VOLUME 6 | ISSUE 4

INTERNATIONAL JOURNAL OF ADVANCED LEGAL RESEARCH

LEGAL CHALLENGES IN STATE-FUNDED FOUNDATIONAL AI MODELS IN INDIA

- P. Iswariya[1]

ABSTRACT

This paper looks at the challenges that come with state-funded foundational artificial intelligence models in India. It does this by looking at the country laws and regulations. The paper starts by examining the role of the Information Technology Act, 2000 the Digital Personal Data Protection Act, 2023 the Indian Penal Code, 2023 and the Copyright Act, 1957 in shaping the environment for artificial intelligence. The study finds some challenges, including unclear liability and accountability tension between data protection principles and training data needs, algorithmic bias and violations of constitutional rights, copyright and ownership conflicts and cross-border data flows and compute sovereignty. All these challenges come from the fact that Indias laws were made before artificial intelligence became big. The paper says that not having a law for artificial intelligence leaves state-funded models open to unclear rules and gaps in accountability. It ends with suggestions for an approach to artificial intelligence governance that aligns with Indias constitutional commitments to equality, dignity and due process.

INTRODUCTION

The world is seeing a change in the digital landscape moving from software services to artificial intelligence. Foundational models, which are scale artificial intelligence systems trained on huge datasets have become the new general-purpose technology of the 21st century. For a country like India that is developing its economy these models offer a big chance for growth and a challenge to national digital sovereignty. The Government of India

[1]LLM- Cyber Space Law and Justice, Tamilnadu Dr.Ambedkar Law University (TNDALU), Chennai, Tamilnadu .

started the IndiaAI Mission in 2024 with a big budget to develop foundational models that meet the country unique needs. This initiative includes projects like BharatGen and the IndiaAI Innovation Centre, which aim to reduce the country reliance on models. However the fast deployment of state funded intelligence models raises complex legal and ethical questions. As the state becomes a developer and funder of intelligence traditional frameworks of accountability are being tested. This research paper looks at the evolution, funding and regulation of state-funded artificial intelligence models in India. It examines the touch regulatory approach adopted by the Ministry of Electronics and Information Technology and assesses whether current regulations are enough to manage the risks of artificial intelligence. The paper argues that India needs a legal framework to ensure that these models remain a public good while protecting against the risks of automated decision-making.

STATE FUNDED FOUNDATIONAL AI MODELS

Foundational artificial intelligence models are large-scale machine learning systems trained on diverse datasets. They are designed to serve as a general-purpose base for downstream applications. Unlike narrow artificial intelligence systems foundation models are built to exhibit broad capabilities that can be adapted across multiple domains.

Core technical characteristics

Foundational models are big both in terms of their training datasets and internal parameters. They are trained on corpora drawn from sources like the open web, digitized books and public-domain archives. This allows them to learn patterns, structures and semantic relationships across a spectrum of human knowledge. The resulting models are usually neural networks, which enable them to attend to long-range dependencies in text and other sequential data. A key feature of foundation models is their two-stage development pattern: first they undergo a pre-training phase on large unstructured datasets, where they learn general linguistic, visual or logical patterns. After pre-training the model can be fine-tuned on task-specific datasets to become more accurate and reliable for specialized applications.

Relationship to Generative AI and GPAI

Foundational models are closely linked to generative intelligence, which refers to systems capable of producing new content like text, images or code. Many foundation models, large language models fall within the generative artificial intelligence family. However not all foundation models are purely generative; some are used for prediction, summarization or

other forms of latent representation learning. The concept of purpose artificial intelligence is sometimes used to describe foundation models that can be adapted to many different tasks and sectors. Regulators and governance frameworks treat GPAI as a category because a single model can influence a broad ecosystem of products and services. If such a model is deployed in sector or state-funded contexts any systemic flaws can propagate across multiple government applications creating systemic risk.

Why Foundation Models Matter in Governance

From a governance perspective foundation models are significant because they function like infrastructure. A state-developed or state-funded foundation model can become a shared resource across ministries and departments enabling language understanding and information-processing capabilities. This infrastructural role offers advantages but it also concentrates legal and ethical risk in a single layer of the artificial intelligence stack.

The Significance of State-Funded Foundational AI in India

The fact that a foundational artificial intelligence model is state-funded significantly alters the constitutional framework within which it operates, especially in Indias rights-based constitutional order. Unlike private-sector artificial intelligence systems state-funded foundational models raise deeper questions of public trust, accountability and the exercise of state power. When the state provides financial support or policy-backed incentives for the development and deployment of these models it effectively becomes a financier and indirect controller of a powerful decision-shaping infrastructure. This dual role means that any biases, opacity or data-related harms produced by a state-funded model can translate into actionable legal wrongs, against citizens.

In India the government is pushing for intelligence models that are made in India. This is a project that is about making sure India has its own digital infrastructure. The government is asking companies and researchers to build intelligence models that are trained on Indian data. These models should be able to handle languages and be fair to all people. The government wants to make sure that India does not have to rely on companies for artificial intelligence. The government is funding these projects so it has to make sure that they are done in a way that's fair and transparent. This means that the government has to be involved in how these modelsre made and used. If the government just lets companies do whatever they want it could lead to problems like models that are not fair to some people. The government has to

make sure that these models are not used to spy on people or to make decisions that're not fair. In India there are laws that say the government has to be fair and transparent when it uses money. If the government is funding intelligence models it has to make sure that these models are used in a way that is fair and transparent. This means that the government has to be able to explain how these models work and how they are used. The government also has to make sure that these models are not used to hurt people or to make decisions that're not fair. When artificial intelligence models are used it can be hard to figure out who is responsible if something goes wrong. If a model is made by a company but funded by the government it can be hard to say who is responsible if the model does something wrong. In India the courts have said that the government can be responsible for things that are done by companies if the government is involved in how they are done. This means that the government has to make sure that artificial intelligence models are used in a way that's fair and transparent. The government needs to make rules for intelligence models that are funded by the government. Now there are laws that talk about artificial intelligence in general but there are not specific rules for models that are funded by the government. The government needs to make rules that say how these models can be used and how they have to be transparent. This will help make sure that these models are used in a way that's fair and good for everyone.

REGULATORY FRAMEWORKS FOR ARTIFICIAL INTELLIGENCE IN INDIA

In India the rules for intelligence are still being made. There are some laws that talk about technology and data but there are not yet specific laws that talk about artificial intelligence. The government is trying to figure out how to balance the need for innovation with the need to protect peoples rights. The government is looking at laws from countries like the European Union to see how they can make rules that work for India. There are some laws in India that already talk about intelligence. These laws include the Information Technology Act, the Digital Personal Data Protection Act and the Indian Penal Code. These laws talk about things like how data can be used and how companies have to protect peoples information. They also talk about what happens if someone uses intelligence to do something wrong.

1. The Information Technology Act

The Information Technology Act is a law in India that talks about digital technology. It was made a time ago but it has been updated to include things about artificial intelligence. The law says that digital documents and signatures are real which is important for intelligence

models that make documents. It also says that companies have to protect peoples information and follow rules when they use data. The law also talks about what happens if someone uses intelligence to do something wrong. If someone uses a model to hack into a computer or steal data they can be punished. The law also says that companies that host artificial intelligence models have to follow rules and make sure that the models are not used to do something wrong.

2. The Digital Personal Data Protection Act, 2023

The Digital Personal Data Protection Act 2023 is a law that protects peoples data in India. This law is very important for AI models that get money from the government. The law says that personal data must be collected, processed and protected in a way. The law applies to companies and organizations that handle data, which are called data fiduciaries. It also applies to individuals whose data is being handled who are called data principals. The Digital Personal Data Protection Act says that personal data can only be used with peoples consent. There are some exceptions. Generally people must agree to have their data used. This creates a problem for AI models that use peoples data. If an AI model uses peoples data without their consent it might be breaking the law. The law also says that data must be kept safe and secure. The law has rules for companies that handle a lot of personal data. These companies are called data fiduciaries. They have to follow rules to keep peoples data safe. The Digital Personal Data Protection Act also gives people the right to know if their data is being used. They can ask to see their data. They can ask to have it deleted. This can be hard for AI models because once data is used to train a model it can't be easily deleted. The law also says that companies must have plans in place to keep peoples data safe. They must also be able to show that they are following the law. The government will make sure that companies follow the law. If companies don't follow the law they can be fined.

3. The Indian Penal Code / Bharatiya Nyaya Sanhita, 2023

The Indian Penal Code and the Bharatiya Nyaya Sanhita are laws that say what is a crime in India. These laws apply to everyone, including people who use AI models.These laws say that it is a crime to cheat or trick people. If an AI model is used to cheat or trick people, the people who made the model can be in trouble. The laws also say that it is a crime to hurt someones reputation. If an AI model is used to spread information about someone, the people who made the model can be in trouble. The laws also say that it is a crime to threaten or

harass people. If an AI model is used to threaten or harass people, the people who made the model can be in trouble. The government will use these laws to make sure that AI models are not used to hurt people.

4. The Copyright Act, 1957

The Copyright Act is a law that protects works, like books, music and movies. This law is important for AI models because AI models often use works to learn. The law says that people who make works have the right to control how their works are used. If an AI model uses works without permission it might be breaking the law. The government will make sure that people follow the law. If people don't follow the law they can be, in trouble. The Copyright Act does have some exceptions like using works for research, study, criticism, review and reporting current events.. These exceptions are not clearly defined for AI training on a large scale. The law does not explicitly allow for "text and data mining" or "AI training". So courts and policymakers are trying to figure out if existing exceptions can be applied to AI model training. Some people think that using training data is transformative and does not harm the creators. They believe it may fit within public interest exceptions.. Others warn that this could undermine the incentive for creators to produce new works. For government-funded models the lack of clarity is a problem. The government cannot rely on interpretations when using public funds and potentially affecting the rights of many creators. Another issue with the Copyright Act is model outputs and derivative works. When a model is trained on copyrighted material its outputs may. Replicate elements of those works. This raises questions about whether the outputs themselves infringe copyright.

- The Act protects the expression of ideas, not ideas themselves.
- It can be hard to draw a clear line between inspiration and copying in AI-generated text, images or code.

If a government-funded model produces outputs to copyrighted works the government or its partners may face liability. This is especially significant in the sector, where the government may publish AI-generated reports, policy drafts or educational materials. The Act also raises questions about ownership and licensing of AI-related works. The Act does not address whether AI-generated outputs are entitled to copyright protection or who the author is.

- Indian courts and copyright offices have not yet fully resolved the status of AI-generated works.

- This leaves uncertainty, about the status of model weights, prompts and outputs.

For government-funded models this uncertainty is a problem. Public funding may be used to develop AI systems that are later commercialized or shared in ways that implicate copyright liability. If the government requires source or open-weight disclosures as part of its funding conditions it must ensure that such disclosures do not expose creators to uncompensated use of their works.

5. Relevant Legal and Policy Frameworks

Besides the main laws there are other legal and policy frameworks that shape the environment for state-funded foundational Artificial Intelligence models in India. These include regulations to certain sectors developments in constitutional law and guidelines that are not legally binding but still important. These frameworks help explain how the law tries to fill gaps left by the laws even if they do not yet provide a clear set of rules for Artificial Intelligence

One important layer is the regimes in areas like finance, telecommunications, health and insurance. These sectors have their authorities that issue rules on data protection, cybersecurity and consumer protection that apply to Artificial Intelligence used in those areas. For example rules in the sector might require Artificial Intelligence based credit risk models to be transparent and auditable while rules in the health sector might govern how Artificial Intelligence assisted diagnostics can be used. These sectoral rules create a patchwork of Artificial Intelligence related standards that vary by domain even though the underlying model technology might be the same.

Another crucial dimension is law and the principles of proportionality, non-arbitrariness and the right to equality and dignity. Indian courts have held that the state must ensure its use of power is fair, transparent and justified and does not discriminate against groups. When foundational Artificial Intelligence models are used in high-risk domains like welfare eligibility screening or policing these constitutional doctrines can be invoked to challenge biased or opaque Artificial Intelligence assisted decisions.

There are also India- Artificial Intelligence policy and governance guidelines that have emerged in recent years. These documents emphasize principles like transparency, fairness, accountability and human oversight. Recommend that Artificial Intelligence systems be

subject to risk-based oversight and public consultation mechanisms. These guidelines are not legally binding. Signal the direction in which Indian regulators are moving.

Finally there is an emerging framework that seeks to integrate Artificial Intelligence regulation with Indias broader digital public infrastructure agenda. This includes initiatives like Aadhaar, UPI, DigiLocker and Bhasini which generate quantities of data that could be used for Artificial Intelligence training or deployment. The government has signalled that it intends to develop a Digital India-style governance framework that will update Indias architecture for the Artificial Intelligence and platform era.

LEGAL CHALLENGES

Ambiguity in Liability and Accountability

When a state-funded foundational Artificial Intelligence model causes harm it is difficult to say who is legally responsible. This is because Artificial Intelligence is a complex system with layers and the states role as funder might be seen as indirect. This creates a risk that no one feels fully accountable and victims of Artificial Intelligence related harm might find it hard to get compensation. Under existing laws Indian law relies on tort, contract and criminal law doctrines rather than a dedicated Artificial Intelligence liability regime. The Information Technology Act and the Digital Personal Data Protection Act mainly target intermediaries and data fiduciaries not Artificial Intelligence model designers. Do not clearly specify how liability should be shared between a state-funded project and its private partners.

Data Protection and Training Data Legality

Foundational Artificial Intelligence models are trained on datasets that may include personal data. Using data without clear rules on consent, anonymization and lawful basis creates a risk that the state is processing personal data in ways that are not fully compliant with privacy norms. At the time insisting on granular consent for every data point makes large-scale training practically impossible. The Digital Personal Data Protection Act requires that personal data be processed on the basis of valid consent or specific statutory grounds and that data fidelity be high while minimizing data retention. For state-funded models this means that training pipelines must be documented and justified and individuals must be able to exercise their rights to confirmation, correction and erasure. However once personal data is absorbed into model weights erasure becomes infeasible and re-identification risks after anonymization remain high.

Algorithmic Bias and Constitutional Rights Violations

Foundational models can. Amplify biases present in their training data, such as underrepresentation of marginalized communities or gender-based stereotypes. When such models are state-funded and used in welfare eligibility policing risk assessment or judiciary-adjacent functions biased outputs can translate into harms. Indian constitutional law requires non-arbitrariness, fairness and due process. No existing law clearly defines what constitutes "fair" Artificial Intelligence assisted decision-making. The Digital Personal Data Protection Act and the Information Technology Act require transparency and security. They do not mandate bias audits, model cards or explainability mechanisms for Artificial Intelligence systems.

Copyright and Training Data Ownership Conflicts

Training datasets for models often consist of copyrighted text, images, code and audiovisual content. When the state funds or supports models that use material at scale it may indirectly authorize large-scale copyright use that right holders can challenge. The absence of a statutory exception for Artificial Intelligence training creates a hostile environment, for both innovation and creators rights. The Copyright Act of 1957 is a law that protects the rights of authors, publishers and creators. This law does not currently recognize an exception for training artificial intelligence models or for mining text and data. Courts have to decide whether existing laws can justify using copyrighted works to train models. There is no agreement on the answer to this question. If a state-funded model is found to rely on copyrighted Indian language content the owners of the content may start legal proceedings against the government. The government may then face pressure to either pay a lot of money or limit the use of the model. The law is also unclear on who owns the outputs generated by intelligence. This creates uncertainty if these outputs are published or used to make money through public sector channels.

Intermediary Liability and AI Governed Speech

There are challenges when state-funded models are used in public sector platforms, media support tools or social media systems. These models can be used to recommend, rank, filter or generate content. This raises questions about whether these platforms should be treated as intermediaries and whether the model itself should be liable for content it helps spread or create. The fact that the model is funded by the state adds another layer of complexity. The

government is. A regulator of speech and a potential amplifier of harmful content. The IT Act of 2000 and the Intermediary Guidelines provide protection for intermediaries that follow diligence and remove unlawful content when notified. However these laws were written for content uploaded by humans not content generated or recommended by intelligence. The law does not clearly specify how artificial intelligence algorithms or moderation tools fit into the framework. This leaves state-funded intelligence systems vulnerable to either overbroad censorship or underregulation. The DPDP Act and the Penal Code add complexity by criminalizing certain forms of online speech and requiring data security. However they do not explain how liability should be allocated between the platform, the artificial intelligence developer and the state funder when artificial intelligence moderation or recommendation systems fail to prevent content.

Cross-Border Data Flows and Compute Sovereignty

There are challenges when state-funded models require large-scale computing infrastructure, including cloud servers and GPU clusters hosted abroad. Training datasets that contain citizens personal data may be processed or stored overseas. This raises concerns about data sovereignty. Creates tension between technical efficiency and national data protection goals. Restricting computing to data centers alone may be technically and financially impractical.

The DPDP Act grants the Central Government power to regulate or restrict cross-border data flows. It requires that such transfers meet safeguards, including standards for recipient jurisdictions. For state-funded models this means that the government must ensure that any foreign hosting or processing of training data complies with these conditions. The IT Act-style security rules and CERT-In guidelines require that critical data infrastructures be protected. However they do not provide intelligence-specific guidance on how to secure training data pipelines that span multiple jurisdictions. This creates a compliance burden and a potential conflict between innovation needs and data protection and sovereignty requirements.

CONCLUSION

The emergence of state-funded intelligence models is a significant step in Indias journey toward sovereign digital capability and inclusive governance. However it also exposes seated legal and constitutional issues that the current regulatory architecture is not equipped to resolve. The combination of existing laws, including the Information Technology Act, the

Digital Personal Data Protection Act, the Indian Penal Code and the Copyright Act creates a patchwork of obligations that govern data processing, cyber offenses, criminal liability and intellectual property use. None of these statutes explicitly addresses the risk character of foundation models. As a result state-funded artificial intelligence projects are caught between the good logic of innovation and the legal realities of fragmented and often contradictory regulation. This may undermine confidence while exposing the state to accountability gaps. The principal challenges, including ambiguity in liability and accountability tension between data protection principles and training data necessity, algorithmic bias and constitutional rights violations, copyright and ownership conflicts, intermediary liability and artificial intelligence governed speech and cross-border data flows and compute sovereignty all stem from the fact that Indias legal framework predates the rise of intelligence infrastructures and has been stretched to fit them. Addressing these issues will require a coordinated approach that combines statutory reform, technical standards and institutional oversight. A future artificial intelligence law should explicitly recognize state-funded models as a distinct category subject to enhanced transparency, bias audits and public scrutiny mechanisms. It should also clarify data protection rules for training data pipelines, -border processing and artificial intelligence-generated works. The government must ensure that its own artificial intelligence projects do not exploit gaps to create opaque or discriminatory systems. Liability should be shared fairly between private actors, in hybrid artificial intelligence ecosystems. By aligning Indias artificial intelligence governance framework with its constitutional commitments to equality, dignity and due process can state-funded foundational artificial intelligence models become trustworthy public good infrastructures.

References

1. Primary Sources – Statutes and Government Documents

1. Information Technology Act, No. 21 of 2000, Acts of Parliament, 2000, India.
2. Digital Personal Data Protection Act, No. 23 of 2023, India.
3. Bharatiya Nyaya Sanhita, 2023, Act No. 41 of 2023, India.
4. Copyright Act, No. 14 of 1957, India.
5. IndiaAIMission , call for proposals to Build Foundational AI Models, Ministry of Electronic and Information Technology

https://indiaai.gov.in/article/inidaai-mission-call-for-proposal-to-build-foundational-ai-models

6. Ministry of Electronic and Information Technology , India AI Governance Guidelines, November 2025, Press Information Bureau,

   http://static.pib.gov.in/WriteReadData/specificdocs/documents/2025/nov/doc2025115685601.pdf.

7. Principle Scientific adviser, Government of India, Strengthening AI Governance Through Techno-Legal Framework, January 2026,

   https://psa.gov.in/CMS/web/sites/default/files/publications/AI-WP_TechnoLegal.pdf.

8. Principal Scientific Adviser, Government of India, Advancing Indigenous Foundation Models, 2026,

   https://psa.gov.in/ai-foundational-model-white-paper

2. Secondary Sources – Law Articles, Reports, and Policy Analyses

9. Lakshminarayan, S. "Governing Generative AI in India: Constitutional Limits and Regulatory Pathways", Indian Journal of Law and Artificial Intelligence,

10. Jain, A. & Yadav, S. *"Why India Needs a Concrete AI Policy Framework"*, The Regulatory Review 2026,

    https://www.theregreview.org/2026/02/20/jain-yadav-why-india-needs-a-concrete-ai-policy-framework/.

11. Ministry of Electronics and Information Technology, India AI Mission and Indigenous Foundational Model – Policy Brief, 2025

    https://indiaai.gov.in/article/building-india-s-fpundational-ai-models-indiaai-innovation-initiatives.

12. LawAsia, "India AI Regulation: Legal Framework Guide", 2025,

    https://law.asia/india-ai-regulation-legal-framework/.

13. White & Case LLP, India's AI Regulation at the crossroads: AComprehensive Law or a Sectoral-Regulatory-Patchwork?, 2025

https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-india.