

The Hidden Circuits of Data Governance

Voids in India's Personal Data Economy

Shishir K Jha & Anjana Nair

www.cps.iitb.ac.in/pbs-2026-008

PBS № 2026-008

This document is made possible by funding from TRUST LAB at IIT Bombay. The views expressed herein are those of the authors and do not reflect the official position of the TRUST LAB, IIT Bombay.

Executive Summary

This policy brief has been drawn from a more comprehensive report on the same theme. The internet has shifted from a democratic space to a core infrastructure for large-scale extraction and monetisation of personal data. In India's rapidly expanding digital economy, everyday mobile-based internet use generates continuous data flows that fuel platform services, alternative credit scoring, and targeted advertising. This policy-brief examines the obscure systems enabling real-time data extraction tracking tools, SDKs, and data brokers and highlights governance gaps that limit accountability. It argues that consent-based laws like the Digital Personal Data Protection (DPDP) Act are not entirely sufficient and calls for structural regulation centred on transparency, proportionality, and stronger regulatory capacity.

Introduction

Nearly 96% of India's 969 million internet users access the internet via mobile phones, turning smartphones and their apps into a pervasive infrastructure for continuous tracking and data collection (TRAI, 2025). Mobile phones, smart devices, IoT systems, digital platforms, and workplace monitoring tools from wearables and e-commerce apps to biometrics and GPS constantly generate data that feeds the digital value chain.

User generated data is collected via embedded tracking infrastructures and techniques like cookies, fingerprinting, pixel tracking, social plugins, SDKs (Software Development Kits), and device identifiers. Beyond direct consumer interactions, data flows through (Business to Business) B2B data supply chains via server-to-server and SDK-to-SDK transfers. Publicly available data from social media, court records, FIRs, university databases, and exam lists are also scraped and appended into dossiers of data brokers.

This cryptic and largely unidirectional data movement creates significant information asymmetry for users. Since the launch of Digital India in 2015, the state has promoted digitisa-

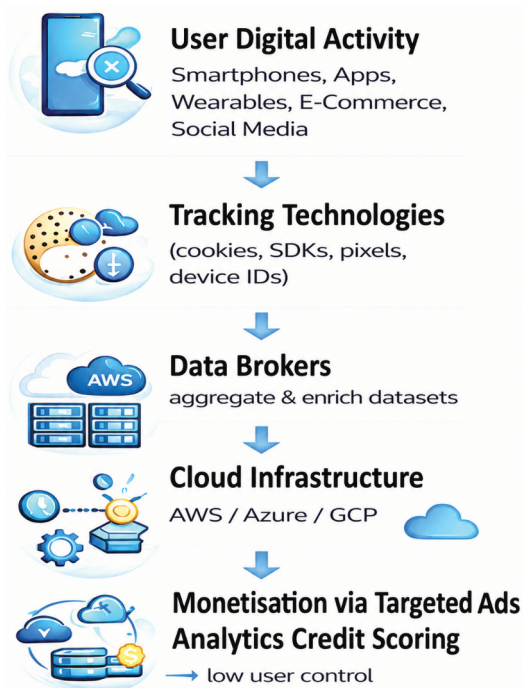


Figure 1: Data Extraction Ecosystem



This content is available under CC BY-NC-SA 4.0 (Attribution-NonCommercial-ShareAlike) License. For more information, see: <https://creativecommons.org/licenses/by-nc-sa/4.0>

tion across public services welfare, and financial systems, while building data centres and digital public goods like India Stack. Advances in hardware miniaturisation large scale data generation and monetisation across both public and private sectors. As shown in Figure 1, user digital activity is tracked and processed through cloud systems for data monetisation.

Data Supply Chains and Intermediary Ecosystems

Routine digital activities generate granular data via tracking tools such as cookies, SDKs, and device identifiers. This data is aggregated, enriched, and monetised by data brokers, cloud providers, and ad tech systems for targeted advertising, credit scoring, and behavioural profiling. Once data enters these systems, it circulates continuously across actors, sharply limiting user visibility and control. Figure 2 highlights how real-time infrastructures, particularly real-time bidding (RTB), institutionalise rapid, large scale data sharing as a core market function.

Regulatory and Governance Gaps

Figure 2 also highlights governance gaps in India's consent-centric data protection framework. Notice-and-consent mechanisms are weakened by information overload, cryptic intermediary exchanges, and users' inability to anticipate downstream data uses. Consent

withdrawal remains largely confined to first party collectors, enabling unchecked data circulation. This misalignment facilitates data maximisation without proportionality or accountability, underscoring the need for stronger intermediary oversight and use case-specific regulation prioritising transparency and public interest.

Targeted Advertisements

Figures 3 and 4 show how personalised advertising aims to deliver the right advertisement to the right user at the moment; it is most likely to generate engagement. RTB, a form of programmatic advertising, enables this process by auctioning ad space in milliseconds each time a user loads a webpage or application. A network of intermediaries coordinates these auctions at massive scale, making RTB a central infrastructure of digital advertising. Online advertising matches publishers offering advertising space with advertisers seeking user attention respectively. Supply-Side Platforms (SSP) manage publisher inventory and connect to Ad Exchanges, while Demand-Side Platforms (DSP) bid on impressions on behalf of advertisers.

Data Management Platforms (DMP) enhance user profiles by aggregating identifiers and linking data across sources through techniques such as cookie syncing. As a result, detailed digital dossiers circulate continuously among intermediaries. In the RTB pipeline, user data flows

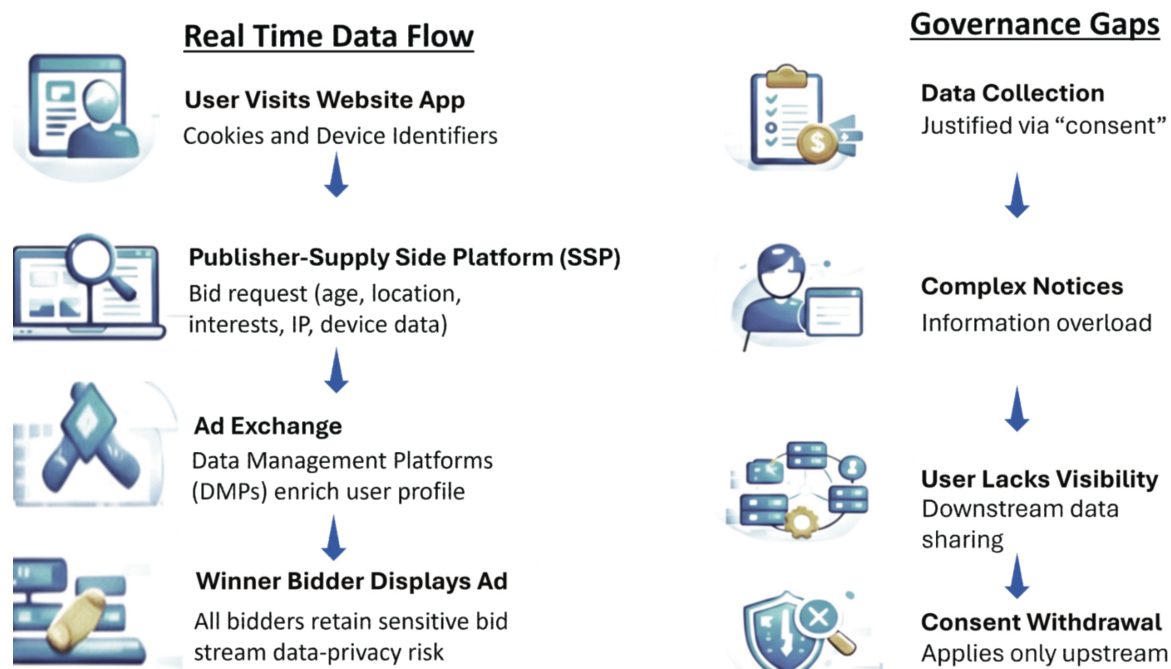
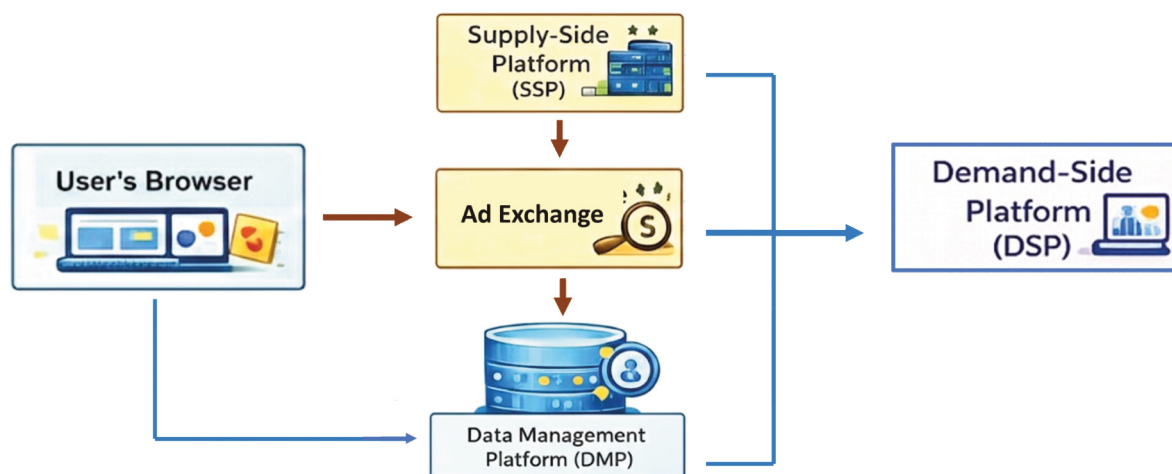
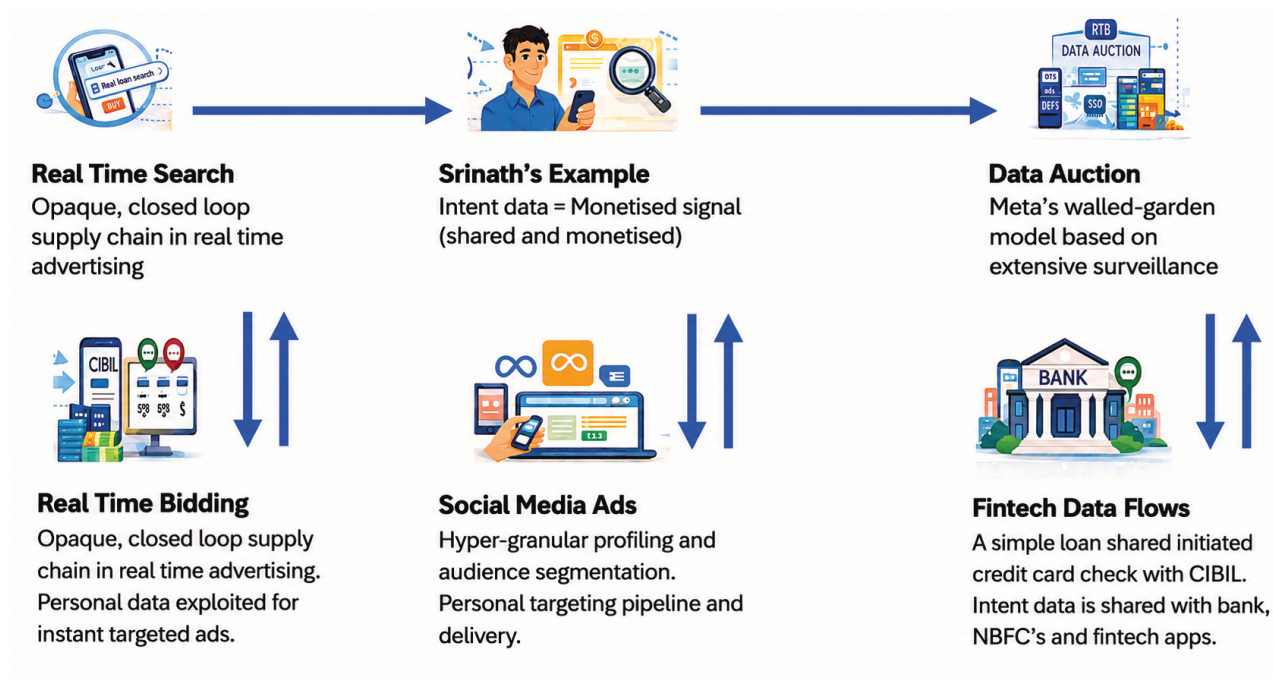


Figure 2: Real Time Data Flow and Governance Gaps Under Consent Based Regulation



Notes: Above figure illustrates the RTB advertising pipeline. A user's browser sends data to the Ad Exchange, where SSPs offer ad inventory and DSPs bid in real time. The DMP supports this process by providing audience data for more targeted ad delivery.

Figure 3: RTB (Real Time Bidding) Advertising Pipeline



Note: CIBIL is Credit Information Bureau (India) Limited. It is the most prominent of the four credit information companies licensed by the Reserve Bank of India (RBI).

Figure 4: Obscure Monetisation of Personal Intent Data in India's Digital Economy.

"India's rapid digitisation has outpaced its data governance, leaving personal data economies weakly regulated and largely invisible to users."

from browsers to platforms as bid-stream data containing sensitive identifiers such as location, device details, and browsing behaviour. This data is broadcast to multiple bidders, including those that do not win the auction, enabling widespread, consent-less data sharing resulting in expanding surveillance and increasing data brokerage across the ecosystem.

The data brokerage ecosystem operates within a fragmented and often weakly regulated landscape, where existing legal frameworks struggle to keep pace with its rapid technological evolution and cross-border nature. In such an environment, governance approach tend to rely on a combination of sector-specific rules, industry-led practices, and technical safeguards. More comprehensive regulatory models, however, emphasise privacy as a fundamental right, embedding stronger state-backed protections and clearer accountability mechanisms. Historical data-sharing arrangements illustrate the persistent tension between enabling seamless data flows and safeguarding data sovereignty. While such frameworks have facilitated uninterrupted data exchange among participating entities, they have also required adherence to conditions such as transparency, user choice, and the protection of data security and integrity.

Key Policy Insights

Data maximisation without accuracy safeguards creates systemic risk.

Large-scale data extraction prioritises volume over quality, amplifying errors that cascade across systems and produce biased, exclusionary outcomes.

Opacity in data ecosystems undermines accountability.

Persistent gaps in transparency around data collection, provenance, and algorithmic decision-making limit accountability. Stronger obligations are needed for traceability, disclosure of downstream data sharing, and algorithmic transparency.

Perpetual data circulation necessitates proportionality.

Once personal data enters intermediary-driven

supply chains, it circulates indefinitely. Consent withdrawal typically applies only to first-party collectors, leaving downstream uses intact, reinforcing the need for strict data minimisation at the point of collection.

Consent-based models are structurally insufficient.

Notice-and-consent frameworks fail in complex systems like RTB, where future data flows are unpredictable. While simplified or vernacular consent mechanisms may help marginally, they cannot replace structural regulation. Act such as DPDP (2023) enhances data protection in India through consent and accountability measures.

Digital inequality is intensified by a cybersecurity divide.

First-time as well as marginalised users face heightened exposure to data exploitation. Digital literacy and cybersecurity awareness must be integral to data protection policy.

Data brokerage remains a regulatory blindspot.

Fragmented and obscure data brokerage ecosystems enables unchecked aggregation and re-sale of personal data, demanding urgent regulatory attention.

Way Forward

Data and information flow is a critical component of internet infrastructure. But, in the era of surveillance capitalism, data as largely an extension of the self, is extracted and commodified at scale by a cohort of companies focused linearly on profits. A network of actors including big techs, intermediaries, data brokers of varying capacities are all part of this multibillion-dollar personal data supply chain. Such intrusive data practices are often obscured from public guise.

However, the social good possibilities offered by the large-scale digitalisation has been seldom explored. Economics and resource allocation determines how the technology is utilised. Allocation of resources for developing regulatory frameworks, implementing proportionality to foster ethical data collection practices and increasing digital literacy are critical policy priorities of the current scenario.

Conclusion

India's personal data economy is increasingly shaped by obscure, intermediary-driven private infrastructures that enable continuous extraction, circulation, and monetisation of personal data. As this brief articulates, consent-based data protection frameworks are structurally ill-equipped to govern ecosystems such as real-time bidding, SDK-driven tracking, and data brokerage, where data flows are rapid, fragmented, and largely invisible to users. Persistent governance gaps around transparency, proportionality, accuracy, and accountability allow data maximisation to outpace regulatory oversight, deepening digital inequality and systemic risk. Addressing these challenges requires a shift from individualised consent towards structural regulation that strengthens intermediary oversight, embeds data minimisation at the point of collection, and treats personal data as a form of critical digital infrastructure. Without such an approach, India's digital transformation risks entrenching surveillance driven growth at the cost of privacy, rights, and democratic trust.

References

- Achara, J. P., Acs, G., & Castelluccia, C. (2015). On the unicity of smartphone applications. DOI:10.48550/arXiv.1507.07851.
- Amnesty International. (2019). Surveillance giants: How the business model of Google and Facebook threatens human rights, <https://www.amnesty.org/en/documents/pol30/1404/2019/en>.
- Armitage, C., Botton, N., Dejeu-Catang, L., & Lemoine, L. (2023). Study on the impact of recent developments in digital advertising on privacy, publishers and advertisers. *Publicatons Office of the European Union* DOI:10.2759/294673.
- Bashir, M. A. (2019). On the privacy implications of real-time bidding (Doctoral dissertation, Northeastern University). <https://repository.library.northeastern.edu/files/neu:m044k3535/fulltext.pdf>.
- Federal Trade Commission. (2023). Lurking beneath the surface: Hidden impacts of pixel tracking. <https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/03/lurking-beneath-surface-hidden-impacts-pixel-tracking>.
- Lamdan, S. (2023). Data cartels: The companies that control and monopolize our information. Stanford University Press. DOI: 10.1080/08963568.2023.2184978.
- Privacy International. (2018). Examples of data points used in profiling. <https://privacyinternational.org/sites/default/files>.
- Ryan, J. (2022). The Biggest Data Breach: ICCL report on scale of Real-Time Bidding data broadcasts in the US and Europe. Dublin, Ireland: Irish Council for Civil Liberties, 10.
- Telecom Regulatory Authority of India (TRAI). (2025). The Indian telecom services yearly performance indicators 2024–2025. https://traigov.in/sites/default/files/2025-07/YIR_08072025_0.pdf.

Previously Published Policy Briefs

- PBS-2026-007** Data Brokering Practices: Conduct and Credibility.
- PBS-2025-006** Towards Sustainable Menstrual Waste Management in India.
- PBS-2025-005** Policy Lessons From India's Privatized Waste-to-Energy Models.
- PBS-2025-004** The Impact of Interface Design on Aadhaar Authentication for 'Non-Dominant' Name Patterns: Lessons from SARAL policy of Maharashtra.
- PBS-2025-003** Transitioning Towards Digital and Data-Driven Decision Support System The Case of Jal Jeevan Mission (JJM) in Maharashtra.
- PBS-2025-002** Jal Jeevan Mission (JJM) Database: From Monitoring to Decision Support Tool.
- PBS-2025-001** Digital Federalism and Service Delivery Governance Transformations in India's Rural Water Sector.



This content is available under CC BY-NC-SA 4.0 (Attribution-NonCommercial-ShareAlike) License.
For more information, see: <https://creativecommons.org/licenses/by-nc-sa/4.0>

Authors

Shishir K Jha , Professor, Ashank Desai Centre of Policy Studies, IIT Bombay.
Anjana Nair, Research Associate, TRUST LAB, IIT Bombay.

Acknowledgement

This document is made possible by funding from the TRUST LAB at IIT Bombay.
The views expressed herein are those of the authors and do not reflect the official position of the TRUST LAB, IIT Bombay.

Suggested Citation

Jha, S. K. & Nair, A. (2026, June).
The Hidden circuits of data governance. (Policy Brief Series No. 2026-008).
<https://www.cps.iitb.ac.in/pbs-2026-008>