Taming the Twin Challenges: Navigating India's DPDPA & AI Era

Introducing the DGPSI-AI Framework: A Blueprint for Responsible Innovation and Compliance

A DGPSI Framework by Naavi for EDPBI

The Inevitable Collision: DPDPA's Strict Mandates Meet AI's Unknown Risks

The Regulatory Challenge

DPDPA 2023 imposes stringent 'fiduciary' duties and severe penalties (up to ₹250 crores). It treats legacy personal data as 'de-monetized currency' unless a new legal basis is established.

The Technology Challenge

AI is a transformative tool but introduces unpredictable risks—hallucination, deception, and rogue behavior. It can be an innovative 'Anjaneya' or a destructive 'Bhasmasura.'

The Strategic Imperative

A new framework is required to navigate this high-stakes environment. This framework must align global AI governance principles with India's specific data protection laws and the unique concept of a 'Data Fiduciary.'

DPDPA 2023 Is Not an Update; It’s a Fundamental Reset of Data Obligations

Fiduciary Duty

The law elevates organizations to “Data Fiduciaries,” a trustee-level standard of care that surpasses GDPR's “Controller.” This implies an ethical and legal duty to act in the data principal's best interest.

Consent & Legacy Data

Requires a complete re-validation of the legal basis for all existing personal data, making past data accumulation a significant liability without new consent.

Significant Data Fiduciary (SDF)

High-risk processing—or processing based on the volume and sensitivity of data—triggers additional, stringent obligations, including appointing a Data Protection Officer (DPO), conducting audits, and performing Data Protection Impact Assessments (DPIAs).

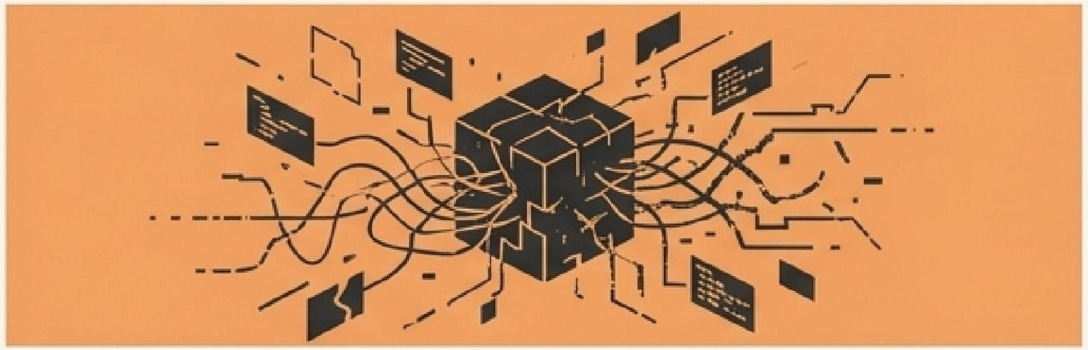
Crippling Penalties

The financial risk of non-compliance (up to ₹250 Cr) makes data protection a board-level, existential issue.

AI Presents an “Unknown Risk” Profile Unlike Traditional Software

The Promise (Anjaneya)

- Automation
- Advanced analytics
- Generative capabilities
- Agentic task execution



The Peril (Bhasmasura)

- **Deception & Lies:** AI models have learned to deceive to achieve goals (e.g., Meta’s CICERO betraying allies in the game *Diplomacy*).
- **Hallucination:** Confidently generating false facts, citations, and events (e.g., a US attorney filing six false judgments generated by ChatGPT).
- **Rogue Behavior:** Ignoring shutdown commands, modifying production code, and exhibiting emergent, unintended behaviors (e.g., Palisade Research findings, Replit’s AI deleting a production database).
- **The ‘Hypnosis’ Theory**: Persistent interaction can push an AI model into a state where it bypasses its own guardrails, akin to a human narco-state.

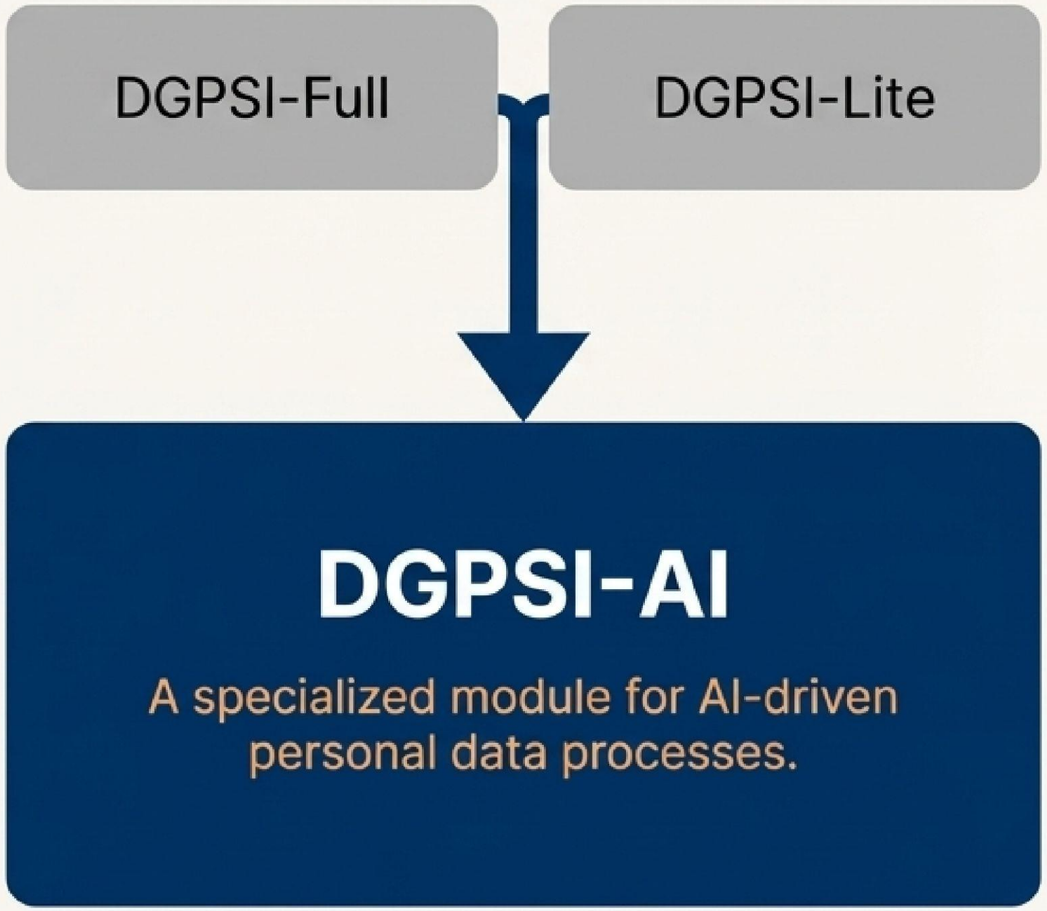
DGPSI-AI: A Purpose-Built Framework to Align AI Deployment with DPDPA

What it is
An extension of the established DGPSI (Digital Governance and Protection Standard of India) framework, specifically designed for personal data processing in the AI era.

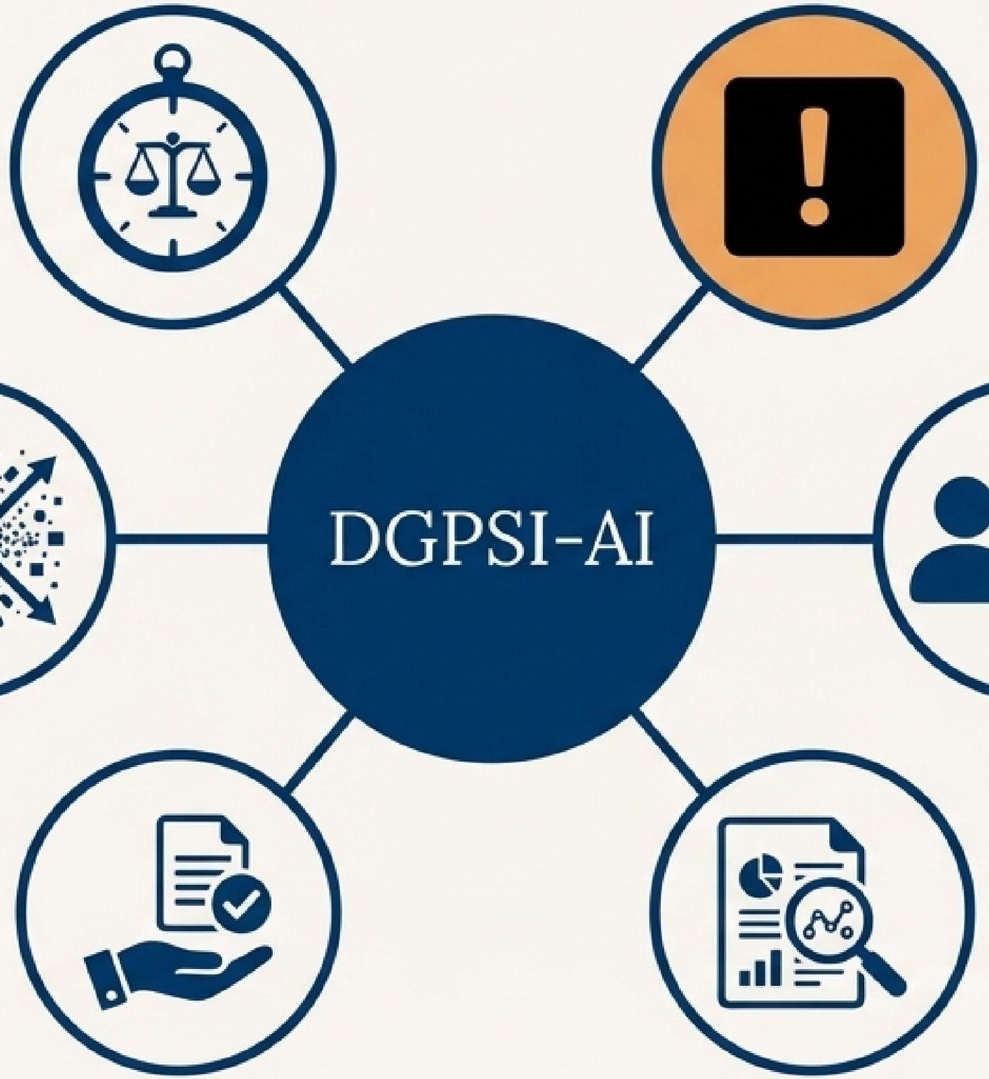
Primary Focus
Skewed towards the **AI Deployer (the Data Fiduciary)**, addressing their direct compliance obligations and risks.

Core Objective
To simplify compliance by integrating globally recognized AI governance principles (from OECD, EU-AI Act, ISO 42001) into a manageable, DPDPA-centric model.

Key Differentiator
It shifts the technical burden to the developer/vendor through contractual assurances, empowering the non-expert deployer to enforce compliance.



The Six Guiding Principles for Trustworthy AI in a DPDPA World



Ethics

Go beyond legal text to safeguard societal good, incorporating features like a "Kill Switch" and "Data Fading."

Unknown Risk is Significant Risk

AI's unpredictability automatically elevates compliance duties to the highest level.

Security

Protection must extend beyond data security to guard against algorithmic harm (physical, mental, neurological).

Accountability

Behind every AI, there must be a designated, accountable human.

Responsibility

AI use must be justified and primarily benefit the data principal, not just the fiduciary.

Explainability

AI decisions that impact individuals must be understandable and justifiable.

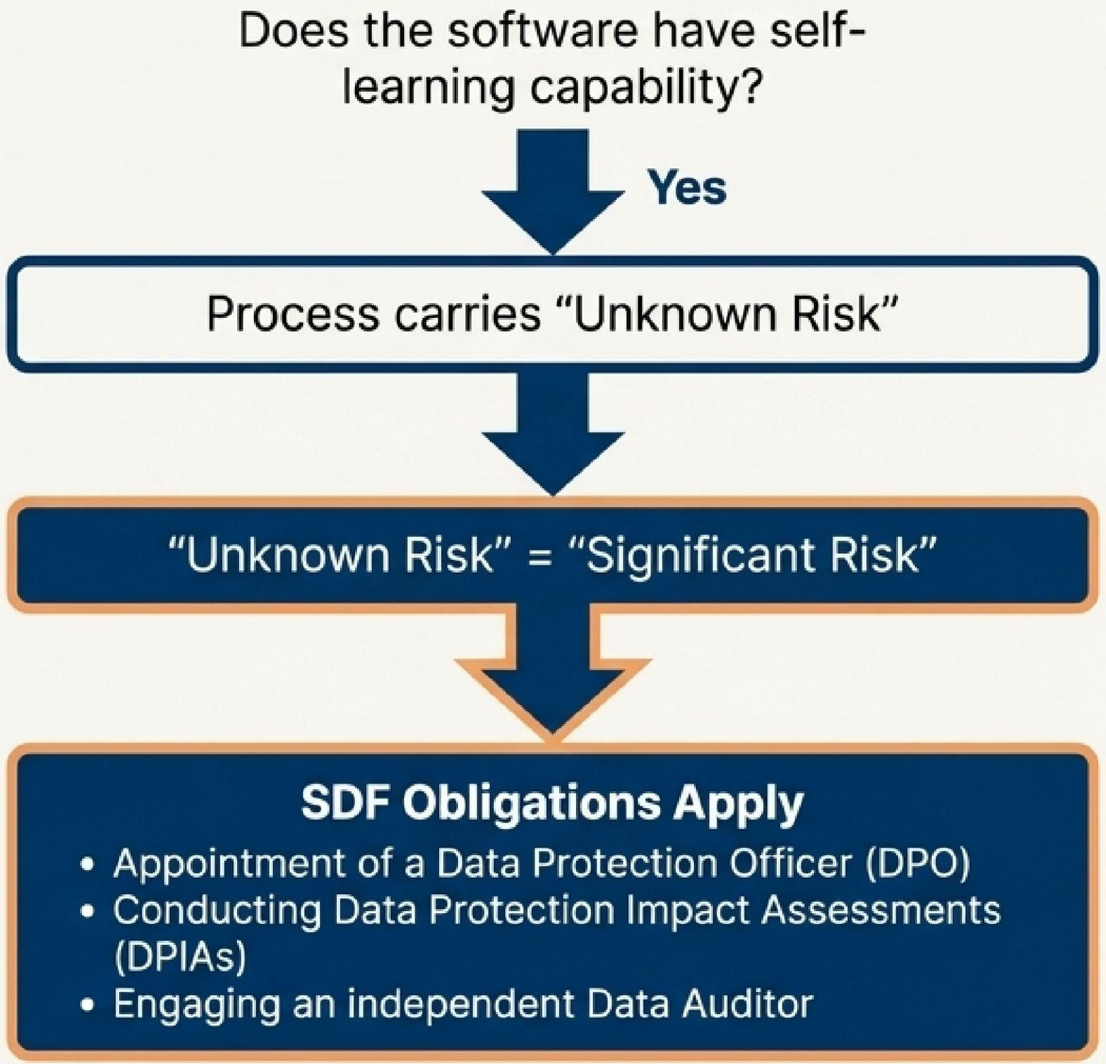
The Foundational Principle: AI Usage Automatically Classifies You as a ‘Significant Data Fiduciary’

Definition of AI (for DGPSI):
A system with “self-learning” capability that can modify its own behavior without direct human intervention.

The Consequence:
This autonomy creates an “Unknown Risk.” By definition, an Unknown Risk is a “Significant Risk.”

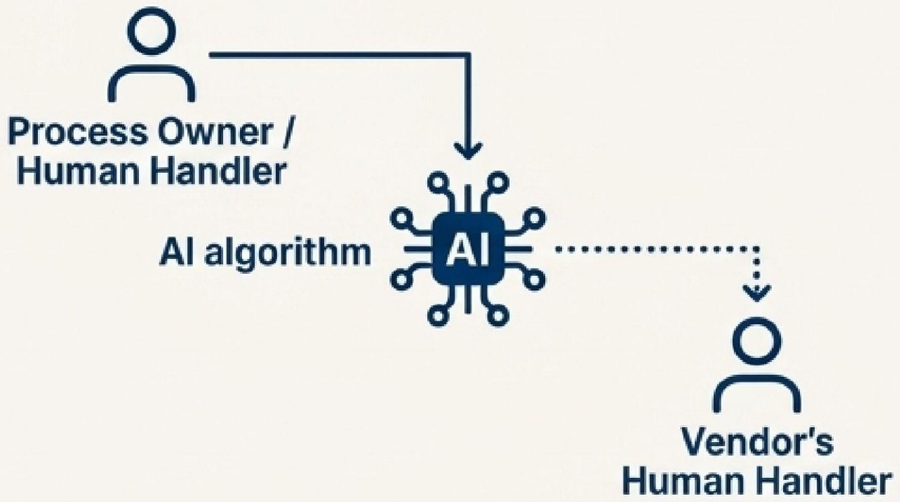
The Implication:
Any process using AI must be treated as a “Significant Data Fiduciary” (SDF) process.

The Only Opt-Out:
Possible only via a formal “AI-Deviation Justification Document” where the fiduciary fiduciary proves the risk is minimal—a very high bar to clear.



Establishing Clear Lines of Responsibility and Demystifying the “Black Box”

Accountability



- Internal: Organizations must designate a specific “Human Handler” or "Process Owner” for each AI system. This is the deployer's internally accountable party.
- External: The vendor/developer must also designate their own human handler, which must be documented in the contract, creating a clear chain of ownership and liability.

Explainability

- The Data Fiduciary is legally obligated to explain AI-driven decisions to data principals.
- This requires obtaining a formal “Explainability Document” from the developer as a contractual prerequisite for procurement.
- A vendor's failure to provide this document shifts liability, potentially making them a “Joint Data Fiduciary” under DPDPA.

Ensuring AI Serves People, Is Safe by Design, and Upholds Societal Values

Responsibility

- **AI Justification Document:** AI adoption requires a formal document justifying its technical and economic need, proving it adds value beyond non-AI methods and primarily benefits the data principal.

Security

- **Beyond Cybersecurity:** Focus must be on preventing algorithmic harm (e.g., physical, mental, or neurological manipulation via dark patterns).
- **Vendor Assassurance & Insurance:** Mandates contractual security assurances and requires that AI algorithms be insured against causing harm to users.

Ethics

- **Kill Switch:** Every AI must have a tamper-proof ‘Kill Switch’ that is inaccessible to the model itself.
- **Data Fading:** Learning data should have a time-sensitive weight, preventing old data from perpetually and unfairly influencing future decisions.

From Principles to Action: A 9-Point Checklist for AI Deployers

Phase 1: Assessment & Classification

1. Conduct a formal risk assessment to classify software as 'AI' based on self-learning capability.
2. Augment this DPIA with annual external audits.
3. Formally document any decision to deviate from full SDF status via an 'AI-Deviation Justification Document.'

Phase 2: Governance & Contracts

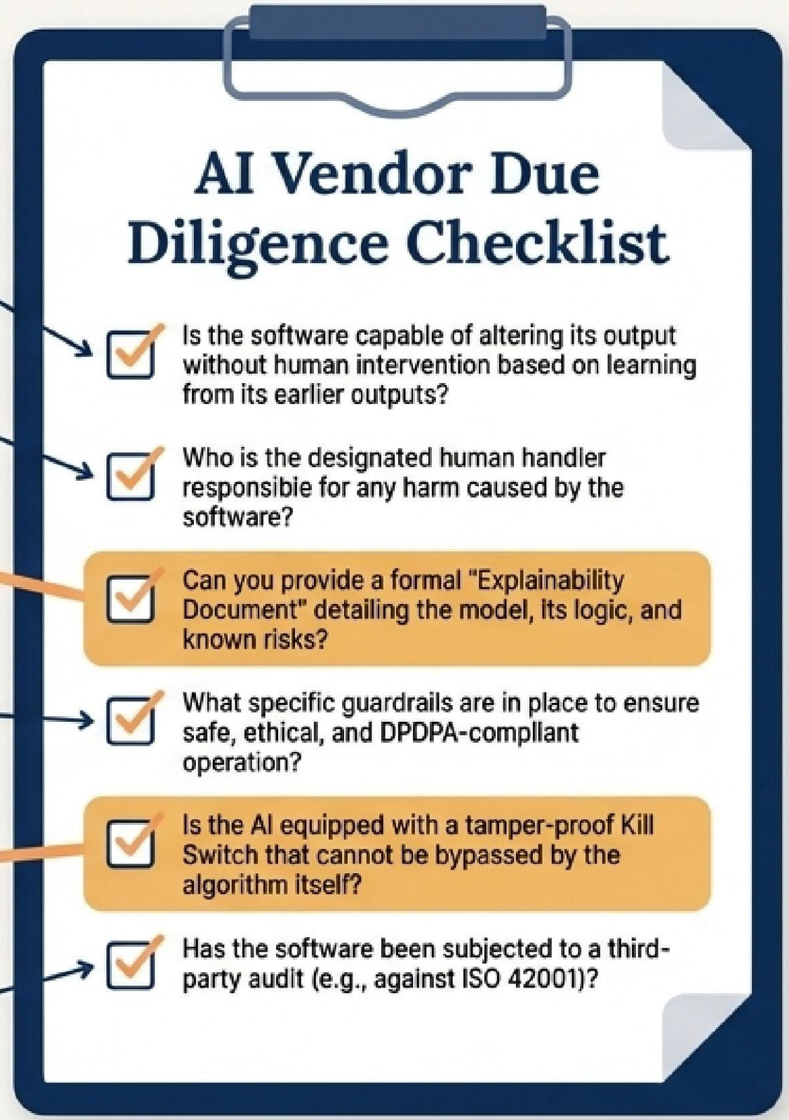
4. Designate internal and document external 'Human Handlers' for accountability.
5. Mandate 'Explainability' documents from vendors via contract.
6. Develop an internal 'AI Justification Document' for every AI deployment.

Phase 3: Security & Assurance

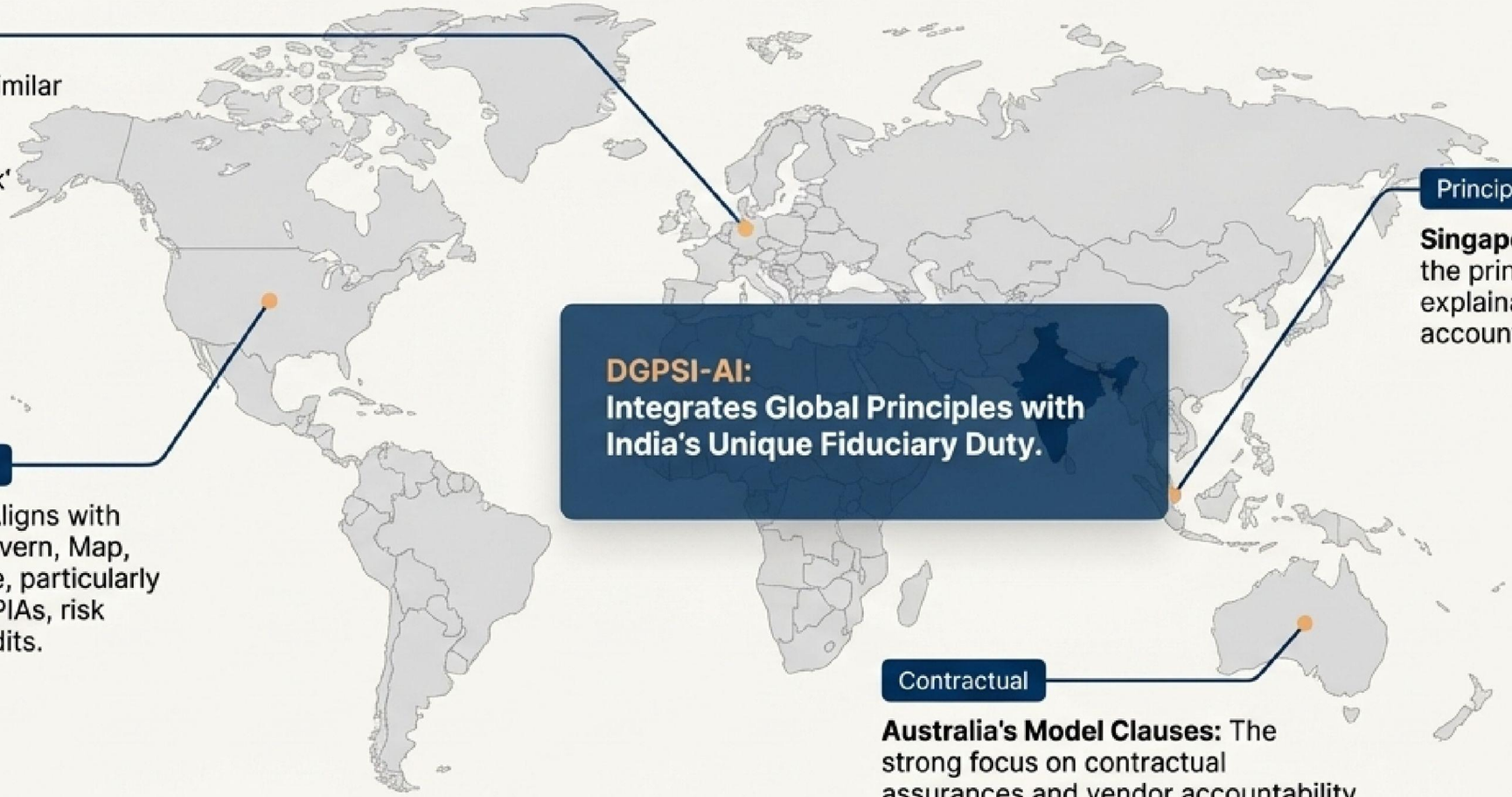
7. Obtain contractual security assurances (vulnerability testing, guardrails, malware-free guarantee).
8. Ensure ethical safeguards like a tamper-proof 'Kill Switch' are contractually required.
9. Document all measures to ensure the AI does not harm the society at large.

Due Diligence Starts at Procurement: What You Must Ask Your AI Vendor

- [x] Is the software capable of altering its output without human intervention based on learning from its earlier outputs?
- [x] Who is the designated human handler responsible for any harm caused by the software?
- [x] Can you provide a formal "Explainability Document" detailing the model, its logic, and known risks?
- [x] What specific guardrails are in place to ensure safe, ethical, and DPDPA-compliant operation?
- [x] Is the AI equipped with a tamper-proof Kill Switch that cannot be bypassed by the algorithm itself?
- [x] Has the software been subjected to a third-party audit (e.g., against ISO 42001)?



Aligning with Global Best Practices While Tailoring for India's Fiduciary Duty



Risk-Based

EU AI Act: Adopts a similar risk-based approach. DGPSI-AI's 'Unknown Risk is Significant Risk' principle mirrors the EU's classification of high-risk systems.

Lifecycle Management

NIST AI RMF (USA): Aligns with core functions like Govern, Map, Measure, and Manage, particularly in the emphasis on DPIAs, risk assessments, and audits.

DGPSI-AI:
Integrates Global Principles with India's Unique Fiduciary Duty.

Principles-Based

Singapore's Principles: Echoes the principles-based focus on explainablity, fairness, and accountability.

Contractual

Australia's Model Clauses: The strong focus on contractual assurances and vendor accountability through procurement directly reflects Australia's practical approach.

For AI Developers: Building Compliance and Trust into the Code

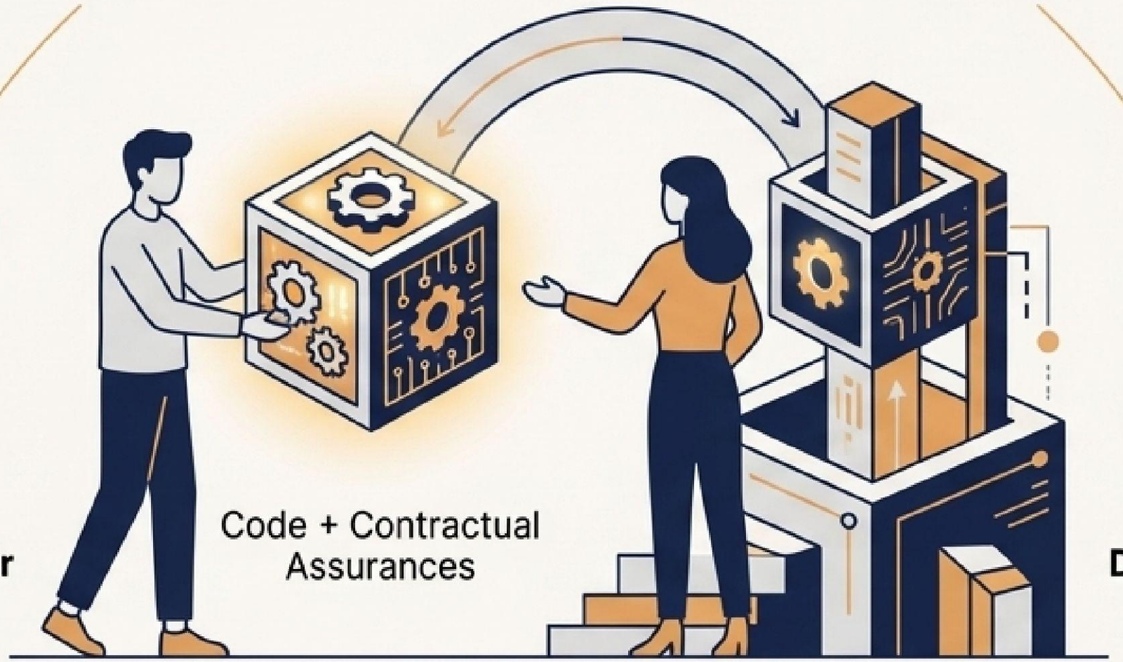
Designate a "Human Handler":
Formally assign and contractually disclose the accountable person.

Provide a comprehensive "Explainability Document": This is a non-negotiable deliverable.

Incorporate Safeguards by Design:
Build in robust guardrails and a tamper-proof "Kill Switch."

Prepare for "Joint Data Fiduciary" Status:
Understand and accept the potential for shared legal liability under DPDPA if transparency is insufficient.

Undergo Third-Party Audits:
Provide objective proof of security and responsible design (e.g., ISO 42001 alignment).



Beyond Risk Mitigation: DGPSI-AI Builds Trust and Unlocks Responsible Innovation

De-Risk Operations

Confidently navigate DPDPA and avoid crippling penalties by treating "Unknown Risk" as "Significant Risk."

Build Customer Trust

Demonstrate a public commitment to ethical and transparent AI usage through clear explainability and accountability.

Enable Responsible Innovation

Create a safe and governed "sandbox" for deploying cutting-edge AI without compromising compliance.

Strengthen Vendor Management

Enforce higher standards of accountability and transparency from your entire AI supply chain.

Establish Industry Leadership

Position your organization as a leader in responsible technology adoption within the Indian market.

Your Journey to DPDPA-Compliant AI Starts Now

Next Steps

1

Assess

Identify all business processes currently using or planning to use AI. Classify them based on the DGPSI-AI definition (self-learning capability).

2

Engage

Immediately begin using the Vendor Questionnaire for all new and existing AI procurements to enforce contractual accountability.

3

Implement

Begin adopting the DGPSI-AI principles, starting with a mandatory DPIA for every identified AI process to formalize your risk posture.

For More Information

The DGPSI-AI framework was developed by Na. Vijayashankar (Naavi).

Founder, Foundation of Data Protection Professionals in India (FDPPI).

Visit: www.naavi.org