Journal of International Research & Multidisciplinary Innovation

Data Protection Risk and Compliance in Startup Investments

A Due Diligence Framework under the Digital Personal Data Protection Act, 2023

An interdisciplinary contribution to venture capital practice and Indian data protection law

Author : Shreya Pandey

Affiliation: University of Allahabad , Prayagraj

Email: 7924shreyapandey@gmail.com

Submitted to: Journal of International Research & Multidisciplinary Innovation (JIRMI)

Date of Submission: 11/06/2026

Abstract

The Digital Personal Data Protection Act of 2023 took its first real bite on 13 November 2025, when the Ministry of Electronics and Information Technology notified the operative Rules and set the clock for full enforcement to 13 May 2027. Indian venture capital was busy with other things. The industry deployed about USD 12 billion across 159,157 DPIIT-recognised startups in 2025 without any standard way of asking whether those startups could survive the new law. This paper builds the missing tool. It is a complete due diligence framework calibrated to the DPDP, designed for use by venture capital firms, private equity funds, angel networks, accelerators, and the lawyers who serve them.

The framework has four parts. A twelve-stage diligence sequence that walks the investor through every place a DPDP problem can hide. A Startup Data Risk Score on one hundred points that converts diligence findings into a single number. A Privacy Compliance Maturity Model with five tiers that lets the investor compare one startup against another. An Investor Risk Rating Framework that turns the score into term-sheet line items: how much to discount the valuation, how large an indemnity to take, how much to keep in escrow, and for how long the survival period should run.

The urgency is not theoretical. The IBM Cost of a Data Breach Report 2025 records that the average breach in India now costs INR 220 million, a thirteen per cent jump over the previous year and the highest figure since IBM began tracking India. 23andMe lost USD 6 billion of enterprise value and entered bankruptcy after a 2023 credential stuffing breach for which mandatory two-factor

authentication would have been enough. Indian incidents at Byju's, Hathway, boAt, BSNL and Mobikwik confirm the same risk lives at every layer of the Indian digital economy. The framework presented here gives the investor a way to detect that risk before the cheque clears.

Keywords. Digital Personal Data Protection Act 2023; DPDP Rules 2025; venture capital diligence; privacy risk; startup investment; Indian regulatory law; data protection compliance; Significant Data Fiduciary; cross-border data transfer.

Methodology

This paper is built on a mixed-method design that combines doctrinal legal analysis, comparative law, and structured empirical synthesis. The choice of method follows the nature of the problem. The DPDP Act 2023 and its 2025 Rules create a set of obligations that must be read closely against the text, against the Constitution, against adjacent statutes, and against the way courts and regulators are likely to interpret them. No single method can do all of that work. The approach used here therefore runs on four tracks that feed into each other.

The first track is doctrinal. The paper reads the DPDP Act section by section, the DPDP Rules rule by rule, and maps each obligation to the specific Data Fiduciary conduct it requires. This is supplemented by close reading of the Information Technology Act 2000 and its SPDI Rules 2011, the CERT-In Directions of 28 April 2022 issued under Section 70B(6), the Reserve Bank of India's Storage of Payment System Data Directions of 6 April 2018, the SEBI LODR Regulations 2015 as amended, the Companies Act 2013, the Consumer Protection Act 2019, and the Bharatiya Nyaya Sanhita 2023. The constitutional anchor is the nine-judge bench decision in Justice K.S. Puttaswamy (Retd.) v. Union of India (2017) 10 SCC 1, read with the Aadhaar judgment in (2019) 1 SCC 1.

The second track is comparative. The DPDP is read against the GDPR, the UK GDPR as modified by the Data (Use and Access) Act 2025, the California Consumer Privacy Act as amended by the CPRA, Singapore's Personal Data Protection Act 2012, Brazil's LGPD, Japan's APPI and Australia's Privacy Act 1988. The comparison is not exhaustive. It is targeted at provisions whose presence in foreign regimes, and whose absence from the DPDP, creates due diligence implications for an investor with multi-jurisdictional exposure.

The third track is empirical. The paper draws on publicly available data from official Indian sources (DPIIT, NASSCOM, CERT-In, the National Crime Records Bureau, the Reserve Bank of India), from reputed industry sources (IBM Cost of a Data Breach Report 2025, Cisco 2025 Data Privacy

Benchmark Study, CMS GDPR Enforcement Tracker, PitchBook Venture Monitor, Tracxn India Startup Reports), and from regulator filings in the United States and Europe. Numbers are reported only where the source can be verified. Case studies are drawn from documented incidents at 23andMe, Meta, LinkedIn, Amazon, Uber, Byju's, Mobikwik, Hathway, boAt and BSNL.

The fourth track is the construction of the original framework. The twelve diligence stages are derived by decomposing the obligations of the Data Fiduciary into atomic compliance requirements, mapping each requirement to its statutory anchor, and overlaying control objectives drawn from ISO/IEC 27001:2022 Annex A and the NIST Privacy Framework v1.0. Governance maturity indicators are imported from COSO ERM (2017) and ISO 31000:2018. The Startup Data Risk Score weights and the Privacy Compliance Maturity Model tiers are calibrated to reflect the relative penalty exposure under the DPDP Schedule and the empirical record of breach causes from the IBM report.

The limitations of this method are owned openly in Section 9. The most important is that the substantive provisions of the DPDP take effect only on 13 May 2027. Indian enforcement data does not yet exist. The framework therefore relies on GDPR and PDPA experience as the closest available comparators and will need recalibration as the Data Protection Board begins to issue decisions.

Literature Review

Indian writing on data protection has run through three phases. The first phase, from the late 1990s to the mid-2010s, was occupied with the question of whether a right to privacy existed at all under the Constitution. M.P. Sharma v. Satish Chandra (1954) and Kharak Singh v. State of Uttar Pradesh (1964) had said it did not. Govind v. State of Madhya Pradesh (1975), R. Rajagopal v. State of Tamil Nadu (1994) and PUCL v. Union of India (1997) had said it did. The position was unresolved.

The second phase opened with the nine-judge bench decision in Justice K.S. Puttaswamy v. Union of India (2017) 10 SCC 1. The Court held that privacy is a fundamental right under Articles 14, 19 and 21 of the Constitution. The judgment overruled M.P. Sharma and Kharak Singh and laid down a three-part proportionality test that has since become canonical. The scholarship in this phase was concerned with what Puttaswamy meant, how it would apply to State surveillance, and how it would shape the data protection statute that was clearly coming. The Justice B.N. Srikrishna Committee report of July 2018, titled A Free and Fair Digital Economy, was the most important policy document of this phase. The Personal Data Protection Bills of 2018, 2019 and 2022 were drafted, debated, and ultimately replaced by the DPDP Bill that was passed in August 2023.

The third phase began with the enactment of the DPDP Act and accelerated after the notification of the DPDP Rules on 13 November 2025. Much of the writing in this phase is exegetical. Practitioner

firms have produced explanatory guides. EY, KPMG, PwC, Deloitte, Cyril Amarchand Mangaldas, AZB and Khaitan have each published practitioner notes on the Act and Rules. The Centre for Internet and Society and the Vidhi Centre for Legal Policy have contributed policy-oriented analysis. The scholarly journals have published on specific issues: the constitutional status of consent (Bhandari and Sane), the architecture of the Data Protection Board (Chinmayi Arun), and the relationship between the DPDP and sectoral regulation (Rishab Bailey).

The literature has two important gaps that this paper addresses. First, almost all Indian writing treats the regulated entity, the Data Fiduciary, as the unit of analysis. The investor in the Data Fiduciary, who inherits the contingent regulatory liabilities of the target through the act of investment, is rarely modelled separately. Some Indian commentaries discuss post-investment risk transfer through warranties and indemnities, but no published Indian scholarship constructs an integrated, pre-investment due diligence methodology calibrated to the DPDP. Second, comparative privacy due diligence work, including the California Lawyers Association's 2025 paper on K-pop startups in California under the CCPA/CPRA, develops methodologies that cannot be transplanted into India without significant adjustment because of structural differences in lawful basis, cross-border transfer architecture, supervisory authority structure and the absence of statutory rights to portability and erasure in the DPDP.

Adjacent literatures inform the framework. Venture capital scholarship, particularly the work of Kaplan and Strömberg on staged financing and contractual mechanisms, and Gompers and Lerner on the venture capital cycle, provides the analytical scaffolding into which a privacy diligence framework must integrate. Risk management literature anchored in COSO ERM, ISO 31000:2018, ISO/IEC 27001:2022 and the NIST Privacy Framework supplies the maturity-modelling and control-mapping tools that the SDRS and PCMM developed in this paper adapt and apply. Empirical literature on breach economics, particularly the IBM Cost of a Data Breach Report and the Cisco Privacy Benchmark Study, supplies the financial calibration that converts diligence findings into term sheet adjustments.

The contribution of this paper sits at the intersection of these literatures. It treats the investor as the protagonist, the DPDP as the binding constraint, and the diligence framework as the operational instrument. It draws methodologically from the comparative work without transplanting its assumptions. It uses the empirical work to calibrate rather than to describe. The result is the first integrated, journal-grade investor due diligence framework calibrated to the Indian regime.

1. Why this paper exists

Indian venture diligence still asks the questions it learned to ask fifteen years ago. Are the revenue numbers real? Is the intellectual property solid? Does the cap table check out? Are the founders honest? These questions matter. They do not catch the risk that defines digital businesses in 2026.

That risk is data. The personal data of Indian residents, processed by Indian startups, governed by a statute that was incomplete on paper for two years and is now operational. The Digital Personal Data Protection Act passed Parliament on 11 August 2023 and then waited. The waiting ended on 13 November 2025 when MeitY notified the Rules. An eighteen-month transition runs to 13 May 2027. From that date the substantive obligations apply in full.

What changes on that date matters for any investor about to write a cheque to an Indian startup. The Data Protection Board of India can impose penalties of up to INR 250 crore per breach for failure to maintain reasonable security safeguards under Section 8(5). Up to INR 200 crore for failure to notify a breach under Section 8(6). Up to INR 200 crore for mishandling children's data under Section 9. These are administrative penalties. They sit alongside civil claims by Data Principals under tort, contract and consumer protection law, and alongside parallel obligations under the Information Technology Act 2000 and the CERT-In Directions of 28 April 2022.

The Indian startup ecosystem is now large enough to make the gap consequential. DPIIT counted 159,157 recognised startups as of January 2025. India is the third-largest startup ecosystem in the world, behind only the United States and China. There are 118 unicorns. Cumulative valuation is north of USD 354 billion. The first nine months of 2025 alone saw about USD 12 billion of capital deployed. Each rupee travels into a portfolio company that is, by default, a Data Fiduciary under the DPDP.

Here is the problem. Conventional diligence does not ask DPDP questions, or it asks them perfunctorily. The investor signs a term sheet that contains a privacy representation, takes comfort from a founder warranty, and discovers eighteen months later that consent flows were bundled, a cloud bucket was misconfigured, an AI model was trained on data without lawful basis, or children's data sat in plain text with no parental verification. The 23andMe case shows what comes next. A USD 6 billion company at SPAC listing in 2021 became a USD 50 million bankruptcy candidate in early 2025 and sold for USD 305 million in July of that year. Multi-factor authentication on customer accounts, which any competent technical diligence would have flagged before the SPAC, would have prevented the breach.

Indian startups are not different. Byju's suffered two breaches in 2020 and 2021, the second exposing children's data of WhiteHat Jr users. Hathway lost 41.5 million records in March 2024 through a CMS vulnerability. boAt lost 7.5 million customer records in February 2024. BSNL was breached in July 2024. Each of these incidents involved a company that had received institutional capital. None of those investments was diligenced for privacy risk in the way this paper proposes.

The chapters that follow build the tool the market lacks. Section 2 explains what the DPDP actually demands, in plain language and with the section numbers an investor counsel will need. Section 3 presents the empirical record of what breaches cost. Section 4 sets out the twelve-stage framework. Section 5 converts diligence findings into a score and the score into term sheet adjustments. Section 6 compares India with the leading foreign regimes. Section 7 catalogues the patterns diligence reliably uncovers in Indian startups but that founders rarely disclose. Section 8 sets out the policy gaps that remain. Section 9 owns the limitations. Section 10 closes.

The paper is written for use. The framework is laid out so a deal lead, a general counsel or external counsel can lift the stages, the scorecard and the indicative term sheet adjustments and apply them to the next investment. The citations are accurate, the numbers are sourced, and the language is plain because the audience is mixed. A scholar reading this paper for the first time and a partner reading it for the third should both find it useful.

2. What the DPDP actually demands

Diligencing a startup against the DPDP requires knowing what the Act asks of the Data Fiduciary. The Data Fiduciary is the entity that decides why and how personal data is processed. Most Indian startups that handle user data are Data Fiduciaries. Many are also Data Processors when they handle data on behalf of business customers. A few will eventually be designated Significant Data Fiduciaries when the Central Government issues a notification under Section 10.

Lawful processing

Section 4 of the Act allows personal data to be processed only on the basis of consent under Section 6, or for a legitimate use listed in Section 7. Section 7 is narrow. It covers employment-related processing, statutory functions, medical emergencies, disaster response and a handful of similar situations. For most consumer startups, consent is the operative ground.

Consent is not whatever the founder chooses to call it. Section 6 requires it to be free, specific, informed, unconditional and unambiguous, given through a clear affirmative action. Section 5 requires that a notice precede or accompany every request for consent. Rule 3 of the 2025 Rules makes the notice concrete. It must be standalone. It must use plain language. It must itemise the categories of personal data being collected and the specific purpose for each category. It must tell the

Data Principal how to withdraw consent, how to exercise rights and how to complain to the Board. A privacy policy buried in a footer with a tick-box reading I agree to the terms will not satisfy this standard.

Security safeguards

Section 8(5) requires the Data Fiduciary to take reasonable security safeguards. Rule 6 makes the standard concrete: encryption and masking, access control, access logging and monitoring, data back-ups, detection of unauthorised access, methods for investigation and remediation, and equivalent flow-through obligations on Data Processors. The maximum penalty for failure is INR 250 crore per breach. That is the largest single number in the Act.

Breach notification

Section 8(6) requires the Data Fiduciary to notify the Board and the affected Data Principal on a personal data breach. Rule 12 specifies the mechanics. The Board must receive an initial notification without delay and a detailed report within seventy-two hours. The Data Principal must be told without delay. This sits on top of an older and tighter obligation under the CERT-In Directions of 28 April 2022, which require any cybersecurity incident to be reported to CERT-In within six hours under Section 70B(6) of the IT Act. The Indian startup that suffers a breach therefore runs three reporting clocks in parallel: six hours to CERT-In, seventy-two hours detailed report to the Board, and immediate notification to affected individuals.

Children's data

Section 9 sets a strict regime for the personal data of children, who under the Act are persons under the age of eighteen. Verifiable parental consent is required for processing. Tracking, behavioural monitoring and targeted advertising directed at children are prohibited unless the Central Government has separately notified the Data Fiduciary as compliant. The Fourth Schedule of the Rules provides limited carve-outs for clinical establishments, healthcare professionals and educational institutions, but commercial edtech operating outside the formal school system gets no shelter. The maximum penalty for failure is INR 200 crore.

Cross-border transfer

Section 16 takes a permissive line. Personal data may be transferred outside India unless the Central Government notifies a specific country as restricted. This is the opposite of the GDPR adequacy model, which begins with restriction and works towards permission. The investor still has three things to check. No negative list notification at the relevant time. No sector-specific restriction. The Reserve

Bank of India's 2018 Storage of Payment System Data Directions require all payment data to stay in India. SEBI imposes its own controls on capital market intermediaries. The Telecom Cyber Security Rules of 2024 add another layer. And no FEMA implication where the cross-border transfer counts as movement of an intangible asset to an overseas affiliate.

Significant Data Fiduciaries

Section 10 lets the Central Government notify any Data Fiduciary or class as Significant. The factors include volume and sensitivity of data, risk to Data Principals, impact on the sovereignty of India, risk to electoral democracy, security of the State and public order. Industry practice expects startups processing more than five million Indian Data Principals, with annual turnover above INR 250 crore, or using AI for decisions affecting individuals, to be in the candidate pool. SDF status adds three obligations: appointment of a Data Protection Officer based in India and reporting to the board, appointment of an independent data auditor, and periodic Data Protection Impact Assessments. The Rules add an algorithmic fairness obligation. The maximum SDF-specific penalty is INR 150 crore.

Phased implementation

The Act and Rules do not all switch on together. Phase I started on 13 November 2025 and made the Data Protection Board operational. Phase II runs from 13 November 2026 and brings the Consent Manager framework into force. Phase III, from 13 May 2027, is the big one. Notice and consent requirements, security safeguards, Data Principal rights, breach notification and SDF obligations all apply from that date. Until then, the IT Act and the SPDI Rules of 2011 continue to govern. For an investor closing a deal in 2026, the diligence question is not only whether the target is compliant today. It is whether the target can be compliant by 13 May 2027.

The penalty arithmetic

The Schedule to the Act is short and concrete.

Table 1. *Maximum administrative penalties under the DPDP Act, 2023*

Provision	Failure	Maximum penalty
Section 8(5)	Failure to take reasonable security safeguards	INR 250 crore
Section 8(6)	Failure to notify a personal data breach	INR 200 crore

Provision	Failure	Maximum penalty
Section 9	Failure to comply with children's data obligations	INR 200 crore
Section 10	Failure on additional SDF obligations	INR 150 crore
Residual	Any other breach of the Act or Rules	INR 50 crore
Section 15	Breach of duty by a Data Principal	INR 10,000

These are administrative penalties only. They sit alongside, not in place of, civil claims by Data Principals under tort, contract and the Consumer Protection Act 2019, criminal exposure under the Bharatiya Nyaya Sanhita 2023 for identity theft and cheating by personation, sectoral regulatory action by RBI, SEBI, IRDAI or TRAI, and reputational loss. An investor inherits all four channels of risk on closing.

3. What real breaches have cost

The legal architecture is one half of the picture. The empirical record is the other. Two questions matter for the investor. How likely is a breach? When one happens, what does it cost?

The IBM Cost of a Data Breach Report 2025 is the most reliable source on the second question. It records that the average data breach in India cost INR 220 million in 2025, a thirteen per cent increase over 2024 and the highest figure since IBM began measuring India. Globally the average dropped to USD 4.44 million, a nine per cent decline, but India bucked the trend. The Indian breach lifecycle is now 263 days, an improvement of fifteen days year on year but still longer than the global average.

The top three initial attack vectors in India are phishing at eighteen per cent, third-party and supply chain compromise at seventeen per cent, and vulnerability exploitation at thirteen per cent. The implication for diligence is that two of the top three vectors run through people and vendors, not through the target's own code. A privacy diligence that only inspects the target's technical stack will miss most of the actual risk.

One specific number is worth remembering. Shadow AI, defined as employee use of AI tools without IT oversight, added INR 17.9 million to the average Indian breach in 2025. Only thirty-seven per cent of Indian organisations had access controls for AI systems. Nearly sixty per cent had no AI governance policy or were still drafting one. AI governance is now the most underdeveloped element

of Indian enterprise security. An investor backing a startup that uses generative AI without governance has a quantifiable problem.

The 23andMe case has become the textbook example of breach economics at the firm level. The company listed via a SPAC merger in June 2021 at an implied valuation of USD 6 billion. Between April and October 2023 a credential stuffing attack compromised the genetic and ancestry data of about 6.9 million customers. Two facts about the attack matter. First, the entry vector was customers reusing passwords from earlier breached sites. The company did not require multi-factor authentication on customer accounts. Second, the attackers curated subsets of customers of Chinese and Ashkenazi Jewish descent, suggesting targeted secondary use.

What happened next is the part that should keep investors awake. Class actions consolidated before Judge Edward Chen in the Northern District of California. A settlement of USD 30 million in September 2024, later increased to USD 50 million. Bankruptcy filing in March 2025 in the Eastern District of Missouri. Asset sale to TTAM Research Institute, an entity led by founder Anne Wojcicki, for USD 305 million in July 2025. Final settlement approval on 30 January 2026. Total customer settlement value of USD 46.75 million, paid in significant part by cyber insurance policies from Allied World, Houston Casualty, Landmark American and Lloyd's syndicates. From USD 6 billion to a USD 305 million bankruptcy sale in less than four years. The breach is not the only cause of the decline but it is the proximate event that converted a struggling business into a terminal one.

The Indian record is less dramatic but the pattern is consistent. Byju's, the edtech company valued at USD 22 billion in 2022, suffered breaches in 2020 and 2021. The second incident involved children's data at WhiteHat Jr. Byju's now sits in insolvency proceedings before the National Company Law Tribunal. The breach is not the sole reason for the collapse but it is part of a compounding pattern that eroded founder credibility and triggered the down-rounds. Mobikwik's 2021 incident allegedly exposed 99 million customer records. Its eventual public listing in 2024 carried a visible discount. Hathway lost 41.5 million customer records in March 2024. boAt lost 7.5 million in February 2024. BSNL lost millions of IMSI and SIM records in July 2024.

Table 2. *Selected breaches and the value they affected*

Entity / Year	Records affected	Cause	Direct cost	Indirect impact
23andMe (2023)	6.9 million	Credential stuffing; no MFA	USD 46.75M settlement	Bankruptcy; valuation from USD 6B to USD 305M

Entity / Year	Records affected	Cause	Direct cost	Indirect impact
Meta / Cambridge Analytica (2018)	87 million	Third-party data harvesting	USD 5B FTC; USD 725M class action; EUR 1.2B Irish DPC	Multi-year governance overhaul
LinkedIn (2024)	Not disclosed	Behavioural advertising basis failure	EUR 310M Irish DPC	Product redesign
Amazon (2021)	Not disclosed	Advertising consent	EUR 746M Luxembourg CNPD	Continued litigation
Byju's / WhiteHat Jr (2021)	Children's data	Misconfigured AWS storage	No formal fine (pre-DPDP)	Down-rounds; later insolvency
Mobikwik (2021)	99 million (alleged)	Vendor / server misconfiguration	No formal fine (pre-DPDP)	Listing discount
Hathway (2024)	41.5 million	CMS vulnerability exploit	Not disclosed	Trust and regulatory pressure
boAt (2024)	7.5 million	Database access; encryption gap	Not disclosed	IPO planning headwind
BSNL (2024)	Multi-million	Internal system compromise	Not disclosed	National security debate

Two observations close this section. First, breaches in India have moved from being rare events to being routine. CERT-In tracked 22.68 lakh cybersecurity incidents in 2024, against 10.29 lakh in 2022. That is a 120 per cent increase in two years. Second, the breaches reach every layer of the digital economy. State-owned telecoms, listed cable operators, direct-to-consumer brands, edtech unicorns and payment companies. Nothing about the Indian environment makes any sub-sector safe. The investor's working assumption must be that any portfolio company will at some point suffer an incident. The question diligence answers is how prepared the target is for the inevitable.

4. The twelve-stage framework

Here is the framework. A sequenced workflow of twelve stages. Each stage covers a specific cluster of DPDP obligations. Each stage feeds a sub-score into the overall risk score. The stages run from

broad to specific. Stages 1 and 2 set the perimeter and the governance baseline. Stages 3 to 6 cover the technical and contractual stack. Stages 7 to 10 examine specific risk concentrations. Stages 11 and 12 cover incident readiness and ongoing oversight.

Stage 1: Preliminary screening

Does the DPDP apply, and how? The investor confirms that the target processes digital personal data of Indian Data Principals. If yes, the target is a Data Fiduciary under Section 3 of the Act. If the target only processes on behalf of customers, it is a Data Processor and the analysis runs through Section 8(2). The investor then confirms the sectoral overlay. Fintech brings RBI. Edtech brings the schools regulator and the POCSO Act. Healthtech brings the National Digital Health Mission and the Clinical Establishments Act. SaaS inherits the sector of its customers. The output is a one-page applicability memorandum. Weight in the score: 5 of 100.

Stage 2: Privacy governance

Who in the target is responsible for privacy, where do they sit, and to whom do they report? Is there a board committee that reviews privacy? Are there written policies? Is there an internal grievance mechanism? The benchmarks are Section 8 of the DPDP, Section 10 for SDF candidates, Section 166 of the Companies Act 2013 on director fiduciary duties, and Regulation 18 of the SEBI LODR. The strongest signal is a Data Protection Officer with a direct line to the board. The weakest is a founder who personally controls every data decision with no internal challenge. Weight: 10.

Stage 3: Technical due diligence

This is the longest stage. Encryption at rest and in transit. Access controls including multi-factor authentication on every system that touches personal data. Network segmentation. Logging and monitoring. Vulnerability management. Penetration testing in the previous twelve months. Secure software development. Cloud security posture. The benchmarks are Section 8(5), Rule 6, ISO/IEC 27001:2022 Annex A and the NIST Privacy Framework. The 23andMe lesson is that a single missing control, in their case mandatory MFA on customer accounts, can collapse a USD 6 billion business. Weight: 15.

Stage 4: Legal documentation

Is the privacy notice itemised and standalone, as Rule 3 demands? Is consent free, specific and unbundled, as Section 6 demands? Is there a working withdrawal mechanism? Is there a grievance officer with published contact details? Are processor contracts in place with proper flow-through clauses under Section 8(2)? The investor asks for current and historical versions of the privacy notice. Historical versions matter because the basis of consent for legacy data must be verified. Weight: 15.

Stage 5: Third-party and processor risk

A modern startup runs on third parties. Payment gateways, marketing automation, email services, customer support tools. Each is a processor under Section 8(2) and each requires a written contract with appropriate flow-through obligations. The diligence asks for the vendor register and checks whether critical vendors have signed processor contracts. It examines breach notification clauses, audit rights, technical and organisational measures and indemnity provisions. It identifies vendors in jurisdictions that might be restricted under Section 16. Weight: 8.

Stage 6: Vendor and data flow mapping

The IBM report tells us that third-party and supply chain compromise is the second most common breach vector in India at seventeen per cent. The diligence asks for an end-to-end data flow map: every system that holds personal data, every transfer between systems, every external destination. Most startups have only a partial map. The diligence identifies the gaps and treats unmapped flows as red flags. Weight: 5.

Stage 7: AI and automated decision-making

Section 8(3) requires that personal data used to make decisions affecting the Data Principal be complete, accurate and consistent. Rule 11 requires SDFs to conduct algorithmic fairness assessments. The diligence asks for an AI inventory. Which models are deployed? What do they decide? What data were they trained on? Was that data lawfully collected? Is there a DPIA for each high-risk model? Is there a bias audit? Given that sixty per cent of Indian organisations have no AI governance, the default expectation is gaps. Weight: 10.

Stage 8: Cross-border transfers

Section 16 is permissive. The negative list is short. But sectoral overlays restrict movement. RBI requires payment data to stay in India. SEBI requires capital market intermediaries to comply with its Cybersecurity Framework. The Telecom Cyber Security Rules of 2024 add telecom-specific storage. The diligence verifies the data residency map and confirms that cloud contracts support the residency requirements. Particular attention to US-based cloud providers because of Schrems II style risk where the target also has GDPR-applicable data. Weight: 8.

Stage 9: Children’s data

Section 9 carries a INR 200 crore penalty and the regulator will enforce children's protections aggressively. The diligence first establishes whether the target processes children's data. For edtech, social platforms and gaming the answer is usually yes. The investor then verifies the age verification mechanism, the verifiable parental consent process, and the tracking and advertising stack. Behavioural advertising to children is effectively prohibited. Weight: 8.

Stage 10: Data monetisation

Does the target make money from data beyond core service revenue? Advertising, data sales, insights products, lead generation. The diligence identifies the revenue streams that depend on personal data, checks that those streams have lawful basis, and looks at Competition Commission of India risk because the CCI now treats data as a competition variable. The 2023 CCI market study on Big Tech is a useful signal. Weight: 6.

Stage 11: Incident response and insurance

What does the target do when a breach happens? Is there a written incident response plan? Has it been exercised in the last twelve months through a tabletop simulation? Is there a CERT-In point of contact registered with the agency? Is there cyber insurance? Does it cover regulatory fines? Does the D&O cover respond to privacy-driven claims against directors? Almost every startup has gaps here. The diligence quantifies them and prices the fix. Weight: 6.

Stage 12: Board governance and ESG

The final stage looks up rather than down. Is privacy a board topic? Is there at least one director with privacy or technology expertise? Does the audit or risk committee have a written charter covering data risk? For pre-IPO companies, does the BRSR disclosure mention privacy? Investor information rights should include a quarterly privacy programme report and an annual right to require an independent audit. Weight: 4.

Table 3. *The twelve diligence stages summarised*

Stage	Focus	Anchor provision	Weight
1	Applicability and perimeter	DPDP s. 3; sectoral rules	5
2	Governance and accountability	DPDP s. 8, 10; Companies Act s. 166	10
3	Technical security	DPDP s. 8(5); Rule 6; ISO 27001	15
4	Legal documentation	DPDP s. 5, 6, 8; Rule 3	15

Stage	Focus	Anchor provision	Weight
5	Processor contracts	DPDP s. 8(2); Rule 7	8
6	Data flow mapping	DPDP s. 8; CERT-In	5
7	AI and profiling	DPDP s. 8(3); Rule 11	10
8	Cross-border transfer	DPDP s. 16; RBI SPSD 2018	8
9	Children's data	DPDP s. 9; Fourth Schedule	8
10	Data monetisation	DPDP s. 6, 7; Competition Act	6
11	Incident response	DPDP s. 8(6); CERT-In	6
12	Board and ESG	DPDP s. 10; SEBI LODR	4
Total			100

Every stage has two flag levels. A yellow flag is something to fix through conditions precedent or post-closing covenants. A red flag is serious enough to renegotiate the term sheet or, in extreme cases, walk away. Table 4 lists the canonical ones.

Table 4. *Yellow and red flags by stage*

Stage	Yellow flag	Red flag
1	Sectoral overlay not mapped	DPDP applicability disputed by founder
2	No DPO; no board oversight	Founder controls every data decision
3	Patchy encryption	No MFA; public cloud buckets
4	Generic privacy notice	Bundled consent; no withdrawal mechanism
5	Some contracts missing	Critical processor in restricted jurisdiction
6	Partial data flow map	Shadow IT or unrecorded data exfiltration
7	GenAI without policy	Automated decisions without DPIA
8	Cloud in single foreign region	Payment data outside India contrary to RBI
9	Age gating without verification	Behavioural tracking of under-18 users
10	Inference-based pricing	Sale of identified personal data
11	Plan exists but never tested	Past breach not notified to CERT-In or Board
12	No privacy expertise on board	ESG disclosure misrepresents data risk

5. The risk score and what it means for the term sheet

Twelve stages produce twelve sub-scores. The sum is the Startup Data Risk Score, on one hundred points. A higher score is better. The bands are simple and the cuts deliberate.

Table 5. *SDRS bands and the recommended investor action*

Score	Band	PCMM tier	Recommended action
85 to 100	Excellent	Tier 5 Optimised	Standard reps and warranties; no privacy escrow
70 to 84	Strong	Tier 4 Managed	Standard indemnity; quarterly monitoring
55 to 69	Adequate	Tier 3 Defined	Conditions precedent; sized indemnity; audit right
40 to 54	Weak	Tier 2 Repeatable	Material conditions; specific indemnity; valuation haircut
Below 40	Critical	Tier 1 Initial	Defer; require remediation; consider declining

The Privacy Compliance Maturity Model adapts the classic capability maturity ladder to the DPDP. Five tiers. Each tier specifies conditions across five domains: governance, processes, technology, people and measurement. The target is placed at the lowest tier in which it meets all five domain conditions. This is deliberately strict. A company with strong technology but no governance is a Tier 1, not a Tier 3. The reason is operational. Privacy weakness in any one domain is enough to enable a serious incident.

Table 6. *Privacy Compliance Maturity Model*

Tier	Designation	Defining marks
1	Initial	No DPO; no written policies; no risk register; reactive only
2	Repeatable	Basic privacy notice; some technical controls; informal incident handling
3	Defined	Documented policies; DPO appointed; DPIAs for high-risk processing; vendor framework
4	Managed	Metrics-driven programme; regular audits; board reporting; insurance aligned
5	Optimised	Privacy by design embedded; AI governance integrated; continuous improvement

Translating the score into the term sheet

This is where the framework earns its keep. The Investor Risk Rating Framework converts the SDRS band into the line items the deal team negotiates.

Table 7. *Investor Risk Rating Framework: SDRS to term sheet*

SDRS band	Valuation discount	Indemnity cap	Privacy escrow	Survival period
85 to 100	0%	10 to 15% of cheque	None	18 months general; 36 months tax / IP
70 to 84	0 to 3%	20%	3 to 5%	24 months general; 48 months privacy
55 to 69	3 to 7%	25 to 30%	7 to 10%	36 months general; 60 months privacy
40 to 54	7 to 15%	40 to 50%	10 to 15%	48 months general; 72 months privacy
Below 40	Re-price required	Up to 100%	20% plus	Indefinite for known issues

A worked example

Take a hypothetical Series B fintech raising INR 200 crore at INR 1,200 crore pre-money. The target serves 2.4 million Indian users with KYC, transaction and credit data. Diligence yields the following sub-scores. Stage 1: five of five. Stage 2: six of ten, because a DPO is in place but with no board oversight. Stage 3: eight of fifteen, because multi-factor authentication is missing on internal admin tools. Stage 4: ten of fifteen, because consent is bundled for marketing. Stage 5: four of eight, because a critical processor sits in a foreign jurisdiction without standard contractual clauses. Stage 6: four of five. Stage 7: five of ten, because the fraud detection model has no DPIA. Stage 8: six of eight. Stage 9: eight of eight, no exposure. Stage 10: five of six. Stage 11: four of six. Stage 12: two of four. The aggregate score is sixty-seven.

That places the target at the upper end of the Adequate band, Tier 3 Defined. The investor action is concrete. Closing is conditional on five items. MFA rolled out within sixty days. Consent unbundled and re-collected within ninety days. Standard contractual clauses signed with the foreign processor before closing. DPIA conducted on the AI model. A tabletop incident exercise scheduled within

ninety days. The valuation discount is five per cent, taking the pre-money to INR 1,140 crore. The indemnity cap is twenty-five per cent of the cheque, or INR 50 crore. The privacy escrow is seven per cent of the consideration, or INR 14 crore, held for sixty months. Semi-annual audit rights are reserved for the investor. The board gets quarterly privacy programme reports.

This is the output the framework is designed to produce. Not a paragraph of comfort in the closing memo. A specific set of contractual terms backed by a defensible score and an audit trail.

Sector-specific weight adjustments

The default weights work for a generic consumer startup. Different sectors require different weights. Fintech tilts towards technical, processor and cross-border risk. Edtech tilts towards children's data. Healthtech tilts towards AI and technical security. D2C tilts towards monetisation. SaaS tilts towards processor and AI risk.

Table 8. *Sector-specific weight adjustments*

Stage	Default	Fintech	Edtech	Healthtech	D2C	SaaS
3 Technical	15	20	12	18	12	15
5 Processor	8	12	6	8	8	14
7 AI	10	8	8	15	8	15
8 Cross-border	8	12	6	6	6	12
9 Children	8	4	18	8	4	2
10 Monetisation	6	4	8	4	12	4

6. How India compares with the rest of the world

The DPDP is not the GDPR. It is closer to Singapore's PDPA in structure and ambition, with elements borrowed from Brazil's LGPD and from the 2018 European debate that produced the GDPR. The differences matter for an investor with multi-jurisdictional exposure.

On lawful basis, the GDPR provides six grounds under Article 6: consent, legitimate interest, contract necessity, legal obligation, vital interests and public task. The DPDP provides two: consent under Section 6 and a closed list of legitimate uses under Section 7. The Indian regime is stricter in its baseline. An investor in a target that relies on the GDPR's legitimate interest basis for marketing or analytics needs to know that the basis does not exist in India. The processing has to be re-engineered to run on consent.

On sensitive data, the GDPR Article 9 has a separate category covering health, biometric, genetic, political opinions and sexual orientation. The DPDP has no separate category. This sounds simpler but creates an interpretive risk. Indian regulators and courts will likely read sensitivity into Section 8(5) when deciding what counts as reasonable security safeguards. Health and financial data will carry an implied higher standard than ordinary personal data.

On rights, both data portability and the right to be forgotten are absent from the DPDP. Both exist in the GDPR (Articles 17 and 20), the LGPD and the CPRA. The Indian Act provides erasure on withdrawal of consent or completion of purpose but not as a standalone right of the Data Principal. This is a legislative choice. Critics argue it weakens the regime. Supporters argue it simplifies compliance. The practical consequence for the investor is that consumer protection litigation under the Consumer Protection Act 2019 may read quasi-portability rights into the duty of service, and tort law may eventually accommodate something resembling the right to be forgotten. The risk channel that does not exist in statute may emerge in case law.

On penalty model, the GDPR caps fines at the higher of EUR 20 million or four per cent of global turnover. The DPDP caps the maximum at INR 250 crore per breach with no turnover linkage. For a small startup the Indian cap is severe in absolute terms. For a unicorn it is more manageable than the GDPR equivalent. Meta has paid EUR 1.2 billion in a single GDPR action. The DPDP equivalent would require multiple breaches to reach that figure. CCPA penalties run at USD 7,500 per intentional violation, which on a multi-million record breach multiplies quickly. Singapore's PDPA caps at SGD 1 million or ten per cent of turnover. LGPD caps at BRL 50 million or two per cent of turnover.

On cross-border transfer, the GDPR uses adequacy decisions, standard contractual clauses and binding corporate rules. The CPRA imposes no specific restriction. The DPDP uses the negative list. The Indian approach is the most permissive in baseline but the least predictable in operation, because the Central Government can change the list by notification. Investors should treat the negative list as a soft constraint that could harden quickly.

On statutory compensation, GDPR Article 82 entitles the Data Subject to compensation for material and non-material damage. The DPDP has no equivalent. Indian Data Principals will seek compensation through tort, contract and consumer protection routes. The implication is that the deepest pockets in the capital structure, namely the institutional investor rather than the founder, will be a litigation target.

What this means for the investor. The DPDP is easier to plan around in some respects and harder in others. The penalty arithmetic is more predictable. The cross-border regime is more permissive. The lack of statutory compensation pushes risk into channels the investor cannot easily price. The

framework in Section 4 is built to handle the Indian-specific permutations, and a global investor should expect to extend it with overlays for the foreign regimes its portfolio touches.

7. What founders rarely tell investors

This section steps away from the formal architecture and reports what diligence consistently uncovers in Indian startups. The pattern is not a list of individual failings. It is a set of structural blind spots that exist almost everywhere.

The first blind spot is consent at scale. A founder who launched in 2018 with a checkbox in the sign-up flow has been collecting personal data for years on the basis of consent that does not meet the Section 6 standard. The fix is consent re-collection at scale. This is operationally painful. It produces measurable user drop-off. It is rarely budgeted for in the operating plan.

The second blind spot is the gap between the privacy notice and the actual data flow. The notice was drafted by external counsel and updated occasionally. The data flow grew through product iteration. New features added new collection. New integrations added new third-party processors. New analytics tools introduced new transfers. The notice never caught up. The diligence finds that the notice describes a company that no longer exists.

The third blind spot is the vendor stack. The target signed up for a CRM, an email service, an analytics tool, a customer support platform, a marketing automation suite and a half-dozen other SaaS subscriptions. Each is a processor under Section 8(2). Each requires a written contract with specific flow-through obligations. In most Indian startups, the contracts in place are the standard click-through terms of the SaaS provider, which were never read and which do not comply with Indian processor requirements.

The fourth blind spot is the AI layer. Generative AI use is now ubiquitous. Diligence will find ChatGPT in the customer support team, Claude in the legal team, Copilot in engineering and a handful of bespoke models in the product. Almost none of this is governed. Shadow AI alone adds INR 17.9 million to the average Indian breach. The fix requires policy, training and technical controls. It cannot be done in a week.

The fifth blind spot is the children's data question. Founders consistently underestimate exposure. Edtech founders forget that their student users are under eighteen. Gaming founders forget that their target demographic skews young. Social platform founders rely on a self-certification age gate that does not meet the verifiable parental consent standard in Section 9.

The sixth blind spot is incident response. The plan exists. It was drafted at some point. Nobody can find the document. Nobody knows who the CERT-In point of contact is. Nobody has run a tabletop

in the last twelve months. When the incident happens the team improvises. The six-hour CERT-In window passes. The seventy-two-hour DPB report is filed late. The penalty arithmetic begins to move.

These six patterns are the high-frequency findings. The framework in Section 4 is calibrated to surface them. The investor who takes the framework seriously and presses the founder for evidence rather than assertion will catch most of the issues that destroy value after closing.

8. Policy gaps and what should change

The DPDP is a workable statute. It is not a perfect one. Six gaps deserve attention from policymakers and, until they are addressed, attention from investor counsel through private ordering.

Statutory compensation. The Act has no equivalent of GDPR Article 82. Data Principals must seek redress through tort, contract and consumer protection. The result is uncertainty and forum-shopping. A statutory compensation provision, even one capped at modest levels, would internalise the cost of breaches within the DPDP rather than displacing it to other statutes.

Data portability. The Act omits the right to data portability that exists in the GDPR, the LGPD and the CPRA. This makes life easier for founders but creates a competition problem because portability is one of the few effective constraints on platform lock-in. A Section 11A introducing a portability right with sectoral phasing would close the gap.

Significant Data Fiduciary threshold clarity. Section 10 leaves SDF designation to executive discretion. A statutory threshold, expressed in volume of Data Principals and revenue, would provide planning certainty. The current uncertainty distorts founder behaviour. Some hide growth metrics to avoid notification. Others over-comply at cost to product investment.

Insolvency and data continuity. The DPDP says nothing about what happens to personal data when a Data Fiduciary enters insolvency or is sold as a going concern. The 23andMe bankruptcy raised this question internationally. The Indian Insolvency and Bankruptcy Code 2016 needs an amendment addressing the continuation of DPDP obligations through the corporate insolvency resolution process and the treatment of personal data as an asset that cannot be transferred without lawful basis.

Safe harbour for good-faith breach response. Section 32 allows the Board to accept voluntary undertakings. This is useful but not predictable. A safe harbour rule, similar to the cooperation discount in competition law, would reward the behaviour the regime exists to encourage. Targets that detect quickly, report promptly and remediate competently would receive a defined discount on the penalty range.

Sectoral integration. RBI, SEBI, IRDAI and TRAI have each issued their own data protection norms. The DPDP runs parallel to these. The result is duplication and occasional contradiction. A

coordinating instrument, perhaps a joint advisory under Section 28 of the DPDP, would reduce the compliance burden without weakening protections.

Independent of statutory change, the venture capital industry can standardise privacy diligence across the market. The leading Indian VC and PE associations should publish a standard diligence template, building on this framework or comparable work. Limited partners should ask their general partners to report portfolio-wide privacy posture as part of standard ESG reporting. Insurance carriers should develop a DPDP-specific cyber product with explicit cover for Section 8(5) and Section 8(6) penalties and a discriminating exclusions framework. None of these requires legislation. All of them require coordination.

9. Limitations and what remains to be done

This paper has limits and it would be dishonest to hide them. The substantive provisions of the DPDP come into effect only on 13 May 2027. The empirical record of enforcement does not yet exist. The framework relies on GDPR, PDPA and CCPA experience as the closest analogues. The Indian Board may turn out to be lighter or heavier than these comparators. The framework will need recalibration once two or three years of enforcement decisions become available.

The Startup Data Risk Score has not been validated against a panel of real diligence files. The weights and band thresholds are derived deductively from the structure of the DPDP and from analogous risk frameworks. They are testable and they are tunable. Validation is the natural next stage of the work and the obvious shape of follow-on research.

The framework is general. Sector tilts are provided in Table 8 but each sector deserves deeper specification. A fintech application would benefit from much more detail on RBI overlays than this paper provides. An edtech application would benefit from school-system integration mapping. A healthtech application needs full integration with the National Digital Health Mission.

The paper is investor-focused by design. The Data Principal, whose rights the regime exists to protect, is implicit rather than explicit. A follow-on paper modelling the Data Principal's perspective, particularly across vulnerable populations and rural users, is the natural complement to this one.

Future research should run in four lines. Empirical validation of the SDRS through application to a panel of recent Indian venture deals. Quantitative measurement of the privacy compliance premium in Indian Series B and later rounds, comparing valuation multiples of high-PCMM and low-PCMM targets across a matched sample. Longitudinal study of the Data Protection Board's enforcement decisions once they begin. Comparative analysis as the DPDP, GDPR, CPRA and LGPD evolve in parallel.

10. Conclusion

Indian venture capital deployed roughly USD 12 billion in 2025 without a standard way to ask whether the recipient startups could survive the law that became operational on 13 November that year. This paper has built the missing tool: a twelve-stage diligence framework, a one-hundred-point risk score, a five-tier maturity model, and an Investor Risk Rating that translates the score into the term sheet.

Three claims hold the analysis together. First, the DPDP makes privacy a balance-sheet issue, not a compliance afterthought. The INR 220 million average breach cost in India in 2025, the INR 250 crore maximum penalty per breach, and the 23andMe sequence from USD 6 billion to bankruptcy in three years make the case beyond argument. Second, privacy compliance maturity can be measured. The framework presented here is one way to measure it. There will be others. Third, the score must translate into the deal. A diligence finding that does not move the term sheet is a diligence finding that did not happen.

The framework is offered for adoption. It can be lifted into a transaction next week. It can be refined by industry associations, by limited partners, by insurance carriers and by sector regulators. The numbers in the score are calibrated but tunable. The stages are stable but extensible. What matters is that the conversation between investor and founder, between counsel and counterparty, between board and limited partner, now has a vocabulary that can hold the weight of the risk.

India is the third-largest startup ecosystem in the world. By May 2027 it will also have one of the most demanding personal data regimes anywhere. Those two facts cannot stay disconnected.

References

Statutes and Rules

1. Constitution of India, 1950, Articles 14, 19 and 21.
2. Digital Personal Data Protection Act, 2023 (Act 22 of 2023).
3. Digital Personal Data Protection Rules, 2025, notified by MeitY on 13 November 2025.
4. Information Technology Act, 2000 (Act 21 of 2000); Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules, 2011.
5. Companies Act, 2013 (Act 18 of 2013); Consumer Protection Act, 2019 (Act 35 of 2019); Competition Act, 2002.
6. CERT-In Directions No. 20(3)/2022-CERT-In dated 28 April 2022 under Section 70B(6) of the IT Act.

7. Reserve Bank of India, Storage of Payment System Data Directions (RBI/2017-18/153) dated 6 April 2018.
8. SEBI (Listing Obligations and Disclosure Requirements) Regulations, 2015, as amended.

Foreign Instruments

1. Regulation (EU) 2016/679 (General Data Protection Regulation); UK GDPR as retained and modified.
2. California Consumer Privacy Act, 2018, as amended by the California Privacy Rights Act, 2020.
3. Personal Data Protection Act, 2012 (Singapore); Lei Geral de Proteção de Dados, Lei No. 13.709/2018 (Brazil).
4. Act on the Protection of Personal Information, Act No. 57 of 2003 (Japan); Privacy Act, 1988 (Australia).

Cases

1. Justice K.S. Puttaswamy (Retd.) v. Union of India (2017) 10 SCC 1.
2. Justice K.S. Puttaswamy (Retd.) v. Union of India (2019) 1 SCC 1 (Aadhaar).
3. Anuradha Bhasin v. Union of India (2020) 3 SCC 637; Internet and Mobile Association of India v. Reserve Bank of India (2020) 10 SCC 274.
4. Schrems v. Data Protection Commissioner C-362/14 EU:C:2015:650; Data Protection Commissioner v. Facebook Ireland Ltd C-311/18 EU:C:2020:559.

Reports and Empirical Sources

1. Justice B.N. Srikrishna Committee, A Free and Fair Digital Economy: Protecting Privacy, Empowering Indians (Government of India, 2018).
2. IBM Security and Ponemon Institute, Cost of a Data Breach Report 2025.
3. Cisco Systems, 2025 Data Privacy Benchmark Study (April 2025).
4. CMS Law, GDPR Enforcement Tracker Report 2024/2025 (6th edition).
5. NASSCOM and Zinnov, India's Tech Start-Up Report 2025.
6. DPIIT, Nine Years of Startup India (Government of India, January 2025).
7. PitchBook, Venture Monitor 2024 and 2025.

Standards

1. ISO 31000:2018 Risk Management Guidelines; ISO/IEC 27001:2022 Information Security Management.
2. ISO/IEC 23894:2023 Guidance on AI Risk Management.
3. NIST Privacy Framework Version 1.0; COSO Enterprise Risk Management (2017).