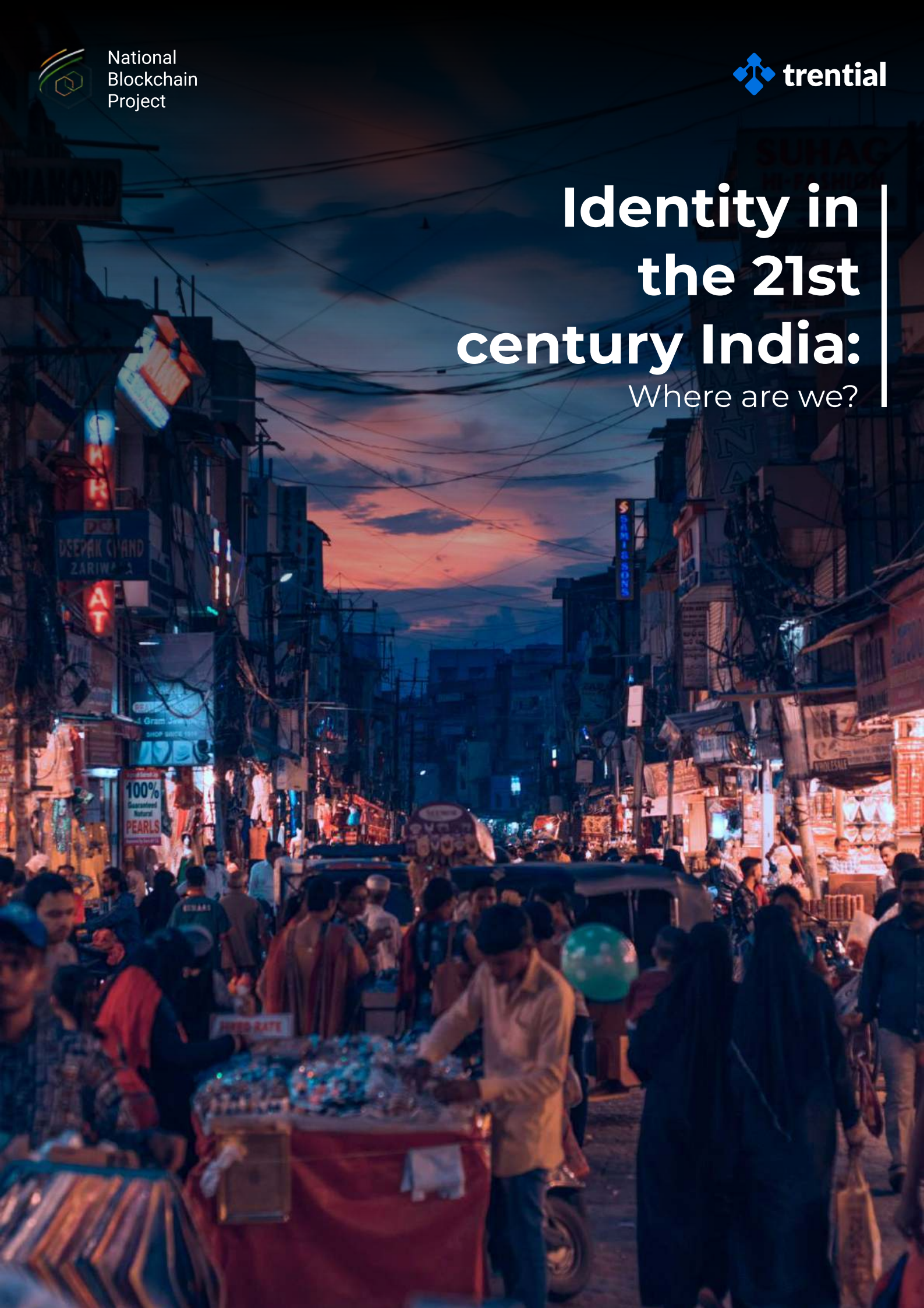
National Blockchain Project

trential

Identity in the 21st century India:

Where are we?



Executive Summary

This whitepaper is an attempt to understand identity from a governance perspective and the various methods of identification used. In particular, it'll analyse India's digital identity infrastructure – its motivations, the benefits it has yielded, and the dangers that might adversely impact it.

The paper comes at a crucial juncture, with India making giant strides toward smart digital governance and facing new challenges surrounding the protection of privacy of individuals and the security of the digital systems. As we move towards a new digital future, our systems must be resilient, adaptive and secure. This paper utilises numerous sources describing the formation of India Stack, judicial review of privacy and Aadhaar, and the recent global trends concerning profiling and cyber-security to develop a critical analysis of India's identity ecosystem.

We briefly introduce identity in the context of governance, and discuss Indian digital identity initiatives- primarily Aadhaar. Launched to plug the welfare leakages by offering unique identification, it has become India's primary channel of government service delivery. Other ventures such as UPI, DigiLocker and newly introduced Account Aggregators have added on to the functionalities of Aadhaar.

However, all of this has not proceeded without dissent and challenges. The 2018 Supreme Court verdict limited Aadhaar authentication only to public service delivery and taxation purposes, prohibiting even its voluntary use in the private sector.

The verdict threw into light the various threats posed to the current digital infrastructure, primarily– privacy, security and single point of failure.

The promise of unique identification by Aadhaar also comes with inevitable trade-offs. The merging of multiple silos with the Aadhaar number as the link to pick out suspect beneficiaries, has presented some dangers to the privacy of individuals, which led to the restriction imposed by the Supreme Court.

With the digitalisation of most services, the question of how the increasing amounts of data are stored gains extreme significance. While large centralised data repositories are gold mines for cyber-attackers, the breach of even small databases makes the citizens vulnerable to phishing and other attack.

Finally, while there are numerous physical documents, Aadhaar is the only sophisticated digital credential that can be easily used for instant and presence-less authentication, making any restrictions on Aadhaar very damaging.

By raising appropriate questions that will help guide policy-makers and leaders, the hope is that this white paper raises important issues and ways to address them that must be considered when designing an identity system for accessing citizen data in a scalable, trustworthy and legally compliant manner.

Contents

Section 1

Introduction

- What is identity and why do we need it?
- Methods of Identification

What is identity?
A simple dictionary search yields multiple meanings: “who or what someone or something is”, “the characteristics, feelings or beliefs that make people different from others”, etc. Thus, it is important to clarify what do we mean by identity in this paper. It is equally important to address other questions such as: why do we need a large-scale identification? How does this identity exist in the real world? And how has the manifestation of identity changed over the years?

Introduction

What is identity and why do we need it?

What is identity?

Identity, in this paper, refers simply to the attributes of individuals that can uniquely and conclusively distinguish one individual from another.

These attributes can represent some basic information about the individual such as- their name, age, gender, address, etc. They can also represent some intricate details such as- the place of work, the assets owned, are they a pre-existing customer of X company. In addition, they can also elaborate the capabilities of an individual, and so on. The number and types of attributes are limitless, and it would be impossible to list them all. They are dynamic in nature, too, as new (types of) attributes are constantly formed.

These attributes are important claims about individuals that help them navigate different spheres of public and private life. They are essential from a governance point of view, allowing government agencies to streamline public service delivery.

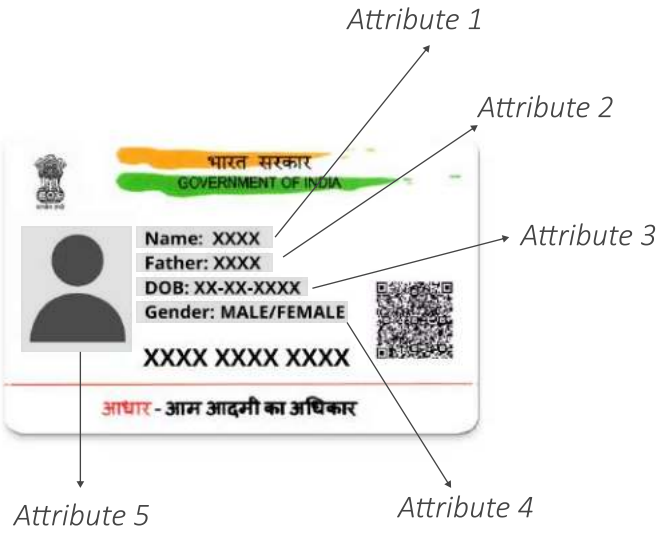


Figure 1: Attributes present in a credential

Why do we need identity?

For governments, one of the major purposes of identification has been distributing public services. Ideally, essential resources like food, fuel, etc., should be provided freely or at a subsidized rate to everyone. However, given the limited resources at hand, pragmatism has dictated that the benefits be provided in a limited amount and a targeted manner considering the lack of privilege of different subsections. This lack of privilege could manifest in terms of class, caste, gender, etc., and thus identifying these characteristics is essential. At other times credentials are required to prove a claim of capability (education, driving ability, ability to provide medical services etc.) or a claim of ownership (land/property ownership, vehicle ownership etc.). As a citizen of the nation, an individual is also entitled to some rights, such as the right to vote. Similarly, there are countless claims that are highly critical in a modern nation.

Thus, to facilitate the wide variety of activities carried out in the public sphere, a form of trust (that the person is who they are claiming to be) is imperative. It is precisely this trust that the different methods of identification seek to deliver. This comes with the trust in the mechanism and technology of identity delivery.

Introduction

Methods of Identification

For many purposes there are only 2 stakeholders, the holder of the credential and the issuer who acts as a verifier too. In a 2 party system, a third party who wants to verify authenticity and correctness of holder's claim-- the third party needs to request the issuer to check veracity and authenticity of the credential presented by the holder to the third party. However, going back to the issuer is not necessary if there are cryptographic methods by which the third party can verify both the veracity and the authenticity of holder's claim. Thus, the identity framework becomes a 3-party one and throughout this whitepaper we will refer to this model.

Issuer

Holder

Verifier

The 3-party identity framework consists of 3 entities, the issuer, the holder and the verifier.

The **issuer** is any organization or a person who 'issues' a credential, i.e., they are the sources of our credentials. Some of the most common issuers are government organizations (driving license), banks (credit cards), financial institutions, corporations (employment credentials) and educational institutes (transcript and degrees) etc. While any one of us can issue or create a credential, the value of the credential is derived from the trust that we have towards the issuer. Hence, the issuer and its trustworthiness play a cruicial role in this framework.

The **holder** is any individual who 'holds' the credentials issued to him/her by the issuer. A credential is usually an attribute of the holder (marks obtained in an exam, a degree obtained, biometry of the holder, or address of the holder etc.). A holder uses its credentials to avail services in most of the cases.

These services are provided by the organizations or people who will verify the holder's credentials and then provide services in return. These entities who 'verify' the credentials (to verify the genuineness of the subject of the credential) are known as **verifiers**. Some common examples of verifiers are government organizations (passport application), corporations (educational credentials), banks (loan application) and so on.

The three major forms of identification we'll study in this whitepaper are:

- Traditional (Physical) method of identification
- Digital method of identification through a centralized system
- Digital method of identification through a decentralized system

The traditional method of identification or the physical method is the oldest form through which identity manifests in the real world. It is also called the physical method of identification because it relies on "physical documents", such as ID cards, ration booklets, property papers, etc. It is still widely used by the billion plus citizens across India.

The rise of computers and the Internet revolution in the early 21st century ushered in a new digital era. Identification methods, too, have started utilising the digital medium for identity transactions, which led to the digital methods of identification. The methods can be crude such as scanning physical documents and using them digitally or they can be highly sophisticated with the entire process flow of issuance to verification to maintenance to storage of records being carried out digitally through specialised applications. This use of digital medium in any form is what we call the digital method of identification.

It is important to note a bifurcation introduced with respect to the digital method – centralized and decentralized identity systems.

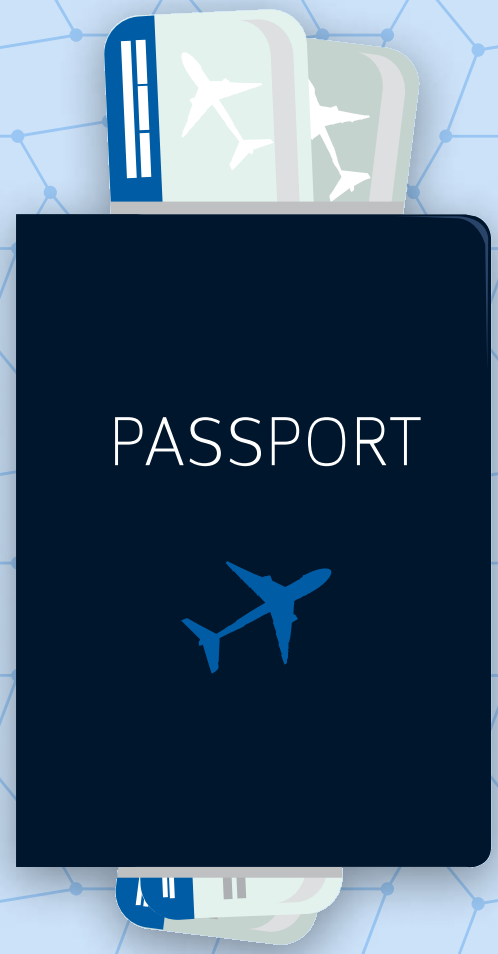
Centralized identity systems are typically characterized by a single organization storing all identity-related data and controlling the various process flows. For example, in the case of a national ID card when all of the personal data and the administration of the identification process is managed and held by the government in one central database. They also control to whom your data must be shared based on your consent.

In a decentralized identity systems, the personal data of individuals are stored on a distributed ledger in an encrypted form or in many cases are stored locally on the personal devices of the individuals. The process flows are also controlled by multiple entities. While previously it would have been impossible to verify the authenticity of such a document, advancements in cryptography and blockchain have made it possible.

Section 2

Traditional Method of Identification

- How easy is the method?
- How robust is the system?



The traditional method of identification refers to the use of physical credentials and physical (manual) processes for identification.

The physical credentials are mostly paper and plastic-based and can be in the form of documents, cards or booklets.

Some of the physical documents issued are used for/as:

Accessing public benefits
Income certificate, Caste certificate, etc.

Proof of capability
Educational records, driving license, medical license, etc.

Proof of ownership
Vehicle RC, property deeds, RoR, etc.

Proof of citizens/ domicile
Domicile certificates, voter ID, etc.

There are a number of documents for other purposes too. And documents are issued to corporations and organizations as well.

The physical nature of identification manifests in the manual process of acquiring and using these documents too. The issuance and verification of these documents often require the physical presence of the individual. An individual either has to be present while the verification is being carried out or has to submit their identity documents at the designated office. In less serious cases though, the documents can also be sent through the post. The process of verifying the authenticity of the document also utilises certain physical characteristics embedded in the document (such as signatures, stamps, etc.)

To analyse the efficiency of the **traditional method of identification**, we can ask **two simple questions.**

- How easy is the method?
- How robust is the system?

Traditional Method of Identification

How easy is the method?

The main bottleneck in the physical method of identification is the step of issuing of credentials. This step involves a complex process of the user filing an application with the requisite documents at the office or sending it through post, further verification of the authenticity of the documents and the identity of the person (which sometimes involves physical presence of the individual), getting requisite approvals from different officials and finally issuing the credential.
The drawbacks are:

01 Administrative Expenses

A lot of effort has to be put in by government offices to facilitate the issuance and filing the details in their records, which eventually translates to a liability on the exchequer.

02 Time Wastage

Both the user and officials spend considerable time on the process. Simple processes take up whole days of the user.

03 Opportunity Cost

This also presents a huge opportunity cost to the individuals and firms, and largely is unnoticed. The effort and time spent in more productive work could lead to significant economic gains.

The verification of documents suffers from similar problems. In fact, some of the problems detailed are actually related to verification. Issuance takes a lot of time and effort precisely because it is strenuous to verify the details and proofs presented by the individual. With photo and document editing tools becoming commonplace in the digital era, careful scrutiny is required. Officials have to spend extra time and effort to verify the authenticity of the document, leading to extra waiting time (sometimes lasting even months and years).

Apart from the issuance and verification, the maintenance of the huge number of documents might be a challenge for the holder and misplacement of documents can again lead to wastage of resources. Similarly the issuers and verifiers have to store the documents and carefully file them in huge file rooms. In many cases (audits, re-issuance of lost credentials, etc.) these documents might even have to be retrieved, which can be extremely strenuous. There is also often no way to record the flow of documents, paving the way for illicit acts by inside officials, identity theft, etc.

Traditional Method of Identification

How robust is the system?

It simply means what kinds of errors can happen and how frequently?

01 Identity Theft:

This is the situation when a person is able to successfully claim that they are someone who they are not (impersonate someone else). This can be the case if someone gets access to a copy of your document and use it for other or the same purpose. For example, in the physical method when only an Aadhaar copy was required, the Aadhaar copy could be used by a person to get another SIM [1] or for some other purposes. This has to do with the transferability and reusability of the document produced by an individual. A person can not only obtain the document but also might claim ownership of the document and consequently the identity of the individual.

02 Forgery and Fraud:

Another source of error in physical identification is the use of forged documents to prove false claims. A vital aspect of the verification process is the authenticity of the document. This depends upon the security features of the document and the ability of the verifier to find flaws (through machine or manually). With manual verification of physical credentials, it is infeasible to always assume careful consideration of the security features by the personnel responsible for verification.

The rise in technology of counterfeiting has further compounded this problem. Even a widely used software like Photoshop, can be used to produce a reasonably authentic looking document. For simple paper-based credentials like say Aadhaar it's very easy to just change the photograph on the Aadhaar card and yet make it look like the original one. The verifier has no way to check if the photo on Aadhaar was tampered or not from a physical verification of the Aadhaar Card. This actually can have very severe consequences. Since Aadhaar, as of now, is widely accepted even in its physical form, a tampered Aadhaar might be used all along the country to avail a number of services.

03 Red Tape:

Another implication of the traditional method of identification is the emergence of red tape. This system of verification for accessing different services, while essential, also introduces a huge layer of bureaucracy. As talked about earlier, in order to ensure that the documents are not forged, it is necessary (in the traditional method) to have officials scrutinize them. Individuals have to go through a set of officials who verified their identity and get approvals from various desks, only after which they can access the services they required. At its best, this leads to long waiting times, adversely affecting citizens. At its worst, it creates space for malpractice by officials (such as bribery). The key aspect here is manual verification of credentials by officials. Often, this problem is handled by having a strong accountability system and a citizen-centric grievance cell. However, systemic problems in numerous administrative units can make it hard to have these. Thus, a good functioning traditional method of identification depends on the existence of strong and resilient institutions.

Thus, we find the physical method of identification to be riddled with problems – chiefly the administration cost, risk of identity theft & forgery, it's susceptibility to insider attacks, and corruption. These problems have ultimately paved the way for a paradigm shift towards digitalisation. All over the world, countries have gradually shifted towards digital management of identity. It is hoped that this will help solve these numerous problems plaguing governance. And it has to some extent. Digital identity management has delivered spectacular gains in terms of increased convenience, reduced time and effort and so on. However, it is yet to solve all the problems faced by the traditional and created new problems too. The rising incidents of cybercrimes and increased fear of surveillance by big tech and governments alike are issues which need deep thought. The following section details the foundations of digital identity to examine these issues.

Section 3

Digital Method of Identification

- Identity in the Digital World
- Siloed Identity Model
- Federated Identity Model
- Decentralized Identity Model

Digital Method of Identification

Identity in the Digital World

> *"The Internet was built without an identity layer."*
>
> Kim Cameron, Chief Architecture of Identity, Microsoft

What is an "identity layer?"

An identity layer refers to the mechanism through which the identity of the humans can be verified. It seeks to answer the question "Is the person who they are claiming to be?"

"The Internet was built without a way to know who and what you are connecting to. This limits what we can do with it and exposes us to growing dangers. If we do nothing, we will face rapidly proliferating episodes of theft and deception that will cumulatively erode public trust in the Internet."

The problem Internet was designed to solve was how to interconnect machines to share information and resources across multiple networks. The solution—packet-based data exchange and the TCP/IP protocol—was so brilliant that it finally enabled a true "network of networks". However, with the internet's TCP/IP protocol, you only know the address of the machine you are connecting to. That tells you nothing about the person, organization, or thing responsible for that machine and who is communicating with you.

Despite all the efforts to solve the Internet identity problem, the lack of a breakthrough solution has proved Kim's prognosis true in spades.

By 2017, the average business user had to keep track of 191 password and the username/ password management had become the most hated consumer experience on the internet. That's just an inconvenience. **The deeper damage has been in cybercrime, fraud, economic friction, and the ever-growing threats to our online privacy.**

We'll next detail the three primary models of digital identification-

1. Centralized Identity Model
2. Federated Identity Model
3. Decentralized Identity Model

Digital Method of Identification

Siloed Identity Model

The siloed model is the original form of internet identity—and the one that, in many cases, we still use today. You establish an identity by registering an account (typically a username and password) with a website, service, or application. For this reason, the model is also called account-based identity. The organisation with whom you register stores all details about you.

We have used this model for almost all identifiers and credentials such as government ID numbers, passports, identity cards, driving licenses, invoices, Facebook logins, Twitter handles, and so on. All of these are issued by governments or service providers like banks or telecom companies.

Governments usually operate using the siloed model of identity. The numerous organs of the government, e.g., revenue, commerce, local bodies, etc., maintain a separate website where an individual has to register themself.

The problems with this model are listed below:

01 Inconvenience

The burden of remembering and managing all the usernames and passwords (and, in some cases, other multi-factor authentication tools such as one-time codes) falls entirely on you.

02 Lack of standards

Every site enforces its own security and privacy policies, and they are all different (a classic example is the maddeningly different rules about passwords: minimum length, special characters allowed, and so on).

03 Non-portable

None of your identity data is portable or reusable anywhere.

04 Unsecure

Because of the inconvenience caused by multiple passwords, people often use generic passwords or use the same password for different websites. This is a grave security threat. A hack of a single account of an individual can lead to multiple hacks (if they use the same password)

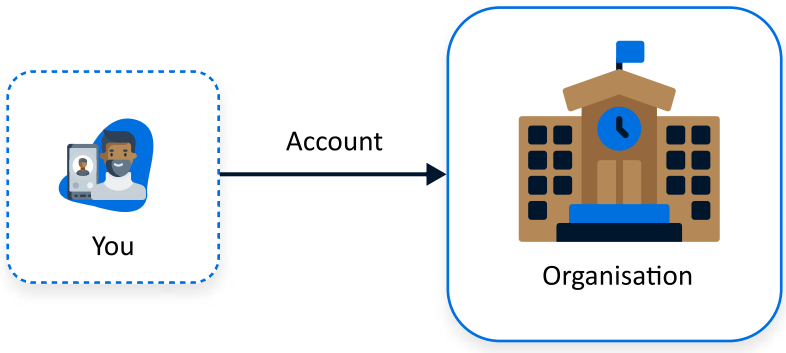


Figure 2: Relationship between Users & Service Provider in a Centralized Identity Model

Digital Method of Identification

Federated Identity Model

To alleviate some of the pain points of siloed identity, the industry developed a new model called federated identity. The basic idea is simple: insert a service provider called an identity provider (IDP) in the middle.

Now you can just have one identity account with the IDP, and it, in turn, can log you in and share some basic identity data with any site, service, or app that uses that IDP.

Federated identity management (FIM) started to catch on in the consumer Internet, where it began to be called user-centric identity. Using protocols like OpenID Connect, social login buttons from Facebook, Google, Twitter, LinkedIn, etc. are now a standard feature on many consumer-facing websites. However, despite all the work that has gone into federated identity since 2005, it has still failed to provide us with the Internet's missing identity layer. Challenegs being faced are

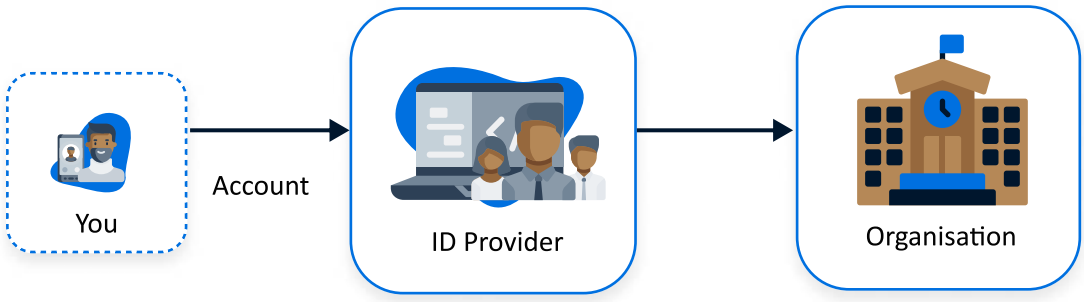


Figure 3: Relationship between Users, Identity Provider & Service Provider in a Federated Identity Model

01 Privacy

Given users use these IDPs for interacting with different service providers, these IDPs can surveil a user's login activity across multiple sites.

02 Security

Large IDPs are some of the biggest gold mines for cybercriminals. Furthermore, because they have to serve so many sites, IDPs must have "lowest common denominator" security and privacy policies.

03 Handling sensitive data

Due to security and privacy concerns, IDPs are not in a position to help users securely share some of their most valuable personal data: passports, government identifiers, health data, financial data, etc. It is difficult to trust Facebook to mediate transfer of one's personally identifiable documents.

04 Sovereignty

The large IDPs are most often foreign companies, who cater to citizens of scores of countries. Often, when the data of citizens is misused, the foreign operating company cannot be sued because of jurisdiction issues. In the absence of any global data protection frameworks, it is often imperative to store and process data of citizens in the country itself. This is increasingly becoming a huge challenge for functioning of IDPs.

Digital Method of Identification

Decentralized Identity Model

The decentralized identity model inspired by blockchain technology first began surfacing in 2015. Since then it has accelerated rapidly, assimilating new developments in cryptography and decentralized identity.

The most important difference in this model is that it is no longer account-based. Instead, it works like identity in the real world: i.e., it is based on a direct relationship between you and another party as peers (Figure 4). Neither entity "provides," "controls," or "owns" the relationship with the other. This is true whether the other party is a person, an organization, or a thing.

In a peer-to-peer relationship, neither entity has an "account" with the other. Rather, both share a connection.

Decentralized identity is a trust framework in which identifiers, such as usernames, can be replaced with IDs that are self-owned, independent, and enable data exchange using cryptography, blockchain and distributed ledger technology to protect privacy and secure transactions. Decentralized Identity can free users from having to use a mess of passwords, emails, text messages and authentication apps to verify identity.

In general, digital identity ecosystem utilize blockchain technology to eliminate the need for any intermediary as an identity provider. User-sensitive information is stored off-chain and can only be accessed by the controlling entity. Also controlling one's own data through cryptographic keys enforces the notion of controlling one's own identity. When compared to a centralized or federated identity system, a decentralized identity network cannot be shut down, use user's data without consent, or block users from using their identities. Moreover, interoperability between systems can be ensured, since users are not locked into one specific identity provider that is unwilling to integrate into services outside its own defined scope. This leads to an independent system that can be integrated by any service or institution.

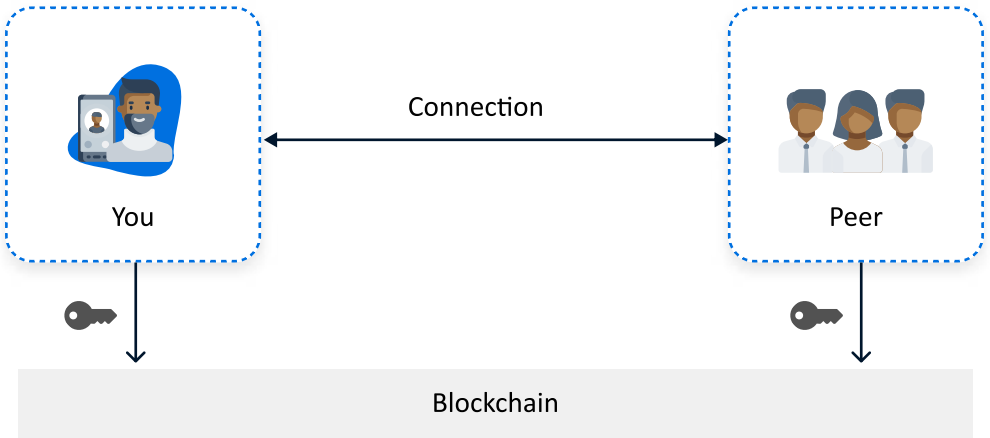


Figure 4: The peer to peer relationship in a decentralized identity model

Section 4

Government Backed Digital Identity Systems

- Presence-less Layer
- Cashless Layer
- Paperless Layer
- Consent Layer
- Economic Potential of Digital India

Welfare Leakages

In 1986-87, India spent over Rs 37 billion in food and fertiliser subsidies, which was over 1.7 per cent of the GDP and over Rs 59 billion on social sectors, especially health and education[2]. Yet, very little of the money allocated under various schemes was reaching the intended beneficiaries. As Rajiv Gandhi put it succinctly: "Out of Rs 100 allocated to an antipoverty project only Rs 15 reaches the people"[3]. The remainder, he said was lost to administration costs and gobbled up by middlemen, power brokers, contractors and the corrupt.

Post liberalisation, between 1991 and 2010, India's spending on subsidies for food, fuel and fertilisers shot up from Rs 122 billion to Rs 1.73 trillion and allocations for the social sector went up from Rs 102 billion to Rs 5.29 trillion [4]. In 2008, Montek Singh Ahluwalia revealed that only 16 paise out of a rupee was reaching the targeted poor [5].

Successive governments crafted public programmes for the needy but found benefits being cornered. The intended beneficiaries of entitlements continued to struggle, running from pillar to post for multiple official signatures and pieces of stamped paper to simply prove their identity.

Exclusion of citizens

Another issue challenging the officialdom was enforcement of customer due diligence (KYC) by financial institutions. On one hand the banks needed to know whether the person they presumed to know was the person he/she claimed to be, on the other opening an account with a bank or a mutual fund was trekking trip through red tape. If someone wanted to open a bank account, he had to find an introducer and fill in a plethora of documents in triplicate. This was repeated every time and for every kind of service, from mutual funds to fixed deposits to bonds.

Finally, the systemic weakness also precluded financial inclusion as nearly two-thirds of Indians were out of the banking system, which impacted the efficacy of government programmes.

All of these factors led to the formation of UIDAI, which till date remains a key player in the Digital India scheme. It developed the presence-less layer (through Aadhaar) which forms the foundation of digitalisation in India. It has also played an active part in the development of the other three layers – cashless layer, paperless layer and consent layer – which constitute the basic stack of Digital India

Government Backed Digital Identity Systems

Presence-less Layer

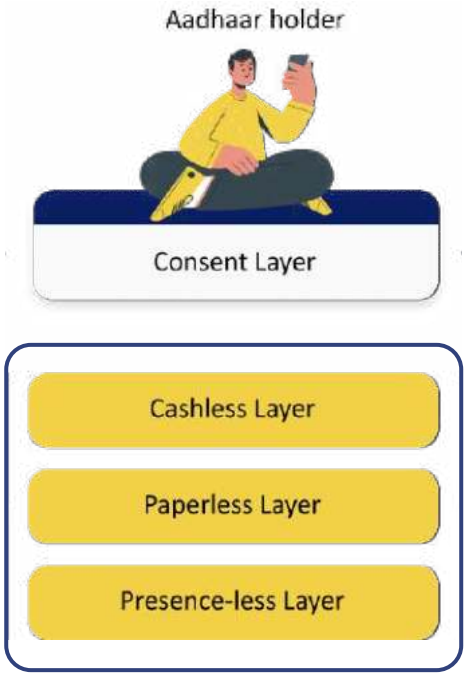


Figure 5: Illustration of the India Stack Model.

Characteristics of Aadhaar

What differentiates Aadhaar from the numerous different other cards and documents? There are two key factors.

- Aadhaar fulfils no one specific purpose (driving, voting, citizenship, etc.), and has very limited personal attributes. Rather it helps individual prove who they are (identification). Aided by a strong back-end infrastructure, Aadhaar can be used to verify identity in multiple use cases in a cost-effective manner.

- Aadhaar is meant to be unique to an individual. This is achieved by capturing biometrics (all ten fingers and iris scan) at the time of enrolment. These biometrics are matched against all the existing biometrics and is rejected if the person's biometrics already exist on the database. This process is known as de-duplication. Thus, a person can have five ration cards but only one Aadhaar number.

Thus, leveraging Aadhaar by linking it to other welfare cards can help in weeding out fake id cards and also limit the ability of power brokers to capture benefits. To this particular end, UIDAI started interlinking UID with partner databases on a continuous basis. The reach of Aadhaar expanded, while the partners leveraged the strong identity proof that Aadhaar provided. However, this approach ran into some problems, especially in the private sector, and will be detailed later (Challenges section).

Government Backed Digital Identity Systems

Cashless Layer

The cashless layer has its underpinnings in the problem of financial inclusion. The problem had two facets: access to banking and access to banks [6] and payment systems.

Access to banking was haunted by poor processes and lack of documents for identification. This problem was tackled by approving Aadhaar as a valid KYC document. (Aadhaar originally had the address field only for correspondence purposes and it was not meant to be an address proof, but was later repurposed to do so). Aided by it, years later, the launch of the Jan Dhan Yojana by the government led to a huge increase in bank accounts most of which were seeded with Aadhaar numbers [7].

The second problem of access to banks and payment systems paved the way for the cashless layer. It was aided not only by the expansion of Aadhaar and the huge number of bank accounts opened under Jan Dhan scheme but also the massive expansion of mobile subscription across India. It came to be known as the JAM trinity (JanDhan, Aadhaar, Mobile).

NPCI
भारतीय राष्ट्रीय भुगतान निगम
NATIONAL PAYMENTS CORPORATION OF INDIA

National Payments Corporation of India (NPCI) which aimed to create the infrastructure for an affordable payment mechanism became the perfect platform to set up mobile applications for expanding inclusion through electronic transfer of funds, subsidies, micro-ATMs and so on.

NPCI established the two main pillars of the cashless layer: Direct Benefit Transfer (DBT) and the Aadhaar enabled Unified Payments Interface (UPI). While DBT aided a cashless transfer of subsidies directly to the bank accounts, cutting leakages and corruption, UPI formed a modernised payment system for easy, electronic payment between any 2 bank accounts.

Government Backed Digital Identity Systems

Paperless Layer

While Aadhaar had looked at solving the problem of unique identification, through the issuance of UIDs (which leveraged biometric deduplication), it didn't replace the myriad identity documents widespread in the ecosystem, e.g., PAN, passport, education certificates, etc. While Aadhaar enabled digital authentication through the use of biometrics or mobile phones, the lack of these other essential documents on a digital platform presented a problem for digitization of numerous processes.

DigiLocker was envisaged as a solution to pave the way for digitalization of the various documents and enabling users to access them on a single platform, the citizen's wallet. Launched in 2015, it forms the paperless layer of governance in the digital stack, aimed at reducing administrative burden by minimizing the use of paper documents.

DigiLocker has tied up with a number of government issuers and includes credentials such as PAN, Driving Licenses, educational marksheets, pension certificates etc. Recently it has also onboarded a number of private issuers, mostly in the education and FinTech sectors.

How is DigiLocker used?

A citizen can sign up on DigiLocker by providing their Aadhaar details. They can subsequently pull documents issued to them by entering some unique identifier of the credential. For example, to pull their CBSE marksheet the user has to enter their CBSE roll number. After successful matching of their demographic details, name and DoB as per Aadhaar, the e-document is pulled into the citizen locker.

A user can sign up without Aadhaar details. However, doing so will take away all the core functionalities of DigiLocker.

There are three important details to note-

- An e-document is a digitally signed electronic document. A unique document URI is mandatory for every document. This unique URI can be resolved to a full URL to access the actual document in appropriate repository
- To make an e-document available on DigiLocker, an issuer has to create a digitally signed document with a unique URI and upload it to a shared repository or tie up with a repository provider for storing the document and making them available online
- When a citizen pulls an e-document, what is made available is the link to the original e-document (called a URI) shared by the issuing agency in the federated repositories [8].

Government Backed Digital Identity Systems

Consent Layer

One of the more recent developments in space of digital identity in India, has been the introduction of the consent layer, primarily in the form of Account Aggregators in the BFSI space regulated by the 4 main regulators, RBI, SEBI, IRDAI and PFRDA.

Breaking open the silos

An individual's or enterprise's data is spread across silos in banks, telcos, healthcare institutions with no framework in place for them to aggregate and share with their benefactors. As a result, there is still friction in accessing data and a large amount of data is not effectively leveraged.

The consent layer aims to provide a mechanism for the user to share their data with service providers, with proper consent. When hitherto closed data becomes open, users will have more freedom to derive empowerment and value from it. From a business perspective, the greater availability of data will lead to a reduction in information asymmetry and a corresponding increase in efficiency.

To this end, RBI has created a new class of regulated entities called Account Aggregators (AAs) to serve as conduits through which data principals will manage their consensual data flows.

An **Account Aggregator (AA) [23]** provides data to a customer or Financial Information User (FIU) from a Financial Information Provider (FIP) based on the user's explicit Electronic/Digital Consent. An AA merely acts as a conduit between FIUs and FIPs and does not process the data. An AA does not and cannot store any user's data – thus, the potential for leakage and misuse of user's data is reduced.

This model is expected to be replicated to other industries opening up marketplaces of data (credit, skill, etc.) benefiting both individuals and organisations.

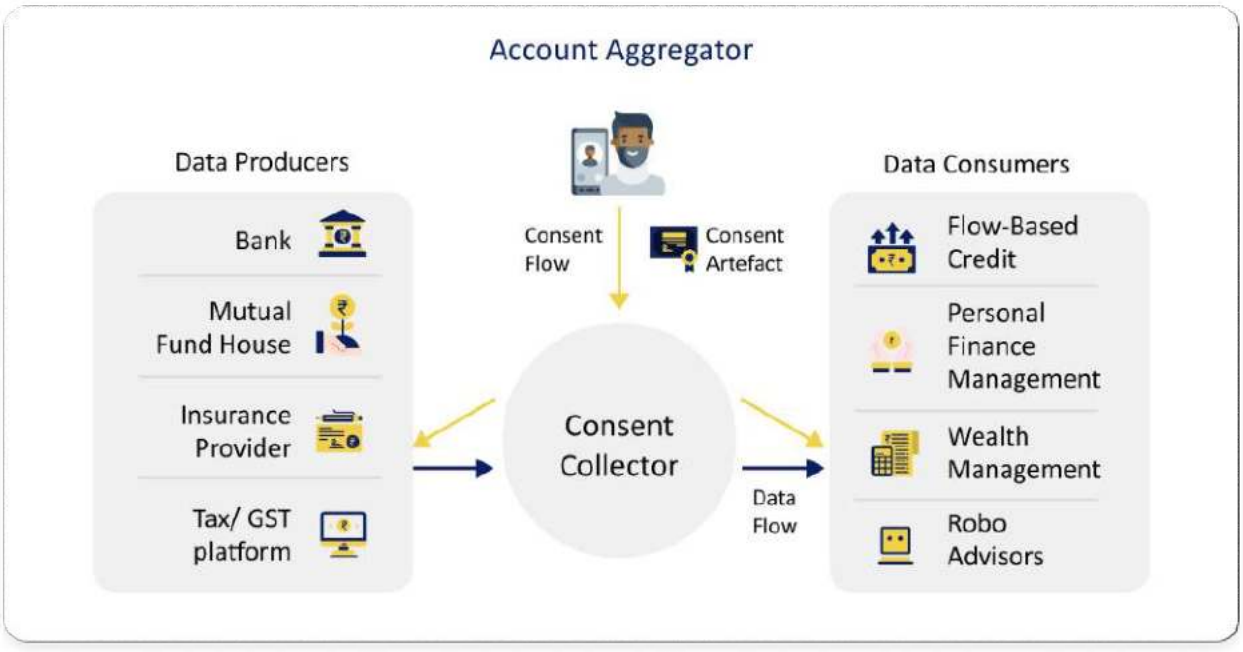


Figure 6: An illustration on how account aggregator system works

Government Backed Digital Identity Systems

Economic Potential of Digital India

The Indian economy is primed for a transformation in its data landscape, with digitization leading to a rapid proliferation in the amount of data generated by each business and individual.

Propelled by the falling cost and rising availability of smartphones and high-speed connectivity, **India is already home to one of the world's largest and fastest-growing bases of digital consumers** and is digitizing faster than many mature and emerging economies.

The public sector, especially the Digital India programme, has been a strong catalyst for India's rapid digitalisation.

1.2 billion Aadhaar has enrolled 1.2 billion people since it was introduced in 2009, making it the single largest digital ID program in the world.

92.28 million DigiLocker has onboarded 1410 issuers and is giving instant digital access of various government issued documents to its registered users.

₹8.3 lakh crore UPI has made giant strides in electronic transactions clocking no less than 461 crore transactions worth ₹8.31 lakh crore in January 2022.

The consent layer aims to replicate this with data, allowing consented data flows outside the walled gardens (silos), unlocking enormous business value. All of this will be backed by Aadhaar, which gives the guarantee of unique identification.

These changes represent a huge economic potential for India. According to a Mckinsey report, India's newly digitizing sectors have the potential to create sizable economic value by 2025: from $130 billion to $170 billion in financial services, including digital payments; $50 billion to $65 billion in agriculture; $25 billion to $35 billion each in retail and e-commerce, logistics and transportation; and $10 billion in energy and healthcare. Digitalisation of more government services and benefit transfers could yield economic value of $20 billion to $40 billion, while digital skill-training and job-market platforms could yield up to $70 billion. However, they cautioned that while these ranges underscore large potential value, realization of this value is not guaranteed: losing momentum on government policies that enable the digital economy would mean India could realize less than half of the potential value by 2025 [9].

Hence, it is essential to study the challenges these government policies face, as they might have significant economic impact. In the next section, we'll discuss the various court verdicts which have restricted the scope of Digital India and their implications for the future. We'll also analyse in detail the various dangers with this new system of identification and the envisioned data & money flows.

Section 5

Challenges

- Legal Challenges
- Privacy
- Security
- Single Point of Failure
- Case Study: Digilocker

Integration of digital ecosystem with the economy

The Digital India programme has without any doubt led to a transformation of the whole identity ecosystem. The delivery of public services to citizens, onboarding of customers by the telecom sector, etc., have all undergone a massive change. The newly emerging FinTech firms, too, have largely depended upon ease provided by Aadhaar authentication to onboard customers, with many even directly based on the data generated by the digital ecosystem, eg., UPI transactions. These trends are likely to be accelerated as the consent layer and paperless layer mature. With the National Digital Health mission, health data is also likely to be integrated with the digital ecosystem with strong linkages to Aadhaar. Similarly, as policymakers and technocrats try to create new opportunities for Digital India, numerous sectors are expected to leverage the digital ecosystem to efficiently handle consumer identity and data.

As digital identity becomes the lifeblood of the communication system between citizens, firms and the government, these entities are likely to base their interactions and future plans on the assumption of a well-functioning identity ecosystem.

Given the stakes involved, we should analyse in advance not only the opportunities but also the dangers to the functioning of our critical infrastructure. This is precisely what we intend to do with regards to the digital identity ecosystem.

Outline of the challenges

We will firstly consider the **legal obstacles** which the ecosystem might face (and even currently faces). We will then look at some **general points of concern** with the identity stack being developed (privacy, security and single points of failure). With the help of a few observable trends seen in the status quo, we'll also analyse **how resilient and adaptive the ecosystem is.** Finally, we will evaluate the paperless layer against the issues we have identified, and find the primary challenges which need to be tackled for its success.

Challenges

Legal Challenges

The Supreme Court in 2018 gave a landmark judgement on the validity and applicability of Aadhaar. While holding Aadhaar (for its immense contribution in public welfare) as legally valid, it severely restricted the scope of Aadhaar.

Restrictions imposed on Aadhaar

The court barred any attempt to interlink unconnected silos of information. Within public welfare, Aadhaar identification was restricted to purposes which were backed by law and cautioned that any such 'law' legislated in the future would be subject to judicial scrutiny. Individuals and corporates were prohibited from collecting Aadhaar data. The court clarified that Aadhaar cannot be used even if there is a contract between the business and the individual [15]. This disallowed even voluntary use of Aadhaar in the private sector. The concern stemmed from a danger of coercion, where it would be voluntary by law, but practically would be the only (convenient) choice. The lack of any other convenient way of verifying identity, might further force them to divulge their Aadhaar details at multiple places (not all safe), putting them at risk. Private companies having lower accountability than the government might be more willing to misuse the data provided to create profiles of individuals.

Test of proportionality

The court instituted the test of proportionality (the means should not be excessive for the ends sought) for Aadhaar [16]. Aadhaar enables efficiency but comes with certain trade-offs. The court deemed the privacy of citizens is compromised with the widespread collection of Aadhaar data, especially the unique ID. This is primarily because the UID can possibly enable interlinking various online and offline behaviour of individuals, the connections they have, the interactions they have with them, etc. Thus, it was ruled that Aadhaar cannot be blindly used for all purposes, allowing it to be used only for public welfare schemes and mandating the use of Aadhaar only when backed by law.

Some other observations of the Supreme Court

"The Constitution does not permit the establishment of an authority that in turn through an invasive programme can chain every Indian citizen/resident to a central data bank and maintain lifelong records and logs of that individual,"

"The Constitution of India, when read as a whole, is designed for a nation of free individuals who enjoy a full range of rights and who are entitled under the Constitution to lead their lives without any monitoring or scrutiny or continuous oversight by the state or any of its organs."

"Informational privacy is a facet of the right to privacy. The dangers to privacy in an age of information can originate not only from the state but from non-state actors as well."[17]

Implications for the digital identity ecosystem

The Supreme Court verdict put strong restrictions on linking of Aadhaar number with other identity documents, limited to use-cases of public welfare along with a law backing it (which would be open to judicial review). Further, any process or product which discloses the Aadhaar number will have the same restrictions placed on it.

The highest court instituting a test of proportionality for Aadhaar use will demand a rethink of the trade-offs involved with the implementation of any digital product built over Aadhaar. Any future development of the digital identity ecosystem in India will have to balance the privacy of citizens in accordance with the ground-rules set by the Supreme Court.

Challenges

Privacy

Right to Privacy

A nine-judge bench of the Supreme Court in Puttaswamy judgement recognised the right to privacy as an intrinsic part of the fundamental right to life and personal liberty under Article 21 of the Constitution of India [10].

However, Big Data has revolutionised how data of individuals is collected and processed. Each and every day individuals using or carrying any connected device are emitting data exhausts for analysers. With an exponential surge in computing power driving data analytics, what we view and click has become a business asset with behaviour analytics being used to profile customers and generate intelligent inferences about preferences and behaviours. The digital trail is rich and deep, almost tailor-made for surveillance. Apart from the discourse about the Orwellian state, there is also a sense of vulnerability as individuals are drawn, even pushed, into the connected world, into digitalisation, ripe for private exploitation.

A nearly $3 trillion economy, India is among the world's largest economies, and has over 560 million internet users. The emphasis to promote data sharing through consent managers, cashless transactions through apps, paperless validation through Digi Lockers, and presence-less governance via online systems will find millions more online [11].

In this new context, it is important to define what constitutes privacy. Classically, it is held that, "**Privacy is the claim of individuals, groups, institutions to determine for themselves when, how and to what extent information about them is communicated to others**"[12].

However, with the ability of private players to correlate the data of individuals present in the digital space back to them and profile them, achieving privacy in the digital space has become a herculean task. Given the expansive nature of the digital world, and existence of players beyond the government's jurisdiction (international companies using data of Indian citizens), for the right to privacy to be upheld, rights would have to be explicitly guaranteed and **data protection must be by design and by default.**

While the digital identity ecosystem provides numerous value additions over the physical identification methods like increased efficiency, lower error rates, reducing leakages, etc., it exponentially increases the threats to data privacy. The replacement of peer-to-peer interactions of the user and organizations by centralized management systems, while increasing efficiency, on the flip side also aggregates the data of the individuals and allows the correlation of the data present in the digital space to form profiles of individuals, for private exploitation or harming them.

The Threat of Correlation

The Aadhaar number, given its design and very purpose (unique identification), can be termed as "perfect correlator". While there is a lot of data already present in the digital space, correlating this data takes effort, correlation unfolds over time as imperfect clues accumulate, requiring maintenance as context triggers re-evaluation, and is often probabilistic. But the introduction of unique identification (like Aadhaar, or even digital signatures) will make correlation far easier, far more accurate, and far cheaper to maintain [14].

The digital prevalence of unique Aadhaar numbers, UPI IDs, Health IDs and the consent layer's aim to break open the silos and create data marketplaces, will have to be re-evaluated in light of this threat to privacy. The sharing of unique IDs should be minimised. The steps taken by UIDAI in this direction are discussed in the Recent Developments section

Merging of silos

This concern was echoed in a draft paper for a legislation for privacy far back in 2011:

> *"One of the inevitable consequences of the UID Project will be that the UID Number will unify multiple databases. As more and more agencies of the government sign on to the UID Project, the UID Number will become the common thread that links all those databases together. Over time, private enterprise could also adopt the UID Number as an identifier for the purposes of the delivery of their services or even for enrolment as a customer. "*
>
> - Approach Paper for a legislation on privacy, DOPT, Govt. of India, September '11 [13]

The fear of the linking of multiple databases stems from the consolidated view of individuals this linking offers, not only to the state but also to other entities such as rogue agents, companies (especially big tech), hackers, etc.

Thus, some separation of databases might be desirable for the privacy and security of the citizens.

Most importantly, the challenge of privacy has to be dealt with effectively to remove the restrictions placed by the Supreme Court on Aadhaar and the wider digital ecosystem dependent on it.

Challenges

Security

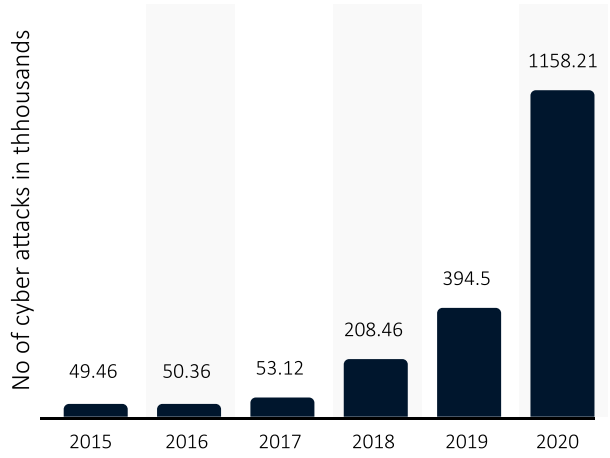
Does it really matter if some data about me is leaked or stolen? I am not a celebrity or politician.

How big is the threat to security? Does it really matter if some data about me is leaked or stolen? Can they create a bank account using my identity, or say buy a SIM card? Can they siphon off money from my accounts? This section will attempt to answer these questions and also explore the measures which can be taken to reduce the vulnerability of individuals and organisations against cyber-attacks, both external and internal.

Cyber-threats

Cyber-security has been a major issue since the dawn of Internet. Misuse of personal identity data predates even cyber-attacks. However, the interconnected nature of the Internet (aptly called the Web) makes it easier to illegally access massive troves of digital data.

Incidents of cyber attacks across India from 2015 to 2020 *(in 1000s)*



And this trend is increasingly becoming worse, with numerous attacks on critical infrastructure like electoral rolls, power, banks, telecoms etc.

Equifax (2017)- Equifax a credit reporting agency, was breached by cybercriminals, who accessed details such as people's names, Social Security numbers, birth dates, addresses and, in some instances, driver's license numbers and credit card numbers of about 145.5 million Americans [24].

Facebook (2019)- A user in a low level hacking forum exposed data of over 533 million Facebook users from 106 countries. It included their phone numbers, Facebook IDs, full names, locations, birthdates, bios, and- in some cases- email addresses. The entire dataset had been posted on the hacking forum for free, making it widely available to anyone with rudimentary data handling skills [25].

CDSL Vulnerability (2021)- A cyber security consultancy startup (CyberX9A) reported a vulnerability at a CDSL subsidiary, CDSL Ventures Limited (CVL) that had exposed personal and financial data (such as investors name, phone number, email address, PAN, income range, father's name, date of birth etc) of over 4 crore Indian investors twice in a period of 10 days. CDSL said that CVL had taken immediate action and mitigated the vulnerability [26].

Vulnerability of Indian identity ecosystem

As more and more personal data of individuals is pumped into the digital economy, there is a considerable threat of misuse of this data. The current digital frameworks in India largely depend on centralised and federated repositories of data. UIDAI's Central Identities Data Repository (CIDR) stores all the Aadhaar data of the 1.2 billion-plus citizens [15]. While highly secure and isolated, other repositories maintained by different central agencies and states with much more connections might be less secure and more vulnerable.

To answer the question of how much damage can security lapses or data leaks can do, we'll have to look at the ways data can be compromised and what are the ways in which this data can be used (which might be harmful to the individual and the society).

While there are numerous forms of cyberattacks, for the sake of brevity, we can classify the attacks being of two types –

- attacks directly aimed at the central data servers of organisations
- attacks which indirectly access information from users on false pretences and are used for malicious purposes.

Threat to centralized servers

Addressing the former first, the attacks can be directed at the issuer (identity issuing organisations), verifier (organisations which store the data of their customers, employees or beneficiaries) or some intermediary which handles the identity transactions. In the status quo, the data in these organisations are usually stored in isolated silos and unconnected to major networks. As the digital ecosystem taps into these data resources, the threat of massive data breaches will increase [27].

Effort in the data protection bill to minimise data held by organisations is an important step in this direction. Insider attacks also will need to be tackled by constraining access of employees and administrators to the personal data of individuals, ensuring compliance by design than by law.

Indirect attacks

Indirect attacks refer to attacks like phishing, identity theft, and spoofing. We are all too aware of the age-old scams of duping people to reveal sensitive information about themselves, like account details, important security questions, etc. With the increasing use of the digital ecosystem to access important services easily and quickly, often requiring just the smartphone to make all kinds of payments and identity transactions, these kinds of attacks are also increasing in scale. It gets easy for fraudsters to establish trust with increase in access to individual's personal information.

Identity Theft

The sharing of digital documents (in the physical or digitally signed format) is rife with issues of identity theft. Once a person shares their credential, the person possessing the document can not only view and exploit the data, they can also claim ownership of it and use it for further purposes not intended by the true owner. This mandates the presence of additional authentication factors like biometric, phone number (through OTP), etc.

KYC Fraud

Phishers, impersonating official bank representatives, often send out messages to users asking them to update their KYC details to prevent their account from being blocked. In many such cases users have divulged their KYC information, or worse their login details to the phishers which leads to their accounts becoming extremely vulnerable (often ending up with huge financial loss for the victims). A lack of standard way of communication is primarily to blame [28].

Pension Fraud

An extension to the KYC fraud, in 2021 fraudsters impersonating pension-disbursing authorities, asked for OTPs from senior citizens for updating Digital Life Certificate failing which they wouldn't receive their pension. Instead they used OTPs to initiate funds transfer from their accounts. They were aided by the access they had to details such as date of retirement, pension pay order, Aadhaar number, communication address, email ID, among others, which made them much more trustable to the victims. Investigation is still underway to find out how the data leaked to the cyber criminals [29].

Core vulnerability

The present diverse mediums of communication and service delivery make the consumers more vulnerable to attacks of such kind. This is because with so many different mechanisms of communication used, people are often confused if the person they are talking to (or the website they are accessing) are what they are claiming to be (look at Pension Fraud). On the other hand, not having diverse media can lead to exclusion of individuals. A balance might be to impose universal standards allowing the existence of multiple players and multiple credentials and yet maintaining a strong trust ecosystem where people can easily verify the identity of the information seekers.

The process of digitalisation has just begun in India. As more and more processes and service deliveries are digitalised, so will the frequency and magnitude of such incidents. Efforts should be directed to address these concerns and implement technologies which are secure and privacy-preserving by design.

The key element in all these different frauds is the ability of the fraudsters to falsely claim themselves to be trusted representatives of a specific organization. Furthermore, each bit of data counts. Data leaked in breaches are often published publicly or traded in the dark web and are further used for pulling off these kinds of spams, by using the data gained on individuals to establish trust with them. Each piece of information they have makes the fraudster more trustable. Imagine a person calling with all your Aadhaar and bank account details and claiming they are a representative of the bank. You will tend to trust them as bank employees

Challenges

Single Points of Failure

> ***"It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so."***
>
> *Mark Twain*

The single point of failure being addressed here is not like the one frequently encountered in IT infrastructure. Rather it refers to the problem of extreme dependency on a single application/ platform. To date, Aadhaar is the only digital credential that can be seamlessly used for online, presence-less and instant authentication of individuals. Other important credentials like PAN, Voter ID, etc., cannot be used in such a manner.

Threat to broad economy

In 2018 after Supreme Court's Aadhaar verdict circulars were released by the Telecom Ministry barring use of Aadhaar for onboarding customers. Similarly, many private banking and FinTech companies were restricted from using Aadhaar for authentication while onboarding customers. This led to huge hue and cry and disruptions in both the industries. Soon, backed by legislation of new laws (and some amendments) these industries were able to use Aadhaar as before. But this disruption gives us a few warnings.

Any such incident in the future will have an adverse impact on digital identification, as firms will be unable to instantly verify documents and, in many cases, might not even have an alternate channel of identification. The government service delivery might be similarly impacted.

Thus we need to create a more flexible platform which can leverage multiple credentials from different institutions.

Why worry about single points of failure?

The world is undergoing a massive change with technological disruptions and structural changes taking place. When we couple this with the heavy dependence of not only the entire identity ecosystem and but the larger business and public ecosystem on critical identity infrastructure, any failure can lead to disastrous consequences.

Given the data privacy, security and judicial issues, there are a number of core challenges that threaten the current ecosystem, and thus the problem of single points of failure has to be addressed. There's a need to reduce the dependency of the digital identity ecosystem on a single credential (Aadhaar) and other central dependencies. There's a need to facilitate instant verification of other credentials independent of Aadhaar.

Challenges

Case Study: Digilocker

This case study attempts to analyse the various challenges of the paperless layer. The paperless layer currently works through DigiLocker, a locker of documents of citizens. It has a federated storage and contains the URIs for the identity documents of Indian citizens [19]. It is much less used than both the presence-less layer (Aadhaar) and the cashless layer (UPI). We'll first discuss the potential reason for its low traction among Indian citizens.

Ease of use

Both Aadhaar and UPI have been tremendous successes because of the ease and convenience they afford to people. The process of onboarding is easy and every transaction which requires them is also simple and can be easily followed by anyone. The paperless layer is not as easy to use. Some of the key issues faced by the different stakeholders are:

- Holders – To pull any issued document, an individual has to enter some information about the credential they are accessing (such as the roll number for CBSE marksheet). This makes the experience troublesome for the user. Often, people require the document through DigiLocker, precisely when they don't have access to the original credential. In such a case, demanding details of the credential presents an impediment to the use of DigiLocker. However, this problem cannot be solved without extensive linking of Aadhaar (serving as the primary key), which comes with its own set of problems.
- Verifiers - Perhaps, because of the numerous issues outlined above, DigiLocker restricts the onboarding of verifiers through a rigorous process, where they have to register with DigiLocker as a verifier, file an application and so on. Apart from that, DigiLocker heavily depends on the API infrastructure, where the verifier has to integrate the APIs with their existing workflows, which is a huge challenge for small and medium enterprises. This is a far cry from the ubiquitous and easy use of UPI by vendors, using simple mobile applications.

Finally, DigiLocker was never able to leverage the network effects which made Aadhaar and UPI so successful. With low usage by citizens and the consequent low levels of acceptance by verifiers (which furthers dampens usage by citizens) the use of DigiLocker is widely limited.

Having talked about the reasons behind low traction of DigiLocker, we'll also talk about some potential security and privacy issues with DigiLocker, and how they restrict the limits within which it can be used.

Centralized storage

As more and more issuers become a part of the ecosystem, the value stored in this federated storage increases, which will make it more enticing for sophisticated hackers. Any compromise of the system, will suddenly have millions of identity documents out in malicious hands. The hacks into the repositories of the tech giants (such as Twitter) and big credit firms (such as Equifax) shows that this is a real possibility if the data stored is valuable enough [20].

Another salient aspect, **is the access management of the data.** The use of identity information for nefarious reasons by insiders can give avenues for targeted discrimination and information leaks [21].

Danger of identity information on digital databases

Further the lack of data minimisation while sharing the identity information with verifiers makes the verifier database a valuable target. In the event an attacker breaches a few companies' databases, the presence of correlated identifiers can help the attackers paint a bigger picture of the individuals. This problem is made more serious by the tendency of the attackers to publish details on the web. Even later attempts at taking this data down, does little as it gets stored offline quickly. Thus, each and every data breach compounds the amount of data of users compromised. This leads to formation of richer profiles of users, which can be further used for malicious purposes. To prevent this, the ecosystem should allow granular sharing of data, where the verifier can get only the information which they require. The platform should also utilise uncorrelated identifiers to curb the profiling of individuals.

There are also, thus, some inherent limits to the kind and amounts of data which can be stored or shared using the current digital stack because of the problems detailed above. It'll be extremely hard to securely store sensitive personal information of individuals on the federated repositories. Further, storage of the credentials generated from all the wide variety of interactions an individual has on a daily basis is fraught with dangers too, as it leads to a much more nuanced profile of an individual. The presence of sensitive information and high frequency data make the system very valuable and thus vulnerable to sophisticated external attacks. Even minor attacks on verifier databases can strengthen the threat of correlation and identification of the individual. This can also include non-malicious mistakes on the part of the verifiers [22].

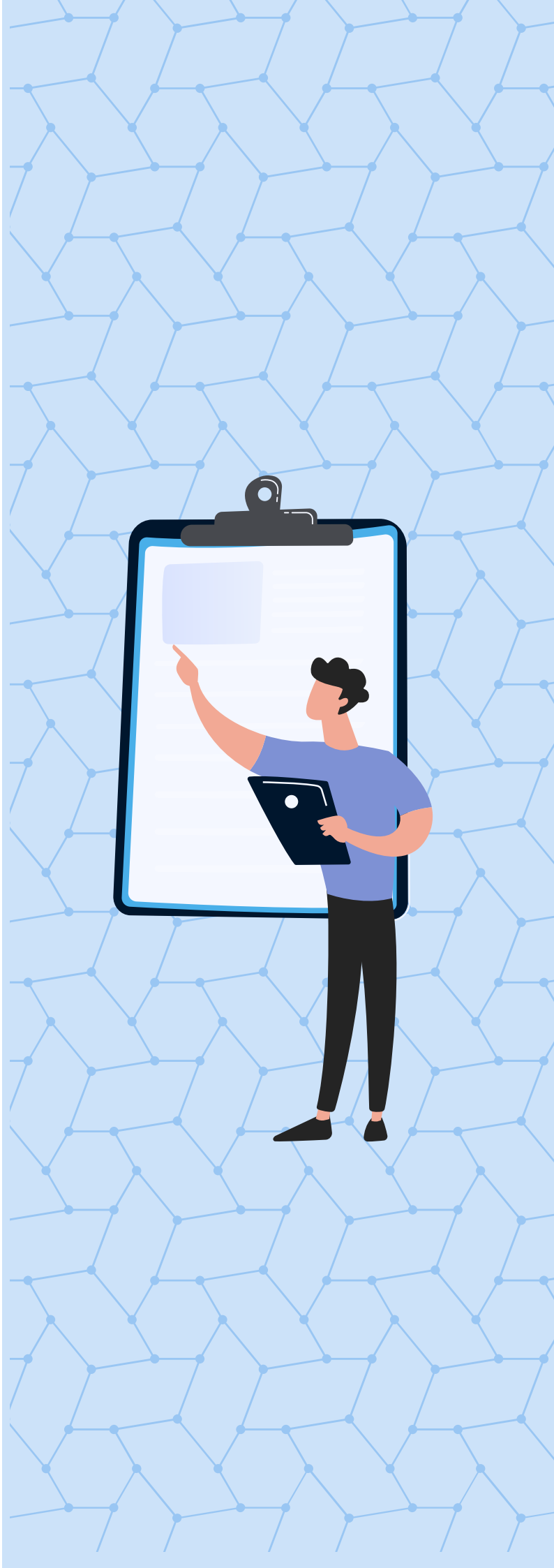
Interaction with private players

Similar challenges can plague the expansion of the digital ecosystem while onboarding the different players in the private economy. Private businesses also hold a lot of data of their customers which if the consumers can safely leverage can completely transform the landscape of data sharing in India. However, apart from the obvious issues talked above, gaining the trust of private players, to become issuers (and not just gain benefits of easy verification) will also be challenging. Right now, they are primarily using the data of customers on DigiLocker to meet certain regulatory requirements. They might not be willing to put up their own data (credentials). They are more comfortable with the direct interactions they typically have with their customers.

Section 6

Future of Digital Identity in India

- Recent Developments
- Guiding Principles
- Final Take-aways



Recent Developments

The Supreme Court privacy verdict in 2017 and Aadhaar verdict in 2018 spurred extensive conversation and actions to better the digital identity ecosystem in India and to strengthen Aadhaar against the various privacy and security concerns. The strategies can be primarily categorized as:

- Immediate steps to mitigate privacy concerns
- Creation of a data protection bill
- Development of an Aadhaar 2.0

Immediate Steps

As pointed out in the "Privacy" section, the existence of a ubiquitous unique identifier constitutes a threat to the privacy (and security) of individuals. Recognizing these issues UIDAI has come up with multiple strategies to address them-:

1. Masked Aadhaar – UIDAI offers an Aadhaar credential in which the Aadhaar no. is masked (only the last 4 digits are shown), very similar to credit card number masking. This credential can then be used in less “secure” places. The firms won’t get access to the UID and the harms may be mitigated. However, the trade-off is that the uniqueness of Aadhaar stems from the unique identification of individuals. Aadhaar can be conceptualized as digital identification system with the UID as the username and the biometrics (or OTP on mobile) as the password. Thus, Aadhaar authentication cannot be performed with a masked Aadhaar. Without UID the core functionality of Aadhaar is lost.

2. Virtual ID (VID) - VID is a temporary number which an Aadhaar holder can generate, which is lost as soon as another VID is generated. The mapping of UID to VID is maintained by UIDAI. An Aadhaar holder can thus give a VID to the verifier as a replacement for UID. The verifier can request for Aadhaar authentication using this VID. UIDAI, which has the mapping for UID to VID, can subsequently find the UID of the individual. This solves the problem of Aadhaar authentication faced by masked Aadhaar while still hiding the Aadhaar number from the verifier. After the authentication, it is useless to store VID as it can be changed by the user. It is yet to be seen how this paradigm will be extended to other unique IDs (such as Health ID, UPI ID, etc.). It might be cumbersome for an individual to generate VIDs before different identity transactions. Thus effort must be made to make the entire process seamless for the citizens. Any complications might make it unfeasible for citizens to use VID, making them vulnerable to various privacy and security threats. Deeper integration of uncorrelatable IDs might offer a solution (by possibly making the generation of these IDs default rather than as an additional step to be undertaken by the user).

Data Protection Bill

The right to privacy verdict led to the constitution of a committee headed by Hon. Rtd. Justice B.N. Srikrishna to help introduce laws related to data protection and privacy, whose recommendations led to the Personal Data Protection Bill (PDPB). This bill aims to alleviate the plethora of privacy and security concerns and to create a proper legal framework of personal data usage which applies to both government and private agencies. The goal is to make Aadhaar accessible to the private sector too, a major impediment to the creation of a unified digital identity ecosystem. A private sector regulated with a data protection bill would mandate only legitimate use of the data and provide a framework for penalizing entities which violate the privacy of individuals.

Ideally, the legislation for data protection should

have come first laying down the foundation for what is privacy and how it is going to be protected, and then the mechanisms for applying it to governance should have been designed. They would then naturally fulfil the basic principles of privacy protection. Since, we already have a massive digital identity infrastructure and the legislation has still not been passed by the Parliament, we are in a position where both the PDPB and the digital identity infrastructure are adapting with each other, which is perhaps inevitable. While technology informs the legal frameworks, legal frameworks typically define the contours within which technology can be applied and also influence development of technology in certain directions. It is within this context we'll have to look at the development of Aadhaar 2.0.

Aadhaar 2.0

So, what are these contours for the digital mechanisms used by Aadhaar? Perhaps the best answer lies in the 2018 Aadhaar verdict, which laid down the ground-rules for any implementation of Aadhaar in India, carefully balancing the objectives of the state (and society) and the rights of individuals, especially the right to privacy. Attention is to be paid also to the minority judgement of the court, aptly pointed out by Hon. Justice A.K. Sikri in his keynote speech at an Aadhaar 2.0 workshop. He said, "Privacy is one of the most cherished fundamental right of any individual. That is why in brief I have recapitulated the contours of this right to privacy. If this right is infringed by the kind of profiling which the dissenting judge expressed apprehension about, that may be a dark day for this nation."

The future developments of Aadhaar and the digital ecosystem at large should attempt to allay these apprehensions. Doing so might not only reduce the chances of the tragedies feared, but might also widen the contours within which Aadhaar can be applied. An Aadhaar 2.0 which more effectively protects the privacy and security of individuals, and reduces the scope for unwanted correlation and profiling might escape the restrictions currently imposed on Aadhaar, for example the prohibition of use in the private sector. Does Aadhaar 2.0 mean a fundamental change in Aadhaar's working? Most definitely not. Aadhaar 2.0 simply refers to a changed context within which Aadhaar is used. While original Aadhaar infrastructure will be the cornerstone of unique identification, development of other layers and new techniques might be utilized to yield a privacy preserving digital ecosystem. A possible solution in the form of VIDs has already been discussed, but is found to be lacking a larger ecosystem, i.e. creating VIDs for the numerous identity transactions might be cumbersome for individuals. The idea of uncorrelatable IDs, however, is something which can be explored more deeply. Implementation of other such mechanisms should be studied to improve the identity management system in India in our view.

What needs to be improved?

Technology will play a key role in the development of Aadhaar 2.0. However, it is important to understand the principal vulnerabilities in the digital ecosystem to direct efforts in a focused manner.

The previous three sections explained these in detail and can be summarized as follows-

- Privacy – The presence of correlated IDs presents a threat to the privacy of individuals as it gives an opportunity of surveillance and profiling.
- Security – Data of individuals stored in huge centralized or federated repositories is a significant vulnerability which can be (and are being) exploited by various malicious actors around the world.
- Single Points of Failure – The dependence on a limited set of service providers and credentials should be reduced.

Advances in technology should look to solve these 3 critical vulnerabilities for a more widespread use of the government backed digital identity ecosystem.

Guiding Principles

Before we conclude this paper, we'd like to talk about some principles, which if followed while designing improvements might help alleviate the challenges being faced by the Indian digital identity ecosystem. While not a solution in itself, these principles might form a guiding light for both identity technologists and policy-makers as they set out on a path to create a long lasting digital identity ecosystem for India. Consideration of these principles will help avert the dangers we have expounded in this paper. They are aimed at solving the issues we have highlighted in the previous section.

The 7 key principles for a resilient digital identity are – **control, access, consent, transparency, minimization, portability and protection.**

Access, Control & Consent

These three principles are intrinsically tied to the holder of the identity and if followed would strongly **protect the privacy of users.** The user must have complete access to all his/her identity data and where they have been shared. They should have complete control over the sharing of this data, and their consent should be necessary for any transfer of claims relating to them (any sharing of identity data relating to them).

Transparency, Minimization & Portability

To protect the **security of the overall ecosystem**, it is essential that the system is transparent. Being open about what decisions are made and making explicit about how decision-making works ensures that all the governance and use of a platform, network, or system are inspectable and understandable by all stakeholders.

Further the sharing of data should be minimized, such that only the data required for any particular process is disclosed. Sensitive details (like DoB) should be shared sparingly. Information and services about identity must be transportable. Identities must not be held by a singular third-party entity. A decentralized storage of identity information reduces security vulnerabilities.

Protection

Finally, for Holders, protection is essential to maintaining individual sovereignty over their digital identities. Holders must be able to (inter)act in a climate of confidence, certain that their digital identities are insulated from circumstantial external pressures (e.g. network failure, system abuse, ambiguous policy changes). The principle of protection revolves around the notion of resilience — an identity management system must shield (the rights of) individual users against arbitrary or inscrutable changes in (whatever) environment their data lives and operates. This can help mitigate the issue of **single points of failure.**

A lot of these principles are not new, and have been previously suggested by committees for data protection and privacy (such as consent and minimization).

It is our firm belief that adhering to these principles will help create an identity management ecosystem which will stand the test of time and create a strong confidence for citizens, businesses and governance entities to interact over a unified identity platform, ushering in a new era of secure, private and easy digital governance.

Future of Digital Identity in India

Final Take-aways

In the preceding sections we have delved deep into the Indian identity ecosystem, and it is time to summarise what the different stakeholders can take away from this whitepaper.

Policymakers have consistently played a strong role in shaping India's identity infrastructure, enabling public service delivery for the billion plus population of India. As India transitions to digital service delivery and digitalisation of the broader economy, the policymakers will face unique opportunities and challenges. To aid them in their efforts, this paper gives a brief review of the Indian identity ecosystem and the digital initiatives undertaken so far, highlighting the challenges overcome and the core functionalities of the different layers of Indian digital identity infrastructure. In addition, it presents a critical analysis of the primary stumbling blocks to establishing a unified identity platform for India. It provides a guide to what policymakers can look for while evaluating technologies and making accommodations in the legal frameworks to allow critical technological advancements to flourish.

The **tech community** has been another important force behind the current digitalisation of identity systems in India. This paper provides a social context for identity within which technology will have to be applied, tackling complex issues like privacy of individuals in this digital age and inclusion of citizens from different socioeconomic backgrounds. It elaborates the directions in which technology will have to develop and the basic principles which have to be considered to build a resilient identity ecosystem in India.

Government agencies and corporates have to take critical decisions regarding which identity systems to adopt and what features to push for from the tech community. As entities transition from the physical method of identification to a digital one, this paper provides a framework for evaluating digital identity systems and understanding what issues they might face before adoption itself. It tries to answer questions like what makes an identity system future resilient and how to deliver the maximum private and public benefits, allowing entities to make better informed decisions.

Finally, the **general public** is perhaps the most important stakeholder. It is in their interest that all the above stakeholders work. In the end, an identity ecosystem is a public good. This paper tries to explore the various aspects of identity and how people are affected by the rapid acceleration of digitalisation. It strives to provide a nuanced picture of digital identity and the various trade-offs involved while designing an identity infrastructure. An informed civil society is the backbone of every endeavour, and it is our hope that this paper will add to intelligent conversations regarding the future of digital identity in India.

In an **upcoming companion white paper** we plan to describe how we envision a digital identity system which satisfies the requirements outlined in this whitepaper.

References

1. *www.news18.com/news/india/fraud-involving-aadhaar-numbers-sim-cards-busted-3-held-in-madhya-pradesh-3056708.html*
2. *www.indiabudget.gov.in/doc/bspeech/bs198687.pdf*
3. *www.hindustantimes.com/india-news/only-15-paise-reaches-the-needy-sc-quotes-rajiv-gandhi-in-its-aadhaar-verdict/story-I8dniDGXF6ksulggTDgb9L.html*
4. *Collated from Annual Economic Surveys*
5. *www.timesofindia.indiatimes.com/india/Rajiv-was-right-Montek-says-only-16p-of-Re-reaches-poor/articleshow/5121893.cms*
6. *Financial Inclusion in India: An Assessment, RBI www.rbidocs.rbi.org.in/rdocs/Speeches/PDFs/MFI101213FS.pdf*
7. *www.business-standard.com/article/economy-policy/so-far-80-bank-accounts-60-mobile-connections-linked-with-aadhaar-uidai-118030400180_1.html*
8. *www.img1.digitallocker.gov.in/assets/img/DigiLocker-Intro.pdf*
9. *https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/digital-india-technology-to-transform-a-connected-nation*
10. *www.thehindu.com/news/national/privacy-is-a-fundamental-right-under-article-21-rules-supreme-court/article19551224.ece*
11. *www.mckinsey.com/business-functions/mckinsey-digital/our-insights/digital-india-technology-to-transform-a-connected-nation*
12. *Alan Westin, "Privacy and Freedom" 1968*
13. *www.documents.doptcirculars.nic.in/D2/D02rti/aproach_paper.pdf*
14. *www.evernym.com/blog/well-be-correlated-anyway/*
15. *www.bloombergquint.com/aadhaar/aadhaar-supreme-court-saves-the-baby-drains-the-bath-water-by-shankkar-aiyar*
16. *www.barandbench.com/columns/proportionality-test-for-aadhaar-the-supreme-courts-two-approaches*
17. *The quotes are taken from https://economictimes.indiatimes.com/news/politics-and-nation/aadhaar-verdict-legal-but-limit-use-to-government-benefits-says-supreme-court/articleshow/65973337.cms?from=mdr*
18. *www.uidai.gov.in/ecosystem/authentication-ecosystem/operation-model.html*
19. *www.img1.digitallocker.gov.in/assets/img/technical-specifications-dlts-ver-2.3.pdf*
20. *www.wired.com/story/inside-twitter-hack-election-plan/ https://www.ftc.gov/enforcement/cases-proceedings/refunds/equifax-data-breach-settlement*
21. *www.edition.cnn.com/2019/01/28/health/hiv-status-data-leak-singapore-intl/index.html*

22. www.economictimes.indiatimes.com/news/politics-and-nation/210-government-sites-found-displaying-aadhaar-details-pp-chaudhary/articleshow/59667922.cms
23. https://sahamati.org.in/
24. https://www.consumer.ftc.gov/blog/2017/09/equifax-data-breach-what-do
25. https://www.businessinsider.in/tech/news/533-million-facebook-users-phone-numbers-and-personal-data-have-been-leaked-online/articleshow/81889315.cms
26. https://www.business-standard.com/article/companies/data-breach-at-cdsl-s-kyc-arm-exposed-43-9-mn-investors-details-cyberx9-121110700431_1.html
27. https://www.ndtv.com/india-news/cbi-probes-pension-fraud-worth-crores-misuse-of-migrants-data-during-covid-pandemic-2537084
28. https://economictimes.indiatimes.com/wealth/save/sbi-alerts-about-online-kyc-fraud-tells-customers-how-to-keep-bank-account-safe/articleshow/84370088.cms
29. https://www.thehindu.com/news/national/railways-alert-banks-of-online-fraud-targeting-pensioners/article35370891.ece

Attributions

1. “https://www.freepik.com/vectors/abstract" Abstract vector created by vectorjuice
2. “https://www.freepik.com/vectors/business" Business vector created by vectorjuice
3. “https://www.freepik.com/vectors/background" Background vector created by freepik
4. “https://www.vecteezy.com/free-vector/card" Card Vectors by Vecteezy
5. “https://storyset.com/illustration/id-card/bro” Illustration by Storyset
7. “https://www.freepik.com/vectors/geometric-illustration" Geometric illustration vector created by GarryKillian
8. “https://www.vecteezy.com/free-vector/greece" Greece Vectors by Vecteezy
9. "https://www.freepik.com/vectors/business" Business vector created by gstudioimagen

About Trential

Incubated at IIT Kanpur, Trential was started by researchers, engineers, and leaders who created, innovated, and adopted cutting-edge blockchain technology under the aegis of the National Blockchain Project. Trential is building innovative products to strengthen trust in the digital world by harnessing the power of blockchain.

Explore Trential at:
https://trential.com

Our social media handles:

National Blockchain Project:
This project has been funded by the National Security Council Secretariat to develop e-governance solutions using blockchain technology.

To know more about National Blockchain project:
https://blockchain.cse.iitk.ac.in/