challenges to the adoption of AI — research that CIS is currently in the process of undertaking.[41]Future reports on Indian AI strategy should pay more attention to the country's unique legal context and to possible defense applications and take the opportunity to establish a forward looking, human rights respecting, and holistic position in global discourse and developments. Reports should also consider infrastructure investment as an important prerequisite for AI development and deployment. Digitised data and connectivity as well as more basic infrastructure, such as rural electricity and well-maintained roads, require more funding to more successfully leverage AI for inclusive economic growth. Although there are important concerns, the discussion paper is an aspirational step toward India's AI strategy.

****

[39] Ibid .Pg. 6

[40] Eg. what works for healthcare might not work for agriculture.

[41] Artificial Intelligence in India a Compendium https://cis-india.org/internet-governance/blog/artificial-intelligence-in-india-a-compendium