Draft National Artificial Intelligence Policy 2023

Digital Rights Foundation

Feedback to the Ministry of Information Technology and Telecommunication

July 6, 2023
https://digitalrightsfoundation.pk/
For more details contact info@digitalrightsfoundation.pk

About

Digital Rights Foundation (DRF) is a not-for-profit, women-led organization based in Pakistan working on digital freedoms. DRF envisions a place where all people, especially women and gender minorities, can exercise their right to access and participate in digital spaces without being threatened. DRF focuses on information and communication technologies to support women's rights, inclusiveness, democratic processes, and digital governance.

EXECUTIVE SUMMARY

The following analysis and feedback on the recently introduced draft of the National Artificial Intelligence Policy (2023) published by the Ministry of Information Technology and Telecommunications (MOITT)[1] seeks to point out lacunas in the current iteration of the policy. The major takeaways from this analysis are:

- The need for more transparency during the consultative process and the necessity of including input from civil society at the formation/drafting stage of the policy to ensure that human rights and citizen-forward perspectives are not overlooked
- At present, the current Personal Data Protection Bill falls short of anticipating the unique privacy and security issues raised by AI and emerging technologies, including a lack of consideration of the privacy aspects of the datasets employed by, used to train and develop AI.
- The policy does not sufficiently incorporate the following defining characteristics: (i) grounded in human rights standards and norms, and (ii) inclusive and centring marginalized communities.
- The section-by-section comments point to specific areas of the policy that overlook key considerations, come across as overly vague and identify policies that require more clarity, detail and specificity.
- An overview of international best practices details the best and most unique aspects of current global AI regulatory regimes, including those in development in the EU, USA, China and India - regimes that are most pertinent to the Pakistani context. Some best practices the policy should consider include system safety obligations; a list of explicitly prohibited AI practices; obligatory human oversight over AI processes; mandatory risk assessments and adherence to AI ethics; obligations on the users and manufacturers; and regulatory requirements of AI regulatory sandboxes.
- A section detailing current legislation in Pakistan that is applicable/relevant to the AI regulatory framework i.e. the Personal Data Protection Bill (2023), the Prevention of Electronic Crimes Act (2016) and SECP Regulatory Sandbox Guidelines (2019) highlights gaps where these laws are silent and that a future AI policy must cover. These include updating definitions to factor in AI technologies, conditions around liability and applicability of caveats in various domestic laws across different sectors, such as finance, health, and intellectual property – among others – to AI technologies.

[1] The policy and call for feedback: https://moitt.gov.pk/Detail/ZTM4NmI3MDAtZmM0OC00MzJlLThhODAtMWVhNWE4MmJmMDU5.

The Digital Rights Foundation's key recommendations for the National AI Policy cover:

- The importance of AI awareness and digital literacy for all citizens, particularly marginalised areas and communities who experience the digital divide.
- The need for explicit criteria for the use of AI, defining unethical and risky applications, especially in places that can result in gross discrimination against vulnerable communities like employment, law enforcement, etc.
- The AI policy should be human rights-centric, reflecting international best practices and Pakistan's international law obligations, and should prioritize the preservation of citizens' rights over institutions.
- Proposed AI regulatory bodies should have clearly defined ambits and duties, not overlap in function, be inclusive, and independent in their operations.
- Inclusion of mandatory human rights audits at the design and development stage of AI technologies with subsequent and ongoing impact assessments across all sectors deploying AI technologies.
- An emphasis on non-discrimination through transparency, accountability, the ability to "opt-out" of AI-based decision making and grievance redressal mechanisms available to the public for the harmful, negligent or inappropriate use of AI.

A. INTRODUCTION

This feedback by Digital Rights Foundation (DRF) seeks to provide procedural and substantive feedback on the **'Draft National Artificial Intelligence Policy**' (henceforth **'Draft Policy**') made public in May 2023. The feedback document has been structured in the same order as the Draft Policy document and includes recommendations from a civil society and human rights perspective. We appreciate the fact that the Ministry of Information Technology and Telecommunications (MOITT) has made the Draft Policy open to public comment and hope that it will continue the consultation and drafting process in an open, transparent and inclusive manner.

B. CONSULTATION PROCESS

The Draft Policy states that it is based "on the evidence collected through more extensive consultations with the stakeholders" as well as a global review of best practices. While the consultative approach is laudable, as is the decision by the MOITT to seek public comment, the consultations have not been inclusive enough to account for a diverse range of stakeholders. The consultations for this Draft Policy were not open to civil society, which has resulted in a lack of focus on human rights concerns within the document. Furthermore, the Draft Policy does not disclose specifics of the consultation process, such as the mode of consultation, the number of meetings conducted and the stakeholders who participated.

We urge the MOITT to make the drafting more inclusive and transparent through the following measures: 1) commit to conducting another round of consultations; 2) take affirmative steps to include civil society actors and stakeholders from marginalised communities; 3) co-lead the consultative process with independent civil society; and 4) make all comments submitted to MOITT public.

It is laudable and extremely important that the Draft Policy seeks to "fundamentally rethink AI adoption in the local context", given that policies need to be grounded in the local context of Pakistan. However, despite its aspirations, it falls short on its own terms. There is an absence of an evidence-based understanding of the unique risks and human rights concerns in the Pakistani context. In fact, the Draft Policy itself points towards the need for interdisciplinary and multi-stakeholder risk assessment to be conducted regarding the potential risks that AI technology poses in Pakistan, particularly in the context of human rights, information integrity and democratic governance. We recommend that an

independent Steering Committee be constituted led by civil society, local industry leaders and academics with the mandate to carry out this assessment and make recommendations to the MOITT. The selection process of this Steering Committee should be based on clearly defined and public criteria with safeguards to ensure its independence. The independent assessment should be made public, as well as the response by the MOITT to the recommendations addressed to it. This process is an essential stepping stone to ensuring that the AI Policy is responsive to the local context and based on input from diverse stakeholders in society. This process will only serve to strengthen the legitimacy and efficacy of the proposed policy.

C. DATA PROTECTION

One essential component of AI is datasets that these technologies draw on in order to train them for real-world applications. In fact, in many cases, it has been posited that AI is only as good as the data it is trained on. This raises two important aspects in the context of the data used to develop, train and make AI functional: a) the privacy and security of the datasets employed by AI; and b) whether the datasets are representative of the Pakistani context, particularly historically excluded and marginalised communities.

The proposed '**Personal Data Protection Bill 2023**'[2] (henceforth '**PDPB 2023**') falls short of anticipating the particular privacy issues raised by AI and emerging technology. Section 29(4) deals with the harms related to "automated processing":

> "The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which results in legal obligations or significantly harms the data subject, unless the data subject has given his explicit consent."

However, it is concerning that the PDPB makes an exception for automated decision-making carried out in the "public interest". Given extensive human rights concerns related to the use of AI-powered automated systems employed by law enforcement and public authorities, which has led to tangible discrimination, exclusion and profiling across the world. This is why regulations regarding AI, such as the EU AI Act, regarding facial recognition, have categorised such uses of AI as "high risk" and imposed complete moratoriums on such usage.[3] It is important that both the Draft Policy and the PDPB reflect these realities.

Additionally, requirements for data localisation regarding Critical Personal Data (Section 31 of PDPB) raise concerns regarding market innovations required for the free flow of data across borders to allow for innovation in fields such as AI. Impediments to data flows anticipated by the PDPB undermine the noble goals of innovation and enabling environment envisioned in this Draft Policy.

---

[2] Latest version of the Bill as of May 2023: https://moitt.gov.pk/SiteImage/Misc/files/Final%20Draft%20Personal%20Data%20Protection%20Bill%20May%202023.pdf.

[3] Jorge Liboreiro and Aida Sanchez Alonso, "MEPs endorse blanket ban on live facial recognition in public spaces, rejecting targeted exemptions," *EuroNews*, May 14, 2023, https://www.euronews.com/my-europe/2023/06/14/meps-endorse-blanket-ban-on-facial-recognition-in-public-spaces-rejecting-targeted-exempti .

D. DEFINING ATTRIBUTES

The Policy posits three, namely "Evidence-Based and Target Oriented; User-Centric and Forward-Looking; [and] Objective and Overarching", attributes as foundational to its focus. It is recommended that the following two attributes be added to the Policy document to make it more comprehensive:

1) Grounded in Human Rights Standards and Norms; and

2) Inclusive and Centering of Marginalised Communities.

For elaboration, standard-setting documents such as **UNESCO's 'Recommendation on the ethics of artificial intelligence'**[4] lay out ethical AI principles which are recommended to be made part of this policy to ensure compliance with international best practices:

I. Proportionality and Do No Harm
II. Safety and security
III. Fairness and non-discrimination
IV. Sustainability
V. Right to Privacy and data protection
VI. Human (and inclusive public) oversight and determination
VII. Transparency and explainability
VIII. Responsibility and accountability
IX. Awareness and literacy
X. Multi-stakeholder and adaptive governance and collaboration

[4] "Recommendation on the ethics of artificial intelligence," UNESCO, https://unesdoc.unesco.org/ark:/48223/pf0000380455/PDF/380455eng.pdf.multi .

E. SECTION BY SECTION COMMENTS

Sec.	Heading	Pg No.	Body Text	Comment(s)
2.2	The State of AI in Pakistan	7	The policy proposes the National AI fund as a perpetual and central fund to support the proposed interventions.	There have been too many "funds" established that have no accountability; lacking transparency and mechanisms for sustainability. These funds often become defunct, misused, evade audits, and lack the appropriate system of checks and balances. Strict guidelines to ensure transparency, independence and sustainability of the proposed fund need to be laid out, along with its functions.
3.4.2	The National AI Targets	10	Target-1: Public Awareness of AI & Allied Technologies	The timeline for public awareness needs to be accelerated since the application of AI is becoming increasingly common, awareness campaigns are essential for preparedness. Additionally, awareness and education-based interventions should combine ethics with skill-based education, particularly education programs geared towards public functionaries.
3.4.2	The National AI Targets	13	Target-9: Proliferating AI and Allied Technologies Responsibly – Current State: The National Commission for Data Protection under	At the time of publishing the AI Policy, Pakistan does not have a Data Protection law in place. As a result, because the NCPDP (National Commission for Personal Data Protection) has not been formulated yet, it is

			Personal Data Protection Act is yet to be established. There is no provision for regulating AI at the moment. However, the NCPDP closely relates in terms of Data Protection Desired State: Establishing an AI Directorate under NCPDP for harnessing AI responsibly through appropriate and need-based regulations	premature to determine whether it would be effective at carrying out its mandate and be able to support an AI Directorate. The powers of the Commission under the current draft fail to provide adequate independence and autonomy from the federal government (see DRF's analysis of the Bill). The policy lacks clear parameters to determine whether the use and proliferation of AI is responsible in line with national and international goals.
		14	Target-11: Harnessing AI through Global Best Practices	International collaborations and adoption of best practices is extremely important, particularly for a developing economy like Pakistan's, however, it is important for any collaboration and adaption to be centered on human rights principles to avoid replication of bad and unethical AI practices into Pakistan. Use of AI to surveil, monitor and discriminate which have manifested in use of facial recognition technology against citizens should not be adopted from other countries where

				these practices are in place.[5]
		14-15	Target-13: Transforming the Public Sector through AI & Allied Technologies	The target refers to the usability and availability of datasets held by public authorities for AI. It is important to raise questions regarding privacy, consent and intellectual property issues that are raised by such usage. For instance, are citizens consenting for their data in the NADRA database to be used to train AI? Are there safeguards to ensure informed consent is taken, and can citizens refuse to be included in AI datasets?
4.1.3.1	Data Standardization and Aggregation for Servicification	18 (incorrectly numbered 3 on the document)	The National and Provincial governments possess heterogeneous datasets through various verifiable and unverifiable sources and require orchestration bottom-up for secure and transparent use. Suppose the social data is standardized and made available by the private sector for public service provisioning in a structured environment. In that case, it can help reduce time to service due to process cluttering and ensure seamless application of AI and allied technologies in an	The section lacks clarification on the following questions: what are the safety protocols in place for such an exhaustive system of private-public data sharing/what protections does the center for excellence in AI have? The CoE-AI should have a mandated protocol in place that stipulates a base standard for data security, privacy and citizen's rights. Furthermore, the assumption that computer vision-based surveillance enhances surveillance of citizens should not be taken at face value, in fact no feasibility study has

[5] Potential points of concern regarding use of AI in urban surveillance without any guardrails or human rights audits. (Reference: 'K-P police launch first AI security control system in Pakistan,' The Express Tribune, June 27, 2023, https://tribune.com.pk/story/2423824/k-p-police-launch-first-ai-security-control-system-in-pakistan).

			integrated manner. I. The CoE-AI shall organize a common and sectoral data collection and processing mechanism for unstructured, semi-structured, and structured datasets for the data available in the public sector and accessible through market regulators for the private sector. … VIII. Smart city-based projects are already in the deployment phase in several cities in Pakistan. CoEAI shall support in indigenization of computer vision technology by standardization targeting high accuracy of person and object detection. Computer Vision-based surveillance application poses significant potential to contribute to the safety of citizens. Therefore, CoE-AI shall accelerate the technology rollout to even smaller cities in Pakistan.?	been published by the Safe City Authorities in the country to suggest a correlation between increased urban digital surveillance and safety. Before such recommendations are made, an independent audit of the Safe City Authority needs to be conducted, particularly focusing on safety of women, gender minorities and marginalised communities along with fundamental rights such as privacy and freedom of assembly.
4.1.3.2	National Health Services Transformation using AI	19	(1) CoE-AI shall support the Ministry of National Health Services Regulation and Coordination in better controlling chronic diseases such as diabetes, hypertension, and high blood cholesterol. CoE-AI shall develop guidelines for	Makes an overestimation of current AI capacities, pushes towards over-reliance on AI in a sector that should be more human-regulated rather than automated. It is unclear how AI can help in “better controlling” chronic conditions such as

			healthcare providers to become more productive and help patients control chronic disease conditions using the latest AI technology.	diabetes. Sets a vague expectation of the application of AI in the Health Services sector. Instead, the Draft Policy should focus more on safeguards that should be implemented to prevent the abuse and misuse of AI in the health space. Furthermore, when taking important decisions regarding healthcare on the basis of aggregated or digitised data, extra care needs to be taken to ensure that the data is representative. For instance, ensure that data from private sector hospitals or urban areas is not over-represented in the healthcare data used for automated decision-making.
4.1.3.3	Intelligent Learning and Assessment using AI	20	Every individual in society has a different capacity to comprehend and retain new information and level of skill in a particular trade. Also, they have different aptitudes and learning needs. Therefore, teachers cannot personalize every individual's learning experience, especially in the subjects where a particular skill is to be taught.	As in the section above, the Draft Policy seeks to solve structural issues, such as lack of resources in the education sector, through AI systems. Catering to different aptitudes and learning needs through AI can be beneficial, but also runs the risk of relying on faulty data regarding issues such as learning disabilities or perpetuating harmful notions regarding learning abilities. If this Draft Policy professes to anticipate harmful AI uses, then

				it needs to account for these risks in outlining these applications.
4.2.1	Public Awareness of AI	22	(1)To empower society to make knowledge-based decisions for personal data sharing and a basic understanding of technologies such as AI, the Ministry of IT & Telecom shall devise a National Awareness Program for Personal AI adaptation. (3)The Ministry of IT & Telecom, through the CoE-AI, shall orchestrate important messages/content such as technical write ups, breach and implication scenarios, preventive and remedial measures, and any other details necessary for informing/educating the citizens. (8) CoE-AI shall devise a national initiative for explainable AI increasing the public's confidence in AIbased services and solutions to improve societal acceptance and modernization.	Emphasis on public awareness is a good initiative, however, the section makes reference to both a “National Awareness Program for Personal AI adaptation” and a “national initiative for explainable AI increasing the public's confidence in AIbased services” – what is the difference between these two initiatives? Also the roles of all relevant regulatory authorities are too broad and overlapping. For example, the MOITT, the CoE-AI and the ADR. The Draft Policy should also envision a role for civil society to participate in public education campaigns. Civil society and community-based organisations often have relationships of trust with communities and can deliver awareness campaigns at the grassroots level more effectively than top-down approaches by the state. Partnerships with civil society organisations are an important channel of communication for dissemination of education on technical subjects such as AI.

4.2.1 (v)	Public Awareness of AI	23	(5) While developing the content, special attention must be given to the silver segment of society, the marginalized section of women and Persons with Disabilities (PWDs).	This is an important consideration, resources should be invested to ensure that accessible content reaches these communities.
4.3.1 (I)	Regulating to Accelerate Socio-Economic Adoption	26	(1) An AI Regulatory Directorate (ARD) shall be constituted under the National Commission for Personal Data Protection (NCPDP), invoking function (i) of 33.2 from the PDPA Act that calls for monitoring of technological developments and commercial practices. It may affect personal protection data and promote measures and undertake research for innovation in personal data protection.	Reference to an Act that is not finalized. The current Personal Data Protection Bill does not contain a framework of how the ARD will be constituted, its membership, and how it will work in coordination with the CoE-AI.
4.3.1	Regulating to Accelerate Socio-Economic Adoption	27	(13) Develop a data-sharing framework and use AI algorithms consistent with social, cultural, and religious norms and international guidelines.	Algorithms being consistent with social, cultural and religious norms are important, however, on the other hand, have the potential to perpetuate parochial biases that prove counter-intuitive for a progressive AI environment. Unless aligned with human rights standards, it can go against sections of the policy that explicitly talks about discrimination caused by AI. It also establishes vague criteria as to what those norms entail which can be prone to abuse.

4.3.1	Regulating to Accelerate Socio-Economic Adoption	27	(9) Encourage local businesses to embrace new AI solutions and provide them with a platform for technical support and some incentives and regulations. Moreover, it should catalyze the creation of new businesses based on AI technology through start-up funds and incubation centers.	At various points in the Draft Policy, there is mention of providing incentives to companies and research institutions for the development of innovative AI products, however, what these incentives look like is never detailed.
4.3.2 (I)	Generative AI	28	Ensuring Ethical Use: Generative AI has the potential to create convincing fake content such as text, images, and videos. Therefore, ARD shall provide regulatory guidelines to address the possible spread of disinformation, data privacy breaches, or fake news.	Ethical use of generative AI is an important issue to address, however, policies and guidelines that target disinformation, fake news etc. should center interventions based on user education and awareness rather than speech-restrictive laws.
4.4.2	Industrial Transformation	30	Industries are turning any country's economic wheels, and their optimization help contributes substantially to GDP. AI and allied technologies can augment industries' capacity building by introducing State-of-the-art technologies such as IoTs and enhancing their efficiency and productivity.	The introduction of AI into industries needs to be done in a labour-centered manner with sufficient provision for economic disruption it can cause in the lives of workers. Potential impact can include, reduced wages, "undermining worker agency as larger numbers of workers compete for deskilled," unemployment, underemployment or precarious employment, and allocation of "the most fulfilling

				tasks in some jobs to algorithms, leaving humans with the remaining drudgery."[6] Similarly, AI has been introduced in industries to increase worker surveillance,[7] reduce break and collectivisation, which impact the physical and mental well-being of workers. These concerns of worker welfare need to be addressed alongside the application of AI to industries, not as an aside. Lastly, the supply chain of AI-related labour needs to be accounted for to ensure the application of fair labour standards. Ethical questions of whether the AI system relies on outsourced labour are important as countries like Pakistan are often on the receiving end of cheap, outsourced labour from the international tech industry. The invisible labour of training datasets is extremely labour-intensive and should be fairly compensated.[8]

[6] "Guidelines for AI and Shared Prosperity: Tools for improving AI's impact on jobs," Partnership on AI, June 2023, https://partnershiponai.org/wp-content/uploads/dlm_uploads/2023/06/pai_guidelines_shared_prosperity.pdf, p. 7.

[7] Jodi Kantor and Arya Sundaram (Produced by Aliza Aufrichtig and Rumsey Taylor), "The Rise of the Worker Productivity Score," The New York Times, August 14, 2022, https://www.nytimes.com/interactive/2022/08/14/business/worker-productivity-tracking.html.

[8] Ibid, p. 23.

4.4.3	Public/Private Sector Evolution	30	COE-AI and ARD need to work in synchronization for a relatively smooth process of AI adoption in the public and private sectors. They will ensure technical, computational, regulatory, and financial assistance as and when required. Furthermore, the National Artificial Intelligence Fund (NAIF), with help from the private sector and international bodies, will fuel the research and innovative endeavors for successfully integrating AI and associated technologies into Pakistan's institutional and industrial fabric. These collaborative investments will bear fruit in commercializing various products produced by the COE-AIenabled AI ecosystem.	Role of each agency/board should be clearly defined. As they currently stand, the functions overlap.

F. OVERVIEW OF INTERNATIONAL BEST PRACTICES

1. European Union

The **EU AI Act** is presently considered the leading human rights-centric legislation of AI. Unlike its US and Chinese counterparts, which are attempting to strike a balance between corporate interest, national security and human rights, the EU regulation places a firm emphasis on preserving fundamental rights through comprehensive guidelines for the use of AI and the obligations therewith for the producers/users/importers. The primary aim of the legislation is to ensure transparency, accountability, safety in the use of AI technology as well as the promotion of non-discrimination and environmental sustainability.[9] The Act includes features such as heavy transparency requirements for AI and strict restrictions on biometric surveillance, facial recognition, generative AI and predictive policing systems. It also categorises AI in terms of low risk, high risk or unacceptable risk; and gives citizens the right to report complaints against AI and seek alternatives to AI.[10]

Of its many provisions, some notable highlights of the Act include:

- Systemic safety obligations

In its preamble, the Act calls attention to "technical robustness" as "a key requirement for high-risk AI systems." It requires that high-risk AI systems "should be resilient against risks connected to the limitations of the system (e.g. errors, faults, inconsistencies, unexpected situations) as well as against malicious actions that may compromise the security of the AI system and result in harmful or otherwise undesirable behavior."[11]

- Prohibited AI practices

Article 5 of Title 2 of the Act comprehensively lists down a schedule of prohibited AI practices that include various discriminatory practices,[12] subliminal messaging,[13] the use of

[9] "AI Act: a step closer to the first rules on Artificial Intelligence" Press Release (European Parliament News, May 2023) https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence

[10] Ibid

[11] Preamble No. 50, EU Artificial Intelligence (AI) Act, 2023, p. 30 https://artificialintelligenceact.eu/the-act/

[12] Title 2, Article 5, Section(s) 1(b) and 1(c), EU AI Act, 2023, p. 43.

[13] Title 2, Article 5, Section 1(a), EU AI Act, 2023, p. 43.

"real-time" remote biometric identification in publicly accessible (not just publicly owned) spaces outside of a very strict specific set of circumstances.[14] Additionally, it regulates the use of such technologies for the purposes of law enforcement and limits it to restricted and specific criteria.[15]

- Obligatory human oversight

The Act requires high-risk AI systems be overseen by a human being throughout its application cycle[16] and additionally stipulates that "human oversight shall aim at preventing or minimizing the risks to health, safety or fundamental rights that may emerge when a high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse."[17]

- Obligations on users of high-risk AI

Article 29 stipulates that "users of high-risk AI systems shall use the information provided under Article 13 to comply with their obligation to carry out a data protection impact assessment."[18]

- Requirements for AI regulatory sandboxes and obligations on stakeholders

The Act calls for AI regulatory sandboxes to have explicit parameters and that they "shall provide a controlled environment that facilitates the development, testing and validation of innovative AI systems for a limited time before their placement on the market or putting into service pursuant to a specific plan.[19] Additionally, it mandates the involvement of all relevant data protection authorities,[20] and calls for the immediate suspension of any development and testing of any such technology that results in "any significant risks to health and safety and fundamental rights identified" unless such risks are mitigated.[21] It also places significant liability on the developers of such an AI technology and for any harms incurred by third parties by the use of the technology.[22]

[14] Title 2, Article 5, Section 1(d), EU AI Act, 2023, p. 43-44
[15] Title 2, Article 5, Section 2, EU AI Act, 2023, p. 44
[16] Chapter 2, Article 14(1), EU Act, 2023, p. 51
[17] Chapter 2, Article 14(2), EU Act, 2023, p. 51
[18] Chapter 2, Article 29, Section 6, EU AI Act, 2023 p. 58
[19] Title 5, Article 53, Section 1, EU AI Act, 2023, p. 69
[20] Title 5, Article 53, Section 2, EU AI Act, 2023, p. 69
[21] Title 5, Article 53, Section 3, EU AI Act, 2023, p. 70
[22] Title 5, Article 53, Section 4, EU AI Act, 2023, p. 70

2. INDIA

India currently does not have a definitive AI regulation. In an approach separate from the EU, India sees AI as a "kinetic enabler" and intends to position itself as a global leader in the field. According to the Ministry of Electronics and IT, the Indian government views "stringent regulation" as potentially stifling innovation. At this stage, the government is "not considering bringing a law or regulating the growth of AI in the country".[23]

Presently, the AI legal framework consists of the **Digital Data Protection Bill 2022,** which introduces "data fiduciary" obligations on certain AI developers. Additionally, in 2018 the prominent Indian think-tank, NITI Aayog, published a **National Strategy for AI**[24] that places an emphasis on using AI for social welfare by facilitating the health, agriculture, education, smart cities and infrastructure and transportation sectors. In 2021, NITI Aayog published an **Approach Document for India on the Principles of Responsible AI**.[25] The document notably highlights system risks that should be considered by any responsible AI policy:

- Lack of understanding of an AI system's functioning can make its deployment unsafe and unreliable;
- Challenges in explaining specific decisions of AI systems can impact trust in the decisions;
- Inherent bias could make the decisions prejudiced and result in discrimination;
- There is high potential for inadvertent exclusion of citizens in AI systems used for delivering important social services and benefits;
- It is difficult to make AI systems accountable; and
- There are high privacy and security risks associated with AI use.

According to the document, the following 7 principles should govern all 'responsible AI' in line with international best practices and human rights standards:

1. Principle of Safety and Reliability
2. Principle of Equality

---

[23] Manish Singh, "India Opts Against Tech Regulation" (Tech Crunch, April 2023) https://techcrunch.com/2023/04/05/india-opts-against-ai-regulation/

[24] "National Strategy for Artificial Intelligence" (NITI Aayog, June 2018) https://niti.gov.in/sites/default/files/2019-01/NationalStrategy-for-AI-Discussion-Paper.pdf

[25] "Responsible AI #AIForAlll – Approach Document for India Part 1 - Principles of Responsible AI", (NITI Aayog, Feb 2021) https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf

3. Principle of Inclusivity and Non-discrimination
4. Principle of Privacy and Security
5. Principle of Transparency
6. Principle of Accountability
7. Principle of protection and reinforcement of positive human values

Additionally, the policy paper emphasizes the need for pre-existing legislation along with a sectoral approach to AI regulation. It cites anti-discrimination laws that are currently silent on AI-facilitated discrimination. According to the document, "it will fall within the jurisdiction of anti-discrimination legislation to regulate decisions arrived at through the use of AI as well, particularly when the decision-making AI is being used by an entity having constitutional or legal obligations to be unbiased."[26] Similarly, in 2019, the Indian SEBI issued a circular Stock Brokers, Depository Participants, Recognized Stock Exchanges and Depositories and in May 2019 to All Mutual Funds (MFs)/ Asset Management companies (AMCs)/ Trustee Companies/ Board of Trustees of Mutual Funds/ Association of Mutual Funds in India (AMFI) on reporting requirements for AI and Machine Learning (ML) applications and systems offered and used. The reporting works towards creating an inventory of AI systems in the market and guide future policies.[27]

3. CHINA

China is currently regulating AI on an application by application basis. Its present AI regulatory framework consists of the **Internet Information Service Algorithmic Recommendation Management Provisions Act (2022),** which regulates recommendation algorithms, the **Administration of Deep Synthesis of Internet Information Services Act (2022)** and **Draft Measures on Generative AI, Draft Administrative Measures for Generative Artificial Intelligence Services (proposed since 2022)** which will cover technologies such as ChatGPT, Dall.E etc.[28]

Most notably, China's **Administration of Deep Synthesis of Internet Information Services Act (2022)** includes much-needed safeguards against deepfake technology focusing especially on consensual use of user data, AI disclosures and the prohibition on the proliferation of fake news. Some of the key provisions include:

[26] Ibid, p. 29
[27] Ibid, p. 31-32
[28] China – Global AI Regulation Tracker (2023) https://www.techieray.com/GlobalAIRegulationTracker.html.

- Users must give consent if their image is to be used in any deep synthesis technology;
- Deep synthesis services cannot use the technology to disseminate fake news;
- Deepfake services need to authenticate the real identity of users;
- Synthetic content must have a notification of some kind to inform users that the image or video has been altered with technology; and
- Content that goes against existing laws is prohibited, as is content that endangers national security and interests, damages the national image or disrupts the economy.

However, the efficacy of these regulations has not been ascertained. Current technology necessary to accurately identify and categorize deep fakes is presently "fairly unreliable"[29] with even as low as a 2% rate of false positives allowing enough harmful content to proliferate online. Additionally, some of the provisions, such as the "watermark" requirement, can easily be edited and removed in posts.

4. United States

There is no overarching AI regulation in place in the US. However, there are some regulations at the state level, with standards varying from state to state. The current AI regulatory framework in the US consists of a federal level **Blueprint of an AI Bill of Rights** issued by the White House, various state-level proposed legislation such as the **California Assembly Bill No. 331**, and the **National Institute of Standards and Technology (NIST) AI Risk Management Framework** which is a framework designed for the management of risks to individuals, organisations, and society associated with AI by being more inclusive of trustworthiness considerations into the design, development, use, and evaluation of AI products, services, and systems.[30] It should be noted that even though the **Blueprint of an AI Bill of Rights** is not a binding legal instrument, it is meant to offer a guide that steers future legislation.

The **Blueprint of an AI Bill of Rights** covers the following five principles, i.e. safe and effective systems; algorithmic discrimination protections; data privacy; notice and explanation; and human alternatives, considerations and fallbacks. Of these, the most

[29] Afiq Fitri, "China Has Just Implemented One Of The World's Strictest Laws On Deepfakes", (Tech Monitor, 2023) https://techmonitor.ai/technology/emerging-technology/china-is-about-to-pass-the-worlds-most-comprehensive-law-on-deepfakes

[30] USA – Global AI Regulation Tracker https://www.techieray.com/GlobalAIRegulationTracker.html.

notable inclusion is for *human alternatives, considerations and fallbacks,* which basically states that:

"You should be able to opt out, where appropriate, and have access to a person who can quickly consider and remedy problems you encounter. You should be able to opt out from automated systems in favor of a human alternative, where appropriate. Appropriateness should be determined based on reasonable expectations in a given context and with a focus on ensuring broad accessibility and protecting the public from especially harmful impacts. In some cases, a human or other alternative may be required by law. " [31]

On the state level, California's Civil Rights Council (CRC) modified employment practices that govern the use of AI in employment decision-making.[32] In the **CRC Draft Modifications to Employment Regulations Regarding Automated-Decision Systems**, the use of ADS (Automated Decision Systems) to discriminate in employment, or result in dissipated treatment of an applicant or employee, especially on the basis of protected characteristics is unlawful unless the employer can show the selection criteria to be job-related and consistent with business necessity.[33] It also imposes liability on the employer for the use of discriminatory ADS by "its supervisors, managers or agents".[34]

California Assembly Bill No. 331 also contains certain key protections, such as providing the necessary disclosure/notification to the subject of an AI-based decision by any entity using ADT (automated-decision technology) to make said "consequential decisions" and such subjects should be entitled to recourse in the form of an alternate decision-making process (i.e human-based). Additionally, it calls for fines for entities using ADT and not furnishing the requisite impact assessment of the use of ADT; requires ADT developers to

---

[31] "The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People" (The White House Office of Science and Technology Policy, Oct 2022), p.46 https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf

[32] Davis Wright Tremaine LLP "California Proposed Employment AI Regulations and Legislation - Proposals from Civil Rights Council and Legislature Concerning the Use of Artificial Intelligence in Employment" (JDSupra, May 2023) https://www.jdsupra.com/legalnews/california-proposed-employment-ai-7288812/#_ftnref1

[33] "Fair Employment & Housing Council Draft Modifications to Employment Regulations Regarding Automated-Decision Systems" (Feb 2023) Article 2. § 11016 (c) (A), p. 8 https://www.dwt.com/-/media/files/2023/05/attachbmodtoemployregautomateddecisionsystems.pdf?la=en&rev=41324b02eadc470eb96c5f51c216aa81&hash=0000D6713DD8F077496020FC0577CDE8

[34] Ibid, Article 1 § 11009 (b), p. 4 https://www.dwt.com/-/media/files/2023/05/attachbmodtoemployregautomateddecisionsystems.pdf?la=en&rev=41324b02eadc470eb96c5f51c216aa81&hash=0000D6713DD8F077496020FC0577CDE8

disclose what data is used to train the ADT; and also design and implement an ADT governance system which includes guardrails, safeguards and other policies.[35]

Additionally, the **United States National Telecommunications and Information Administration (NTIA)** is currently[36] collecting feedback from the public on policies the government should employ to regulate AI and facilitate its accountability.[37]

[35] USA – Global AI Regulation Tracker https://www.techieray.com/GlobalAIRegulationTracker.html
[36] As of July 2023.
[37] Ibid.

F. COMPLEMENTARY LEGISLATION

Name of Legislation	Relevant Section(s)	Comment(s)
Personal Data Protection Bill (PDPB), 2023	***S.2 Definitions:*** (dd) "profiling" means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to the data subject in particular to analyse or predict aspects concerning that data subject's attributes related to employment, social preferences, religious beliefs, economic situation, health, reliability, behaviour, location or movements; **S.29 *Right to data portability and automated processing:*** (4) The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which results in legal obligations or significantly harms the data subject, unless the data subject has given his explicit consent. (5) The data subject shall have the right to obtain from the data controller: (a) specific information about automated decision-making including profiling, (b) human intervention	Liability for data breaches and data privacy violations through the use of AI or such technologies should be explicitly included in the Act. AI should also be defined within the Act, including via reference to any future AI policy/Act that may more comprehensively define AI. S29. (6) states that "the data subject rights mentioned in sub-section (4) shall not apply to the extent where processing is necessary for the performance of a task carried out in the public interest." This has the potential to be problematic as public interest is too vague and broad a criteria to effectively prevent human rights abuses and ill-intentioned violation of a data subject's consent.

Prevention of Electronic Crimes Act (PECA), 2016	***S.2 Definitions:*** (i)(b) causing an act to be done by a person either directly or through an automated information system or automated mechanism or self-executing, adaptive or autonomous device ard whether having temporary or permanent impact; (xix) "information" includes text, message, data, voice, sound, database, video, signals, software, computer programmes, any forms of intelligence as defined under the Pakistan Telecommunication (Reorganization) Act, 1996 (xv[of 1996) and codes including object code and source code; (xx) "information system" means an electronic system for creating, generating, sending, receiving, storing. reproducing. displaying, recording or processing any information (x)(v) “device” includes – (d) automated, self-executing, adaptive or autonomous devices, programs or information systems ***S.15 Making, obtaining, or supplying device for use in offence:*** Whoever produces, makes,, generates; adapts, exports,	**S.39 (Real-time collection and recording of information)** does not fully capture the high potential for harm that AI-powered real-time biometric surveillance systems can cause. While the section does set out limitations and criteria by which a court may grant an order for such real-time collection of information, it does not include penalties for the abuse of this information. Such penalties should be codified either in a future AI regulation or in the Personal Data Protection Bill. Additionally, should any person feel as if this provision has been abused, they should have a mechanism for legal recourse made available to them.

	supplies, offers to supply or imports for use any information system, data or device, with the intent to be used or believing that it is primarily to be used to commit or to assist in the commission of an offence under this Act shall, without prejudice to any other liability that he may incur in this behalf be punished with imprisonment for a term which may extend to six months or with fine which may extend to fifty thousand rupees or with both.	
SECP Regulatory Sandbox Guidelines, 2019	The SECP Regulatory Sandbox Guidelines sets out the process by which companies can test out various new innovative products the regulation of which may be out of the scope of current existing laws. Companies, start-ups etc. who are working in the field of AI in its various forms through various sectors fall within the approved Potential Partners list provided by the Guidelines **(S.2(ii)(a))** and lists Artificial Intelligence explicitly as a separate category for which potential candidates can apply for **(S.3(b))** ***S.3(c) Preliminary Screening of Applications(iii)(3) Consumer Benefit (iii):*** Put in place suitable mitigation plans to manage risks and ensure protection to	Provisions in the SECP Regulatory Sandbox Guidelines that cover mitigation of risks, consumer protection and complaint redressal mechanisms fall in line with some of the best international practices on the development and regulation of new technologies and particularly AI. Additionally, necessary infrastructural support is an important component to mitigate the harms that could arise from the potential use of new technologies and shifting the burden of proof to the companies/tester of providing these resources before they can be granted approval to test these products is a positive

	customers through a complaints redressal mechanism. ***S.3(c) Preliminary Screening of Applications(iii)(4) Readiness for testing (i)*:** Has adequate and necessary resources to support the testing in the sandbox including human capital and technology infrastructure	development.
Article 25, Constitution of Pakistan	**25 Equality of citizens.** (1) All citizens are equal before law and are entitled to equal protection of law. (2) There shall be no discrimination on the basis of sex. (3) Nothing in this Article shall prevent the State from making any special provision for the protection of women and children.	Article 25 provides Constitutional protections against discrimination, and should be extended to automated and AI-facilitated decision-making. Reference should be made to Article 25 in the Draft Policy to place a positive obligation on public bodies to ensure that AI systems do not discriminate against citizens, particularly vulnerable communities, on the basis of their protected characteristics.[38]

Given that AI technologies have the potential to be used in every sector, where numerous laws apply, it is particularly important to stress that either those laws should be amended to account for the potential use, application and harms of AI specific to that field or a provision should be added to any/all AI guidelines and policies that say that all AI

[38] Uzma Nazir Chaudhry, "Algorithmic Decision-Making in Pakistan: A Challenge to Right to Equality & Non-Discrimination," *Centre for Human Rights*, https://cfhr.com.pk/wp-content/uploads/2022/04/Algorithmic-Decision-Making-in-Pakistan.pdf.

technologies should follow the principle and spirit of/standards set by the laws (federal and provincial) enacted for health, labour, employment, finance, social services, trademark, copyright, etc. In addition, where laws do not cover certain parts of the AI value chain, such as for example contracted employees or outsourced freelance labour, measures should be taken to extend protections to these people in accordance with best international practices and rights under domestic law. Public bodies and courts are encouraged to adopt a purposive approach and interpret existing human rights laws and protections to extend to AI and automated technologies.

G. RECOMMENDATIONS

1. The national policy should define AI and related technologies and determine the scope of the policy in relation to present and future technologies.
2. A human rights-compliant Personal Data Protection Act must be passed that provides extensive protections against automated decision-making and consent rights against the use of personal data in datasets to train and maintain AI systems. The current version of the PDPB must be amended to provide definitions for AI and related technologies, remove 'public interest' exceptions for automated decisions, and protect personal data along the AI supply change.
3. AI awareness, digital literacy and online safety should be included in school, college and university curriculums so that education regarding the subject can begin reaching children at the primary school level and extend to higher education. The relevant AI bodies envisioned by this Draft Policy should work with the Ministry of Federal Education and Professional Training, School Education Departments, and the Higher Education Commission (HEC) to ensure AI education, with an emphasis on ethics and privacy, is integrated into curriculums.
4. SOPs and guidelines should be developed that demarcate and define the jurisdictions of multiple regulatory stakeholders overseeing AI development, regulation and implementation in Pakistan. It is recommended that rather than creating more regulatory authorities, the focus be shifted to improving coordination and collaboration between these agencies.
5. Further, the roles of the various regulatory agencies, such as MOITT, COE-AI, NIDP, and ADR, should be clearly defined.
6. In the establishment of AI-related funds, emphasis should be placed on transparency, accountability and sustainability. Annual audits by an independent auditing firm of international repute should be instituted to ensure transparency and accountability.
7. Government contracts negotiated with private companies for the development of AI should come with a declaration of non-competing interest or no conflict of interest to ensure transparency and accountability of all parties involved.
8. The policy should provide linkages between fundamental rights enshrined in the Constitution of Pakistan, such as Article 14 (right to privacy) and Article 25 (non-discrimination), to contextually ground frameworks of ethical and non-discriminatory AI.
9. Human rights audits should be mandated at the design stage, which must include an impact assessment on the development and potential use of that technology.

10. An annexe of prohibited AI practices, including classifications for unacceptable risk, high risk, low or minimal risk with an emphasis on the preservation of fundamental rights and protection of vulnerable groups, like women, gender minorities, children and PWDs, should be included.
11. Conditions on the use of AI in public service delivery and by government authorities, especially for purposes of law enforcement, should be covered. For example, the use of 'real time' remote biometric identification systems should be prohibited unless certain limited exceptions apply.
12. An accessible and independent reporting mechanism should be made available to citizens who feel aggrieved by inappropriate, harmful, negligent or discriminatory use of AI technologies.
13. The policy should feature certain baseline protections that form the blueprint for all subsequent industry-specific guidelines developed. For example, safeguards such as human moderation/override systems, especially in critical infrastructure such as health services, social services, and services that can potentially impact an individual's safety and livelihood. Further, all applications of decision-making AI should come with the option of "opting out" for a human assessment.
14. Obligations should be placed on AI providers, users and participants in the AI supply/value chain to uphold human rights and ethical AI practices.
15. Labour laws and protection mechanisms need to be updated and strengthened to ensure that adequate protections are in place for workers impacted, displaced and employed in sectors, industries and entities that employ AI technologies.
16. Requirements should apply to high-risk AI systems regarding the quality of data sets used, technical documentation and record-keeping, transparency and the provision of information to users, human oversight, robustness, accuracy and cybersecurity.

H. CONCLUSION

In its current iteration, the National AI Policy, whilst a positive step, is still lacking some necessary provisions and fails to account for the practical, human rights and safety implications of a world that may become heavily reliant on AI in a variety of applications. The policy places a disproportionate emphasis on the creation of subsidiary regulatory bodies with unclear jurisdictional boundaries and potentially overlapping mandates. Even if the functions of these regulatory bodies were more distinctly defined, the policy is silent on mechanisms of coordination between these various agencies. The burden of the enactment of positive AI literacy, innovation, research and development is placed on the AI Fund, the effectiveness of which can be impacted due to issues such as lack of resources, misappropriation of funds, non-transparent practices and absence of sustainability.

The policy envisions some proactive measures, such as placing special emphasis on targeting women and underprivileged groups in the government's AI literacy and awareness efforts and protections against discriminatory practices. However, subsequent versions of the policy need to be more detailed in the protection of principles such as safety, security, privacy, non-discrimination, transparency, and accountability and must prioritize the rights of citizens over the rights of institutions and corporations in line with international standards of human rights.