To,

Prof. Ajay Kumar Sood

Principle Scientific Advisor, MeiTY

27th February 2025

In Re: Submission of Comments from Public on the *Report on AI Governance Guidelines Development.*

Respected Sir,

This letter is in reference to the report dated 6th January 2025, issued by the Ministry of Electronic and IT, Government of India, inviting comments from the public on the *Report on AI Governance Guidelines Development*. In furtherance of our commitment to contributing towards the legal and technological discourse in the country and working for public welfare in the capacity of law students, the team at the Cell for Law and Technology ("CLT") hereby submits its suggestions in response to the Press Release.

At the very outset, we would like to express our appreciation for this progressive initiative taken by the Ministry to seek public participation in shaping the regulatory framework for Artificial Intelligence (AI). The inclusive approach towards policy formulation will ensure a balanced, transparent, and effective governance structure for AI development and deployment in India. As AI continues to play an increasingly crucial role across industries, implementing robust and adaptable regulations will be vital to addressing ethical concerns, ensuring accountability, and fostering innovation.

The proposed framework has been thoroughly analyzed, and the team comprising members of CLT has identified various points that we would like to highlight as suggestions and comments. We hope that our submissions will contribute meaningfully to the ongoing discussions surrounding AI governance in India.

Thank you for your consideration.

Warm Regards,

Atul Kumar Pandey

(Professor of Cyber Law, Head, Department of Cyber Law

Faculty in Charge, Cell for Law and Technology

National Law Institute University, Bhopal)

The team which has been instrumental in putting forward this suggestion comprises of the following members of the Cell for Law and Technology, NLIU Bhopal:

1. Rishita Sethi (Convenor) (Final Year BALLB Student, NLIU Bhopal)
2. Hussain (Co-Convenor) (Final Year BALLB Student, NLIU Bhopal)
3. Lavya Bhasin (3rd Year BSCLLB Student, NLIU Bhopal)
4. Suvansh Shanker (3rd Year BSCLLB Student, NLIU Bhopal)
5. Vibhuti Sharma (2nd Year BSCLLB Student, NLIU Bhopal)
6. Arpit Dhadhich (2nd Year BSCLLB Student, NLIU Bhopal)
7. Siddhant Samaiya (2nd Year BSLLB Student, NLIU Bhopal)

Comments submitted by the Cell for Law and Technology (CLT), National Law Institute University (NLIU), Bhopal, Madhya Pradesh

REPORT ON AI GOVERNANCE GUIDELINES DEVELOPMENT

Comments on Chapter II: Governance of AI

A. AI Governance Principles

S. No.	Concept	Issues	Suggestions	Summary and Conclusion
1.	Transparency	While it is correctly mentioned that AI systems should be accompanied with meaningful information behind their functioning, the crucial nuance between 'transparency of AI' and 'explainability of AI' has been missed.	However, where explainability is all about providing clear, understandable reasons for the decisions made by an AI system, i.e., the 'why' behind decisions; transparency is about openness and accessibility of information regarding the AI system, i.e., the 'how'. After explaining such a difference in	It is true that both the methods of 'transparency of AI' and 'explainability of AI' aim to make AI systems more understandable and trustworthy. After explaining the difference between the above principles, the report must also talk about the kind of tools that can achieve explainability.

			the principles, the report must also talk about the kind of tools that can achieve explainability. For example, model-agnostic tools like LIME (*Local Interpretable Model-agnostic Explanations*) and SHAP (*SHapley Additive exPlanations*, help in breaking down complex models to show how different features contribute to a specific decision, and may even use visualizations, such as decision trees and heat maps to present data in a format where users can easily derive how AI reached to a particular solution.	
2.	Accountability	This principle has talked about having 'mechanisms' in place that clarify accountability, but it has not explained whether AI in	Mostly, AI governance for accountability has three parties, i.e., the developer, the deployer, and the integrator. This has also been	The accountability principle has talked about having 'mechanisms' in place that clarify accountability, but it has

		the Indian context is envisioned to be governed in a two-party, or a three-party framework.	outlined in the Global AI Policy Recommendations of 2021, which suggest that AI accountability should be shared by actors all across the AI value chain. A 'developer' is the entity that produces or develops the AI model or system, and a 'deployer' is the entity that puts the AI system into use, decides the purpose for which the AI system is used, and uses the system to make decisions that impact end-users. An 'integrator' is an intermediate actor in the supply chain, which may take appropriate steps to facilitate the developer or the deployer, depending on the context. As such, integrators should not be viewed monolithically - there is no set of static responsibilities that will be appropriate for every integrator to	not explained whether AI in the Indian context is envisioned to be governed in a two-party, or a three-party framework. It is important to include the actors involved in the AI value chain in the AI governance framework to ensure better accountability at all steps.

			undertake in every circumstance, however, it still becomes important to include these actors in the AI governance framework to ensure better accountability at all steps.	
3.	Safety, Reliability & Robustness	Two things that have not been clarified are: a) who has the responsibility of monitoring AI systems? Does this work in the same way as 'AI audits', and b) what happens when during the course of monitoring, an adverse event is spotted or encountered?	The term 'regulatory monitored' is too broad and poses several questions. These questions must be answered in the principles itself so that the subsequent framework has greater clarity and better chances of implementation.	This principle mentions that AI systems should be 'regularly monitored' to ensure that they operate in accordance with their specifications and perform their intended functions. The term 'regulatory monitored' is too broad and poses several questions. These questions must be answered in the principles itself so that the subsequent framework has greater clarity and better chances of implementation.
4.	Privacy & Security	The words 'security by design'	A suggestion here would be to either	The term 'security by design' is

		have been used. However, the rationale behind using 'security by design' instead of 'privacy by design' is not clear, which is crucial, since it is the latter that is commonly used for data privacy standards in the EU- GDPR.	use one term, which is also in compliance with global standards, or explain both the terms being used with the differences and individual functionalities of both, to ensure a fool-proof mechanism for ensuring privacy and security. Further, the issue of an opt-out mechanism for data usage is also absent. When AI developers use publicly available datasets for training, there should be a clear mechanism for individuals or organizations to opt-out, particularly if they have authority over the data.	not explained in the report. This term resonates to "privacy by design' which is globally used for the data privacy standards. It is crucial to explain this term to ensure a fool-proof mechanism for ensuring privacy and security.
5.	Fairness and non-discrimination	The concept of 'perpetuation of biases' is not elaborated. Further, the fact that bias may creep into all stages of the lifecycle of an AI system is missed out.	The concept of 'perpetuation of biases' needs to be elaborated on. Many AI models are trained on datasets predominantly from Western countries, resulting in a westernized	Many AI models are trained on datasets predominantly from Western countries, resulting in a westernized perspective in their outputs. Given the rapid pace of

			perspective in their outputs. Given the rapid pace of AI advancements, a temporary solution is urgently needed to address this representational imbalance until a comprehensive repository of diverse datasets can be developed. Failure to address biased data will perpetuate systemic inequalities in AI-generated responses. Steps like pre-processing (training data and model outputs), in-processing, and post-processing should have been mentioned, along with the techniques that are used to mitigate bias at each stage.	AI advancements, a temporary solution is urgently needed to address this representational imbalance until a comprehensive repository of diverse datasets can be developed.
6.	Human-centred values & 'do no harm'	The words 'complex ethical dilemmas' and 'adverse outcomes' remain ambiguous.	The words 'complex ethical dilemmas' and 'adverse outcomes' remain ambiguous. The effect of vague terms is that it leads to	The words 'complex ethical dilemmas' and 'adverse outcomes' remain ambiguous, which leads to uncertainty in the

			uncertainty in the applicability or the scope of the law. Therefore, thresholds for what qualifies as a 'complex ethical dilemma' or 'adverse outcome' must be specified to the best possible extent, so that the judiciary while adjudicating disputes has some point of reference while exercising judicial discretion. A suggestion here would be to conduct an 'ethical impact assessment' which includes identification of concerns and risks of AI systems, as well as appropriate risk prevention, mitigation and monitoring measures, among other assurance mechanisms, which will help identify impacts on human rights and fundamental freedoms, in particular	applicability or the scope of the law. Therefore, thresholds for what qualifies as a 'complex ethical dilemma' or 'adverse outcome' must be specified to the best possible extent.

			but not limited to the rights of marginalized and vulnerable people or people in vulnerable situations, labour rights, the environment and ecosystems and ethical and social implications. Such an impact assessment has also been suggested in the UNESCO Recommendations on the Ethics of Artificial Intelligence, and will help one classify a particular AI use impact as a 'complex ethical dilemma'.	

B. Considerations to operationalise the principles

S. No.	Concept	Issues	Suggestions	Summary and Conclusion
1.	Examining AI Systems using a lifecycle approach	The third stage of the lifecycle is mentioned to be 'diffusion'. Such usage may not be	Diffusion models are advanced machine learning algorithms that generate high-quality data by gradually adding Gaussian noise to a	The third stage of the lifecycle is mentioned to be 'diffusion'. Such usage may not be appropriate as it conflicts with the well-

		appropriate as it conflicts with the well-established terminology of "Diffusion Modelling" in the field of AI.	dataset and then learning to reverse this process. This approach allows for the creation of highly accurate and detailed outputs. Using "Diffusion" in this context can cause confusion, as it does not align with its technical meaning. International standards define the final stage of the AI lifecycle as "Machine Learning Operations" (MLOps), a term that is more precise and widely recognized. Replacing "Diffusion" with "Machine Learning Operations" would enhance clarity and align with established practices in the field.	established terminology of "Diffusion Modelling" in the field of AI. This approach allows for the creation of highly accurate and detailed outputs. Replacing "Diffusion" with "Machine Learning Operations" would enhance clarity and align with established practices in the field.
2.	Taking an ecosystem-view of AI actors:	The ecosystem (consisting of five actors as per the framework as of now) is incomplete.	It is true that we need an ecosystem-view of actors to look at distribution of responsibilities better. However, the ecosystem (consisting of five actors as per the framework as of	We need an ecosystem-view of actors to look at distribution of responsibilities better. However, the ecosystem (consisting of five actors as per the framework as of

			now) is incomplete, and should add two more actors, a) government and regulatory bodies, and b) investors. This is because a) government and regulatory bodies play a key role in setting policies, ensuring ethical AI development, and protecting public interests, including privacy, security, and fairness. They shape the regulatory environment to ensure responsible AI deployment while mitigating potential harm or misuse, thus they must also be counted as an AI actor, and not merely as a legislative body that controls other AI actors. Further, b) Investors hold influence over the direction of AI innovation and business models. Their funding decisions impact which AI technologies and companies emerge	now) is incomplete, and should add two more actors, a) government and regulatory bodies, and b) investors.

			and scale. Investors have a responsibility to prioritize ethical considerations and sustainability, ensuring that AI development aligns with broader societal goals rather than purely financial interests, thus, they must also be added to this envisioned AI ecosystem as an AI actor.	
3.	Leveraging technology for governance	Vague terms such as 'unlawful information' and 'security incidents' need to be more sharply and exhaustively defined.	On page 6, second-last paragraph talks about a techno-legal approach for the purposes of tracing unlawful information using AI developers and deployers after a valid request from the Government on grounds such as prevention, detection, investigation or prosecution of harms, crimes, and security incidents. It is also correctly mentioned that use of such automated tools will have	Vague terms such as 'unlawful information' and 'security incidents' will have to be more sharply and exhaustively defined, so that users can foresee liability from the consequences of their conduct. If what conduct/ content over the internet qualifies as a 'security incident' or 'unlawful information' is clearly defined, only then will the law be

			bearing on fundamental rights. Given that India has limited jurisprudence on intersection of free speech and algorithmic regulation of content, the recent US Supreme Court judgement of *Moody v Netchoice* may be looked at, among others. Further, such measures also need to be consistent with the The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 if the AI system in question qualifies as an intermediary. Lastly, vague terms such as 'unlawful information' and 'security incidents' will have to be more sharply and exhaustively defined, so that users are can foresee liability from the consequences of their conduct.	sufficiently precise and consistent of adequate safeguards. And it is also correctly mentioned that use of such automated tools will have bearing on fundamental rights.

			If what conduct/ content over the internet qualifies as a 'security incident' or 'unlawful information' is clearly defined, only then will the law be sufficiently precise and consistent of adequate safeguards.	

Comments on Chapter III: Gap Analysis

A. The need to enable effective compliance and enforcement of existing laws

S. No.	Concept	Issues	Suggestions	Summary and Conclusion
1.	Deepfakes/ fakes/ malicious content	The report has not taken into account the number of deepfakes that peaked at the time of elections in India and the absence of a specific law/ measure to address the same as a 'gap', in order to preserve the sanctity of	The report has correctly mentioned a number of criminal as well as civil laws that may apply in order to detect, prevent, remove and prosecute the creation and distribution of malicious synthetic media. However, the report has not taken into account the number of deepfakes that peaked at the time of	The report must take into account the number of deepfakes that peaked at the time of elections in India and the absence of a specific law/ measure to address the same as a 'gap', in order to preserve the sanctity of the democratic nature of India. The report should also consider

		the democratic nature of India. The report has also not commented on whether India requires a specific deepfake legislation, or at least a set of rules for the same, under s 66D of the IT Act.	elections in India and the absence of a specific law/ measure to address the same as a 'gap', in order to preserve the sanctity of the democratic nature of India. The report has also not commented on whether India requires a specific deepfake legislation, or at least a set of rules for the same, under s 66D of the IT Act. The report must consider the above mentioned points.	whether India requires a specific deepfake legislation, or at least a set of rules for the same, under s 66D of the IT Act.
2.	Cyber security	Instead of vaguely mentioning that there is a need for upgrading compliance to deal with rapid development of AI, there should be a clear mention of suggested legislation.	The report mentions that AI enables non-technical specialists to carry out sophisticated measures, which may lead to heightened risks. However, instead of vaguely mentioning that there is a need for upgrading compliance to deal with rapid development of AI, there should be a	The report correctly mentions the various legislations and mechanisms that already exist to ensure cybersecurity of computer systems. However, instead of vaguely mentioning that there is a need for upgrading compliance to deal with rapid development of

		Further, the report has not made a mention of whether the Consumer Protection Act, 2019 (chapter of product liability) can extend to AI system developers and distributors.	clear mention of suggested legislation. Inspiration can be taken from the EU AI Act, which clearly mentions obligations of Providers, Product Manufacturers, Deployers, Importers and Distributors. It also groups systems into banned AI and high-risk AI, wherein the latter is also expected to carry out additional obligations.	AI, there should be a clear mention of suggested legislation.
3.	Intellectual property rights	While the report considers the interaction between copyright law and AI, it neglects to address the implications for patent and trademark	AI advancements, especially the evolution of autonomous AI agents that require minimal human oversight, expand AI capabilities beyond content creation to include the development of new inventions.	While the report considers the interaction between copyright law and AI, it neglects to address the implications in patent and trademark law, both of which are crucial in understanding the

		law, both of which are crucial in understanding the broader impact of AI. The "inventive step" requirement assumes human inventorship, creating ambiguity regarding whether AI can qualify as an inventor under existing frameworks. In trademark law, the integration of AI highlights significant gaps and challenges.	This raises critical challenges for patent law, which mandates that inventions meet criteria such as novelty, utility, and industrial application. Moreover, assessing innovation and ingenuity in AI-generated outputs—particularly for mechanical or algorithmic creations—is inherently complex. While India's patent laws are evolving to include software patents, the lack of clarity on handling AI-driven inventions poses challenges. Thus, addressing these ambiguities is essential to balance promoting AI innovation with protecting intellectual property. Traditional concepts such as "imperfect recollection" and "confusion" are becoming less relevant as AI-driven platforms	broader impact of AI. AI advancements raises critical challenges in patent law, and trademark law.

			personalize consumer choices, reducing reliance on human perception. AI-generated trademarks raise unresolved questions about distinctiveness, eligibility for protection, and ownership, as current laws assume human involvement in creation. Additionally, the rise of AI-generated content complicates the detection and enforcement of trademark infringement. Thus, there arises a need for the report to also take into consideration, and comment on these issues.	
4.	AI led bias and discrimination	The report mentions that only biases that are 'legally or socially prohibited' need to be protected against. The	The report mentions that only biases that are 'legally or socially prohibited' need to be protected against. The same is vague and ambiguous, and needs more clarity.	It is true that AI systems can perpetuate biases when they are trained on historical data that reflects societal prejudices, stereotypes, or discriminatory

		same is vague and ambiguous. The report also does not talk about smaller bias that may not have a direct, but indirect social or legal impact. There is a need for transparency and responsibility across the AI ecosystem in India.	A suggestion would be to define the scope of what is legally prohibited as something that is against 'fundamental and legal rights under all laws in force in India'. However, the scope of socially prohibited biases, especially in a culturally and economically developing country such as India is impossible to be defined, and thus, this word must either be replaced or qualified. Further, the report also needs to explain the rationale behind why other subtler and more technical biases, such as selection bias, sampling bias, historical bias, coverage bias, group attribution bias, etc, need not be protected against. The report highlighted the importance of transparency and adopting a "whole-of-government"	practices. There is a need to define the scope of what is legally prohibited as something that is against 'fundamental and legal rights under all laws in force in India'. The report highlighted the importance of transparency and adopting a "whole-of government" approach in detecting biases in AI systems.

			approach in detecting biases in AI systems. For instance, it acknowledged that individuals may not recognize discrimination, or, even if they do, proving intent can be challenging, allowing such biases to remain undetected. While the importance of transparency was noted, the report must also explore specific tools, strategies, and the practicality of implementing such approaches for detecting biases. As has been stated above (see pg 1 of the comments), the report needs to draw a difference between, and also consider as a factor 'explainability of AI' along with transparency. Further, the recommendation for a baseline framework is promising. To strengthen this suggestion, the report	

			could outline potential components of such a framework, including standardized risk assessment protocols and clear rules for liability assignment. The words "cross-cutting" issues need to be elaborated upon.	