Data Empowerment And Protection Architecture

Draft for Discussion

A Secure Consent-Based Data Sharing Framework To Accelerate Financial Inclusion



Draft Document For Discussion

All stakeholders are requested to provide comments by 30th November, 2020. Comments may be submitted through e-mail: **annaroy**@**nic.in**

Acknowledgements

I would like to acknowledge the iSPIRT team for their initiative and effort in preparing this paper. Special mention is warranted for Siddharth Shetty and Kamya Chandra. Others who contributed towards this paper include Arnab Kumar, Pritish Mishra, Ankan De, Aaryaman Vir, Gayatri VS, and Ayna Agarwal. The report was designed by a team at Iconscout, led by Dalpat Prajapati, Jemis Mali, and Falak Mehta.

This paper reflects the on-ground implementation of the Data Empowerment and Protection Architecture (DEPA) framework set to launch in 2020, and thus would not have been possible without the key players orchestrating rollout. This includes a number of departments of the Government of India (including the four major financial sector regulators (the Reserve Bank of India (RBI), Securities & Exchanges Board of India (SEBI), Provident Fund Regulatory & Development Agency (PFRDA), and Insurance Regulatory and Development Agency India (IRDAI)), the Ministry of Finance (including the Department of Revenue, the Department of Economic Affairs, the Department of Financial Services, and the Financial Sector Development Committee), the Ministry of Health and Family Welfare, the National Health Authority, the Ministry of Information Technology (MeiTY), and the Telecom Regulatory Authority of India. The list also includes representatives from non profit organisations (such as iSPIRT Foundation, DICE India, Sahamati, and CredAll); individual thought leaders on financial inclusion, data, and privacy (including Nandan Nilekani, Justice Srikrishna, Arundathi Bhatacharya, and Rahul Matthan); and key financial sector market players (including the top leadership of State Bank of India, IDFC First, HDFC Bank, ICICI Bank, IndusInd Bank, Axis Bank, and Kotak Bank amongst others). DEPA is truly an ecosystem-wide, joint public-private effort for a new and improved data governance approach.

This draft is intended to be a dynamic document that continues a vibrant discourse. This paper invites actionable recommendations from individuals and institutions who are passionate about refining and co-creating DEPA as it evolves.

Anna Roy Senior Adviser NITI Aayog

Foreword

In an evolving and fast paced digital landscape, headlines world-over have squarely placed data protection, privacy, and unauthorised data sharing or misuse in the limelight. Yet this lens is incomplete. In India, we not only need stronger data protection, but also data empowerment: everyday Indians need control over their own personal data to improve their lives. They should be able to leverage their digital history to access growth opportunities offered by different institutions. Imagine if a small business could use business invoices submitted to GST to digitally prove capacity to repay a working capital loan, and thus access cheaper credit.

With the Data Empowerment and Protection Architecture, India will be taking a historic step towards empowering individuals with control over their personal data, by operationalising an evolvable regulatory, institutional, and technology design for secure data sharing. Just as the launch of UPI transformed India's digital payments world irreversibly, it is expected that the RBI-driven Account Aggregator (AA) model will transform the way financial services are delivered through a unique architecture for consent-based data sharing. In the AA model, individuals can seamlessly share their financial data for the first time across banks, insurers, investors, tax collectors, and pension funds in a safe, secure, and consented manner. This has the power to transform the availability and affordability of financial products. Beyond the financial sector, DEPA also presents opportunities in health, jobs, and urban data.

This has become more exigent in a post-COVID world. Small businesses across the country urgently need a suite of financial products providing working capital support, and lending models that did not function effectively before the crisis will not serve us today. Solving these challenges at an infrastructure level -- asking what we can unlock to change the business and operating model for all financial and technology institutions in this time -- is the only way to achieve change at scale amidst a crisis. The recently announced Open Credit Enablement Network will also leverage AA infrastructure to democratise access to credit.

DEPA builds the right infrastructure. It inverts the traditional Western model where data is simply used to advertise and sell products, to one where data can be used to empower a billion Indians. It can show a new India Way on data governance that allows us to offer inclusive and affordable financial products that help businesses recover from the crisis and chart a path towards sustainable growth.

Amitabh Kant CEO, NITI Aayog

Executive Summary

India's Data Empowerment and Protection Architecture (DEPA) is predicated on the notion that individuals should have control over how their personal data is used and shared. It is designed with the belief that agency over data could empower Indians with opportunities to improve their own lives.

Today millions of Indians are creating electronic transaction histories and becoming 'data-rich' at historic rates, even before becoming economically rich or even financially stable. Personal data helps people inform and build trust with key institutions providing life-altering services, such as hospitals, banks, or future employers. Knowing this, it is unreasonable not to give individuals agency over their data. DEPA is founded on the premise that individuals themselves are the best judges of the 'right' uses of their personal data, rather than competing institutional interests. They should not struggle to access and share their data.

Orchestrating a paradigm shift to empower individuals with their data requires three key building blocks: enabling **regulations**, cutting edge **technology standards**, and new types of public and private **organisations** with incentives closely aligned to those of individuals. DEPA seeks to provide a foundation for all three in India. It will not be a static policy or product; instead, **DEPA is designed as an evolvable and agile framework** for good data governance, given the rapid pace of change in this arena.

In a nutshell, DEPA empowers people to seamlessly and securely access their data and share it with third party institutions. A new type of private Consent Manager institution ensures that individuals can provide consent as per an innovative digital standard for every granular piece of data shared securely (using newly created standard APIs). These Consent Managers should also work to protect your data rights. This architecture replaces costly and cumbersome data access and sharing practices that disempower individuals, such as bulk printout notarisation and physical submission, screen scraping, username/password sharing, and terms and conditions forms providing blanket consent. DEPA combines public digital infrastructure and private market-led innovation: it creates a competitive ecosystem where any new Consent Manager can plug in to a network of information providers and users without setting up expensive, duplicative, and exclusive bilateral data sharing rails. And it ensures that data sharing occurs by default with granular, revocable, auditable, and secure consent. Consent managers can compete to reach different customer segments with accessible and inclusive modes of obtaining informed consent. They can also experiment with different business models. While consent cannot be the only backstop, it is a powerful first step to empowering individuals with data.



The Data Empowerment and Protection Architecture

DEPA needs to be flexibly applied to various sectors, and in each context led by institutions who tailor its implementation. Its first application is in the financial sector towards greater financial inclusion and economic growth. Even pre-COVID-19, 92% of small businesses in India lacked access to formal credit. Consented data sharing can reduce the cost and risk premium of offering loans to small entrepreneurs, by creating frictionless and secure access to data used to establish creditworthiness with individual consent. Most such loans today are offered based on collateral. Instead, offering short term working capital loans based on evidence of past turnover (eg through GST) that indicate a future capacity to repay (referred to as Cash Flow-based lending in the seminal RBI MSME Committee Report) is critical to solving the 20-25 trillion rupee credit gap faced by MSMEs in the country. Using DEPA, individuals and small businesses can use their digital footprints to access not just affordable loans, but also insurance, savings, and better financial management products. DEPA makes this possible only together with the other layers of India Stack built since 2010 (eg Aadhar, Aadhar based eKYC and Aadhaar based eSign for digital contracts; UPI for cashless payments; DigiLocker, etc.); and Open Credit Enablement Network for lending. DEPA marks another step in a decade-long journey building digital infrastructure designed to improve private service delivery.

Digital Footprints

Virtuous Cycle for an MSME Entrepreneur

DEPA is going **live in the financial sector** in 2020 under the joint leadership of the Ministry of Finance, RBI, PFRDA, IRDAI, and SEBI.

RBI issued a Master Directive creating Consent Managers in the financial sector to be known as Account Aggregators (AAs), and seven AAs have already received in-principle regulatory licenes. Entrepreneurial energy has been building in incumbent and new market participants, who are now innovating to compete on new roles, products, and services. A July 2020 AA Hackathon attracted over 1250+ applicants. A newly created non-profit collective of Account Aggregators - the DigiSahamati Foundation (known as 'Sahamati') - is mobilising support to existing financial institutions to adopt the technical standards. They are also establishing open data governance and legal working groups to innovate on the technology architecture to further protect data rights and drive empowerment - those keen to shape the space are encouraged to join. DEPA is also being piloted in the health sector in 2020: On August 15, Prime Minister Modi announced the National Digital Health Mission, which includes a Health ID and a data sharing framework for personal health records. This is based on the National Digital Health Blueprint (July 2019) published by the Ministry of Health which in turn builds on the National Health Stack Strategy paper published by NITI Aayog in July 2018. DEPA is also being launched in the telecom sector following a TRAI consultation report on privacy released in July 2018 and a workshop held and a by TRAI Chairman RS Sharma in August 2020 with major industry players announcing the partnership allowing telcos to become financial information providers and users in AA. The first major government department to become a Government Information Provider (GIP) will be GST; future departments with data on individuals and MSMEs could adopt the specifications to improve the ease of doing business or create greater data portability of individual education, jobs, or transaction data.

DEPA, together with other layers of India Stack, could do for India's data ecosystem what the TCP/IP Internet protocol or GPS - both powerful examples of American public digital infrastructure - did for communication and navigation respectively: introduce a new possibility that creates a Cambrian explosion of novel products and services that empower people. Breaking data silos and monopolies allows fintech or healthtech companies to compete on product design, analytics, and value creation, rather than data access, and simultaneously furthers objectives like financial inclusion which increase the total addressable market for all. Based on the Personal Data Protection Bill 2019 and the planned Data Protection Authority, DEPA is on the road to be applied in other sectors. This could empower individuals with not just financial and healthcare data, but also telecom, educational, or jobs data to better improve access to opportunities. DEPA is a new Indian model of data governance that can be shared with the world -- one that is evolving, and targets individual empowerment, economic recovery and growth, and a competitive data democracy.

The Data Empowerment and Protection Architecture



Financial Exclusion

A well designed pool of financial products is instrumental in pulling individuals out of poverty traps and stimulating the growth of micro, small, and medium enterprises (MSMEs). Yet currently, most of India's rural and urban poor population faces significant exclusion from accessing appropriate financial products for themselves and their enterprises – partly due to a lack of trust and asymmetry of data.

Financial **products** can enable increased **prosperity** (through savings and investment), greater security and **resilience** to income or health shocks (through insurance), and new **aspirations** (through credit for business or learning opportunities). However, India's poor **struggle to access** appropriately sized, priced, and timed financial products. For instance, many micro entrepreneurs or small businesses struggle to get short term working capital loans to cover liquidity shortfalls in running their livelihood businesses.

This is primarily due to the **high costs** formal financial institutions face in **reposing trust** in individuals or businesses with a largely **undocumented financial background**, and thus no digital trail to reference. Read more>>



02 Digital Opportunity

An opportunity is emerging: The lower socioeconomic strata in India has been rapidly accessing and adopting digital services over the last decade.

Across platforms such as **Aadhaar** (offering unique digital identification) and the **Unified Payments Interface (UPI)** (a mobile-based digital payments system based on a common technology standard), as well as through increasing **mobile connectivity**, tele-density and **internet usage**, members of the lower socioeconomic strata are for the first time becoming **data rich even before becoming economically wealthier.**

Small shop owners, farmers, traders, MSME entrepreneurs, rural Self Help Groups, and gig economy workers are increasingly generating a digital transaction history that could be use to inform and **build trust** with financial institutions. For instance, the total number of registered businesses under the previous tax regime was around 6.5 million (FY 2015-16), while in 2020 the number of formally registered businesses filing invoices and returns is around 10 million. A **survey** of 2700 MSMEs across 20 industries highlighted that over 60% of MSME owner respondents were digital users. Read more>>



03 **Data Silos**

Despite increasing digitisation and the tremendous value data could have for individuals to build trust with institutions, personal data (and particularly financial data) continues to remain in silos today. The custodian-centric data sharing model will struggle to scale to address Indians' emerging data access needs.

In a world where an **exponentially increasing** number of companies and institutions control an individual's data as custodians or fiduciaries, going to each actor individually to access and move data interoperably across data users is a model that will not scale. Gathering your own data directly from various financial institutions, for instance, is a **cumbersome** task - typically involving physical branch visits or call center engagements, sharing **physical documents** using browser uploads or USB sticks, or **screen scraping** and sharing of confidential username and password data with a third party. For those living in rural or semi-urban areas, the challenges around data access are further exacerbated.

Moreover, data is stored in different formats and porting specific data (proportional to the need) from one database to another service provider is not a standardised process. These issues, described here in the context of personal data, also apply to other forms of data such as derived data (eg. credit scores) and public or anonymized data. Finally, there is a **lack of harmonisation** around the regulations for data sharing within and across sectors. These factors together mean that **individuals and small businesses lack control** over their own data. Read more>>



Risks of Inaction

Unless an evolvable, interoperable, and secure data sharing framework is implemented, newly generated data on Indians will at best remain in silos without benefiting individuals who urgently require it to access better services, and at worst be misused without individuals' knowledge and consent.

In today's world, personal data is used to create **deep profiles**, **walled gardens**, or **barriers to exit** due to business needs of large corporations controlling user data. Making it simple and secure to share this data with the individuals' consent would empower them to use data to improve their well-being themselves (via ease of access to new financial products and services), or to contribute data to research and better-designed machine learning models that benefit them. However, this is only possible if action is taken to ensure ease of data flows between siloed data fiduciaries housing information (e.g. different banks, NBFCs, insurance companies, government departments, etc.) with user consent. Moreover, incidents such as the Cambridge Analytica data breach have highlighted global gaps in data sharing and consent approaches - and with an increasing penetration of public digital services in India, risks such as **data farming** and other malpractice that arise from a lack of a robust and privacy protecting data sharing framework will grow exponentially over time. Read more>>



05 **A Global Challenge**

Other countries have responded to these challenges by implementing efforts to improve data protection and consent-based sharing (such as Open Banking in the UK or **General Data Protection Regulation** (GDPR) in the EU), which India can learn from. However, these approaches have not addressed the issue in a manner that is fully relevant to India's scale and diversity, and to our objectives around accelerating financial inclusion, economic growth, and data democracy.

The United States to date does not have a nation-wide data protection law in place. **The EU** has opted for strong data protection laws (through policies encouraging the right to be forgotten and data minimisation), but fundamentally approaches the issue from a prevention-of-harm perspective rather than through the lens of individual empowerment through data.

And although **Open Banking** exists as a regulatory mandate in the UK, it has not been implemented at scale partly due to misaligned incentives between market forces and regulatory authorities, as well as the lack of a shared technology architecture adopted by banks. Some learnings from these approaches on strong data protection ought to be captured in India's Personal Data Protection Bill, but replicating their data sharing strategies would not go far enough to achieve India's objectives in our national context: that of individual **empowerment** and financial inclusion through data, of encouraging a vibrant and competitive **data democracy**, and of building an environment for small and large **businesses to thrive** based on legitimate and high value **use cases** for data sharing that ultimately help individuals and MSMEs prosper. Read more>>

A Paradigm Shift towards Data Empowerment

India needs a paradigm shift in personal data management that transforms the current organisation-centric data sharing system to an individual centric approach that promotes user control on data sharing for empowerment.



The problem is not that companies are benefiting from individuals' data; the problem is that **individuals and small firms do not benefit**. The mission of the Data Empowerment and Protection Architecture is therefore to provide individuals and small businesses with the practical means to **access, control, and selectively share personal data** that they have stored across multiple institutional datasets – to maximise the benefits of data sharing for individual empowerment whilst minimising **privacy** risks and data misuse. By giving people the power to decide how their data can be used, DEPA enables an individual to control the flow of and benefit from the value of her personal data, relying on not only institutional data protection measures but also restoring **individual agency** over data use. Read more>>

An Evolving DEPA Framework

The Data Empowerment and Protection Architecture (DEPA) is a strategy for data empowerment towards economic well being for all. Based on an underlying legal and regulatory framework, DEPA introduces new types of institutions, and cutting edge and evolvable technological building blocks to enable true data empowerment.

Legal, market, and technology infrastructure are all needed to bring this to life, with a 'jugalbandi' of public and private players: governments offering digital infrastructure as a public good, and private players innovating on it to offer better services. Moreover, because the generation of vast amounts of data and its related storage, analytics, sharing, and overall management norms are rapidly emerging fields, our legal and policy framework, technology architecture, and institutional data governance will need to **dynamically change** over time to meet new and emerging needs.

The DEPA framework is not a static policy, product, or infrastructure; rather, it is an **evolvable program** that offers a process and structure for an evolution of data policy by building a **dynamic technology foundation** based on shared standards that can be upgraded over time, and **institutional arrangements** which realign incentives and empower experts who care deeply about building next generation data governance to co-create the future. DEPA can be adopted sector by sector based on potential value add and readiness. Read more>>



08 **Regulatory Foundation**



cial institutions required for DEPA's application in the financial sector was provided through a Supreme Court Judgement on the fundamental **Right to Privacy** (Aug 2017), the **Personal Data Protection Bill** (PDP) 2019 and its precursor the Justice Srikrishna Committee Report, and (for the financial sector) the **RBI Master Direction on NBFC-Account Aggregators** of September 2016. In the financial sector, four regulators across banking, securities, insurance, and pensions (RBI, SEBI, IRDAI, PFRDA) and the Ministry of Finance have come together to implement this model. This regulatory foundation is also expected to evolve with time (eg. with the forthcoming Data Protection Authority) as India's experience and public discourse around data protection and sharing grows richer. Regulation for data empowerment likely needs to be sector-specific, so TRAI and policymakers in spaces such as health and urban who have also indicated intent to adopt will devise a sector-specific regulatory architecture under the aegis of the PDP and Privacy Bills. Read more>>



and many more...

A New Class of Institutions

DEPA's Institutional Architecture will involve the creation of new market players whose incentives align more closely with individuals: user Consent Managers These Consent Managers in the financial sector will be known as Account Aggregators, and a non-profit collective or alliance of these players will be created called the DigiSahamati Foundation ('Sahamati').



The PDP Bill introduces the concept of "consent managers" to manage a data principal's **consent for data sharing** through an accessible, transparent and interoperable platform. These consent managers are **'data blind'** and will not see or use personal data themselves; rather they will serve as a conduit for encrypted data flows. In the future they could also help individuals and small businesses **protect and enforce their data rights**. Consent Managers in the financial sector will be known as **Account Aggregators (AAs)**. A non-profit collective for the AA Ecosystem called **Sahamati** will provide procedural and best practice guidelines for all participating institutions, support organisations to adopt and go live, and continue to foster innovation in protecting data rights across the AA network through new shared technology building blocks. Read more>>

10 Technology Foundation

DEPA's technology architecture is a first of its kind interoperable, secure, and privacy preserving digital framework for data sharing through

The Consent Artefact: is a technology Standard for programmable consent to replace the all-permissive terms and conditions forms. The consent individuals provide is designed on principles acronymed ORGANS: Open standards (ensuring all institutions use the same approach interoperably); Revocable (by individuals); Granular (provided for each time you share data, stipulates how long data can be accessed, etc.); Auditable (in machine readable logs of consent provided), provide Notice to all parties, and Secure by design.

	TERMS AND CONDITIONS		
Before →		After →	Share Data X with ABC untill X?
	І АССЕРТ		NO YES

- Open APIs for Data Sharing: allow many new consent managers to 'plug in' to a common sharing system rather than having to build bilateral relationships with information providers to access data.
- **3** Financial Information Standards: allow a data recipient to quickly interpret and understand information from a new institution.



These are the first key building blocks of the technology framework; further elements (for instance, tools to prevent over-consent or a lack of informed consent) are evolving under the leadership of Sahamati and will be implemented within and across sectors as the market develops. Since data security and protection is a critical prerequisite for empowerment, DEPA also relies on the adoption of related standards for data storage and processing techniques. Read more>>

Guiding Principles

DEPA's model and architecture choices are guided by a set of key design principles. Because DEPA is an evolving framework, these principles are intended to steer future technology or institutional decisions.

These principles are: restoring individual agency; promoting informed consent for every data transaction (rather than blanket consent for data use); building in accountability for institutional data controllers (i.e. consent as not the only backstop); building an open infrastructure for data sharing (minimising bilateral or closed-loop networks); building incentive alignment between new public or private institutions and the needs of individuals around their data; ensuring accessibility and affordability of data sharing; remaining technology agnostic (through open standards); supporting data minimisation; ensuring reciprocity of data use and data provision (institutions cannot be users of data in the system without also being providers); enabling other key data rights; ensuring evolvability of technology and institutions by design; and using penalties as deterrents to data misuse where required. New institutional and market players will continue to bring these principles to life through their innovations as the DEPA framework matures. Those interested in helping DEPA evolve as an ecosystem could join the Sahamati **Data Governance Working Group.** Read more>>



Combinatorial Layered Innovation

DEPA as a layer of secure digital data sharing through consent forms the final layer of India Stack – a series of digital public goods designed to enable private market innovators to improve digital services for India across a range of sectors.

The other key layers of India Stack include unique and digitally verifiable proof of **identity** (Aadhaar, launched 2010), a low cost and interoperable mobile digital payments **platform** (the Unified Payments Interface, launched 2016). Read more>>

IDENTITY LAYER	Giving every resident a unique id and enabling them to prove"I am who I clain to be"	 Aadhaar eKYC eSign
PAYMENTS LAYER	Enabling Interoperable, Instant & cheap payments	 Unified Payments Interface Aadhaar Payments Bridge Aadhaar Enabled Payment Service
DATA EMPOWERMENT	To enable secure sharing of data	DigiLockerConsent ArtefactAccount Aggregator
	2009 2011 2013	2015 2017 2019



Impact on Kirana Storeowner

DEPA, India Stack, and other digital public goods such as the Open Credit Enablement Network (OCEN) and the Public Credit Registry could change the life of a small business owner through new and tailored financial products, such as cash flow based lending.

Even pre-COVID, only about 8% of the total MSMEs in the country had access to formal finance; the other 92% are likely taking loans at onerous terms from ad-hoc sources, and are regularly facing **working capital shortages**. these small businesses are increasingly transacting digitally. If portability and control of data could allow an MSME owner to digitally share proof of the business' regular historic tax (GST) payments or receivables invoices easily, a bank could design and offer regular small ticket working capital loans based on demonstrated ability to repay (known as Flow based lending) rather than only offering bank loans backed by assets or collateral. Flow based lending is the norm for individuals providing proof of salary to access home and car loans, yet these types of products are yet to take off at scale for MSMEs, partly due to frictions in accessing required data. The Account Aggregator framework could transform access to much-needed working capital credit for micro enterprises, particularly when bundled with **OCEN APIs** for Lending. Similarly, DEPA could also enable better personal financial management services, wealth management, robo advisory, or different types of lending, insurance, and investment use cases and products that we may not be able to foresee today. Read more>>



Roadmap

DEPA roll-out has already begun in the financial sector, with a closed user group (CUG) launch by major banks in July 2019 and a public launch expected in Fall 2020. This will be followed by launches in healthcare and telecom.



An RBI Master Directive first created AAs as an NBFC in Sept. 2016, and the RBI MSME Committee recommended implementing AA to facilitate Cash Flow Lending in June 2019. DEPA's CUG launch in July 2019 saw the first demo of consented financial data flows, a commitment of CEOs of major banks & NBFCs to the technical standards, and the introduction of the non-profit Sahamati as a facilitator of adoption. Since the Nov. 2019 publication of AA technical standards, seven AAs received in-principle approval from RBI, and 10 major banks and NBFCs are in different stages of integration working towards a public launch in Fall 2020. Finally, adoption of the DEPA approach is also being planned by other sectors - for instance in **healthcare**, **telecom**, and **skills** data. In health, COVID-19 has re-emphasised the urgency of creating digital infrastructure to share medical data. The National Health Authority is tasked with implementing the National Digital Health Blueprint (including piloting the DEPA architecture for electronic health records) later this year. Read more>>

15 **Co Creation**

Now that the DEPA platform is available as a public good, tremendous entrepreneurial energy has been building in market participants who are leveraging the opportunity to innovate across the various new roles in the ecosystem.

Many different market players can co-create and innovate on this public good: Financial institutions can continue to adopt the public APIs and become financial information providers and users (the new nonprofit Collective of AAs called **Sahamati** is helping with this transition); entrepreneurs and fintechs can start up Account Aggregators catering to diverse users and/ or innovate on new modes of gathering informed consent and protecting data rights; banks, NBFCs, and fintechs across the ecosystem can build innovative products and services to serve previously difficult to access populations (such as cash flow-based lending for micro businesses, improved personal financial decision management apps, etc.), which leverage the new data sharing possibilities. All players in the ecosystem could work to build awareness around informed consent, and continue to contribute to the evolving technology framework for better data protection and empowerment. Moreover, innovation is not restricted to market players: within the government, a Data Protection Authority to enforce data protection legislation can be created as per the Srikrishna Bill recommendations that strives to ensure data is secure and protected. Regulators and other government ministries can consider operationalising consent managers in their respective sectors to empower users with access to data in telecom,



Govt Officials

education, or jobs data. Finally, government departments could adopt a 'Government Information Provider' technology module that allows secure sharing of data with consent of individuals or small businesses; this could significantly improve the ease of accessing government services and the ease of doing business. Read more>>

An "India way" for the World on Data

We are confident DEPA will be a transformative platform that shows a new India model on data protection, sharing, consent and privacy quite distinct from other models around the world.

The India model of data governance is one that is inclusive, sensitive to the needs of the poor, technologically innovative and robust, and continues to drive and stimulate economic and business value. DEPA adoption in any sector is likely to nudge the market forward towards higher quality services: simple access or control of user data will no longer be a source of competitive advantage; institutions will have to create value through better analysis and more sophisticated

predictions based on data, as well as improve accessibility for users. Finally, because the stadards underpinning DEPA are open, the architecture can be applied to other countries - an institutional framework can be designed to globalise this standard and apply it to other markets facing similar challenges. Read more>>

Table of Contents

1. Context: The Emerging Data Revolution and Financial Inclusion		
2. Global Approaches to Data Protection & Sharing	28	
3. Introducing India's Data Empowerment and Protection Architecture	30	
a. Guiding Principles of DEPA	30	
b. DEPA's Legal & Regulatory Framework	33	
c. DEPA Institutional Architecture: Introducing the Consent Manager	34	
Business Models for Consent Managers	36	
d. DEPA's Technology Architecture	38	
Electronic Consent Architecture as a Foundation of DEPA	38	
APIs for Data Sharing	40	
4. DEPA for the Financial Sector: The Account Aggregator Model		
a. Regulatory Framework for Data Sharing in the Financial Sector	42	
b. Market Architecture: Account Aggregators and Sahamati	43	
c. Impact of Data Sharing: The Potential for Cash Flow Lending	44	
d. Impact of Data Sharing: A Cambrian Explosion of Financial Products	46	
e. Implementation and Rollout: Financial Data and Beyond	47	
f. Future Innovation Based on DEPA	48	
5. An Opportunity for the Ecosystem to Co-Creat	50	
Annex: List of Acronyms		

How to Read this Document

Given the complexity of this approach as one encompassing legal, regulatory, institutional, business, and technology domains, topic experts of different functions can choose to focus on specific sections of this document. In addition to the Executive Brief,

- Current and future financial sector players (including banks, NBFCs, and fintechs, as well as new entrepreneurs considering innovating in this space) could focus on Sections III, IV, and V;
- II. Global data sharing and privacy experts or lawyers could focus on I, II, III, and V;
- **III. Policymakers** interested in data sharing for a specific sector or country could focus on I, III (especially a, b, and c), IV (especially a, b, and d), and V.

1. Context: The Emerging Data Revolution and Financial Inclusion

One could use many lenses to describe India's radical shift towards a digital economy. Connectivity has soared: we now have over 687 million internet subscribers (the second largest in the world), up over 300% from circa 200 million just five years ago. The number of mobile users, too, has seen a similar trajectory - today we have 1.2 billion mobile connections, around 600 million of which are unique users - almost double the 349 million unique users five years ago. This connectivity has come hand in glove with access to new public-private platforms and services - especially those essential to financial inclusion. Over a billion people have a unique and immediately verifiable digital identity in the form of Aadhaar. At least 647 million individuals have a formal bank account; As per World Bank estimates, over half the total accounts opened in the world between 2014-17 were in India. Individual transactions are increasingly cashless and digital: after just 27 months of launch, over 1.3 billion transactions take place per month over UPI, a seamless digital payments platform that enables tech players like Google, PhonePe, or Whatsapp to become payment service providers on a shared interoperable public payment rails. A survey of 2700 micro, small, and medium enterprises (MSMEs) across 20 industries highlighted that over 60% of business owner respondents were digital users. The total number of registered businesses as of 2016 was around 6.5 million, while in 2020 the number of formally registered businesses filing invoices and returns into the digital indirect tax system is around 10 million. India has always been a nation of large numbers. But it is the rapid scale of change of these numbers, as well as the increasing penetration of services into traditionally vulnerable groups that seemed far from it – the urban and rural poor populations for instance – that merits attention.

India is going digital, and fast.

With an increase digitisation different data types are created. Some data generated is **secret personal data** – raw personally identifiable information (PII), biometric or demographic information, KYC data, transaction or browsing logs, etc. Other data is **derived**, meaning raw data is manipulated or analysed by a company's proprietary algorithms, indexes, or models to generate useful information (for instance, a credit score). Some combination of personal and derived data is also sometimes **anonymised** and converted to publicly available datasets.

Particularly for personal and derived information, the most recent global push has been towards **data security and protection**. Incidents such as the Cambridge Analytica scandal

have clearly highlighted the need to protect data - especially personal data - from misuse, unauthorised sharing, and violations of privacy. Data protection is a worthy goal, and one that has firmly taken root in India through the Personal Data Protection Bill. However, while a security oriented lens is necessary, it may not suffice to address the needs in the Indian context. India is a nation where poorer and more vulnerable populations are for the first time becoming data rich. People are generating a digital footprint of activity - even before overcoming poverty to become financially independent and secure. Small shop owners, farmers, traders, MSME entrepreneurs, rural Self Help Groups, and gig economy workers are increasingly generating a digital footprint that could be used, for the first time, to provide evidence that builds trust with institutions. This could also enable better access to services that could meaningfully improve people's lives. A well designed data governance framework for the Indian context would enable, not just secure data protection, but also grant users control over data through a safe and seamless protocol to share data across institutions, leading to individual empowerment and well being.

Uniquely for India, the objectives of strong data governance and financial inclusion are inextricably linked. India has made significant strides in access to financial institutions in recent years - for instance, the Pradhan Mantri Jan Dhan Yojana program launched in 2014 was instrumental in bringing 370 million individuals into the formal banking system. Yet we still need to do more. The next wave of digital financial inclusion in India ought to shift focus from access to institutions to access to formal financial products of the right size, at the right cost, and at the right points in an individual's life. To increase penetration of key financial products like insurance, saving instruments such as mutual funds, provident funds, access to capital markets, pensions and other investment opportunities, it is critical to establish a data sharing framework that gives back user control over data. An individual or small business (referred to as a 'data principal' in the Srikrishna Committee bill) should be able to safely and easily share their digital transaction history held in one system (for instance, with a new loan provider) to easily qualify for cheaper and formal credit.

Opening up an API-based data sharing framework would bring significant innovation by new fintech entities, whose participation the RBI has recognised as a 'potentially transformative force in financial markets'. We have already seen this play out: the layered digital service and open API framework known as India Stack enabling verifiable identity (Aadhaar), eKYC data sharing, and an interoperable Unified Payments Interface (UPI) saw mass adoption across new and existing businesses. This brought individuals into the formal financial system by making it easier to open new accounts and conduct cashless transactions. Fintechs leveraged IndiaStack to provide financial services at a reduced cost, increased trust and greater convenience. In conjunction with platforms such as the RBI Public Credit Registry, a secure and privacy protecting data sharing framework could accelerate adoption of new types

of more suitable financial products, such as Cash Flow-based Lending for MSMEs. This is an alternative to asset backed loans; it focuses instead on actual revenue generation capability and creates a flexible repayment schedule based on incoming cash flows.

India's current data governance approach would not scale to achieve these outcomes around financial well being. Like many other countries around the world, India today has a data fiduciary-centric model, meaning individuals or small businesses must only go to the original custodian of data (referred to as a data fiduciary in the Srikrishna Bill) to access or share information. Even as India becomes increasingly digital, based on current trends newly created data will stay locked in the systems of large data fiduciaries who provide most key services and store transactions. These large data fiduciaries are typically major public and private banks or insurers (for example SBI, LIC, or others), ecommerce players (for example Flipkart or Amazon) or technology companies (for example Google Pay. PhonePe, PayTM, UrbanClap, Uber, Ola, Whatsapp, or others).

As an increasing number of applications or systems start to hold our digital information, needing to go to each data fiduciary individually to access or share data becomes a lengthy and tedious exercise. Data is stored in different formats and porting specific data from databases to share with another another service provider is not a standardised process, so **individuals today are compelled to rely on a patchwork of workaround solutions to access data**. Often providing account usernames and passwords to third-party providers to scrape online data (known as screen-scraping), navigating lengthy phone operators only to receive a physical document of your records which needs to be notarised or shared in person with the third party, using an email attachment/USB stick/browser upload, or authenticating through an access delegation (for example oAuth).

The rarity of APIs means that **consumers currently have limited control over their precise information**. Data is usually shared in all or nothing form (sometimes without user permission) with limited options for granularity in a data sharing request. Thus, the current technology framework for data sharing is not designed to scale, nor does it effectively protect privacy. Looking ahead, if India's data governance framework focuses solely on increasing protection and at the cost of enabling secure and granular data sharing only with user consent, we will start to further entrench the data silos controlled by large data fiduciary companies, effectively allowing them to use our data in their competitive interests rather than ours. This comes at the cost of data disenfranchisement and continued financial exclusion.

The world of data use and sharing is a new and emerging landscape, and one which has only really taken shape in the last 5-7 years. To build a data governance framework that can handle new and emerging needs, India will need an **evolvable regulatory, institutional, and** **technological framework** to enable user-controlled and secure data sharing, rather than a static product or policy. For data sharing to evolve to support emerging needs, the framework needs to engage the right experts and stakeholder communities on data sharing policy and regulation through institutions, upgrade APIs appropriately, and move forward on a continuum as technology evolves. *DEPA is thus an institutional arrangement, a process, and a structure for that evolution to happen rather than a baked set of specifications written in stone.* Moreover, the framework will need to enable a non-uniform and scalable set of solutions for all, ranging from the third of Indians with smartphones and to the third Indians without mobile phones, to enable all residents to improve their financial well-being whilst protecting their right to privacy. Without an appropriate data governance framework, at best we will continue to live in a world with data silos where business interests drive the uses of personal data, and individuals lose the ability to use their own data to improve their well being. At worst, we prop up data silos and the power held by companies managing them, and open ourselves up to data farming and unchecked and unauthorised misuse.

2. Global Approaches to Data Protection & Sharing

Data governance is a global challenge, and other nations have mobilised several efforts to improve data security as well as data sharing. On data protection, India can learn from a variety of other nations: the EU's GDPR introduces strong data protection laws (through policies such as the right to be forgotten, and the emphasis on gathering minimum data), while China and some US states also have strong cybersecurity measures (the US Government has no national data protection regulation and has adopted more of a laissez faire approach). The Srikrishna Committee Draft Data Protection Bill incorporates some of the learnings from these approaches, in particular by highlighting the key responsibilities of data fiduciaries (those who store and use your data). However, these strong data protection efforts have come with a variety of spillover effects in each context: in the EU, small and large businesses have complained that GDPR hurts their short and medium term profitability, whilst India would seek to establish a framework that encourages business and economic growth. Meanwhile, China's approach has been to create a tightly controlled internet that prioritises national security over user control, which conflicts with India's objective of building a vibrant data democracy.

Looking beyond data protection to data sharing, the approaches adopted by other nations still may not be fully suited to Indian objectives and context. For instance, Estonia has long advocated for the free flow of data, but their X-Road platform is a single channel for all data sharing flow - and such a model could not scale up effectively to serve the needs of India's billion-strong population. Moreover, these flows are governed based on standard regulations. Organisations decide on appropriate use of data without individual agency over these flows. In a nation of India's size and diversity, it is critical to ensure that a consent-based sharing approach is adopted (building on a framework of institutional accountability and regulation) to restore individual agency in data sharing transactions.

The UK's Open Banking data sharing framework does operate based on consent: it takes a restoration of competition perspective, and mandates that banks work with Account Information Service Providers (AISPs) to gather individual consent to share data. However, it also has a few elements which may not work for the Indian context: there is no unbundling of the institution collecting data and the institution collecting consent, which may not work to address India's scale and diversity. To reach our full population, we will need multiple institutions specialised in consent management innovating to provide multiple modes of obtaining informed consent (for example various form factors - audio, visual or video, or assisted with an agent). Indian institutions will also need to compete with a sustainable business model to prevent data creators from maintaining a monopoly on the data sharing market. Moreover, much of the software behind Open Banking is designed to be web based rather than mobile first - with a low penetration of computers but a higher penetration of mobile in India (smart and feature phones), the latter is a more suitable for much of the population to access to provide consent. And finally, because Open Banking was born out of a competition perspective and made mandatory, its approach does not necessarily encourage fintech innovation and market development.

Other noted data sharing approaches are those of Australia's My Health Record, which has an opt-out system rather than a consent-to-share system for health records and the Australian Consumer Data Right for the banking sector, which will provider users with access to, and the ability to safely transfer, their banking data to trusted parties from July 2020. Similarly, parts of the US have adopted the Blue Button API for data sharing.

As Mary Meeker notes, the US, Europe, and China have all approached internet and data regulation differently: while the US has opted for almost no regulation of the internet, China has opted for a tightly controlled internet. Meanwhile the EU has opted for strong data protection laws but fundamentally approaches the issue from a prevention-of-harm perspective rather than through the lens of user empowerment and individual agency to improve socioeconomic status through data. *This is largely because the EU, like other countries that have established data protection or sharing frameworks, have already had much of their population become economically wealthy prior to becoming 'data-rich'.* Therefore data was primarily used for advertising to target consumption. Although learnings from global data security approaches have been captured in India's draft Data Protection Law, replicating other nations' data sharing strategies would not go far enough to achieve India's objectives: those of individual empowerment and financial inclusion through data, of encouraging a vibrant data democracy, and of building an environment for businesses to thrive based on legitimate and high value use cases for data sharing.

These global trends illustrate that actors worldwide have been grappling with the same data governance issues for the last 5-10 years. Even the technology trend towards areas such as blockchain are rooted in a desire to build greater trust and ownership. The key learning from existing global efforts is clear: strong data governance needs a combination of a legal and regulatory framework, the right institutional arrangements, and a robust technology architecture encompassing both data protection as well as data sharing. India will need to bring all of these elements together to create an evolvable framework that is secure, empowering, and scalable for a diverse population, and suited to a vibrant and diverse democracy.

3. Introducing India's Data Empowerment and Protection Architecture

DEPA rests on the premise that individuals have the right to collect, share and access data pertaining to them in an accessible and easily understandable manner. Based on the consent philosophy codified by the Draft Personal Data Protection Bill, 2018 (described in the regulatory framework section below), the aim is to provide individuals with the practical means to access, share, and use datasets containing their personal information. This includes purchase data, traffic data, telecommunications data, medical records, financial information and data derived from various online services. It is hoped that this will encourage organisations holding personal data to give individuals control over this data extending beyond their minimum legal requirements to do so.

DEPA is a paradigm shift in personal data management and processing that seeks to tranform the current organisation centric data sharing approach to an individual centric system.

By giving people the power to decide how their data can be used, DEPA enables the collection and use of personal data in ways that maximise the benefits gained while minimising the privacy lost; DEPA enables an individual or a small business to control and benefit from the value of their personal data. Moreover, the technology behind DEPA allows organisations to implement data protection and privacy measures, and provides individuals and entrepreneurs the means to share their consent for data sharing and transparently view how their data is collected and processed.

Guiding Principles of DEPA

3

- **Restoring Agency and User Control:** Individuals are empowered actors, not passive targets, in the management of their personal lives (both online and offline) they should have both the right and the practical means to manage their data and privacy.
- 2 Informed consent: Consent is an expression of human autonomy. For such an expression to be genuine, it must be informed and meaningful. Personal data should never be shared without consent.

Institutional and Data Controller Accountability: While customers are in control and can

consent to various uses of their data, individual consent does not absolve institutions holding data (data fiduciaries) of responsibility to protect, manage, and minimise data misuse. They can and will be penalised under governing laws (for example the RBI Act, or the upcoming Personal Data Protection Bill) for misusing data, not taking appropriate measures to ensure data security, and misusing the consent framework.

- Accessibility and Affordability: It is essential that personal data is technically easy to access and use – it is accessible in machine-readable open formats via secure, standardised APIs (Application Programming Interfaces) which can be leveraged by various organisations to present information in a user-friendly and virtual form. DEPA enables the break-down of closed silos enabling personal data to become an important, reusable resource accessible to all with appropriate permissions. Moreover, the objective is to allow for data accessibility and empowerment in a broad and inclusive manner across the population, not just for the wealthiest or the most technologically savvy. This requires market players to innovate on the business model (for instance, through assisted modes of obtaining consent).
- **5** Shared open infrastructure: A shared infrastructure and set of standards enables decentralised management of personal data, and allows interoperability across the many decentralised players (allowing individuals to change service providers without proprietary data lock-ins). It also makes it easier for companies to comply with tightening data protection regulations.
- 6 Incentive alignment: For DEPA to be successful, it is critical that the incentives of individuals are aligned with institutions operationalising their data rights. Under the status quo, data fiduciaries have very different incentives around data use. Therefore, it will be necessary to create new institutions that have incentives more closely aligned with those of individuals, in order to help empower individuals with their data.
- Reciprocity: Although initial market players will want to only be information users rather than providers, for the ecosystem to thrive players will need to be both information providers and users. Therefore, the DEPA market architecture will function on the reciprocity principle; all data user agencies must also adopt the technology standards required to be information providers to ensure sustainability of the ecosystem.
- 8 Technology agnosticism & interoperability: The architecture must be technology agnostic. It must be flexible enough to take into account evolving technologies and standards of compliance. The technical specifications for data flows and consent flows moreover will be agnostic to the kind of data that flows (for example, specifications not particular to a sector or type of data).

- Data minimisation: Data that is processed and shared ought to be minimal and necessary for the purposes for which such data is sought and other compatible purposes beneficial for the data subject.
- **Enabling other data rights:** DEPA ought to make it easier for individual users to operationalise (through market structures and technology tools, for example) the right to know how your data is being used, the right to share only purpose-specific data, and the right to be forgotten or to have your data be deleted. However, this is premised on the existence of a legal framework that calls out the importance of these rights.
- **Evolvability:** The final principle is that of evolvability. Recognising that this more than almost any other area of regulation, governance, or service delivery is an emerging space shaped by rapidly advancing technology possibilities and evolving market dynamics, DEPA's architecture and building blocks must be built to change in order to stay current.

Legal & Regulatory Framework as a Foundation for DEPA

In view of the growing concerns over the misuse and exploitation of personal data, and following on from a historic Supreme Court Judgement in August 2016 declaring privacy as a fundamental right, the Government drafted a Draft Data Protection Bill in 2018 under the Chairmanship of Justice B.N. Srikrishna. The Committee also shared its recommendations in a report titled "A Free and Fair Digital Economy Protecting Privacy, Empowering Indians" ("Srikrishna Committee Report"). The Bill envisages key concepts such as the right to be forgotten, data portability, and anonymisation of data which will be key in enabling digital empowerment in India. The Bill also suggests creation of a Data Protection Authority, to ensure that these rights are upheld by institutions.

Most importantly, the Bill places consent at the bedrock of data sharing, collection, and destruction and advocates for an electronic consent dashboard which would enable data principals to keep track of consent for processing in real time and allow them to operationalise the right accorded to them under the data protection law. Under the law, unless an individual has provided explicit and informed consent no later than at the commencement of the processing, personal data cannot be shared or processed. In addition, the consent must be **free**, **informed**, **specific**, **clear**, and **revocable**. Moreover, a data fiduciary cannot make provision of goods or services or performance of a contract conditional on consent. The Bill also introduces the concept of consent managers. These entities are responsible for managing consent of data principals across multiple fiduciaries through an accessible, transparent and interoperable platform.

As of December 2019, the PDP Bill stood approved by the Union Cabinet and was sent to a Joint Select Committee (comprising of members of both Houses of Parliament). This review was recommended before it could be passed into law. However, even while the overarching data protection bill that applies to all sectors is yet to obtain final approval, there are domain specific laws and regulations that enable the DEPA framework. For instance, the application of DEPA for the financial sector – known as the Account Aggregator framework – is built on the foundation of the Banking Regulation Act of RBI and an RBI Master Directive (an official notice).

DEPA's Institutional Architecture: Introducing the Consent Manager

To ensure individual data rights around privacy and portability are protected, a new class of institutions must be created that have economic incentives aligned with those of the users when it comes to the sharing of personal data. Under DEPA, the interaction between an individual, a potential data user, and the data fiduciary holding a users information will be mediated through consent managers – organisations maintaining the 'electronic consent dashboard' for users as articulated in the Srikrishna Bill. Consent Managers will be in the business of making sure individual ual data is not shared without user consent.



DEPA Institutional Architecture

DEPA's market architecture will be based on several competing interoperable consent managers. This model provides the data principals (individuals or small businesses) seamless control over their personal data with a single view, even while the data is created, stored, and processed by hundreds of different services. Consent managers can proactively look out for individual data interests (for example, by making sure you have consented to data shared, innovating on modes of obtaining consent for a diverse population, and creatively designing means to grapple with consent fatigue) independently of the data fiduciaries (those custodians storing your data) and should compete to do so.

For developers at data user and provider organisations, the account model facilitates access to data and removes dependencies on specific data aggregators - by leveraging open APIs and
Registries (discussed in further detail in the Technology Architecture section).

In this institutional model, the flow of consents or permissions is separate from the actual flow of data. The Consent Manager should not be confused with personal data storage (PDS) solutions, that enable storage of data in a secure place under the direct control of an individual custodian. The primary function of a Consent Manager is to allow users to access and share data, but the **data itself is not necessarily streamed through the servers where the account is hosted**. Consent managers are data blind by design; they are not permitted to store user data.

The Consent Manager approach works in practice as follows:

- Consent Managers hold consent logs that determine how data can flow from data sources to data users in an authorised system.
- Consent Managers are data blind. They only enable the transaction, but are unable to read, store or analyse the data.
- For personal data management, it is sufficient for the authorisation consents to be centralised in the account. Data can flow directly between the source and the user.
- Due to account portability, individuals can easily choose and change their Consent Manager operator service. The service provider lock-in is minimal.

For government data, the first created example of an aggregator of digital data from multiple official sources with individual consent was DigiLocker. DigiLocker could transition to becoming an official consent manager if multiple government departments become Government Information Providers (GIPs). Similar regulated consent management organisations will need to be designed for each sector. The rollout for the financial sector and plans for sectors such as health, telecom, and skilling is discussed in further detail in Section IV.

Finally, to ease the burden on regulators in certain sectors, a non-profit collective of Consent Managers, data providers, and consumers could be created as a self-regulatory organisation (SRO) to look out for user interests, design procedural data sharing guidelines specific to the sector, enable data providers and users to operationalise the framework quickly, and monitor adherence and compliance by all players. For the financial sector, a non profit collective titled 'Sahamati' (meaning consent or agreement in Sanskrit) has been formed by market participants. For further detail, see Section IVb on Market Architecture.

Business Models for Consent Managers

For the DEPA ecosystem to flourish, it is crucial that there are viable business models for the new Consent Managers, as well as for data users and providers. Globally, personal data is regularly shared or sold (often without user consent) by the data fiduciary company so that the individual gets a "free" service. The DEPA infrastructure provides a simple and transparent mechanism for making data exchange visible and explicit in ways that benefit all parties – either through enhanced services or direct monetary profits from sharing data. **Consent Managers can facilitate a data exchange by charging a nominal fee.** However, since most individuals are accustomed to free services, Consent Managers could subsidise or relinquish the service fee charged to the data principals by charging the data users (much like a subscription model). Information Providers could go on to charge a service fee in the future, but in the financial sector they have agreed to provide data without a charge for the time being. Finally, a competitive ecosystem of Consent Managers in each sector could keep prices manageable but cover costs to ensure profits.

Theoretically, different or alternate business operating models for Consent Managers could exist:

- **Consent Management accounts or operators:** This model entails the operator to be an independent entity that just acts as a consent manager. They merely allow and manage data and consent flows to the data principal and data user. This is the model adopted by DEPA in the financial sector.
- In house model: Here the operator and data user is combined. The data user understands the need for access to personal data and incorporates a consent manager along with the other services it provides to the data principal. This model has been adopted in the UK, but would not be suitable for the diversity of the Indian context, which would require constant innovation by consent managers to reach diverse user groups.
- **Public Sector Model:** Public sector entities could offer a subsided, low cost consent management service. This model could be appropriate for some sectors.
- **Privacy based model:** Some Consent Managers may offer additional services with regard to data privacy and security. This could be a future avatar of Consent Managers in the DEPA framework.

DEPA as a framework could apply to personal data (data with personally identifying information) and to derived data (data with masked personally identifiable information but could reveal confidential data of a company). When sharing the latter, care ought to be taken to maintain a company's earned competitive advantage – although personally identifying information may be masked, proprietary company algorithms or techniques may be revealed through data sharing. Therefore, the application of DEPA to new data sub-categories will need to be decided in an evolving manner through more specific procedural guidelines with sector-specific nuances.

DEPA's Technology Architecture

In order to enable a thriving ecosystem of data access fiduciaries, a variety of digital public goods have been created:



An Electronic Consent Framework, with a specification for a consent artefact managed by MeitY: http://dla.gov.in/sites/default/files/pdf/MeitY-Consent-Tech-Framework%20v1.1. pdf

2 Data Sharing API Standards to enable an encrypted flow of data between data providers and users

3 **Data Information Standard** for the launch of DEPA that is sector specific. For the financial sector, this is the Financial Information Standard which explains the required shared elements of a bank statement across institutions for instance.

Critically, these are the basic building blocks of a DEPA technology framework - all necessary but not sufficient. As DEPA evolves, other technology modules should be added which better preserve privacy and data rights - through a combination of public and private players.

Electronic Consent Architecture as a Foundation of DEPA

A shared specification to communicate consent is a critical foundation of the DEPA technology architecture. Standardising consent brings several benefits, including:

Providing a clear process for obtaining consent to share. This makes it difficult for companies to share data without obtaining consent, or obtain a blanket consent that could enable misuse of data.

2 Identifying why the data is being used in a particular context (rather than a blanket authorisation) in standard form. A consent artefact means a standardised, codified purpose of data sharing that can be used in future audits.

3 Enabling users to choose how long their data is shared for, specify consent for granular data elements, and decide whether data can be shared further to third parties.

Simplifying jargon on consent forms and allowing users to make meaningful comparisons between privacy policies of products Consent with regard to data empowerment is defined as the ability of an individual to collect and aggregate data about themselves across multiple data sources and their ability to share their own data for access to goods or services. Consent encapsulates the entitlements given to an individual over their data. Although the electronic consent specification that has been designed can be used for both consent to collect data and the consent to share, **in the context of DEPA the consent artefact is used to enable permission to share data**.

Consent ought to have the following characteristics:

- Consent should be freely given, informed and specific to the purpose of processing.
- All transactions do not warrant the same standards of consent. The validity of consent needs
 to be carefully determined based on the sensitivity of the data and the size/reversibility of the
 transaction on a sector-wise basis. Mechanisms that enable consent can exist on a spectrum,
 on one end it can be protected heavily by-laws/regulations and on the other end, for more
 straightforward use cases control can be given entirely to the individual to grant and revoke
 consent as seen fit.

Consent To Collect

Consent with regard to data is not a monolithic blanket yes or no, but a much more granular, contextual process. Thus to frame the concept of a good technology framework to enable informed consent, DEPA uses the ORGANS framework:

- **Open Standards:** The consent architecture must follow the principles of open standards.
- **Revocable:** The consent given should be revocable by the user at any stage.
- **Granular:** The consent given must be presented in granular level, where the data is broken down in terms of its characteristics and each characteristic has its own time and sharing privileges.
- Auditable: All events in the consent flow and data flow must be digitally signed and logged using the MeitY Consent Log artifact. These non-repudiable transaction trails shall lead to higher trust
- **Notice:** The user must be informed and given due notice through Email, SMS, In-App Notice, and other notification mechanisms when consent is created or revoked and when data has been requested, sent or denied.
- Security By Design: The internal and external software and systems must be designed from the ground up to be secure. There must be end-to-end security of data (PKI, DSC, tamper detection) and it must be network agnostic and data-centric.

It goes without saying that consent alone cannot be the only backstop to prevent data misuse. There is a strong role for a future Data Protection Authority, a data protection regulation framework to manage data use on the Information Provider side, as well as non-profit collectives to build sector-specific procedural guidelines for data management to prevent unauthorised use (the Call to Action section covers this in further detail). However, consent is a crucial tool to restore individual agency over data sharing.

APIs for Data Sharing

Application Programming Interfaces (APIs) enable seamless interaction flow of and encrypted data flow between data providers and data users through a consent manager. Institutions adopting DEPA APIs can provide data in a machine readable format to all licensed consent managers. As a result, it is possible to build a centralised dashboard where the individual may grant access and give or cancel permissions for multiple data sources and services. Any service provider can build a consent manager API and enable their service to be connected with the accounts directly.

A standardised Consent Management architecture makes the accounts interoperable and allows individuals to easily switch operators. This is a major element contributing to DEPA's trustworthiness. Interoperability is the core advantage provided by a consent manager, but it is also the core challenge: interoperability within the data management system can be understood as functioning similarly to interoperability in mobile telephone networks. Both systems require a common network that connects distributed nodes.

Data Protection and Processing Standards

DEPA also relies on adoption of related technology standards around data storage and processing techniques. Some of these are outlined in the Personal Data Protection Bill - which for instance states that all processing of sensitive and critical data must occur within India. Further technology standards around data storage, based on the sensitivity of data, ought to be designed and regulated by the forthcoming Data Protection Authority.

4. Building DEPA for the Financial Sector: The Account Aggregator Model

The implications for rolling out DEPA across the financial sector could be transformative for financial inclusion. The sector already generates a high amount of structured and interoperable data regulated by strict standards - making it an ideal place to begin operationalising a consent based data sharing framework, and see the impact of consented data sharing on availability and accessibility of financial products to new, more vulnerable markets. This section provides an overview of the regulatory framework, market architecture, and potential combinatorial impact of DEPA for the financial sector.

Regulatory Framework for Data Sharing in the Financial Sector

The Reserve Bank of India (RBI) has recognised fintech as a 'potentially transformative force in the financial markets': the regulator believes fintech can play a pivotal role in efficiency improvements, risk reduction and greater financial inclusion. RBI has also recommended collaboration between Banks and FinTechs to overhaul manual time-consuming traditional banking processes to empower customers. Recognising the need for an electronic consent framework in financial data sharing to catalyse fintech innovation in the sector, RBI in 2016 published a notification announcing Account Aggregators, the financial sector Consent Managers, electronic consent dashboards for the banking industry, titled "Master Direction- Non-Banking Financial Company - Account Aggregator (Reserve Bank) Directions, 2016" The Account Aggregator (AA) model allows regulated entities under their control to share data with user consent. This was with a view to revolutionise areas such as lending, wealth management, and financial management by allowing an AA to securely share information about bank, insurance, pension, securities, and eventually income tax and GST data at the discretion of the data principal. This followed on from December 2014, when the key financial sector regulators - the RBI, the Securities and Exchanges Board of India (SEBI), the Insurance Regulatory and Development Agency (IRDA) and the Provident Fund Regulatory and Development Agency (PFRDA), came together in an Inter-Regulatory Technical Group (IRTG) to recommend creating an account aggregation facility to view information from multiple accounts in one view.

Market Architecture: Account Aggregators and Sahamati

Account Aggregators (AAs) will act as Consent Managers for the financial sector, working with Financial Information Providers (FIPs) to share the data of an individual or small business with their consent to a Financial Information User (FIU).



As of this writing, seven AAs have received in principle approval from RBI to begin operations, and two have received operational licenses. This number is expected to grow over time as AAs emerge to target different customer segments with novel approaches of communicating data requests and obtaining informed consent.

The API architecture is designed so that individual AAs will not need to integrate afresh with every new FIU. To ensure sustainability and competitive service delivery, AAs can charge FIUs or users themselves per transaction. Financial information providers have committed to provide free service today could charge a nominal fee in the future.

As with other consent managers, AAs are designed to be data blind: the data that flows through an AA is encrypted and can be processed only by the FIU intended by the user. Moreover, the AA regulations do not allow them to store user data, to minimise risk of data leaks and misuse.

To operationalise the AA framework quickly, market players have come together to create a new organisation to support the rollout of best practices for the AA ecosystem: a non-profit called Sahamati. Sahamati will educate new financial information providers, users, and potential AAs

about the DEPA architecture, provide technical support for institutions to go live, design procedural guidelines and best practices to support the ecosystem. It will also support the update of specifications and guidelines over time. Just as the WiFi or Bluetooth alliances ensure individual members are complying to the standards, Sahamati can be a collective working for the benefit of the AA ecosystem.

Sahamati will have an independent board, with members from industry, academia, and regulators, and ensure that users interests are also represented on the board. Specific activities could include:

- Raising awareness on the AA model and providing technical support to new AAs, FIPs, and FIUs
- Publishing a code of conduct, audit guidelines, and interoperability standards for members
- Establishing data standards for reporting to regulators
- Creating a grievance redressal framework for all customer complaints
- Ensuring members adopt regulatory tools for self reporting of data and granular, automated, and data-based audit
- Monitoring member compliance

Impact of Data Sharing for Financial Inclusion: The Potential for Cash Flow Lending

One of the key bottlenecks to access to financial products has been the friction associated with accessing data necessary to underwrite and price products, especially for less affluent customers. Introducing the AA model as a seamless, secure, and consent based micro-data sharing tool certainly makes existing processes easier (for instance, onboarding a new individual or MSME to a bank, or NBFC or underwriting for a collateral-backed loan). However, it also allows new and different types of financial products to scale and gain traction in the ecosystem. One such example is Cash Flow based Lending.

The recent **report** by the Expert Committee on Micro, Small and Medium Enterprises under the chairmanship of UK Sinha explained the transformative potential of cash flow lending for MSMEs: Only about 8 percent of the total MSMEs in the country have access to formal finance; the other 92% are likely taking loans at onerous terms from ad-hoc sources. Unmet MSME credit demand was estimated at a staggering Rs. 25 lakh crore (~USD \$35 billion) even pre-COVID by independent credit rating information agency ICRA. A majority of loans to MSMEs today

are offered backed by on assets and collateral owned, which excludes a significant portion of businesses. Cash Flow Lending is a type of small ticket working capital loan which is not asset backed; instead, it provides credit based on the revenue generation and repayment capability of an MSME. It creates a short and flexible tenure and repayment schedule based on incoming cash flows. It has not become a mainstream mode of credit, in part because trusted data about invoices that indicate a close-to-certain future cash flow is difficult to access. AA opens up access to many different types of data that could be shared by MSMEs to inform banks and NBFCs of their cash flows and creditworthiness: GST data which is trusted information on turnover or future receivables, invoices on government procurement platforms such as GeM (Government eMarketplace), e-commerce invoices and transactions on private aggregators such as Flipkart or Amazon, or other kinds of digital sales records from trusted sources. A registered seller using an online sales platform could use transaction data to provide a history of cash flows to a potential lender. This could create a virtuous cycle where data could be used to improve the finances and growth of a small business owner, as shown below. Moreover, as outlined in the Committee report, for lending institutions CFL ensures a reduction in credit risk, reduced monitoring costs for banks, and a reduction in TAT and ability to serve entities without adequate collateral.



Successfully solving for a democratisation of access to credit along the full cycle requires additional public digital infrastructure:

- **The Public Credit Registry** announced in June 2018 by RBI will allow lending institutions to confirm the existing loans of any customer, to improve underwriting and prevent fraud (eg double-financing against the same invoice);
- Currently, MSMEs must apply only to a regulated financial institution to access financial products. However, a new set of APIs could enable any institution with a touch point with customers to become a **Loan Service Providers**. For instance, in the future startups such as

Ola, Flipkart, Swiggy, or PhonePe could allow frictionless loan applications. A new Collective for Loan Service Providers, Credall, is developing and updating a set of APIs to be known as **'LSP Bridge'** here: https://www.credall.org/resources;

 NPCI manages eMandates and e-NACH which allow anyone with a bank account to enable automatic recurring payments (for many use cases, but in this case for loan repayments).
 eMandates are currently used by a few banks, but need to be enabled over UPI to increase reach.

Impact of Data Sharing on Financial Inclusion: Credit Scoring, Better Financial Management, and a Cambrian Explosion of Financial Products!

Beyond cash flow based lending in particular, DEPA also makes several other possibilities come to life. Some examples:

- **1** More lending products and better credit scoring: The AA framework in conjunction with other platforms like the Public Credit Registry allow for companies to diversify their credit products (for example through sachet sized and regular flow-based lending to MSMEs or individuals as discussed above). It also allows for more sophisticated and cost effective credit scoring.
- 2 Better Personal financial management: By some estimates, individuals make an average of 200 financial decisions in their lifetime starting from which bank to go to for a savings or investment account to which type of loan, insurance, or mutual fund to buy. Consented and secure sharing of financial data could allow for a quantum leap in the quality of personal financial management support by providing tailored recommendations based on overall financial status and history.
- 3 Improved Wealth Management and robo advisory: Through the DEPA platform, users will be able to easily grant and provide wealth managers and digital advice firms with access to data on their banking and financial circumstances, enabling these firms to quickly tailor financial services and advice to match the user's needs.
- Unforseen Use Cases and Products: The power of DEPA as a platform is to provide a single service, and allow many players to innovate on new products and services that the original platform designers may not even have considered. Just as the creators of GPS could never have envisioned Uber as a service that leveraged its technology, a great deal of 'unforseen innovation' and new types of services are expected based on the AA framework.

DEPA Implementation and Rollout: Financial Data and Beyond

A sector-wise approach to DEPA implementation allows for flexibility in operational models for different types of data in different sectors.

In the financial sector, the RBI has already taken major steps forward towards operationalising DEPA through adoption of the MeiTY Electronic Consent Framework and creation of a new entity - the NBFC Account Aggregator (NBFC-AA) - in its Master Directive of September 2016. In June 2019, the Expert Committee on MSMEs published a report which dedicated a full chapter to recommending that MSMEs leverage AAs to seamlessly share their financial data (including transactions, payments, GST invoices, financial statements, etc.) to accelerate access to credit through cash-flow based lending. In July 2019, a closed user group of regulators, CEOs of major banks and NBFCs, and prospective AAs launched DEPA in the financial sector. This launch showcased a demo of consented data flows across different institutions, announced financial institutions' intent to co-create and implement the standards to share financial asset data, and also introduced Sahamati as the non-profit collective of Account Aggregators. Sahamati would serve as a source of information and provide support on adoption, technical standards, and kindle continuous innovation in the ecosystem. In November 2019, RBI published the key technical specifications required for the AA ecosystem. Since then, 7 AAs have received in principle licenses from RBI, of which two have received operational licenses, and approximately 10 banks and NBFCs are in various stages of adoption of the FIP and FIU technical modules which allow them to work with AAs across the ecosystem. A competitive AA Hackathon with 550+ participants took place from July-August 2020 enabling startups, fintechs, and product teams at financial institutions to innovate and build on consent management or FIU designs. A full public launch which allows sharing of key financial sector data to access better credit products for individuals and MSMEs is planned for Fall 2020.

Similarly, the healthcare sector has also taken on a leadership role in addressing data related challenges. NITI Aayog's strategy document on the National Health Stack published in July 2018 recommends a federated Personal Health Record system that could leverage Health Data Fiduciaries to enable consent-based data sharing. The National Digital Health Blueprint published by the Ministry of Health and Family Welfare on April 2019 also proposes piloting DEPA, positing that India needs a "Federated National Health Information Architecture, to roll-out and link systems across public and private health providers at State and National levels consistent with Metadata and Data Standards (MDDS) & Electronic Health Record (EHR)". This effort has gained renewed vigour in light of COVID 19; the National Health Authority has been tasked with implementing the National Digital Health Mission, and is piloting the DEPA architecture for healthcare data in Fall 2020.

Exchange, and could need to leverage a data sharing framework.

Other sectors too have taken important steps towards data empowerment. The telecom space is also planning its adoption of DEPA. In a seminal paper by the Telecom Regulatory Authority of India (TRAI) on telecom data use, the regulator suggests in Recommendation 3.3C that "The Right to Choice, Notice, Consent, Data Portability, and Right to be Forgotten should be conferred upon the telecommunication consumers [...] For the benefit of telecommunication users, a framework, on the basis of the Electronic Consent Framework developed by MeitY and the master direction for data fiduciary (account aggregator) issued by Reserve Bank of India, should be notified for telecommunication sector also." On 28 August, 2020 the TRAI Chairman RS Sharma convened a workshop on 'Telecom Subscriber Empowerment' with major industry players present and announced the TRAI-RBI partnership which would allow telecom companies to become FIPs in the Account Aggregator system. Telecom data is often the first digital footprint generated by a low-income household, and a steady history of on-time recharges could contribute to a budding credit history. The Ministry of Skill Development and Entrepreneurship has published a report encouraging adoption of a digital skill credential that could be used to address low data portability in employment by sharing verified information on work experience or educational training. The Srikrishna Report addresses data sharing of private data in its bill (page 39, chapter 3F). Finally, the Urban governance authorities are interested in building an Urban Data

A lot of good work has already been done by various public institutions in moving towards a data sharing and empowerment model. The DEPA framework could work to accelerate such existing efforts. For regulated sectors (for example, financial, health, or telecom) the sector regulators can create new Consent Managers. Whereas for currently unregulated sectors (for instance social media, e-commerce, education & jobs, etc.) the new Data Protection Authority may need to create consent managers for each sector and design incentives for institutions in the sector to work through these consent managers to share data. For data held by various government institutions, each department could adopt the technology standards required to become a Government Information Provider, allowing individuals and businesses to access and share their data housed within different departments. To accelerate Cash Flow lending, GST is in the finishing stages of becoming the first GIP, and as of August 2020 had written formally to the RBI to join the Account Aggregator network. Adoption by other departments could enable substantial improvements to the Ease of Doing Business.

Future Required Public Innovation based on DEPA

The financial sector use case highlights further technology or process building blocks which should be designed (both by individual Account Aggregators competitively, and by the ecosytem as standards) to improve data privacy and further empowerment. Many of these are being developed in the **data governance working group** led by Sahamati (the team is looking for more volunteers and suggestions). Looking beyond the consent artefact, and APIs for sharing, and data protection standards:

- A strong overall data governance enforcement, and certification framework. Strong FIU data governance would need to be defended by both the RBI, a future Data Protection Authority, and by Sahamati. Sahamati could support the design of a world class certification framework for every agent in the ecosystem.
- 2 Minimalist data request templates for different purposes (with clear definition of a bad template): These could ensure that for a specific purpose, there is a standard set of minimal data that is requested. Adoption of these templates could be encouraged by the collective Sahamati or led by AAs.
- 3 **Data/Institutional trust scores**, as recommended in the Srikrishna report, can only come to life using technology standards. Consent Managers in each sector, or a collective of these can start to build multidimensional trust rating scores for information providers and users.
- Standards for nudge tactics towards more informed consent. These could include planned speedbumps in consent prior to data sharing, cooling off periods or time lags for consent for particularly sensitive data, etc.
- Enhancement of access permissions: The "Access permissions" supported by consent artefacts need to be enhanced to enable different levels of privacy protection when data about users is shared by one entity to another. Currently, the Consent Artefact provides clarity only on how to implement the STORE permission. The ecosystem needs to lay down a technical standard or a set of guiding principles for implementing the VIEW, QUERY and STREAM permissions. These standards will invoke either existing privacy-preserving utilities or create new ones along the way. Other permission types may also need to be enabled.
- Secure Multi-Party Computation (SMPC) Standards: SMPC will allow data consumers to query or to compute on certain "aspects" of the data (e.g., did the customer have enough of a cash flow for the past 6 months to be eligible for a loan of X rupees) without getting a full copy of the data. SMPC is designed to be a way to securely share data between multiple parties however, we may need to use only two-party versions of it in most of our use cases.
- Secure Vault Technology: Secure vault technology will allow data consumers (FIUs) to view the data in a secure environment without the ability to copy the data into their domain. Variations of this will be needed to implement the VIEW access permission. An industry collective, RBI, or the Data Protection Authority will need to lay down a set of standards on how this should be used by data providers and users.

5. An Opportunity for the Ecosystem to Co-Create

Now that the DEPA platform is available as a public good, market players across the financial and technology ecosystems as well as new entrepreneurs have an opportunity to leverage and build on this digital platform by innovating across the various new roles that have been created. Rather than thinking of a traditional government versus private sector model in terms of taking on leadership for collective progress, it is helpful to consider a 'Relay Race' approach where all these actors co-create societal outcomes. In such a model, the government takes a first step by designing a digital public good, the industry and market players then work to leverage the public platform to compete to design value added services for users. Media and civil society then plays a role to hold industry players accountable in the public eye and build awareness of the changes amongst the population.

In the financial sector, the first critical opportunity for new and existing market players will be to create new Account Aggregator enterprises that address various segments of the market. Some of these could focus on being consent managers for individual consumers of financial products, whilst others could focus on consent management for small businesses or micro enterprises. Some could target upper socioeconomic groups by building user-friendly smartphone applications, whilst others could focus on the often excluded majority of the nation: financially constrained urban or rural populations (or disadvantaged groups). For these groups, AAs could design competitive product experiences with innovation around modes of gathering consent (eg. audio or voice enabled, visual images or video supported, or assisted through a customer service center or a banking correspondent role). A hackathon for teams to design new AA apps or FIU use cases organised by Sahamati in July 2020 attracted over 1250 applicants, of which over 550 were selected to participate.

Similarly, entrepreneurs will need to create Consent Manager institutions for data across sectors such as telecom, health, or jobs as these sectors roll out their own adapted version of DEPA.

All actors across the ecosystem should allocate budget where possible for public awareness on what comprise meaningful consent, and the power of data aggregation and sharing for empowerment.

Some suggested actions other players could take to leverage this evolving opportunity include:

Current Financial Institutions will need to become FIPs and FIUs to make the most of this new possibility and cement their role in the changing ecosystem.

- **2** Financial Sector Regulators (RBI, SBI, IRDAI, and PFRDA) could to work together to create a competitive ecosystem of Account Aggregators and customer protection in their sectors to drive adoption and successful rollout of AA for all data around assets and liabilities banking, non banking, securities, pension funds, etc.
- 3 **The Ministry of Finance** could steer the sector-wide rollout with common interoperable/ harmonised adoption across banking, securities, insurance, and pensions – without AA silos within each sector.
- The Ministry of Electronics and Information Technology could continue to manage and revise the Electronic Consent Artefact as needed, and ensure the notification of the Personal Data Protection Bill.
- Civil Society can serve as an alert watchdog for data sharing, and build awareness around the existence of informed consent and data sharing possibilities. They will also need to ensure protection of the rights outlined in the Srikrishna bill (especially data portability).
- Fintechs and InsurTechs (NBFCs, Payment Wallets, Online Financial Marketplace and Insurance Web Aggregators) will need to first ensure integration as FIUs to leverage the AA framework, and also start designing new products based on these possibilities. The prevous business model which relied on a monetisation of raw data will have to shift to monetisation of predictions, analytics, decisions, and scoring – forms of 'value add' on the data.
- A New Data Protection Authority could be created as outlined in the Justice Srikrishna Bill to enable the creation and regulation of Consent Managers in other sectors. Moreover, they could create new data flow auditor institutions who perform functions such as trust score mapping.

This is the beginning of a new uniquely Indian journey on data empowerment and financial inclusion. An open and vibrant data democracy can be created if we can enable a billion individuals to thrive in an increasingly digital economy, based on foundational digital public goods that are designed to scale to meet the needs of a diverse population. Moreover, because the technology standards underpinning DEPA are open and now publicly available, the technical and institutional architecture can also be applied to other countries. An institutional body could even be designed to help globalise this standard and apply it to other nations facing similar

challenges as appropriate.

We hope that DEPA's regulatory, institutional, and technology architecture will be a transformative data governance approach. This approach shows the new '**India Way**' that is quite distinct from the other models around the world with respect to on data protection, sharing, consent and privacy. The Indian approach is specially designed to inclusively cater to the needs of a developing economy, to be technologically cutting edge and innovative, to drive and stimulate economic and business value, and lastly to evolve over time to meet ever emerging "new" applications of data.

Annex I: List of Acronyms

PMJDY	Pradhan Mantri Jan Dhan Yojana
ΑΡΙ	Application Programming Interface
DEPA	Data Empowerment and Protection Architecture
AA	Account Aggregator
MSME	Micro Medium and Small Enterprises
PSD2	Revised Payment Services Directive
RTS	Regulatory Technical Standard
NETS	Network for Electronic Transfers
FI	Financial Institution
ICT	Information and Communications Technology Systems
B2B	Business to Business
UPI	Unified Payments Interface
RBI	The Reserve Bank of India
SEBI	Securities and Exchange Board of India
IRDAI	Insurance Regulatory and Development Authority of India
SRO	Self Regulatory Organisation
ERP	Enterprise Resource Planning
PAN	Permanent Account Number
GDPR	General Data Protection Legislation
GST	Goods and Services Tax
TPP	Third Party Provider
PDPC	Personal Data Protection Commission
SDLC	System Development Life Cycle
MFA	Multi Factor Authentication
MAS	Monetary Authority of Singapore
EU	European Union
AISP	Account Information Service Provider
PISP	Payment Initiation Service Provider
ASPSP	Account Servicing Payment Service Provider
P2P	Peer to Peer
NPA	New Payments Architecture
PSP	Payment Service Provider