DSCI
PROMOTING DATA PROTECTION

Brief on

Office of the PSA's Discussion Paper

"Enabling Cross-Border Data Interoperability for AI Systems"

(September 2026)

DSCI Brief: Office of the PSA's Discussion Paper, *"Enabling Cross-Border Data Interoperability for AI Systems"* (September 2026)

The Office of Principal Scientific Adviser to the Government of India released a discussion paper on Enabling Cross-Border Data Interoperability for AI Systems on 17th September 2026. This paper places and addresses cross-border data interoperability for AI at the core - an issue few jurisdictions have addressed with such clarity and structure, even as the global AI industry, regulators, and research institutions grapple with it simultaneously -squarely on the national policy agenda at a time when it could not be more consequential.

This paper's treatment of why India needs a risk-based approach to international data interoperability (Chapter 2), how other major jurisdictions and regional blocs have approached the same problem (Chapter 4), and the AI-specific framework the paper proposes as a way forward (Chapter 5). Throughout, it situates these arguments against India's existing regulatory landscape - the Digital Personal Data Protection Act, 2023 (DPDP Act), the RBI's payment-data localisation mandate, the National Data Governance Committee (NDGC), and emerging bilateral arrangements such as the ICMR–France Health Data Hub collaboration - to show where the paper's recommendations reinforce, extend or challenge current Indian practice.

A Case for Risk-Based Approach to Data Flows

The case for a risk-based approach begins with a sober accounting of the costs at both extremes of the localization-openness spectrum. Strict data localization, though often justified by appeals to national security and regulatory oversight, imposes real economic costs: it raises compliance burdens disproportionately on startups and MSMEs, fragments the datasets available to domestic AI researchers, and can invite retaliatory trade measures from partner countries. At the other end, unrestricted data flows create their own vulnerabilities - exposure to foreign surveillance, diminished regulatory oversight once data leaves Indian jurisdiction, and the concentration of data (and therefore market power) in the hands of a small number of global technology firms.

The paper cites modelled estimates suggesting that balanced, safeguarded data flows could lift global GDP by roughly 1.77% and exports by about 3.6%, whereas highly restrictive or fragmented regimes could depress global GDP by close to 4.5% and exports by around 8.5%[1]. While these are computable general-equilibrium projections rather than observed outcomes, they underline a point that has real force for Indian policymakers: uncoordinated national approaches to data governance risk contributing to a fragmented *splinternet*[2], raising compliance costs and constraining the very cross-border research collaboration on which advanced AI development increasingly depends.

[1] OECD/WTO (2025), *Economic Implications of Data Regulation: Balancing Openness and Trust*, OECD Publishing, Paris, https://doi.org/10.1787/aa285504-en

[2] Global Commission on Internet Governance. (2016). *Our Internet: The collected research papers of the Global Commission on Internet Governance* (Vol. 1). Centre for International Governance Innovation.

This scale is precisely what makes India's data both a strategic asset and a strategic vulnerability. The paper suggests a middle path built around risk-based and sector-specific governance - trusted data corridors, mutual recognition frameworks, standards-based interoperability using privacy-enhancing technologies, and hybrid models such as data embassies or sovereign clouds - with the intensity of control calibrated to the sensitivity of the sector and the data involved.

POLICY INSIGHT - I

India already straddles this spectrum in practice rather than committing to either extremes. The RBI's 2018 payment-data localisation direction reflects a localization-first instinct for a systemically sensitive sector, while the DPDP Act, 2023 adopts a comparatively liberal negative-list model for personal data generally.

The paper's argument that uniform, all-or-nothing rules are economically costly lends support to India's continued reliance on sectoral regulators (RBI, MeitY, health authorities) to calibrate localization requirements by sector rather than through a single blanket law - a structure the DPDP Act's Section 16 explicitly preserves by allowing sector-specific rules to sit alongside the general cross-border transfer regime.

The paper's caution about the compliance burden on startups and MSMEs is directly relevant to how India designs the notified restricted-country list under Section 16: an overly broad or frequently changing list could reproduce the very costs the paper warns against.

Data Sovereignty as a Thread Running Through the Paper

Although data sovereignty is introduced early in the paper as a foundational concept, it is not treated as a static idea confined to a single section - it recurs and evolves across the discussion, functioning as the constant against which every proposal for interoperability is tested. The paper defines data sovereignty as a country's effort to exercise national control over data generated within its territory, covering its storage, processing and cross-border transfer, and ties this closely to strategic autonomy in digital infrastructure, protection of citizens' personal data, safeguarding of critical information assets, and mitigation of risks from foreign surveillance and external dependence. This is framed not as an abstract principle but as a practical necessity underpinning public administration, national security and socio-economic growth.

The paper is candid about the tension this creates. It acknowledges that interoperability, by its very design, asks jurisdictions to make their data usable and trustworthy beyond their own borders - and that this can come at the cost of exactly the control that sovereignty is meant to preserve. **Four specific risks are identified:**

(i) a loss of domestic control over data once it leaves the jurisdiction, particularly where foreign legal protocols or private contractual arrangements come to determine how the data is treated;

(ii) heightened sovereignty risk in strategic sectors such as finance, health, telecommunications and other critical functions, where the case for openness is far weaker than for low-sensitivity commercial data;

(iii) asymmetry in digital dependence, where interoperability may disproportionately benefit countries and firms that already hold outsized power in the global data economy; and

(iv) downstream or onward-transfer risk, where the country of origin loses visibility over how its data is subsequently processed, combined or accessed once it has moved across the first border.

The paper adds a distinctly Indian dimension to this last point, noting that **semantic interoperability — the push toward standardized definitions, taxonomies and metadata** — may not fully capture India's linguistic diversity, administrative structure and socio-cultural context, so that harmonization for cross-border comparability could come at the cost of contextual accuracy in Indian data.

Rather than treating these risks as reasons to retreat from interoperability altogether, the paper resolves the tension by embedding sovereignty preservation directly into the architecture of its proposed framework. The **risk-based classification layer** reserves a distinct 'sovereign' tier for strategic or national-security-related data, under which raw data does not leave the country at all — suggesting that only outputs from secure, domestic model-to-data or federated processing may be released, and only after strict scrutiny. The **compatibility pillar** is deliberately designed so that shared technical and semantic standards do not require any country to replace its own domestic taxonomies or classification systems; where exact equivalence between two countries' categories is not possible, the paper calls for the mapping to document contextual differences rather than force artificial alignment, so that data does not lose or distort its nationally relevant meaning. The **accountability pillar** builds in periodic review and reclassification triggers, so that a change in model capability, data combination or downstream use can prompt a fresh sovereignty assessment rather than treating an initial transfer approval as permanent. And the **coordination pillar** is explicit that any institutional mechanism for cross-border cooperation is meant to function alongside domestic regulators, not in place of them — countries are described throughout as *retaining 'sovereign flexibility'* even as they participate in structured bilateral, plurilateral or multilateral arrangements.

In this sense, the paper's treatment of data sovereignty is less about resolving the tension between control and openness than about making that tension governable. It repeatedly returns to the idea that interoperability does not require identical laws, systems or standards across countries - a formulation that allows India to engage in cross-border data and AI cooperation while still reserving the classification power, the notification power, and the localization power that its own domestic instruments already provide for. The paper's proposed framework is, in effect, an attempt to give India a similarly targeted toolkit — one calibrated closely enough to sovereignty concerns that openness becomes safe to extend, sector by sector and dataset by dataset, rather than something to be granted or withheld wholesale.

POLICY INSIGHT II — Data sovereignty and India's regulatory toolkit

The paper's four sovereignty risks map onto specific instruments India already has, or is building: loss of downstream control is precisely what DPDP Act Section 16's notification power and the yet-to-be-operationalized cross-border transfer rules are designed to pre-empt; sectoral sovereignty risk is what justifies the RBI's payment-data localization direction and analogous sectoral carve-outs the DPDP Act preserves.

The paper's 'sovereign tier' concept gives India a ready conceptual basis for deciding what belongs on the Section 16 restricted/notified list — strategic, critical-infrastructure or national-security-linked datasets -

while keeping lower-risk categories on a more open, negative-list footing, consistent with the Act's general design.

The concern about semantic interoperability eroding India's contextual and linguistic specificity is a caution the National Data Governance Committee and sectoral standard-setters (e.g., in health, via ABDM/EHR standards) would need to actively guard against as they adopt international metadata and taxonomy standards - favoring documented 'mapping with contextual notes' over forced one-to-one equivalence.

The emphasis on sovereignty as an ongoing, reviewable state rather than a one-time transfer approval reinforces the case for the paper's recommended continuous, risk-adaptive assurance mechanism - an idea current Indian instruments (DPDP Act, RBI circulars) do not yet require, since compliance is generally assessed at the point of transfer rather than monitored thereafter.

Learning from Global Approaches to Cross-Border Data Flows

Chapter 4 of the paper undertakes a comparative survey of how India, the European Union, the United States and China regulate cross-border data flows, alongside the more framework-based, soft-law approaches taken by regional groupings such as ASEAN, the African Union and APEC. What emerges is not a single dominant model but a family of related techniques - *negative lists, adequacy decisions, security assessments, certification schemes and access-based restrictions* - each reflecting a different balance between sovereign control and openness.

India's own approach, as described in the paper, is one of controlled and selective interoperability. The DPDP Act, 2023 permits personal data transfers by default but allows the central government to restrict transfers to specific notified countries or territories under Section 16 - a negative-list model that is more permissive in structure than the EU's adequacy-based system, but one whose substantive cross-border provisions are not yet in force *(they are scheduled to commence in May 2027).* Alongside this general regime, India retains a localization-first posture for payment system data under the RBI's 2018 direction, under which data may be processed abroad only in narrow circumstances and must ultimately be stored, and deleted from abroad, within India. The National Data Governance Committee adds an institutional layer, coordinating data-sharing platforms, standardized metadata, APIs, auditability and legal compliance across government - though, as the paper notes, this mechanism was not originally designed with AI-specific concerns such as downstream reuse, model outputs or re-identification risk in mind.

By contrast, the European Union anchors its regime in a principle of continuity of protection: transfers are permitted either where the European Commission has made an adequacy determination (Article 45 GDPR) or where the transferring party has put in place "appropriate safeguards" such as standard contractual clauses or binding corporate rules (Article 46). The United States, lacking a comparable omnibus federal framework, instead relies on an access-based national-security model - Executive Order 14117 and the resulting Data Security Program restrict specific categories of "bulk sensitive personal data" and government-related data from reaching countries of concern, while the US CLOUD Act creates a separate executive-agreement mechanism for lawful law-enforcement access to data held by US-linked providers. China occupies a third position: a differentiated but generally more restrictive regime combining localization for critical information infrastructure operators, mandatory security assessments

for higher-risk transfers, and - since 2024 - relaxed compliance requirements for routine, lower-risk commercial transfers.

The regional models reviewed in the chapter (ASEAN, the African Union, APEC) are instructive for a different reason: they show that binding, identical domestic law is not a precondition for functional interoperability. ASEAN's Model Contractual Clauses and Framework on Digital Data Governance, APEC's Cross-Border Privacy Rules certification system, and the AU's reciprocity-based Data Policy Framework all operate through voluntary convergence, mutual recognition and readiness-based participation rather than a single supranational statute. From this comparative exercise, the paper distils a set of common principles: cross-border governance is increasingly risk-differentiated rather than uniform; countries seek to preserve domestic regulatory authority even while participating in interoperability arrangements; controlled access can substitute for unrestricted transfer where data is sensitive; and institutional coordination becomes essential wherever multiple regulators and standards bodies are involved.

POLICY INSIGHT III — How this maps onto current Indian regulatory provisions

DPDP Act, Section 16 (cross-border transfer): India's negative-list approach is structurally closer to China's differentiated model than to the EU's adequacy system - transfers are the default and restriction is the exception, rather than the reverse. The paper's comparative analysis suggests India could still borrow selectively from the EU's 'appropriate safeguards' toolkit (e.g., standard contractual clauses) for sectors or partners where a blanket notification approach is too blunt, once the Section 16 provisions come into force in 2027.

RBI's Payment System Data localisation direction (2018) exemplifies the 'strict data localisation' end of the spectrum discussed in Chapter 2 and is comparable in spirit to China's localizations requirements for critical information infrastructure - both treat certain data classes as too systemically sensitive for the general cross-border regime.

The National Data Governance Committee is India's closest institutional analogue to the coordination mechanisms the paper highlights internationally (e.g., the EU's Interoperable Europe Board, the Global CBPR Forum), but the paper is explicit that it currently lacks a dedicated mandate for AI-specific interoperability questions - a gap Chapter 5's recommendations directly target.

The ICMR–French Health Data Hub collaboration, using the Data Empowerment and Protection Architecture (DEPA), is cited as India's own emerging example of a 'trusted data corridor' - a bilateral, purpose-specific arrangement of exactly the kind Chapter 4 identifies as a middle path between localizations and open transfer.

Towards an AI-Specific Framework for Cross-Border Interoperability

Building on this comparative foundation, Chapter 5 sets out the paper's central contribution: a proposed framework for AI-specific cross-border data interoperability, **anchored in three pillars - compatibility, accountability and coordination** - all supported by a cross-cutting, risk-based classification and access layer. Crucially, the paper frames this not as a new data-governance regime but as an interoperability layer to be built on top of existing mechanisms, because AI raises a distinct governance question that ordinary data-transfer rules do not fully answer: what crosses the border may not be raw data at all, but

model parameters, analytical outputs, or a fully trained model that has internalized patterns from sensitive training data.

The risk-based access layer is the framework's organizing device. It proposes a **four-tier matrix** running from **aggregate or statistical data** (which can move through open, low-friction channels), through **functional** and **controlled data** (subject to differential privacy or strict contractual and audit safeguards, respectively), to **sovereign-tier data** - strategic or national-security-related information that should not leave the country at all, with only approved, federated-processing outputs permitted to cross borders. This tiering allows India, in principle, to open the door wider for lower-risk AI training data while keeping strategic and security-sensitive data firmly under domestic control, addressing the very trade-off Chapter 2 identifies between innovation and sovereignty.

The **compatibility pillar** calls for shared technical standards (drawing on ISO/IEC AI standards, IETF protocols such as **TLS 1.3** for secure communications and **OAuth 2.1** for delegated authorization, alongside identity standards such as **OpenID Federation 1.0** and the **W3C Verifiable Credentials Data Model 2.1**) and shared semantic standards and metadata, so that data retains consistent meaning as it moves across systems and jurisdictions — without requiring India or any partner country to abandon its own domestic taxonomies. The **accountability pillar** proposes a minimum certification profile (building on ISO/IEC 27001, 27701 and 42001) combined with mutual recognition of certifications and, importantly, a shift from static, point-in-time certification toward continuous, risk-adaptive assurance - recognizing that an AI system's capabilities and risk profile can change substantially after the initial data transfer or model training event. The **coordination pillar** calls for designated nodal authorities, issue-specific working groups, and periodic review mechanisms, explicitly modelled on international precedents such as the Interoperable Europe Act's governance board and the International Network of AI Safety Institutes.

The paper tests this architecture against India's health-data ecosystem, chosen precisely because it already has relatively mature technical and semantic infrastructure - **EHR Standards for India**, **HL7 FHIR** as implemented through the **ABDM Implementation Guide**, **DICOM** for medical imaging and associated information, and clinical terminology standards such as **SNOMED CT** and **LOINC** - while still carrying significant privacy and ethical stakes. It shows that even in this comparatively advanced sector, the pieces required for a fully coordinated cross-border AI pathway (technical standards, legal accountability instruments such as HMSC clearance and ICMR's AI ethics guidance, and institutional coordination among ICMR, DHR and ABDM) exist but are not yet linked into a single operational pathway for AI-specific secondary use, downstream transfer or benefit-sharing.

POLICY INSIGHT IV - Interplay with India's regulatory architecture going forward

The framework's proposed 'central interdisciplinary AI-data interoperability coordination body' is explicitly recommended to sit within the existing National Data Governance Committee rather than as a new statutory authority — a design choice that keeps the recommendation compatible with India's current institutional architecture rather than requiring fresh legislation.

The four-tier risk matrix (Aggregate / Functional / Controlled / Sovereign) offers a ready-made template for operationalizing DPDP Act Section 16 once it takes effect: the 'sovereign' tier maps naturally onto categories

the government may wish to exclude from the general cross-border regime, while 'aggregate' and 'functional' tiers could justify a lighter-touch approach for de-identified or research-oriented datasets.

For health data specifically, the paper's call for standard data-sharing/data-use agreements linked to technical safeguards would formalize what is currently handled through a patchwork of instruments - HMSC clearance, ICMR's 2023 AI ethics guidelines, DPDP Section 16, and case-by-case Material Transfer Agreements - into a single, AI-aware approval pathway, directly relevant to scaling arrangements like the ICMR-France DEPA pilot beyond a single bilateral case.

The proposed shift to continuous, risk-adaptive certification (rather than static certification) has a direct bearing on how Indian regulators might supervise AI models trained on cross-border data after deployment - an area where current Indian instruments (DPDP Act, sectoral circulars) are largely silent on post-transfer monitoring obligations.

Conclusion

Overall, the discussion paper builds a cumulative argument: the economic and strategic case against both extremes of the data-governance spectrum (Chapter 2) is reinforced by comparative evidence that every major jurisdiction has, in practice, settled on some risk-differentiated middle ground (Chapter 4), and this evidence is then translated into a concrete, AI-specific architecture of compatibility, accountability and coordination that India could build on top of its existing instruments rather than replace (Chapter 5).

For India, the practical significance lies less in any single recommendation than in the overall orientation: the DPDP Act's cross-border provisions, the RBI's sectoral localizations rules, and the National Data Governance Committee are not treated as obstacles to be swept aside, but as a foundation on which a more AI-aware, risk-calibrated interoperability layer can be constructed - provided India moves deliberately to close the specific gaps the paper identifies around AI-specific classification, continuous assurance, and dedicated institutional coordination.

*****