TQH

THE DPI DECADE

A review of research on India's Digital Public Infrastructure

May 2025

Mapping the evidence gap in DPI impact: what we know, what's missing, and why open access to granular data is essential

AUTHORS
Paavi Kulshreshth, Shubham Mudgil

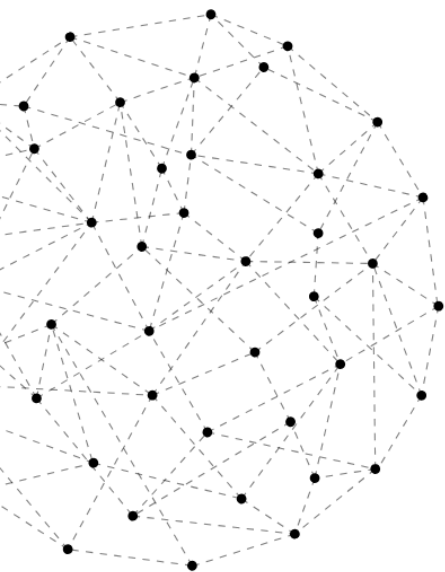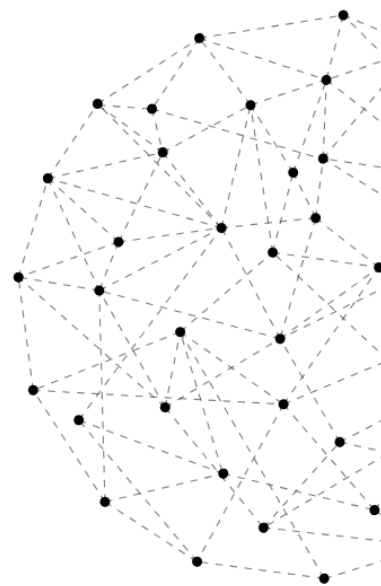**RESEARCH ADVISORS**
Arun Sudarsan, Deepro Guha

Table of Contents

Acknowledgements

Acronyms

Preface

Executive Summary

Acknowledgments

Acronyms

AA	: Account Aggregator
ABDM	: Ayushman Bharat Digital Mission
AePS	: Aadhaar Enabled Payment System
API	: Application Programming Interface
B2B	: Business to Business
B2C	: Business to Consumer
BHIM	: Bharat Interface for Money
CAFRAL	: Centre for Advanced Financial Research and Learning
CIBIL	: Credit Information Bureau (India) Limited
CMA	: Competition and Markets Authority
DaaS	: Data as a Service
DBT	: Direct Benefits Transfer
DEPA	: Data Empowerment and Protection Architecture
DPI	: Digital Public Infrastructure
eKYC	: Electronic Know Your Customer
FIP	: Financial Information Provider
FIU	: Financial Information User
G20	: Group of Twenty
GDP	: Gross Domestic Product
GDPIR	: Global Digital Public Infrastructure Repository
GST	: Goods and Services Tax
GSTN	: GST Network
ICRIER	: Indian Council for Research on International Economic Relations
IIIT	: Indian Institute of Information Technology
IMF	: International Monetary Fund
MGNREGA	: Mahatma Gandhi National Rural Employment Guarantee Scheme
MOSIP	: Modular Open-Source Identity Platform
NASSCOM	: National Association of Software and Service Companies
NBFC	: Non-Banking Financial Company
NDSAP	: National Data Sharing and Accessibility Policy
NPCI	: National Payments Corporation of India
NSAP	: National Social Assistance Programme
OTP	: One-time Password
PMAY-G	: Pradhan Mantri Awas Yojana Gramin

PMJDY   : Pradhan Mantri Jan Dhan Yojana
PoS   : Point of Sale
PRICE   : People Research on India's Consumer Economy
RBI   : Reserve Bank of India
RFID   : Radio Frequency Identification
SIF   : Social Impact Fund
UIDAI   : Unique Identification Authority of India
UNDP   : United Nations Development Programme
UPI   : Unified Payments Interface
USD   : United States Dollar
USSD   : Unstructured Supplementary Device Data
WHO   : World Health Organization

Preface

The rapid advancement of technology, with greater affordability and accessibility, is creating a generation of people whose identity and interactions with the world around them are primarily digital. This ubiquity of technology has encouraged the introduction of digital tools as essential components of governance across the world. These tools are now collectively known as DPIs, or Digital Public Infrastructure. **DPIs are foundational digital platforms, and include accompanying institutional and legal frameworks, that underpin society-wide functions and services.** The world is bullish on the wide-ranging potential of DPIs across sectors and use cases[1]; which can help **increase economic growth, advance access to justice, and reduce emissions in low and middle income countries.**

The Indian example is often cited as a successful model of the development and deployment of DPIs for large-scale impact. India's digital ID program, Aadhaar, has near universal coverage, with over 1.42 billion Aadhaar numbers issued as of March 2025. Authentication of identity through Aadhaar is plugging leakages in targeted subsidy schemes of the government, and reducing the costs incurred by businesses in customer acquisition and verification[2]. The peer-to-peer payment system, UPI (Unified Payments Interface) recorded a peak of 18 billion monthly transactions in March 2025.[3] Other services are being built on top of the underlying DPI protocols and consequently improving residents' access to government and commercial services. India has also launched the Global Digital Public Infrastructure Repository to export the systems that underpin IndiaStack with the promise of transformational change.[4]

While the aggregate platform adoption/use statistics are impressive, the impact of DPIs has not been studied and evaluated extensively, particularly its contribution to financial inclusion and overall economic growth. Our conversations with researchers reveal a lack of granular and disaggregated data on DPIs that can facilitate the estimation of its overall impact on society. In this report, we provide an overview of the existing research on DPIs, available data, and advocate strongly for more data to be made available in the open.

---

[1] United Nations Development Programme. (2023). Accelerating the SDGs through digital public infrastructure: A compendium of the potential of digital public infrastructure. https://www.undp.org/publications/accelerating-sdgs-through-digital-public-infrastructure-compendium-potential-digital-public-infrastructure
[2] Mishra, D., Kedia, M., et al. (2023). State of India's digital economy. Indian Council for Research on International Economic Relations (ICRIER). https://icrier.org/pdf/State_of_India_Digital_Economy_Report_2023.pdf
[3] National Payments Corporation of India. (2024). UPI Product Statistics. https://www.npci.org.in/what-we-do/upi/product-statistics
[4] Modular Open-Source Identity Platform. The MOSIP Project. https://mosip.io/mosip_project

Executive Summary

Over the past decade, India has emerged as a global pioneer in the development and implementation of Digital Public Infrastructure (DPI). By leveraging foundational technologies such as Aadhaar, the Unified Payments Interface (UPI), and the Data Empowerment and Protection Architecture (DEPA), India has created a digital ecosystem that drives financial inclusion, enhances governance, and stimulates economic growth. **This report examines the evolution, usage, and impact of India's DPIs, while identifying challenges, and offers a vision for the future.**

The rapid expansion of technology in India, particularly following the telecom revolution, set the stage for a transformative digital era. Affordable mobile phones and internet services reached millions, making digital connectivity more accessible. This environment gave rise to India Stack, a set of interoperable digital systems that support critical societal functions, ranging from identity verification to payments and data sharing. **As one of the most widely adopted digital identity systems globally, Aadhaar became a cornerstone of this infrastructure. It has provided near universal identity coverage, enabling residents to access government services, financial products, and commercial services securely and efficiently.** Aadhaar has reduced costs and inefficiencies in the delivery of government subsidies by eliminating fraud, while also streamlining private-sector operations like customer onboarding.

Following the success of Aadhaar, the introduction of UPI in 2016 marked a watershed moment in the digital payments landscape. **UPI enabled real-time peer-to-peer and merchant transactions, revolutionizing how money moves across India. Its usage surged after the demonetization of high-value currency notes in 2016, and it now processes well over 18 billion transactions per month.** UPI's widespread adoption has been a critical driver of financial access, allowing millions of people to participate in the digital economy for the first time. Despite this, the coexistence of high cash usage alongside digital transactions reflects the complex nature of consumer behaviour, underscoring the ongoing need for more research and policy interventions.

While financial access remains a key success story, DPIs have also made substantial contributions to India's economic growth. UPI, in particular, has been instrumental in promoting digital payments across sectors. It has not only facilitated greater transactional efficiency but has also supported small businesses by lowering transaction costs and expanding market access. **A recent study by NASSCOM estimates that DPIs like Aadhaar and UPI contributed ₹2.6 lakh crore to the economy in 2022, accounting for approximately 0.9% of India's GDP. Projections indicate that the digital economy could grow to USD 1 trillion by 2030, driven in large part by DPIs and their integration into multiple sectors.**

DEPA represents a significant innovation in data management, allowing individuals to securely control and share their data. This framework gives users the agency to consent to sharing their financial, health, or educational data with service providers, enabling easier access to credit, insurance, and other essential services. The Account Aggregator (AA) system, a key component of DEPA, empowers individuals and small businesses to use their financial data to access credit

without relying on traditional credit scores, which can be particularly beneficial for new-to-credit or underserved customers.

Internationally, India's DPI model has garnered attention and is being viewed as a template for other developing countries. **The launch of the Global Digital Public Infrastructure Repository (GDPIR) underscores India's role in exporting its digital infrastructure expertise. Nations across Africa and Southeast Asia have expressed interest in replicating India's success with Aadhaar and UPI.** Modular Open-Source Identity Platform (MOSIP) is helping countries like Morocco, Sri Lanka, and Tunisia develop their own digital identity systems based on India's experience. The export of DPI expertise not only fosters international collaboration but also positions India as a leader in digital governance, with the potential to accelerate global progress toward sustainable development goals (SDGs).

However, to succeed in the internalisation of India's DPIs, the availability of comprehensive, granular data is critical. While there is significant aggregate data on Aadhaar, UPI, and DBT usage, disaggregated data—such as gender, income, or regional breakdowns—are often missing, making it difficult to fully understand how DPIs are affecting different demographics. Transaction failure rates, particularly in systems like Aadhaar-enabled payment services (AePS), are not regularly disclosed, which hampers efforts to improve service delivery and user experience. Additionally, despite the growth in digital payments, cash remains a dominant form of transaction, highlighting a paradox that requires further investigation into user behaviour and barriers to fully embracing digital systems.

Looking ahead, the expansion and refinement of DPIs will be critical. As India continues to scale its digital infrastructure, the focus must shift toward closing inclusion gaps, particularly in rural and underserved regions. Increased transparency in data collection and reporting will enable better monitoring of DPI performance, guiding interventions to address challenges such as transaction failures or uneven regional adoption. Further integration of DPIs with emerging sectors like healthcare (through platforms like eSanjeevani) and education will deepen their impact and broaden the scope of digital inclusion.

The future of India's DPIs will also depend on balancing innovation with privacy and data security. As more individuals share their data to access services, ensuring robust safeguards will be paramount. The Data Empowerment and Protection Architecture will play a crucial role in this by giving users control over their personal information while maintaining secure data flows between individuals and service providers.

India's DPIs have set a global benchmark for how digital infrastructure can drive inclusive growth, improve governance, and foster economic development. The next phase will require continued innovation, improved data transparency, and stronger privacy frameworks. **By addressing these challenges, India can further extend the benefits of its DPIs to every corner of society while solidifying its position as a global leader in digital transformation.**

Chapter 1
DPIs in India: A Decadal Review

Rooted in the larger *Digital India* vision, Digital Public Infrastructures (DPIs) are a key priority of the Indian government. As the name suggests, these are public infrastructure – like roads and railways – that are essential for governments, businesses, and individuals to deliver and access services in digital societies. **Built on open protocols and specifications, DPIs are meant to be interoperable, federated, and decentralised. This means they allow different digital devices or platforms to work together and enable data to be stored where it is collected.**[5] India's most popular digital payments system, Unified Payment Interface (UPI), is a prime example. Its interoperability enables seamless payments across platforms, apps, and devices, while its federated architecture ensures that user data gets stored in the UPI app where it is collected. This has decentralised data ownership, empowering users to own, access, and reuse their data for the first time in India's digital history.

What has worked for India's veteran DPIs — Aadhaar and UPI — is their timely alignment with policy interventions and technological advancements. UPI, for example, leveraged the telecommunication boom driven by affordable smartphones and low-cost, high-speed internet. Its launch in 2016 also coincided with the demonetization of INR 500 and INR 1000 currency notes, which further boosted its prominence.[6] UPI transactions surged to over 10 billion, setting a world record of 3729 transactions per second in 2023, less than a decade after its introduction.[7] In contrast, Aadhaar's policy alignment has been more deliberate. Government benefit transfers were linked to Aadhaar in 2013 to reduce fraud and corruption, and have reportedly resulted in fiscal savings of INR 8,000 crore in 9 years.[8]

The use of Aadhaar and UPI has also created a digital trail — data that has to be securely stored, and shared if necessary. The Data Empowerment and Protection Architecture (DEPA) achieves this by giving users control over their data through consent-based mechanisms. Newer DPIs, such as DigiLocker[9] and the Account Aggregator (AA) framework[10], use this architecture to facilitate secure, consent-based data sharing.

---

[5] https://docs.cdpi.dev/the-dpi-wiki/dpi-tech-architecture-principles
[6] 2016 Indian banknote demonetisation - Wikipedia
[7] https://www.livemint.com/money/personal-finance/upi-the-world-s-favourite-payment-method-hits-964-billion-in-record-time-digital-payments-credit-cards-11724997818372.html
[8] https://timesofindia.indiatimes.com/city/lucknow/aadhaar-helped-up-save-8k-cr-in-nine-yrs/articleshow/94522305.cms
[9] DigiLocker. DigiLocker: A Digital India Initiative that allows users to store their official documents (eg: birth certificates, mark sheets, driving licences) in their "digital document wallet", and facilitates consent-based sharing with third-parties. https://www.digilocker.gov.in/
[10] AA Framework is a financial data sharing network that allows users to digitally access and share their financial data across institutions. https://sahamati.org.in/what-is-account-aggregator/

Together, Aadhaar (identity), UPI (payments) and DEPA (data) are popularly referred to as the India Stack.[11] A more recent addition to Stack has been Open Networks, which allow participants in a specific ecosystem to connect easily without the need for custom APIs or infrastructure. The National Digital Health Mission acts as an open network for consumers, allowing them to use their Health IDs, which is linked to their medical records, to securely access healthcare services across different apps and providers.[12]

Both the Stack and India's broader DPI approach have gained global relevance for their role in tech-enabled development. **The Indian G20 Presidency, saw a strong emphasis on DPIs, with the UN highlighting their potential to accelerate realisation of SDGs.[13] The Presidency also culminated in the creation of a Global DPI Repository (GDPIR) for knowledge sharing on development and deployment of population-scale DPIs.**[14] Earlier, a 2019 analysis by the Bank of International Settlements (BIS) had recognised India's digital financial infrastructure as a valuable case study relevant and applicable for all economies, irrespective of their stage of development.[15]

Recognition of DPI-led development has also resulted in enhanced investment in DPI adoption. The World Bank and the Indian IT Ministry are now looking to build a state-level DPI index to analyse adoption gaps and boost digital economy and financial inclusion.[16] Indian technologists are also looking to pioneer the concept of 'DPI as a service' (DaaS), which extends beyond domestic borders, allowing other nations to replicate the Indian DPIs.[17] MOSIP, an Indian homegrown startup seen as an early service provider in this space, is currently assisting 26 countries to pilot Aadhaar-like national identification systems.[18] The Indian government, meanwhile, has signed several MoUs with Global South countries, including Papua New Guinea, Trinidad and Tobago, Tanzania, Kenya, Cuba and Colombia, to aid implementation of population scale DPIs.[19]

1.1 What we know about DPI Impact

The positioning of DPIs as both domestic success stories and emerging India-to-the-world solutions is supported by their perceived positive impact on inclusion and growth. However, the push for wider adoption is accompanied by a growing interest in assessment of DPI impact. Analyses by national and international institutes have highlighted certain trends. **ICRIER's State of Digital Economy Report 2023 suggests that UPI's share of total digital transactions in India catapulted from a mere 2% in 2016-17 to 52% in 2021-22.**[20] A 2024 mobile survey conducted with over 2,000 respondents from 20 states and UTs further revealed that two out of three Indians use

---

[11] https://indiastack.org/index.html

[12] https://www.youtube.com/watch?v=rF6t6mTmxNc

[13] https://www.undp.org/publications/accelerating-sdgs-through-digital-public-infrastructure-compendium-potential-digital-public-infrastructure

[14] https://www.dpi.global/

[15] https://www.bis.org/publ/bppdf/bispap106.pdf

[16] https://www.business-standard.com/economy/news/world-bank-meity-to-come-up-with-state-digital-public-infra-adoption-index-124042500986_1.html

[17] https://economictimes.indiatimes.com/tech/technology/after-saas-daas-may-soon-catch-on-for-faster-rollout-of-digital-public-infrastructure/articleshow/107830026.cms?from=mdr

[18] https://www.mosip.io/country_partners

[19] https://pib.gov.in/PressReleasePage.aspx?PRID=2037598

[20] Mishra, D., Kedia, M., et al. (2023). State of India's Digital Economy. Indian Council for Research on International Economic Relations (ICRIER). https://icrier.org/pdf/State_of_India_Digital_Economy_Report_2023.pdf

UPI.[21] Similarly, Aadhaar is reported to have achieved near-universal adult coverage, according to government estimates.[22]

As India's primary ID verification system, linked to bank accounts, tax receipts, and various government and private services, studies on Aadhaar impact often note significant shifts in access and inclusion. The JAM trinity[23] is frequently credited with large-scale digital financial inclusion in India. A World Bank report suggests that this inclusion is likely as high as 80%[24] — a dramatic rise from 25% in 2008.[25] It goes on to propose that without DPIs, it would have taken India nearly five decades to achieve such inclusion.

Aadhaar has also been described as a "digital multiplier of the Indian economy", with its contributions linked to the growth of India's startup ecosystem, comprising 40 unicorns valued at $94.77 billion in 2021.[26] **A 2024 analysis by NASSCOM estimates that Aadhaar, UPI, and FasTag[27] added INR 1.2 lakh crores in direct, and INR 1.4 lakh crores in indirect impact to the economy in 2022. This accounted for nearly 0.9% of the country's GDP. The report forecasts that the digital economy will reach USD 1 trillion by 2030, or 12.5% of a USD 8 trillion economy.[28]**

1.2 What we must know about DPI Impact

Despite high penetration rates and optimistic projections, the growing research interest in DPI impact has underscored the need for data that captures critical, granular realities. Data on UPI transactions indicates skyrocketing volumes and value. However, cash in circulation has more than doubled since the UPI launch, even after the phasing out of ₹2000 notes.[29] This "cash paradox" may reflect unequal adoption, particularly in rural areas and among populations with limited digital literacy or poor internet connectivity. Alternatively, higher economic activity, which often drives up the demand for liquidity across formal and informal sectors, might be the key determinant of increased demand. To better understand these dynamics and the interplay between digital payments and cash usage, granular and comprehensive data on transaction patterns and demographic adoption trends are essential.

Further, while digital systems are essential enablers of access and inclusion, without adjustments for equity and equality challenges, they risk exacerbating the very gaps they aim to bridge. The gender digital divide remains significant in India – in 2024, 10% fewer women than men owned mobile phones and 16% fewer women had internet access on their mobile phones.[30] If DPI adoption, predominantly made accessible through mobile phones, neglects this disparity, it could

[21] https://artha.global/op-ed/opinion-more-than-two-in-three-indians-use-upi-survey/
[22] Unique Identification Authority of India. (2023). State/UT wise Aadhaar Saturation.
https://uidai.gov.in/images/StateWiseAge_AadhaarSat_Rep_31122023_Projected-2023.pdf
[23] Pradhan Mantri Jan Dhan Yojana (PMJDY), which provided bank accounts to unbanked adults, Aadhaar – the digitized national identification system, and access to affordable mobile phones, together form the JAM trinity
[24] https://documents1.worldbank.org/curated/en/099092023121016458/pdf/P178703046f82d07c0bbc60b5e474ea7841.pdf
[25] https://pib.gov.in/PressReleasePage.aspx?PRID=1955439
[26]
https://www.isb.edu/en/research-thought-leadership/research-centres-institutes/cbi/research-insights/case-study/aadhar--the-digital-multiplier-of-the-indian-economy-by-professo.html
[27] FASTag is a device that uses Radio Frequency Identification (RFID) technology for making toll payments directly while the vehicle is in motion. This is part of the larger National Electronic Toll Collection (NETC) program, developed by the National Payments Corporation of India (NPCI).
[28] NASSCOM. (2023, February). India's Digital Public Infrastructure – Accelerating India's Digital Inclusion
[29] https://economictimes.indiatimes.com/news/economy/finance/cash-is-king-along-with-upi/articleshow/109670328.cms?from=mdr
[30] https://www.gsma.com/r/wp-content/uploads/2024/05/The-Mobile-Gender-Gap-Report-2024.pdf

widen the access gap rather than close it. This makes access to user-level data that helps identify such gaps crucial to inclusivity.

To evaluate the current state of data availability, we engaged with researchers and experts from Indian and international institutions. These conversations highlighted significant gaps in data accessibility, posing major challenges in evaluating the granular, user-level impact of DPIs. **In this brief, we analyse existing literature and data gaps through two case studies — Aadhaar and UPI. We also explore the viability of an open data policy in enhancing access to critical datasets, and conclude by proposing first principles for data sharing that prioritize both usability and security.**

Chapter 2
Bridging Evidence and Unknowns

India's DPI approach has garnered relevance both domestically within the country and globally. Public and private actors alike are now adopting the DPI approach. Google launched 'DPI in a box' – a 'Digital India' toolkit for the world – at the Google for India event in October 2024.[31] Indian firm MOSIP is helping 26 countries develop their national identity systems similar to the Aadhaar. Global organisations like the World Bank, BIS, UN have also endorsed the Indian DPI approach, with the UN emphasising its utility and viability for achieving SDGs.

However, to further bolster the credibility of the Indian DPI approach and bridge any gaps in adoption or efficiency, we need data that provides insight into DPI impact on macro as well as micro economic indicators. **However, our consultations with researchers from domestic and international institutions highlight a significant lack of viable datasets required to evaluate DPI impact, particularly on micro-level indicators such as financial inclusion. This data gap arises from two key challenges: the limited availability of publicly accessible datasets necessary for comprehensive assessments, and the reluctance of relevant authorities to share proprietary data, often citing privacy concerns.**

Despite these challenges, researchers have used extrapolation and proxies to estimate DPI impact to some extent. However, most of the available literature focuses on two of India's most well-integrated DPIs — Aadhaar and UPI. While both our analysis and consultations outline common gaps in data availability across the DPI ecosystem, this chapter presents Aadhaar and UPI as case studies to demonstrate how improved data availability could strengthen impact assessment.

2.1 Aadhaar and Aadhaar-enabled Services

Launched in 2009, Aadhaar is the oldest and most ubiquitous among Indian DPIs. Through individual-consent, Aadhaar can be used for identity authentication with any supported government or private service provider. In this section, we assess existing literature and estimates, and outline gaps in data availability for three key services built on top of Aadhaar — (i) eKYC and Aadhaar authentication, direct benefit transfers (DBT), and Aadhaar-enable payment system (AePS).

2.1.1 eKYC and Aadhaar authentication

The Reserve Bank of India (RBI) requires financial institutions and service providers to verify customer identities through its KYC (know your customer) guidelines[32]. Previously, customers had to submit physical document copies to confirm their identity. Today, the Government and UIDAI[33] have enabled Aadhaar based eKYC[34] for accessing various public and private services. Research indicates that eKYC has lowered customer acquisition costs when we juxtapose against traditional methods of verification[35], with the cost of verification and authentication dropping from USD 12 (INR 1000) to 6 cents (INR 6) per customer.[36] This reduction is likely to have improved the commercial viability and incentives for service providers to expand into untapped markets and reach underserved residents. Such causal relationships are not made evident with current data availability.
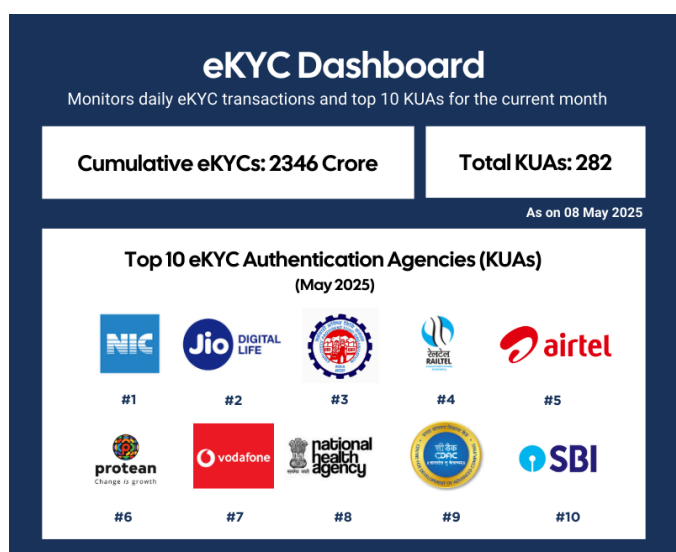
Figure 1
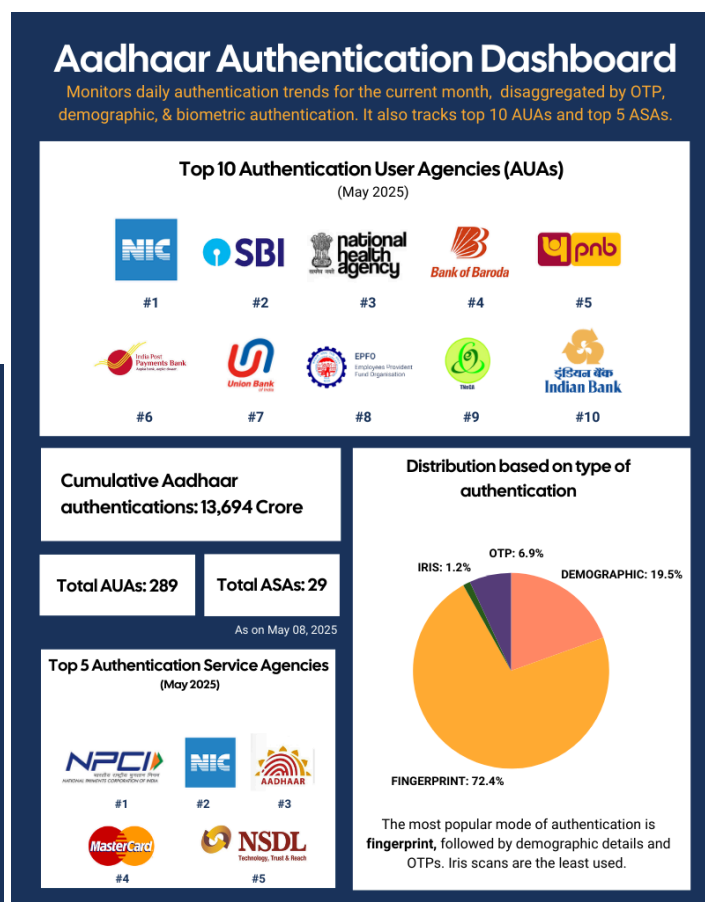
Figure 2

UIDAI maintains the Aadhaar dashboard to provide information on eKYC and authentication transactions. These are mostly limited to user registration metrics, such as counts of daily eKYC and biometric authentications, and monthly Aadhaar updates. The dashboard also provides some level of disaggregation, including number of Aadhaar cards generated at the state, district, and sub-district levels; the percentage of Aadhaar generation for male, female, and transgender individuals at the state and district levels; and Aadhaar authentications categorized by method —

[32] Reserve Bank of India. (2016, February 25). Master Direction - Know Your Customer (KYC) Direction, 2016. Reserve Bank of India. https://www.rbi.org.in/scripts/BS_ViewMasDirections.aspx?id=10292
[33] Unique Identification Authority of India
[34] Electronic Know Your Customer (eKYC) allows service providers to digitally authenticate customer identity.
[35] Mishra, D., Kedia, M., et al. (2023). State of India's digital economy. Indian Council for Research on International Economic Relations (ICRIER).https://icrier.org/pdf/State_of_India_Digital_Economy_Report_2023.pdf
[36] https://www.moneycontrol.com/technology/e-kyc-cost-down-from-rs-1000-to-less-than-rs-6-economic-survey-2024-article-12774537.html

demographic, biometric, or OTP-based. While eKYC and authentication data for different sectors, such as banking and health, is not available on the dashboard, it does highlight monthly top service providers based on the number of eKYC and authentication transactions. These monthly estimates available for the incumbent calendar year (including the current month) - Figures 1 and 2 - provide an illustrative example for May 2025.

To establish the causal relationship between Aadhaar eKYCs and access, these registration metrics must include additional levels of disaggregation that provide insight into user experience. The Aadhaar Dashboard includes the cumulative number of eKYC transactions for every authentication agency. This can be used to identify certain sector-wise authentications percentages by clubbing all banking agencies or all telecom ones. While this can help assess which sectors see the highest to lowest shares of authentication, it does not provide any real insight into access. Tracking detailed eKYC transactions across telecom, healthcare, banking and other sectors—with further breakdown by device type and network metrics—would provide valuable insights that are currently unavailable. If made available, this data could reveal sector-specific usage patterns and serve as proxies for participation by different user groups, particularly low-income users through affordable smartphone metrics. **Such disaggregated data would enable identification of access gaps, for instance by correlating network latency with transaction failures to determine if connectivity issues impede successful authentication.**

Transaction failure rate is an example of user journey data, i.e., data that provides insight into what occurs once a user registers/enters a digital ecosystem. In this case, transaction failure data reveals what happens once a user initiates eKYC or authentication. Notably, as defined by UIDAI, Aadhaar authentication can encounter over 90 types of authentication errors.[37] Metrics like the average time taken for transactions also provide similar insight. **When combined with data on device type, network latency, user demographics (gender, age, etc.), residence type (rural or urban), authentication method (e.g., biometric), and sectoral distribution, user journey data can provide nuanced insights into the experience of different user categories.** A high rate of transaction failures in banking eKYCs from districts with low network latency can highlight access disparities, signalling the need for targeted market or governance interventions.

At present, such comprehensive data is publicly inaccessible for all Aadhaar-enabled ecosystems. The following sections highlight these gaps in the context of DBTs and AePS.

2.1.2 Direct Benefits Transfer (DBT)

Under the DBT Scheme, commenced in 2013, individuals can link their unique Aadhaar ID to their bank accounts, and government authorities can carry out digital cash transfers of government benefits directly into the accounts of eligible beneficiaries. The government claims that this Aadhaar-based payment system (ABPS) helps reduce administrative costs and minimizes leakages by eliminating ghost beneficiaries. The Ministry of Rural Development further argues that ABPS DBTs help avoid delays since the transfer of benefits is intrinsically linked to Aadhaar IDs, and changes in beneficiary location or bank account number does not result in delayed transfer of benefits.[38] Based on this premise, the DBT Bharat portal evaluates the performance of states/UTs

---

[37] Unified Identification Authority of India. What are error codes?
https://uidai.gov.in/en/305-faqs/aadhaar-online-services/aadhaar-authentication-history/10779-what-are-error-codes.html
[38] https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1953438

and ministries, ranking them by the volume and value of DBT transactions. These are available on the [DBT portal](#), which also provides macro estimates offering an overview of coverage. These include national and state-wise total DBT transactions and amounts disbursed per fiscal year (available up to 2019-20); estimated cost savings for each fiscal year (since 2019-20); and the cumulative number of non-unique beneficiaries since the scheme's inception (2013-14). Additionally, the DBT portal provides ministry-wise aggregates, such as cumulative estimated gains or cost savings for specific ministries,[39] and links to portals of various government schemes for scheme-specific DBT details.

Some scheme portals such as those for Pradhan Mantri Awas Yojana Gramin (PMAY-G) and the Mahatma Gandhi National Rural Employment Guarantee Scheme (MGNREGS) also maintain details of DBT transactions across states. **However, estimates remain high-level, similar to the DBT portal. Inclusion of baseline disaggregation based on demographic indicators such as gender, age, and disability can significantly enhance the usability of these high-level aggregates.** For example, analysing the proportion of male to female beneficiaries who receive seed DBTs[40] can help assess the gender-responsive impact of the scheme on the farm sector.

During our consultations, researchers evaluating the DBT scheme impact highlighted the absence of user journey data, particularly average transactional delays and failure rates. One intended benefit of Aadhaar-linked DBTs is to reduce payment delays for beneficiaries, including migrant and contract workers, who often relocate. However, without tracking the average monthly or periodic delays in DBTs, this intended consequence cannot be measured effectively. **Researchers also confirmed that while essential for estimating the scheme's impact, access to transaction failure data is severely restricted. Details of failed transactions, recorded in DBT-09 reports, are accessible only through RTI requests, which are typically subject to delays or denials.[41]** Periodic audit reports provide some insight into irregularities, but may not be sufficient for timely action or root cause analysis. For example, a 2017-2021 audit report for Punjab noted multiple transaction failures caused by unauthorized changes to beneficiaries' bank account details by DSSOs (district social security officers). Real-time or monthly dashboards could help promptly detect such irregularities, enabling timely interventions by authorities and supporting researchers in conducting root cause assessments. Including transaction failure metrics in state rankings on the DBT portal could incentivize administrative action to address irregularities or reduce high failure rates.

Aadhaar DBTs and their impact, however, should not be viewed in isolation. India still lacks adequate physical banking infrastructure, particularly in rural pockets, needed to access these DBTs. AePS aims to bridge this gap through a *phygital* approach, combining physical and digital solutions. However, limited data availability continues to constrain our understanding of its effectiveness.

---

[39] https://dbtbharat.gov.in/static-page-content/spagecont?id=18
[40] https://seednet.gov.in/SMIS/Indentor/DBTSeed/Dashboard.aspx
[41] Kodali, S. (2020, April 16). COVID-19, Aadhaar-DBT and a reminder of the issues with transaction failure data. The Wire. https://thewire.in/government/covid-19-aadhaar-dbt-and-a-reminder-of-the-issues-with-transaction-failure-data

2.1.3 Aadhaar-enabled Payment System (AePS)

AePS, introduced in 2016 and developed by the National Payment Corporation of India (NPCI), uses Aadhaar authentication to provide basic banking services like cash withdrawal, deposit, fund transfer, and balance inquiry. **A customer only needs their bank name, Aadhaar number, and biometric fingerprint (submitted during registration) to complete a transaction[42].** Similar to DBT, AePS is used to enhance access to government payments under various central and state schemes such as MGNREGS, PMAY, and National Social Assistance Program (NSAP).

The NPCI publishes monthly statistics that track:

- Total approved transactions – monthly average of over 468 million/month in FY 2024-25
- Interbank transactions[43] (volume and value) – monthly average of nearly 200 million transactions worth over INR 23,000 crores in FY 2024-25
- Intra-bank transactions[44] (volume) – monthly average of over 300 million transactions in FY 2024-25
- BHIM Aadhaar Pay transactions (volume and value) – monthly average of over 1.9 million transactions worth over INR 570 crore in FY 2024-25
- Successful eKYCs – monthly average of nearly 40 million in FY 2024-25

To enable financial inclusion, AePS relies on phygital access, allowing residents to access digitally transferred government payments through physical point of sale (PoS) devices. According to the Indian government, rural residents would typically travel 3-25 km on average to access banking services, spending Rs.125 to deposit Rs.100.[45] To reduce this travel time to 15 minutes, it intended to deploy 2 million AePS devices by September 2017 (initial target).[46] **Despite the centrality of physical PoS devices to AePS, monthly statistics published by the NPCI do not include the total number of deployed PoS devices or MicroATMs at the grassroots.**

Similar to authentication and DBTs, AePS transaction failure data is not published by the NPCI, creating a blindspot when evaluating user experience. Inclusion of such data in monthly statistics can highlight potential exclusions in accessing government benefits and wage payments across central and state government schemes. A 2016 finance ministry committee report indicated a 60% failure rate for AePS transactions, followed by a 2019 RBI committee report confirming high failure rates.[47] **An academic analysis by the Indian School of Business (ISB) suggested that AePS had a ~34% transaction failure rate between 2014 and 2018 (17% biometric identification, 3.7% technical and 13% non-technical failure).[48]** The increase in total approved transactions — from ~190 million in April 2019 to ~470 million in April 2024 — marks increased adoption, which may be

---

[42] National Payments Corporation of India. Aadhaar Enabled Payment System – Aadhaar Pay. https://www.npci.org.in/what-we-do/aeps/product-overview

[43] Available as OFF-US transactions – technical for interbank transactions that require interbank settlement for transferring funds from one bank to another.

[44] Available under ON-US transaction – technical term for intra-bank transactions

[45] National Payments Corporation of India. To create a significant social impact: At least 2 million AePS-enabled devices. MyGov. https://transformingindia.mygov.in/to-create-a-significant-social-impact-at-least-2-million-aeps-enabled-devices-li/

[46] https://www.embassyofindiatunis.gov.in/pdf/Aadhaar_Overview_African%20nations_180717_v03.pdf

[47] https://dvararesearch.com/wp-content/uploads/2023/12/Transaction-failure-rates-in-the-Aadhaar-enabled-Payment-System-Urgent-issues-for-consideration-and-proposed-solutions.pdf

[48] https://dvararesearch.com/wp-content/uploads/2023/12/Transaction-failure-rates-in-the-Aadhaar-enabled-Payment-System-Urgent-issues-for-consideration-and-proposed-solutions.pdf

driven by reduced transaction failure. However, without regular transaction failure statistics, such estimations remain speculative. Geographical disaggregation of AePS failure data could further support targeted interventions.

Finally, our analysis and stakeholder consultations revealed a lack of consistency across datasets and publishing agencies, particularly when assessed for timeliness and the level of disaggregation. This inconsistency in siloed datasets reduces the usability of even openly accessible public data. The live ABPS dashboard tracks MGNREGS workers' eligibility for Aadhaar-based payments at the national, state, district, and block levels. However, its usefulness in assessing real-world access remains limited without data on how many workers actually use AePS to withdraw their wages at these levels. While some workers may rely on other modes, AePS data can be critical for assessing gaps between eligibility and actual access. Therefore, datasets must interact cohesively to yield insights that enable targeted interventions to address systemic barriers. Data availability gaps discussed in this case study have been noted across the DPI ecosystem. In the next section, we explore these in the context of UPI.

2.2 Unified Payments Interface (UPI)

Aadhaar has achieved near 100% penetration among the adult population. However, the ubiquity of DPI is likely more tangible for UPI users, given its everyday, interactive payment functionality. The government estimated in 2023 that more than 30 crore individuals, and 5 crore merchants, are registered on UPI.[49] The decadal growth of UPI has been remarkable, and the share of UPI-linked services are rapidly increasing in the market – making it a go-to payments option for most Indians.

Similar to AePS, NPCI publishes usage metrics for UPI, such as daily and monthly volume and value of UPI transactions and the number of banks that offer UPI services (see Figure 3).[50] It also publishes statistics that provide insight into the larger UPI ecosystem, including usage metrics for different banks, UPI TPAPs[51], types of transaction[52], and merchant categories (see Figure 4).

Month	No. of Banks live on UPI	Volume (in Mn)	Value (in Cr.)
Apr-25	668	17,893.42	23,94,925.87
Mar-25	661	18,301.51	24,77,221.61
Feb-25	653	16,106.19	21,96,481.69
Jan-25	647	16,996.00	23,48,037.12

Figure 3

---

[49] Press Information Bureau. (2023, June 19). India among top countries in innovation index, surpasses China, Thailand, and South Korea. https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1973082&ref=indiatech.com
[50] National Payments Corporation of India. Monthly Metrics. https://www.npci.org.in/statistics/monthly-metrics
[51] Third party application provider
[52] Peer-to-peer (P2P) and peer-to-merchant (P2M)

UPI Apps (Apr'25)

Sr.No.	Application Name	Customer Initiated Transactions		B2C Transactions		B2B Transactions		Total	
		Volume (Mn)	Value (Cr)	Volume (Mn)	Value (Cr)	Volume (Mn)	Value (Cr)	Volume (Mn)	Value (Cr)
1	Phone Pe #	8,362.31	12,05,047.30	0.00	0.00	0.00	0.00	8,362.31	12,05,047.30
2	Google Pay #	6,488.55	8,42,628.56	0.00	0.00	0.00	0.00	6,488.55	8,42,628.56
3	Paytm #	1,210.24	1,31,153.97	0.00	0.00	0.00	0.00	1,210.24	1,31,153.97

Figure 4

NPCI statistics, however, do not include any user metrics, such as total number of users or active monthly users. This omission creates a critical gap in understanding the relationship between user and transaction patterns. If both transaction volume and the user base increase month-on-month, it suggests greater UPI adoption and usage among a growing user set. Conversely, an increase in transaction volume without a corresponding rise in user numbers points to enhanced usage among existing users. Similarly, a decrease in the total transaction value alongside an increase in the number of users indicates a shift toward more frequent, smaller-value transactions across a growing user base. Access to user metrics is essential to uncover these relationships. Further disaggregation of these metrics can provide more in-depth insights. For example, geographical disaggregation could identify areas with the highest UPI penetration and those lagging. Such data insight can guide on-ground assessment, which in turn can help target interventions. While NPCI currently lacks this granularity, PhonePe, a dominant TPAP that comprised ~50% of the total market share[53] in 2024, bridges these gaps to some extent. It publishes quarterly user metrics and region-specific estimates through its [public data portal](#) (available since 2018).

Despite data constraints, our review of existing literature shows that researchers have managed to estimate UPI impact on economic indicators like financial inclusion, credit access, and overall economic growth. **In the following sections, we highlight a few prominent analyses and explore how improved data availability could enhance research.** This assessment incorporates insights from authors of the cited research as well as perspectives from researchers who were unable to conduct studies due to inadequate data access.

2.2.1 Financial Inclusion

To assess UPI's impact on financial inclusion, disaggregated user metrics are essential. These metrics would offer insights into UPI adoption trends over time and among different user groups. Currently, only PhonePe publishes periodic, geographically disaggregated user statistics, offering some insight into regional financial inclusion. Existing literature also provides limited insight into the demographic distribution of UPI users. For example, in 2023, a senior Google Pay executive

---

[53] Both in terms of volume and value of UPI transactions (source: NPCI)

noted that only 30% of UPI users were women.[54] Similarly, an analysis by PhonePe found that their male users consistently displayed higher transaction frequency than their female users across Generations X, Y, and Z.[55] While valuable, these insights are insufficient for understanding gaps in UPI adoption. Demographic and geographical disaggregation can not only enhance usability of user metrics, but also help target interventions to improve UPI adoption among specific population groups across underserved regions.

In the absence of such data, however, researchers have used monthly usage estimates (eg: transaction volume) from NPCI to understand inclusion. NPCI's transaction volume and value estimates have been used to calculate the average ticket size (ATS)[56] of UPI transactions over time. Subsequently, an evident decline in this ATS has been attributed to a more inclusive digital payments' ecosystem. Specifically, smaller peer-to-merchant (P2M) and peer-to-peer (P2P) transactions have been attributed to adoption of UPI by small businesses such as street vendors[57] and increased use by poorer users.[58] While a best resort given the absence of user metrics, such estimations do not offer conclusive results.

Declining ATS reflected in NPCI data has been attributed to adoption among small businesses and poorer users. However, without periodic and disaggregated user metrics, such as monthly active users across user segments, such attribution may not be the most reliable. A more viable attribution would be increased use of UPI for smaller, everyday purchases,[59] and can be supported by NPCI insights on UPI merchant categories.[60] In 2024 (up to November), the merchant category that saw the highest volume of transactions was "groceries and supermarkets" — an everyday-use retail category that rarely comprises high-value goods. Therefore, to establish a more credible link between UPI transactions and inclusive adoption by poorer users, we need disaggregated user data.

An alternative suggested by a researcher from an international funding institute during our consultations is to disaggregate UPI transactions by phone type or operating system (OS). This data could be clubbed into proxies for richer and poorer user segments based on affordability of different phone and OS types, and provide more credible insight into financial inclusion. However, such disaggregation is currently as absent from NPCI datasets as user metrics.

2.2.2 UPI Impact on Credit Access

Unlike financial inclusion, in our assessment, existing literature establishes a credible link between UPI use and improved credit access. Enhanced credit access has largely been attributed to lenders' ability to make use of digital financial histories to assess borrowers. CAFRAL[61] assessment, based on CIBIL[62] data, noted a 4.6% increase in FinTech lending per capita with a 10%

[54] https://economictimes.indiatimes.com/industry/banking/finance/upi-use-among-women-low-assisted-onboarding-can-drive-uptake-gpays-arati-deo/articleshow/100497224.cms?from=mdr
[55] https://www.phonepe.com/pulse/articles/unlocking_the_essence_of_the_bharatiya_consumer/
[56] Average value per UPI transaction; calculated by dividing total value of transactions by total number of transactions
[57] https://economictimes.indiatimes.com/tech/tech-bytes/infographic-insight-how-is-india-using-upi-and-where/articleshow/104992207.cms?from=mdr
[58] https://www.moneycontrol.com/news/opinion/how-indias-upi-and-brazils-pix-are-fast-pushing-the-financial-inclusion-pedal-12352961.html
[59] www.business-standard.com/finance/personal-finance/how-india-spends-upi-reigns-supreme-credit-cards-witness-highest-growth-124040400094_1.html
[60] https://www.npci.org.in/what-we-do/upi/upi-ecosystem-statistics
[61] Centre for Advanced Financial Research and Learning (CAFRAL) set up by the RBI
[62] TransUnion CIBIL Ltd is a credit information company operating in India that maintains credit files on millions of individuals and businesses

increase in UPI transactions per capita.[63] Similarly, a 2024 analysis based on proprietary CIBIL, SBI and RBI datasets noted a 10% increase in UPI transactions leads to a 7% increase in credit availability.[64] This growth was attributed to UPI's ability to create verifiable digital financial histories, which allowed banks and fintechs to lend more. Further noting a 10x increase in fintech lending to new-to-credit and subprime borrowers, the analysis established a positive correlation between UPI and financial inclusion. **However, as acknowledged by the researchers, this analysis was based on proprietary datasets rarely made available to researchers.** Such exclusive access restricts the research ecosystem's ability to popularize UPI impact and benefits.

2.2.3 UPI Impact on Economic Activity

A BIS assessment of over a hundred emerging economies, including India, in 2014-19, found that a 1.0 percentage-point increase in the use of digital payments is linked to a 0.10 percentage-point rise in GDP per capita growth over a two-year period.[65] Economic theory also emphasises this direct link between adoption of digital payments and economic growth. Given this premise, establishing the merit of an easy-to-use digital payment system like UPI becomes a simple task. However, assessment of specific impact on different economic indicators requires disaggregated data. For example, UPI adoption data disaggregated by geographies can be used for comparative analysis between high and low adoption districts. As discussed, currently such geographically disaggregated data is only published by PhonePe.

Using PhonePe data from 500 districts, a 2023 analysis established a causal link between UPI adoption and increased economic activity. This was measured across three indicators for 2,00,000 households between 2014 and 2022: (a) overall household income, (b) business ownership, and (c) business income.[66] The analysis found that, after UPI's launch in 2016, households in districts with higher adoption of cashless payments saw significantly higher overall income, business ownership, and business income. While valuable, these insights are only based on PhonePe data, which may limit its representation. Geographically disaggregated statistics published by NPCI can enable a more representative assessment of UPI impact.

As evidenced from these case studies, there are common gaps in data availability across Aadhaar and UPI ecosystems. These gaps also persist across other, emerging sector-specific DPIs. For instance, India's Ayushman Bharat Digital Mission (ABDM), designed to create digital health infrastructure, includes *eSanjeevani*, a cloud-based telemedicine system for virtual consultations. However, publicly available data on its impact is limited to state-level aggregates and registration metrics such as total healthcare providers onboarded, patients served (cumulative and daily aggregates), total number of hubs and spokes operationalised, and number of specialities available. As discussed, data privacy and security concerns are often cited as an overlapping justification by appropriate authorities for missing datasets. **While these concerns are valid, estimation of DPI impact requires non-personal datasets. These can be entirely anonymized,**

[63] Centre of Excellence for Policy Research and Advanced Learning. (2023). India Finance Report 2023. https://www.cafral.org.in/sfControl/content/NewsEvent/CAFRAL_Report.pdf

[64] https://www.nber.org/system/files/working_papers/w33259/w33259.pdf

[65] https://www.bis.org/publ/work1196.pdf

[66] Singh Dubey, T., & Purnanandam, A. (2023). Can cashless payments spur economic growth? SSRN. https://www.jbs.cam.ac.uk/wp-content/uploads/2023/07/2023-ccaf-conference-paper-3-dubey-purnanandam.pdf

pseudonymized, or, in some cases, limited to aggregates. However, they must encompass the following attributes to unlock both access and reusability.

<u>**Essential Attributes for DPI Datasets**</u>

User Journey
This data provides an insight into what happens once the user is in the DPI ecosystem. In the case of eSanjeevani, Figure 5 illustrates different steps in the eSanjeevani consultation process; without understanding what happens to the user at each of them, the impact of eSanjeevani remains unclear. High registration rates or login counts accompanied by high dropout rates at stages 3 and 4 indicate strong demand for telemedicine but limited access. Such data, even if aggregated, can pinpoint challenges and inform targeted policy interventions. In instances, where data reaffirms the effectiveness of digital systems, impact gets validated and can encourage state-level adoption, offer a replicable model to other countries, and attract investment.
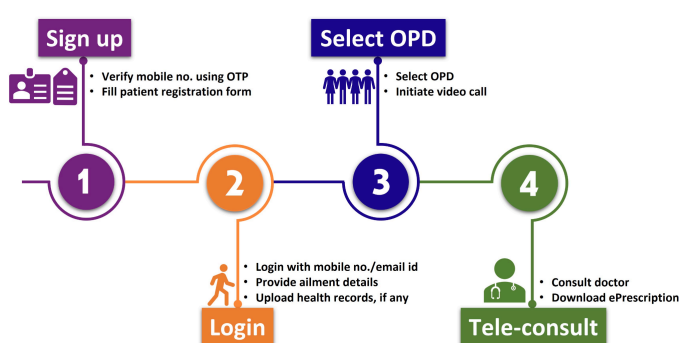

Figure 5: eSanjeevani consultation process (Source: eSanjeevani)

Disaggregation
During our consultations, researchers across the board confirmed the need for disaggregated data for understanding DPI impact on inclusion. Adoption of financial DPIs like UPI must be analysed across factors that influence the socio-economic status of individuals. Typically, these can include disaggregation by demographics (gender, age, income level, education level), type of residence (rural v. urban, metros v. tier II cities), or nature of employment (self-employed, salaried).

Granularity
Granularity here refers to geographical disaggregation; we treat it as a distinct category for clarity. Both government and private actors rely on DPIs such as Aadhaar, UPI and FasTag for service delivery, enabling access, inclusion and economic activity. With granular, geolocation tagged data, comparative impact assessments can help identify important relationships between DPIs and access, inclusion and economic activity. For example, district-level UPI data could be analysed alongside nighttime luminosity (a proxy for economic activity) in regions with similar socio-economic conditions but varying UPI adoption levels. This could reveal correlations or causal links between UPI usage and economic growth. Similarly, breaking down government datasets into national, state, district, and block/village levels can not only enable impact evaluation, but also reduce friction within governments. For example, the ABPS dashboard includes the total number of MGNREGS workers eligible for Aadhaar-based payments across these levels, eliminating the need for internal bureaucratic tracking. As noted previously, if the

Dashboard begins covering user journey estimates beyond registration metrics, it can also be used to block inefficiencies at the appropriate level of governance.

Panel Data

Panel data is multidimensional data that tracks a variable over time. The NPCI tracks the value of UPI transactions over time, with monthly statistics available since 2016, which allows us to assess the growth in UPI transaction value. Similarly, the Pradhan Mantri Awas Yojana-Grameen (PMAY-G)[67] dashboard provides yearly estimates for multiple variables such as the total number of houses completed against sanctions made. During our consultations, researchers stressed the importance of ensuring DPI datasets are both disaggregated or granular and available over time to capture changes across socio-economic segments or geographies, facilitating accurate identification of causal links.

The lack of DPI datasets calls for an open data approach tailored to the DPI ecosystem. In the final chapter, we explore the benefits of such an approach and outline key principles for implementing an open data policy. This policy can enable research that empowers individuals, businesses, and policymakers to design targeted solutions for enhanced access, service delivery and inclusion through DPIs.

---

[67] A national rural housing welfare scheme

Chapter 3
The Promise of Open Data

Data availability gaps across the DPI ecosystem underscore the need for open data – i.e., data that is publicly accessible and reusable. The benefits of such data have been mapped across sectors, with multi-year, multi-country studies indicating significant economic benefits.[68] In India's rapidly developing economy, the GDP contribution of financial open data can be as high as 4-5% in 2030.[69] However, realising this economic potential would require open datasets for financial digital infrastructures such as the veteran ABPS and UPI and newer AA Framework and Open Credit Enablement Network (OCEN)[70] ecosystems.

Case studies from around the world have reinforced the merit of building open data ecosystems that can be leveraged by individuals, businesses, civil society, and governments themselves. In 2005, Denmark's Building and Dwelling Register began publishing its address datasets.[71] Previously, this data could only be accessed through individual municipalities for a fee, making real-time use virtually unattainable. Multiple agencies, such as the Land Registry, collected their own address data, leading to discrepancies across government sources. Introducing a common open address data source improved internal efficiency and record keeping. The Danish government estimates that efficiency gains between 2005 and 2009 resulted in financial benefits of EUR 62 million, with fiscal expenditure of only EUR 2 million. These open datasets are used by the police, transport and postal services, and the Global Positioning System (GPS). In fact, GPS presents a stellar example of what happens when government data is made available for reuse. A 2016 report suggests that economies could suffer losses worth $96 billion without access to the geospatial data collected by GPS.[72]

Within governments, open data can also be used to boost accountability and competition. The India Innovation Index makes public states' innovation ranking and creates an incentive for local governments to improve outcomes. As noted in the Aadhaar case study, state rankings are also available on the DBT portal. However, these are only based on the volume and value of transactions in every state – insufficient metrics for introducing a sense of competition. In contrast, the innovation index gives states scores on more tangible, improvable parameters such as human capital, knowledge output and diffusion, business environment, among others.[73]

---

[68] https://opendatatoolkit.worldbank.org/en/data/opendatatoolkit/starting
[69] https://www.mckinsey.com/industries/financial-services/our-insights/financial-data-unbound-the-value-of-open-data-for-individuals-and-institutions
[70] An open network that connects different participants in the credit ecosystem without the need for any separate infrastructure and APIs. For instance, it can allow credit firms to offer order-financing services to end consumers and working-capital financing to suppliers without any bespoke infrastructure (Source: Indiastack.org)
[71] https://odimpact.org/files/case-study-denmark.pdf
[72] https://odimpact.org/files/case-studies-gps.pdf
[73] https://www.niti.gov.in/sites/default/files/2021-01/IndiaInnovationReport2020Book.pdf

Open data and the impact assessment it enables is crucial to enhance access and service delivery – a benefit to individuals, public and private service providers, and society at large. In India, the Electricity Supply Monitoring Initiative (ESMI) presents an important example of such impact at scale. India was plagued with poor power quality, particularly voltage fluctuations, when the initiative was introduced by the Prayas Energy Group (NGO). ESMI solves this by providing access to real-time power supply monitoring data to multiple users. Active in 200 locations across 18 states by 2017, this data served as critical evidence that drove advocacy for improving quality power supply at state and central levels.[74]

Open data can also establish trust and attract investment in digital systems designed for inclusion and development. In February 2024, Sahamati, a market-led industry alliance for the AA ecosystem, shared disaggregated impact estimates for the AA Framework with TQH. These estimates included data from an AA-enabled lending insurer, showing a 7x increase in the proportion of low-income consumers[75] and a 2x increase in self-employed consumers between FY 2017-18 and FY 2022-23. Such evidence of positive impacts on inclusion can strengthen trust in DPIs and encourage greater investment. Additionally, an open data policy that identifies gaps in access or efficiency enables timely remediation, further bolstering trust and driving both innovation and investment.

Further, as India gears to take its DPIs to international markets and expand their presence domestically, the need to substantiate their impact with open data becomes more relevant than ever. A track record demonstrating DPIs' positive influence can instil global confidence in India's push for scaling DPI(s). This can assist wider adoption by countries and increase investments by international financial institutions such as The World Bank and IMF.

That said, our consultation with Sahamati revealed a cautious approach to data collection and sharing, emphasising a need to carefully assess what data gets collected and made publicly accessible. Uniform security and privacy standards are crucial to the success of any potential open data policy designed for the DPI ecosystem. In the following section, we explore guiding principles for developing such standards.

3.1 Way Forward

Participation from DPI developers and regulators is crucial to move towards greater sharing of data. They can encourage collection of user journey, disaggregated, granular, panel data, and facilitate greater access. However, it is also pertinent to acknowledge concerns related to privacy, data protection, and unrestricted use that accompany the demand for more data. While our analysis indicates that an open data policy for DPIs does not require publication of personal data, it is important to establish a principles-based approach to guide secure and trusted access. Through our secondary research and stakeholder consultations, we have identified the following key principles to guide collection, release, and reuse of DPI-generated data.

---

[74] https://odimpact.org/files/case-indias-esmi.pdf
[75] Income is less than 5 lakh per annum

Key Data Sharing Principles for the DPI Ecosystem

Open and Collaborative Development of new DPIs

To ensure that future DPIs are built with responsible data collection and transparent data sharing mechanisms from their inception, it is imperative to incorporate inclusive design principles and extensive consultations at the design stage. By engaging policymakers, technologists, researchers, civil society, and end-users early on, DPI developers can identify key data needs and ensure that data collection aligns with ethical standards. These inclusive discussions would foster trust, accountability, and user-centricity, leading to DPIs that reflect the diverse needs and preferences of their stakeholders. Moreover, by integrating inclusive design principles from the outset, DPIs can proactively address privacy concerns, mitigate risks, and optimize impact assessments. Embracing stakeholder consultations facilitates iterative improvements, allowing DPIs to adapt to evolving challenges and user requirements. Ultimately, this collaborative approach not only enhances the effectiveness of DPIs, but also fosters innovation and maximizes their potential for positive societal impact.

Augmenting Data Collection for Impact Assessment

Some data collection and reporting practices of existing DPIs fail to include key parameters that are essential for comprehensively estimating their impact. This gap is highlighted by the unavailability of specific data sets: we note the absence of data on transaction failure data in Aadhaar and the lack of detailed transaction and user data, broken down by gender, location (whether urban or rural), and income group for UPI. The lack of transaction failure data for Aadhaar makes it difficult to accurately measure the effectiveness of Aadhaar-based services. Similarly, the absence of detailed user demographics for UPI hampers our ability to gain insights, especially those that consider the interplay between different social and economic factors. Thus, it is pertinent for the governing institutions of existing DPIs to identify such indicators via engagements with researchers, service providers and the industry to promote the collection and publication of data on the identified parameters. For example, Brazil's Pix payment service provides data about the transactions settled by municipality, by individuals and legal entities, highlighting the perspective of the payer and the recipient.[76]

In 2016, the Competition and Markets Authority (CMA) in the United Kingdom published a report on their investigation into competition and innovation in the UK's retail banking industry.[77] The report found that established banks don't have to compete hard enough to gain customers' business, while newer banks find it difficult to access the market and grow.[78] One of the CMA's recommendations to tackle this lack of competition was open banking. The CMA and UK government mandated nine of the largest banks to implement common standards for open banking and tasked OBL with the charge of implementing its order. To ensure secure data sharing amongst these banks, OBL prescribed and implemented common standards related to open APIs, data formats, authorisation and authentication standards, and standardised permission

[76] Banco Central do Brasil. (2023, June 2). Estatísticas do Pix - Transações Pix por Município. Portal de Dados Abertos do Banco Central do Brasil. https://dadosabertos.bcb.gov.br/dataset/pix/resource/42e0c55a-ab4e-4f9a-88f1-c5893df8d47b

[77] Open Banking. 'Regulatory.' Open Banking. https://www.openbanking.org.uk/regulatory/

[78] Open Banking Limited. 'About Open Banking Limited.' Open Banking. https://www.openbanking.org.uk/about-us/

frameworks.[79] As a result, the number of open banking users reached 8.03 million by December 2023, with the Application Programming Interfaces (APIs)[80] designed to facilitate data sharing across banks achieving a success rate of 99.45%.[81]

Accessible Data Release

DPI-led digitalisation has helped capture and generate a wide range of data at scale. Trends in such data can highlight problems, support research and analysis, aid impact assessment, prompt discovery of new opportunities and enable evidence-based decision-making for the future. However, for such benefits to accrue, it is crucial for the data to be made publicly available.

The National Data Sharing and Accessibility Policy, 2012 clearly states that non-personal data generated using public funds should be made accessible to the public for scientific, economic and development purposes.[82] Furthermore, the 2020 Report by the Committee of Experts on Non-Personal Data Governance Framework, recommends identification, creation, and sharing of datasets that serve public good purpose and benefit the society at large as "high-value datasets".[83] The reports suggest that data that can be useful for policymaking, improving public service, creating new jobs and businesses, and help in research and education, should be classified as high-value datasets.

Globally, several countries have taken a strong stride towards making such data accessible. For example, established under Finland's Act on the Secondary Use of Health and Social Data[84], the Findata platform hosts social and health care sector data from 11 Finnish authorities[85] at a single place. For the European Union's TARGET Instant Payment Settlement, the European Central Bank hosts country wise datasets based on the value and number of credit transfers processed via TIPS on its data portal.[86] These initiatives spotlight the power of accessible data repositories in fostering collaboration, innovation, and evidence-based decision-making.

Preserving Privacy

Privacy is a paramount concern in the realm of data sharing, with significant implications for individuals' personal data rights, especially as we recommend greater data sharing by DPIs. Onboarding users on DPIs like Aadhaar, UPI, ABDM, et al. involves the collection of sensitive personal data such as biometric details, mobile numbers, dates of birth, residence details, and bank account information. The unique user IDs are inextricably linked to sensitive personal data, necessitating utmost priority on preserving individual privacy and data protection during both data collection and publishing. To safeguard privacy in data-sharing initiatives, several key principles

[79] Competition and Markets Authority. (2017, February 2). Retail Banking Market Investigation Order 2017. https://assets.publishing.service.gov.uk/media/5893063bed915d06e1000000/retail-banking-market-investigation-order-2017.pdf

[80] The term "Application Programming Interface" (API) means any mechanism that allows a system or service to access data or functionality provided by another system or service. The API is generally used to interact (like query, list, search, sometimes submit & update) directly with the specific information on a system, to trigger some action on other systems, or to perform some other action on other systems.

[81] Open Banking. API Performance Stats. Open Banking. https://www.openbanking.org.uk/api-performance/

[82] Department of Science & Technology, Government of India. (2012). National Data Sharing and Accessibility Policy (NDSAP) 2012. The Gazette of India. https://dst.gov.in/sites/default/files/nsdi_gazette_0.pdf.

[83] Gopalakrishnan, Kris, et al. (2020, December). Revised Report on Non-Personal Data Governance Framework. https://ourgovdotin.files.wordpress.com/2020/12/revised-report-kris-gopalakrishnan-committee-report-on-non-personal-data-governance-framework.pdf

[84] Ministry of Social Affairs and Health Finland. (2019). The Act on the Secondary Use of Health and Social Data. https://stm.fi/documents/1271139/1365571/The+Act+on+the+Secondary+Use+of+Health+and+Social+Data/a2bca08c-d067-3e54-45d1-18096de0ed76/The+Act+on+the+Secondary+Use+of+Health+and+Social+Data.pdf.

[85] Findata. Data. Finnish Social and Health Data Permit Authority. https://findata.fi/en/data/#what-data-are-available-via-findata.

[86] European Central Bank. Search Results. European Central Bank Data Portal. https://data.ecb.europa.eu/search-results?searchTerm=TIPS

can be implemented. Firstly, adopting a tiered access system, as outlined in the National Data Sharing and Accessibility Policy 2012, which categorizes datasets into open access, registered access, and restricted access tiers based on their public good potential and privacy risks. Secondly, employing advanced anonymization techniques like tokenization and generalization before releasing datasets can help protect individual identities and minimize unauthorized access and misuse. Thirdly, for high-risk datasets requiring registration, defining the purpose for data usage could ensure compliance with privacy regulations, exemplified by Finland's Findata platform[87]. The Findata platform that provides social and health care sector data requires for a "brief description of the intended use for the data" while making the application to obtain a permit to use the data. Additionally, compliance with the Digital Personal Data Protection Act, 2023 can be enforced through submission of statements affirming adherence to privacy principles.

3.2 Conclusion

This report's recommendations underscore the critical importance of publishing data about India's DPIs to substantiate their impact and foster innovation, investment, and policy choices. While acknowledging the need to address concerns regarding privacy and data protection, it is evident that a principled approach to data sharing is essential for fostering trust and enabling wider adoption of DPIs. By promoting greater transparency, accessibility, and accountability in data sharing practices, DPI developers and regulators can enhance the effectiveness of impact assessments and drive positive societal outcomes.

Immediate steps can be undertaken to expedite this process. **DPIs developers and regulators can initiate a consultation process with researchers and other relevant stakeholders to understand specific data requirements.** By opening a call for researchers to request needed data, DPIs can then prioritize the publication of requested data as per feasibility. Additionally, conducting public consultations and actively incorporating public feedback will encourage greater participation and align the data sharing strategy with the needs and concerns of a broader audience. Sharing metadata of DPIs would aid in comprehending the nature of collected data, while drafting a policy for data sharing would streamline the process, ensuring that data is released in a standard format and procedure. Ultimately, by embracing inclusive design principles and stakeholder consultations from the outset, future DPIs can be tailored to meet the diverse needs of their stakeholders. This can help maximize their potential for positive impact, and pave the way for a more digitally inclusive society.

---

[87] Finland's Findata platform that provides social and health care sector data requires for a *"brief description of the intended use for the data"* while making the application to obtain a permit to use the data.

This page has been intentionally left blank

This page has been intentionally left blank