



ITUKALEIDOSCOPE

NANJING 2017

Challenges for a data-driven society

Open Data & Digital Identity: *Lessons for Aadhaar*

Amba Kak

Mozilla Foundation (amba@mozillafoundation.org)

Smriti Parsheera

National Institute of Public Finance & Policy

(smriti.parsheera@nipfp.org.in)

Nanjing, China

27-29 November 2017





The case for more open data

- What is “open data”? : *openness is on a spectrum*
- Open data as a public good - accountability, research, business decisions
- Moving beyond government data BUT - *Are there incentives to “open”?*
 - Government: right to information laws, open data policies
 - Private sector: voluntary frameworks, MiData case study
- **Why is Aadhaar - *the largest biometric identification system in the world* - an appropriate case study?**



An introduction to Aadhaar

- What is Aadhaar?
- Enrolment of residents:
 - 1.18 billion people enrolled (Nov)
 - 85% of Indian population in 2017*
- Applications of Aadhaar:
 - Distribution of welfare services/ payments
 - Financial sector (the “JAM” trinity)
 - Telecom sector
 - Emerging uses: health, education.

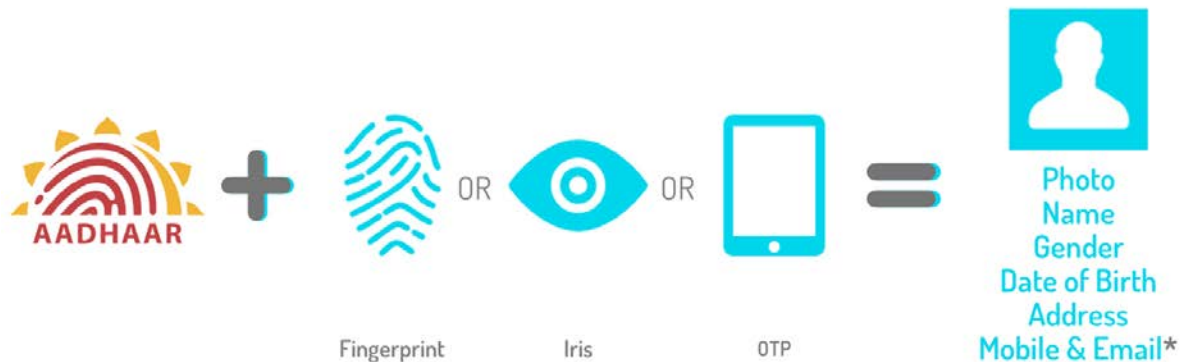
Source: *State of Aadhaar report

INFORMATION	REQUIREMENT
DEMOGRAPHIC	
Name	Mandatory
Gender	Mandatory
Date of birth	Mandatory
Residential Address	Mandatory
Mobile Number	Optional
Email ID	Optional
BIOMETRIC	
Photograph of face	Mandatory
Fingerprints (all 10 fingers)	Mandatory, if available
Iris captures of both eyes	Mandatory, if available

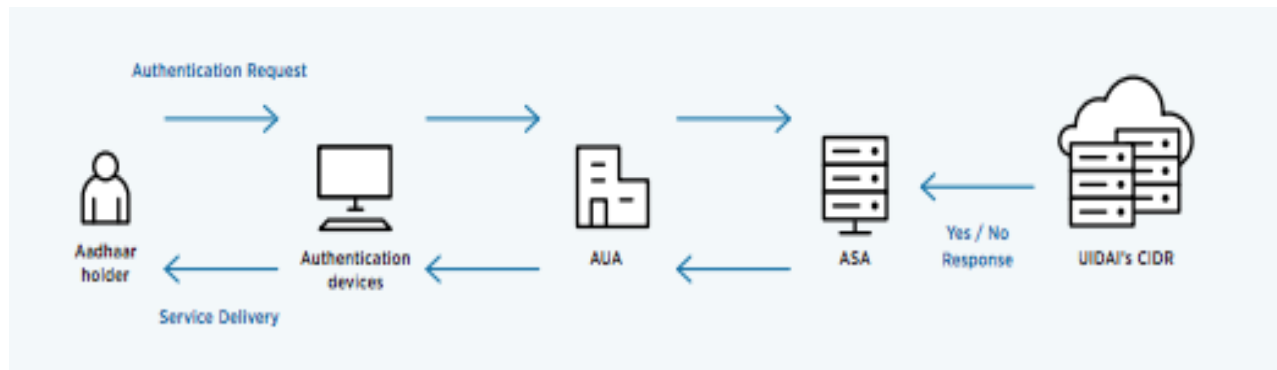


How it works: Data flows for authentication & eKYC

eKYC*



Authentication**



* <https://aadhaarapi.com/>
 **State of Aadhaar Report



Open data & privacy

- Privacy safeguards already hardcoded in law:
 - Provisions in Information Technology Act
 - Exception of privacy in RTI law
 - Aadhaar law contains confidentiality provisions
- The global debate on open data, privacy, anonymity
- Privacy debate as it is unfolding in India:
 - Aug 2017 Supreme Court decision on privacy
 - Committee on data protection law

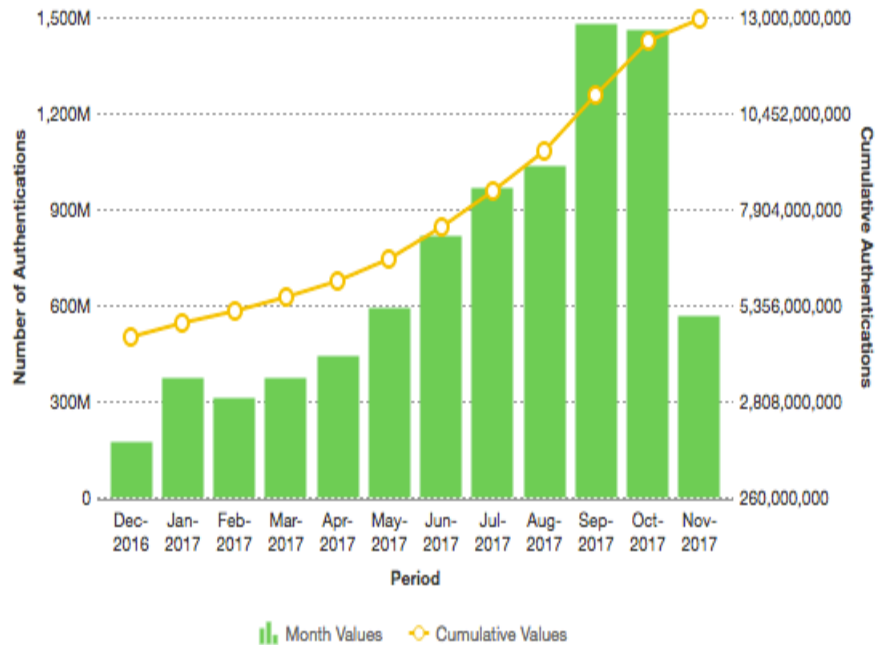
Where should the balance lie?



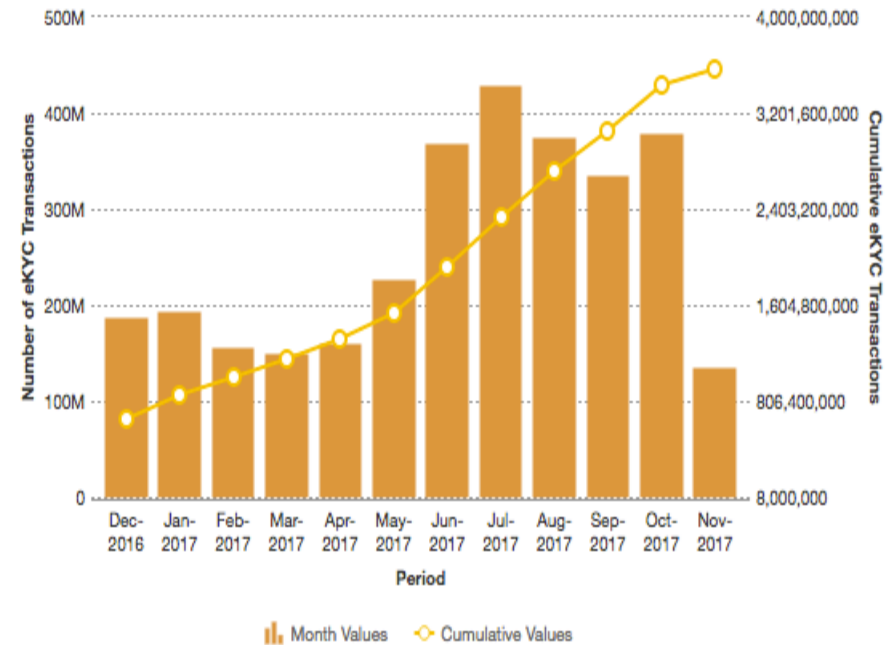
Sources of open data in Aadhaar

- UIDAI: enrolment, authentication, failure rates
- User agencies: public and private

Authentication Trend



eKYC Trend





Some potential applications

- Telecom subscriber data
 - *gender divide in technology?*
 - *distribution across districts?*
- Hospital registration data
 - *what is the incidence of disease?*
 - *age and gender of patients?*
- Transgender community
 - *extent of exclusion?*



Key principles and implementation

- Key principles:
 - Personally identifiable information must not be disclosed
 - Begin with aggregated, anonymised data
 - Develop further standards with time
- Who decides what is to be shared and how?
 - *Sector regulators, UIDAI, data collectors themselves?*
- Our proposal (For Aadhaar and beyond)
 - An independent mechanism
 - Multi-stakeholder approach
 - Oversight of implementation processes



Proposed framework

- Role of the open data committee
 - Recommend contractual commitments - *will vary from sector to sector*
 - Develop principles and standards for data sharing
 - Follow open, transparent processes
 - Monitor compliance by agencies
 - Communications strategy for disseminating information
- UIDAI makes decisions, informed by open data committee



Objective

Open data for better research, enhanced accountability and design improvements in the Aadhaar ecosystem.

Thank you