SoK: A Billion Souls: A Security & Privacy Review of India's "Aadhaar" Biometric ID

1st Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

4th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

5th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

6th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address

***Abstract*—India's Aadhaar is the largest biometric identity system in history. The Unique Identification Authority of India (UIDAI) is responsible for providing each Indian resident with a distinct identity—a 12-digit Aadhaar number—using their biometric and demographic details. Aadhaar is designed to assist in the efficient, transparent, and targeted delivery of subsidies, benefits, and services to India's 1.36 billion residents. However, with increasing global awareness on data security and privacy, public trust in Aadhaar is crucial to its efficacy. With this in mind, we aim to highlight and catalogue the existing technical and structural vulnerabilities in the Aadhaar infrastructure and provide mitigation strategies for the same. We do so by drawing the first detailed snapshot of Aadhaar's technical, structural, and policy infrastructure. We examine the legitimacy of alleged security breaches reported by Indian media outlets based on the standard benchmark for information security—the Confidentiality, Integrity, and Availability (CIA) triad. Moreover, we categorise the feasibility of these breaches based on the threat actor involved, cost of carrying out the breach (time and resources) and the level of security provided by the Aadhaar infrastructure. Finally, we also consider threat actors and privacy breaches to complete our analysis.**

***Index Terms*—Aadhaar, UIDAI, digital identity, security, privacy**

1. Introduction

In September 2019, a massive data breach compromised the phone numbers of around 400 million Facebook users [1]. Sadly, this was nothing special: in just the last few years, retailers such as Target and eBay, government agencies such as the US Internal Revenue Service and the German Bundestag (including personal emails belonging to Chancellor Angela Merkel), and companies as diverse as Ashley Madison (a Canadian extramarital dating company) and JP Morgan Chase (the largest bank in the US) have all been hacked [2]. This relentless surge of cyber-attacks has resulted in a newfound awareness amongst the public. As individuals with unique digital footprints, we are becoming more aware of the worth of our personal information as well as the high level of security (and policy) required to protect this data from hacks and data leaks.

Data security and privacy are no longer just concerns of individual users. This is apparent in the European Union's introduction of the General Data Protection Regulation (GDPR) [3]. In August 2017, a unanimous judgment by the Supreme Court of India declared that "The right to privacy is protected as an intrinsic part of the right to life and personal liberty under Article 21" [4]. This landmark judgment supported the opinion that privacy is a fundamental and inalienable right inherent to human dignity and liberty. Beyond this ruling, India's own Personal Data Protection bill [5] and Digital Information Security in Healthcare Act (DISHA) [6] are currently in the works.

Although a legal framework is necessary, it does not guarantee effective implementation – it is critical to periodically verify if our data is secure and used ethically. This is especially true for **Aadhaar**—a 12-digit unique ID issued by the Indian government to each Indian resident, using their demographic and biometric information. Civil activists [7] and media outlets [8] have alleged that the Aadhaar is vulnerable to numerous types of breaches. Unfortunately, corroborating these claims becomes a difficult task due to the lack of a single comprehensive resource detailing Aadhaar's technical and structural security framework.

We believe that Aadhaar has great potential in dispensing social benefits to every Indian resident in an efficient and accessible manner. However, for Aadhaar to be effective in the long-term, without being hobbled by civil suits, residents must trust the infrastructure. This can only happen if residents are made aware of, and are allowed to help improve, the rules and technologies that protect and restrict the collection of their biometric and

demographic information.

1.1. Outcomes

From our review of the Aadhaar security and privacy infrastructure, we find the following:

- Aadhaar is fairly secure against data Availability breaches. We find these to be rare and occurs only in cases of insider attacks. The primary database (CIDR) itself is secure and removing/editing information is hard to do illegally.
- Aadhaar is vulnerable to both breaches of data Confidentiality and Integrity. This can affect both the access and quality of resident data. Out of 29 reported security breaches, 12 are integrity breaches while 15 are confidentiality breaches.
- The security breaches can primarily be attributed to the following: server hacking, infrastructural loopholes and the use of sub-par hardware. Although we find the CIDR to be secure, there have been multiple reports of Aadhaar data being stored by UIDAIs partner organisations in insecure databases.
- The most significant threat to the Aadhaar with respect to damage caused are insider attacks. As per our knowledge of the infrastructure, there are not enough robust mechanisms to prevent such breaches from occurring.
- Although most security breaches occur within the Enrolment Ecosystem, privacy breaches are more evident in the Authentication Ecosystem.
- Privacy breaches primarily occur due to the following: illegal storage by UIDAI's partner organisations, database access and log access without appropriate privacy conserving methods and access to physical records with the same personal information as Aadhaar.

1.2. Paper Overview

Documentation about the Aadhaar system in the public domain is outdated or ambiguous, and *no detailed description of the entire infrastructure exists*. As a result, one is forced to collate information from multiple (often unreliable) sources, resulting in a confusing and contradictory overview of the system.

Our first step is to create a detailed overview of the current Aadhaar system's infrastructure and policies regarding data privacy and security (Section 3). We collate information from a myriad of technical reports, policy documents, Memoranda of Understanding (MoUs), and circulars published and signed by UIDAI and other organizations that constitute the Aadhaar infrastructure[1]. We call this overview a "Snapshot" of the Aadhaar system. For comprehensibility, the snapshot is divided into four sections: the Enrollment Ecosystem (Section 3.1), the Authentication Ecosystem (Section 3.2), the Central Identities Data Repository or CIDR (Section 3.3), and Biometric De-duplication (Section 3.4).

In section 4, we introduce our definitions of "Security Breaches" and "Privacy Breaches" with regard to Aadhaar, using standard information security benchmarks. We follow these definitions in all our subsequent analyses.

In Section 5, we use the Snapshot and the definitions to analyze Aadhaar in 4 steps. First, we **filter legitimate breaches** from our database of media allegations (Section 5.1). Second, we **categorise the feasibility** of these breaches based on the threat actor involved, cost (time and resources) and the level of security provided by Aadhaar (Section 5.2). We tag each of these features as *low*, *medium*, *high*, or *unsure*. Moreover, we provide technical and structural mitigation strategies for each type of breach. Third, we perform a threat-actor analysis on possible future breaches where we examine the **capability**, **motivation** and **damage caused** by each **threat actor** (Section 5.3). Finally, similar to our "Security Breaches" analysis, we provide our findings on "Privacy Breaches" in the Aadhaar ecosystem along with individual mitigation strategies (Section 5.4).

2. Background

Between 1991 and 2010, India's spending on subsidies and welfare shot up from **INR 122 billion to INR 1.73 trillion** [9]. This, however, did not translate into tangible outcomes. High rates of corruption and logistical incompetence during the implementation of these subsidy programmes led Nandan Nilekani (Founding Chairman, UIDAI) to comment on the need to establish a system that inducted technology to issue a distinct identity number that could be validated across the country [10]. This laid the foundations for Aadhaar – a (digital) demographic and biometric-based unique identity for the residents of India. In short, the vision of Aadhaar is to answer the question, "Who am I?" for every Indian resident [10].

The Unique Identification Authority of India (UIDAI) was established in January 2009. Its mission was to issue a unique identification (UID) number, also known as an "Aadhaar Number", to every resident of the country. The purpose of the UID was to be a one stop identification that is, eventually, linked to every social service to make disbursement of welfare services effective and efficient. The bill that provides legal backing to Aadhaar is called the "Aadhaar (Targeted Delivery of Financial and other Subsidies, benefits and services) Act".

Apart from providing Indian residents with a unique identity — an Aadhaar number —the UIDAI is also responsible for providing a platform to authenticate their physical presence [11] at a point of service. Since its inception, Aadhaar's policies regarding its vision, data security, and privacy have been under intense scrutiny. Indian and international [12] media have reported possible loopholes, vulnerabilities, and policy violations, not to mention the ethical issues of having a national biometric ID in the first place.

Our work builds upon previous examination of Aadhaar by Agrawal et al. [7], which highlights security vulnerabilities within the Aadhaar ecosystem. However, it does not address the issues of data security and privacy. It is also important to note that the Aadhaar system has evolved ever since, and some highlighted breaches are no longer possible. Agrawal et al. [7] also do not discuss vulnerabilities in Aadhaar due to the involvement of other

1. We make all these reports available here: <Temporarily removed for anonymity>

(non-UIDAI) organisations that access some form of Aadhaar data. We expand more on the role of UIDAI's partner organisations within the security and privacy infrastructure of Aadhaar.

*We emphasize that throughout this research, our focus remains on the **strengths** and **vulnerabilities** of the security (technology, structure, and policy) and privacy of Aadhaar, and **not** the consequences of the idea of such a system in India, or biometric ID schemes in general.*

3. Snapshot of the Aadhaar System

We divide the Aadhaar system into the **enrolment ecosystem**, the **authentication ecosystem**, the **CIDR** (Central Identities Data Repository), and **biometric de-duplication**. The first handles onboarding and allocation of unique identities (UID); the second provides verification services such as "e-KYC"; CIDR is the database that stores all collected information; and the last involves technical provisions to prevent duplicate biometrics in the CIDR. Refer figure 1 for a summary of the workflow of the Aadhaar system.

3.1. Enrolment Ecosystem

The Enrolment Ecosystem handles onboarding of residents into Aadhaar with the objective to provide each resident with a valid UID; it also updates demographic and biometric details of existing UID holders. Residents are allowed to **enrol only once** but may perform updates throughout their lifetime. This ecosystem is designed to work offline as well. "Enrolment Agencies" and "Registrars" are the major actors in this ecosystem.

3.1.1. Enrolment Agency.

Function. UIDAI employs third party vendors called Enrolment Agencies (EA) to carry out enrolment services and provide them with specific tools and procedures [13]. EAs are hired by UIDAI (Specifically, they are hired by Registrars, as discussed in Section 3.1.2.)

EAs collect the demographic and biometric data of each resident and are responsible for providing operators and supervisors for each enrolment centre (such as public and private sector bank branches, post offices etc.) [14]. They must also provide the necessary equipment for enroling residents [15].

EAs are also in-charge of updating existing resident details. While certain types of data such as date of birth (or age) and gender will usually remain unchanged, other demographic details may undergo multiple changes. The onus of this update falls on the EAs who, with the assistance of Registrars, execute the initial enrollment process and subsequent updating using an application called "Enrolment Client".

"Verifiers" are officials who are authorized by Registrars (and in turn, UIDAI) to authenticate the resident's documents submitted to the EA. UIDAI specifies that "any serving /retired official both from Government (including Armed forces and CPMFs) and PSU's including Banks not below the rank of Group 'C'/class III employees may be allowed to be deployed as Verifiers." Verifiers are required to follow regulations set by UIDAI [16].

Security (Technical).

- **Enrolment Client and Equipment** – UIDAI mandates all Registrars to follow guidelines to set up the technical environment for enrolment. Only certified equipment [17] (such as devices for biometric data capture) are allowed; the most prominent being the Enrolment Client (for more details, refer to 3.3).
- **Data Validation** – The biometric capture occurs after all demographic data is entered into the client by the Enrolment Operator (EO). The client is equipped to work under "Indian conditions"—although this is not properly defined, we assume that this entails low lighting, lack of internet connectivity, dusty environments, etc. [18]. Most of the onboarding happens offline—data and metadata are periodically synced with CIDR [15]. The resident is allowed to **validate** their data after all the information is filled.
- **Operator Activity Tracking** – Every EO in charge of the enrolment client must sign each enrolment and updation with their own biometrics to ensure security (i.e. CIDR integrity). EO login involves a username, password and EO's biometric, allowing UIDAI to track untoward activity [15].

Security (Policy and Logs). EAs must ensure continued empanelment with UIDAI in order to be engaged by Registrars. If non-empanelled agencies are employed by Registrars, they are also subject to the same terms as empanelled agencies (the difference between empanelled and non-empanelled agencies is unclear since non-empanelled agencies can also become EAs.) Any organisation can apply to be an EA if they meet UIDAI's requirements, regardless of empanelment status. EAs are classified into different categories based on technical and financial competence according to the RFE (Request for Empanelment) [19] published by UIDAI. UIDAI mandates that when a Registrar hires an EA, the EOs working at the agency need to be certified. The Registrar and EA are responsible for their training and certification. For this, the UIDAI has provided a questionnaire [20] and a presentation to ensure minimal and adequate training. A team called the "Training, Testing and Certification" [21] designs lessons to help potential operators who have no prior qualifications to ensure EOs can recognize the necessary documents (like proof of identity) as they will be performing the first check. Periodically, UIDAI alongside it's Registrars, organise "Mega Training and Certification Programs" [19] to facilitate mass onboarding of operators when there is high demand. Refresher courses are also organised to ensure operators stay up to date.

Accountability. EAs are accountable to the Registrar that hired them (though these are often the same entity). The Registrar in turn is accountable to the UIDAI. Failure to adhere to the guidelines set by UIDAI usually results in termination of the contract [22].

3.1.2. Registrar.

Function. UIDAI partners with state and central ministries, banks, public sector organizations, and other agencies [23]. These entities are known as "Registrars" and their main purpose is to facilitate issuing of Aadhaar numbers by enrolling residents, and validating resident data during enrolment and updation. Registrars upload encrypted resident data to the CIDR for de-duplication.

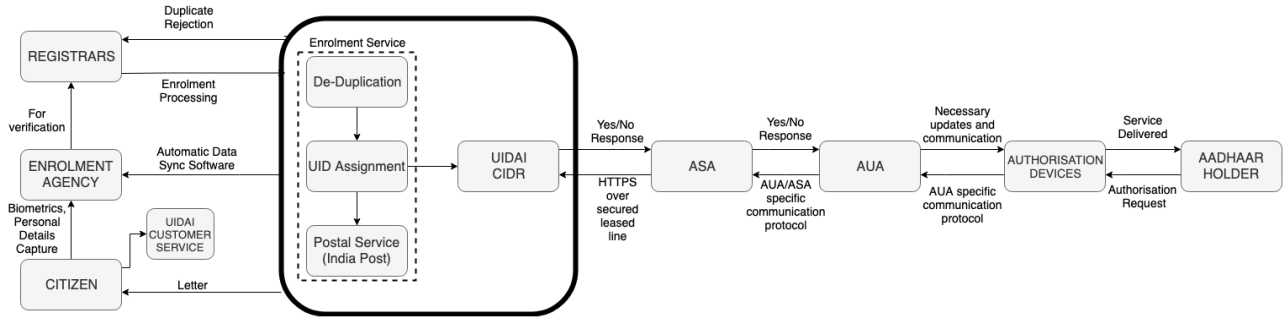


Figure 1. Flowchart depicting the architecture of the Aadhaar system.

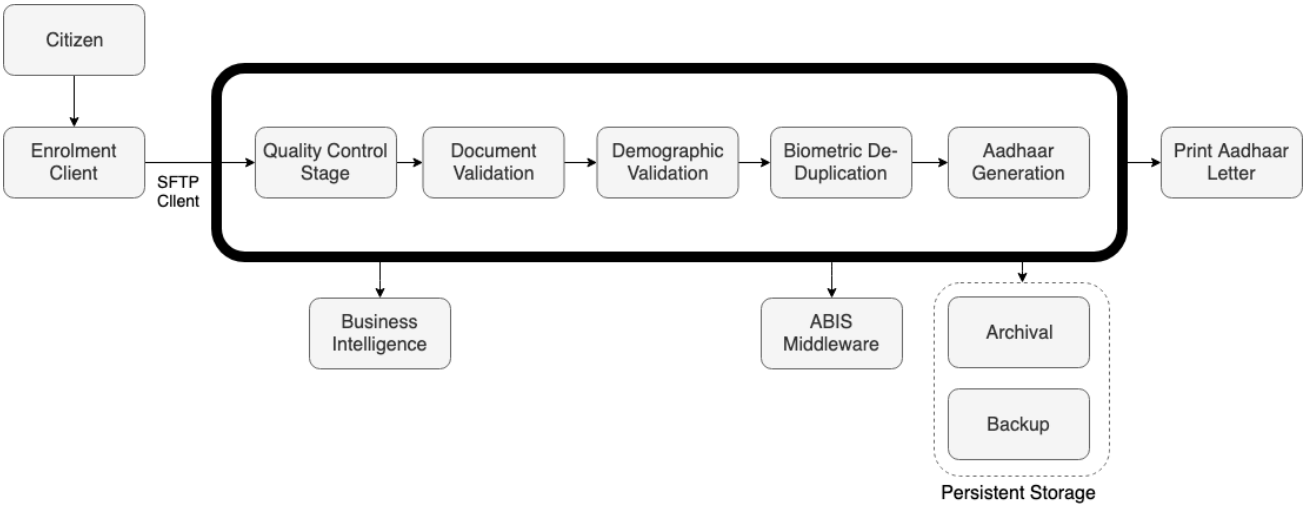


Figure 2. Flowchart of the Aadhaar Enrolment Ecosystem.

Registrars usually hire EAs to perform the aforementioned tasks on their behalf and are responsible for the correct functioning of these EAs. However, there is no mention of Registrars having to inform UIDAI about entities whom they hire. Registrars must follow protocols, and standards prescribed by the UIDAI, and submit periodic reports (details about the frequency of said "periodic" reports aren't publicly available.)

Registrars must also take special measures to enrol women, children, persons with disabilities, unskilled workers, nomadic tribes, and people belonging to marginalised groups who cannot produce a valid Proof of Identity (PoI) and/or Proof of Address (PoA) [23].

"Introducers" are individuals (such as Registrar employees, members of local administrative and elected bodies, postal staff, etc.) who are recognized by Registrars to confirm the identity and address of residents who do not have a valid PoI or PoA. Introducers are not allowed to charge residents, however, Registrars can offer them some fees for their service. Introducers are required to take part in Aadhaar awareness workshops (describing their roles, responsibilities and liabilities) organized by their Registrar and UIDAI [16].

Security (Technical). For the purpose of enrolment, a Registrar uses UIDAI developed Enrolment Client to enroll residents; follows DDSVP (Demographic Data Standards and Verification Procedure) [24] prescribed standards and procedures for the collection of data, uses only those IT systems and devices [17] whose specifications have been approved by the UIDAI for the purpose of collecting data, follows UIDAI protocols for record keeping and maintenance, and adheres to UIDAI confidentiality, security, and privacy protocols.

Security (Policy and Logs). The MoUs [22] between Registrars and UIDAI specify that UIDAI periodically audits the Registrars and EAs (frequency is not specified). The MoUs state that if a Registrar fails to follow the security mandates, willfully or otherwise, then UIDAI will **only** make reasonable attempts to discuss and resolve difficulties with the Registrar. Organisations have been penalized in the past. Specifically, UIDAI terminated a contract citing "enormous number of complaints of corruption and enrolment process violations against Aadhaar Enrolment/Update Centres under CSC e-Gov" [25]. *At the same time, we note that the contract documents do* ***not*** *detail any standard penalties for Registrars that fail to adhere to security and privacy guidelines; nor are they present in other publicly available documents to the best of our knowledge.*

3.2. Authentication Ecosystem

3.2.1. AUAs and KUAs. The Aadhaar Act 2016 [26] defines a requesting entity as "an agency or a person that submits an Aadhaar number and demographic information or biometric information, of an individual to the Central Identities Data Repository (CIDR)

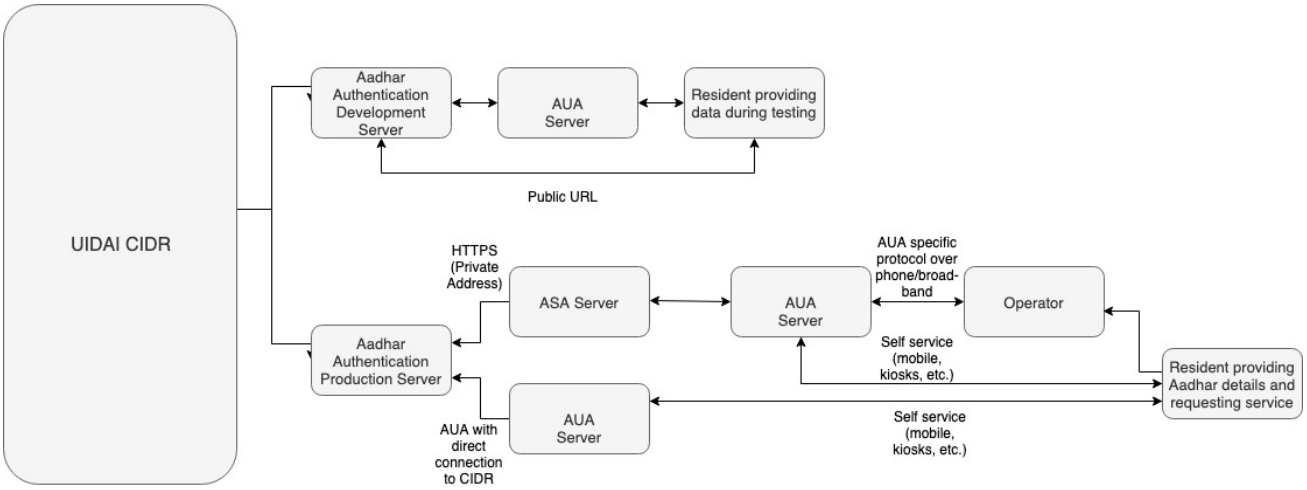


Figure 3. Flowchart of the Aadhaar Authentication Ecosystem.

for authentication". There are two types: Authentication User Agency (AUA) [27] and Know-Your-Customer User Agency (KUA) [28].

Function. AUA/KUA may be a government, public, or private legal agency registered in India which uses UIDAI's Aadhaar authentication and e-KYC services [29]. A requesting entity connects to the CIDR through an "Authentication Service Agency" (either by becoming an ASA or by contracting the services of an existing ASA). AUAs/KUAs use Aadhaar authentication to provide services such as opening of bank accounts, LPG connections, etc. to residents. To streamline AUA/KUA enrolment, UIDAI has created 3 broad categories based on which entities are onboarded. The categories and the type of entities are listed in Table 5 (see Appendix A). These categories are further divided on the basis of financial and technical requirements (see Table 6 in Appendix A).

1) **AUA** – The primary role of an AUA is to verify if the user is who they claim to be. This is done by sending Aadhaar authentication requests to the CIDR through an ASA. (The AUA receives a Yes/No response from the CIDR.)
   Sub-AUAs are agencies that use Aadhaar authentication to enable its services through an existing AUA/KUA.
2) **KUA** – A Know-Your-Customer User Agency (KUA) is a requesting entity which, *in addition to being an AUA*, uses e-KYC authentication facility to match the submitted biometric information and/or OTP and Aadhaar number with the data available in the CIDR. When an Aadhaar holder wants to submit their KYC details to a KUA, they download a copy of their e-KYC in XML or QR Code format from the Aadhaar website. This file is encrypted with a "Share Code" that the holder decides at the time of downloading. They can then submit this file to the KUA requesting the KYC. To verify the submitted file, the KUA requests a copy of the holder's e-KYC from the UIDAI. When requesting for an e-KYC, the KUA receives a "digitally signed [machine readable XML] e-KYC authentication response with encrypted e-KYC data" [30] along with other technical details related to the authentication transaction. The KUA can use this copy of the holder's KYC data retrieved from UIDAI to verify the offline copy that the resident submitted, to complete the KYC process.
   The encrypted XML file contains the resident name, download reference number, address, photo, gender, DoB/YoB, hash of mobile number, hash of email.

In January 2018, UIDAI categorized all AUAs and KUAs into Local and Global on the basis of the kind of e-KYC data available to them. Global AUAs/KUAs have access to the entire e-KYC service using Aadhaar numbers whereas, Local AUAs/KUAs have access to "limited KYC" and are not allowed to store Aadhaar numbers internally. A Local AUA/KUA is must get an agency specific UID token to authenticate and access "limited KYC" [31]. UIDAI reserves the right to decide the type of biometric data that can be accessed by the Local AUA as needed.

Security (Technical) [32, p. 37-46]. . We have mentioned a few security measures below, while an extensive list can be found in the *Compendium* [32].

1) **Aadhaar Security Vault** [32, p. 115] – Aadhaar numbers collected by any agencies for purposes under the *Aadhaar Act and Regulations, 2016* [26] are stored locally in a "Aadhaar Data Vault" [33]. It is located within the organization's infrastructure and is accessed only on a need-to-know basis. Interestingly, UIDAI does not mandate an audit of the this vault. Moreover, **there are no repercussions** if the data vault fails to store information in an encrypted format.
   As per UIDAI guidelines for "Aadhaar Data Vault", the Aadhaar number is to be encrypted at all times. While the encryption algorithm for has not been specified, the standard RSA 2048 for public key encryption and AES 256 for symmetric encryption is recommended [34]. UIDAI guidelines also state that the keys for the encryption must be stored in a Hardware Security Module (HSM). The implementation of this is usually outsourced as many third party vendors [35] offer their own variants of an "Aadhaar Data Vault".
2) **Cryptography** – An AUA/KUA can transmit core biometric information over a network only after

creating an encrypted Personal Identity Data (PID) block. This has to be in accordance to the specifications set by UIDAI [34]. The AUA/KUA cannot store the encrypted PID block unless in case of buffered authentication. Even for this purpose, the PID block cannot be stored for more than 24 hours, after which it must be deleted from local storage [32].

3) **Communication Security** – Communications between the biometric and encrypting devices must be "secured against all security threats/attacks". It is not clear what kinds of threats have been considered here [32]. Every terminal device has a unique terminal ID that is transmitted during each transaction alongside the UIDAI-assigned institution code.
AUAs/KUAs may permit any other agency to perform Yes/No authentication by generating and sharing a separate license key for each entity through the portal provided by UIDAI to the said AUA/KUA. As such, AUAs/KUAs use separate license keys to send requests to and receive responses from the CIDR through an ASA over a secure network. Information is sent through leased lines or secure private lines. If an AUA/KUA must use a public network, then a secure channel (SSL, VPN) is used [36].

Security (Policy and Logs).

1) **Access Control** – Access to information facilities like the Authentication application, audit logs, servers, source code etc. is only given to authorized personnel [32]. However, the documentation does not state the basis on which a person becomes authorized, and the extent of access.
2) **Password Policy** – UIDAI states that the password of each user must remain confidential. Additionally, if these passwords are required to be stored, then they must be encrypted in an unspecified encryption standard and the plaintext must be discarded.
3) **Logs** – AUAs/KUAs are required to maintain logs of each authentication transaction for 2 years. These logs can be accessed by UIDAI or the requesting entity solely for grievance and dispute redressal. The logs contain the following information: Aadhaar number, parameters of the authentication request, parameters received from the CIDR as a response, information disclosed to the person being authenticated at the time of authentication, and a record of the person's consent for authentication [32, p. 12]. Logs do not contain PID information. There is no specification of the encryption standards, if any, imposed on the storage of the logs by the AUA/KUA. Upon the 2 year expiry period of each log, they are archived for either 5 more years or for a period specified by law, after which they must be deleted. The only exception is for logs retained by a court or pending disputes [32]. In the event that the contract is terminated, the obligations for authentication logs are still binding. The security protocols set by the MoUs between AUAs and the UIDAI are unclear, since existing MoUs are inaccessible.
4) **Virtual ID** – To protect the privacy of the Aadhaar card holder and to prevent Authenticating agencies from gaining access to resident Aadhaar number, Virtual IDs (VID) can be given. VIDs are "temporary, revocable 16 digit random numbers that are mapped to the Aadhaar number" such that it is impossible to obtain the original Aadhaar number [31] from the VID. Authenticating agencies are not allowed to generate VIDs for any Aadhaar card holder.

Accountability. AUAs/KUAs are required to ensure that its operations are audited including information security controls and technical testing like vulnerability assessment, penetration test of Information Systems and new technologies or delivery channel introduced [32]. This audit must be done by an "information system auditor certified by a recognised body annually and on a need basis" [32, p. 60] or by UIDAI itself to ensure compliance with UIDAI standards and specifications. Although we are unsure what constitutes a "recognised body", we believe that it refers to empanelled auditors recognized by the government. The list of empanelled auditors by CERT-IN is available [37]. These audit reports must be shared with UIDAI upon request. The Information Technology Act, 2000, states that "The Certifying Authority must ensure that its audit logs are reviewed by its personnel at least once every two weeks and all significant events are detailed in an audit log summary". The rights to access information facilities processing UIDAI data need to be reviewed and audited every quarter. Although UIDAI states that only authorized personnel can access the audit trails, selection criteria and security policies for each level of authority are unspecified.

3.2.2. ASAs and KSAs. Authentication Service Agencies (ASA) and KYC Service Agency (KSA) are government and private agencies that have an "established secured leased line connectivity with the CIDR" [29], in accordance with UIDAI's standards and specifications [32]. There are 27 live ASAs as of 31-Aug-2018 [38].

Function. ASAs provide secure CIDR access to requesting entities for authentication; KSAs are essentially ASAs with KYC permissions. Only the ASA/KSA can interact directly with the CIDR in the Authentication Ecosystem. Hence, ASAs/KSAs act as enabling intermediaries between an AUA/KUA and the CIDR as shown in figures 1 and 3. Entities seeking to be an ASA/KSA are classified into 5 categories based on the type of institution (see Table 7 in Appendix B). Further, these organizations must meet minimum financial and technical criteria (see Table 8 in Appendix B).

Security (Technical) [32, p. 52-61]. We list some critical security procedures below; an extensive list can be found in the *Compendium* [32].

1) Servers that are used by ASAs to connect to the CIDR must be stored within data centers within India.
2) ASA server host must be within a segregated network segment. It should be isolated from the rest of the network of the ASA. The ASA server host is solely dedicated for Aadhaar authentication.
3) **Cryptography** – The PID block comprises of the keys generated by the ASAs. The task of key generation, distribution, and storage is that of ASAs. The PID blocks received by the ASAs can never be stored anywhere, since it is sensitive information.

Security (Policy and Logs). Access control, password, communication policies, log maintenance and expi-

ration are same as that of AUAs/KUAs (Refer to Section 3.2.1). The logs can be accessed by UIDAI or the requesting entity solely for grievance and dispute redressal and contain the following information: identity of the requesting entity, parameters of authentication request submitted and parameters received as authentication response.

Accountability. The audit protocols are the same as that of AUAs/KUAs (see Section 3.2.1).

3.3. CIDR (Central Identities Data Repository)

Function. Central Identities Data Repository (CIDR) is a centralised database run and maintained by UIDAI. It contains all Aadhaar numbers ever issued and the corresponding demographic and biometric information. It is at the core of the Aadhaar infrastructure and interacts with both, the enrollment and authentication ecosystems. The database is distributed throughout the country over multiple servers. EAs use the Enrolment Client to connect to the CIDR. The client stays updated by periodically pushing and pulling data from the CIDR. It is also (indirectly) responsible for de-duplication as de-duplication servers access biometric data residing in the CIDR to check for matches before enrolling a new resident. After enrolment, data is accessed from the CIDR mainly for two reasons: Authentication and e-KYC.

- **Authentication** – The process uses an Aadhaar number and an OTP (or biometrics) as a second factor to authenticate an individual on the vendors (AUA/Sub-AUA) request. In response, the CIDR returns a digitally signed "Yes/No" depending on the success of the authentication request [29].
- **e-KYC** – In order to receive KYC data, KUAs must send a request to KSAs (ASAs with KYC permissions) – the only entities allowed to use the e-KYC connection endpoints. In response, the CIDR returns a digitally signed and encrypted demographic record (see e-KYC details given in Section 3.2.1).

Security (Technical).

- **Enrolment Client** – In the Enrolment Ecosystem, the CIDR is connected to the Enrolment Client and performs periodic data syncs (following the "Mandated Sync Frequency" [18]). Since the connection is not active continuously, the Enrolment Client pushes data as required from the CIDR only after establishing an internet connection. The connection between the CIDR and the Client is protected using SSL. The enrolment data (XML) is POST*ed* [39] to the CIDR [18]. To ensure only certified operators and enrolment clients connect to the CIDR, each time an operator logs into the Client, an XML document containing the machine identifier, enrolment agency code, and station number is sent to the CIDR to validate the connection. On successful validation, the CIDR sends back a security token (used to send subsequent enrolment data). The XML document containing the enrolment data is sent in the form of packets to the CIDR, *each* of which is encrypted using a public key published by UIDAI. This "packet encryption" phase is handled by the "Client Security" module (a subpart of the Enrolment Client). The packets are signed before being transmitted to the CIDR. This is done to avoid wasting resources on extracting packets without a valid signature [18]. The Client Security module stores certificates and manages keys to ensure a valid connection with the CIDR. The key management uses a public key style encryption where *two* sets of public keys are maintained—one for data exchange between the Enrolment Client and the CIDR, and another for data exchange between the Registrar and the CIDR. The CIDR is classified as a Protected System under the IT Act and the link between the CIDR and the Enrolment Client is encrypted using 2048 bit PKI.
- **De-duplication** – The Aadhaar Technology Architecture documentation (2014) [18] says that "Since de-duplication at this scale (1.2 billion residents) had not been previously attempted anywhere in the world, UIDAI decided to procure 3 ABIS (Automatic Biometric Identification System) software solution to perform biometric de-duplication as a risk mitigation strategy." For a new enrolment, Aadhaar first does a demographic and reduced biometric check for matches. The Aadhaar Enrolment Server needs to integrate with all three solutions using the standard interface API (ABIS API) and allocate de-duplication requests to the 3 ABIS servers as per UIDAI policy of dynamic allocation. After enrolment, the ABIS (Automatic Biometric Identification System) de-duplication servers are sent packages of size 3-5 MB. The enrolment packet (containing all demographic, biometric, and metadata) is encrypted at the client side and then sent to CIDR; the CIDR interacts with the ABIS servers and sends them these packages. Only the Enrolment Server (maintained by CIDR) can decrypt the enrolment packet. It does this in memory; the decrypted packet is never sent to storage. Original biometric data is archived and sent to offline storage, and is not available on an online network. At every step, data is 2048-bit PKI encrypted. With 10 fingerprints and facial image, 95% de-duplication rate could be achieved over a population of 50 million. To increase the de-duplication rate to 99%, usage of iris biometrics was proposed. However, there is no documentation about the exact matching algorithms running at the ABIS and how well they perform.
- When a registered device is called by the application, it captures, processes, and encodes the digitally signed biometric record. The biometric data that is received by the UIDAI/CIDR is essentially a Base-64 of the DSA signature of a hash (SHA-256) of the biometric data along with a time stamp, device code and device private key.

Security (Policy). The CIDR receives a log of the Enrolment Client's performance which is later used to debug the application. The audit information is also maintained by the Enrolment Client which is used to keep track of the application usage—analytics on the number of enrolments and updates at any given centre. Every enrolment is first reviewed by a field supervisor who heads the centre and then signs the packet. This allows UIDAI to track "who, when, where, under which agency, under which registrar, who reviewed it" so that every packet

can be traced back to individuals who worked on it. This metadata is pushed along with the enrolment data in the same XML document.

Morpho produces the Aadhaar enrolment kit [40]. We note that the financial contract signed between the UIDAI and Morpho for manufacturing these biometric devices is not publicly available. There may be a possibility of a supply chain attack on the kits produced by Morpho, since the sources of the internal hardware components that are used to make the biometric devices are unknown.

3.4. Biometric Deduplication

When a user enrols into the Aadhaar database, their biometric templates are stored. Using these biometric templates, UIDAI checks for uniqueness of the identity among the already enrolled population to make sure that each person enrols only once.

Since Aadhaar has the face, fingerprint and iris biometrics for all enrolled citizens, combining all of these for de-duplication gives more distinctive features compared to using a subset of these. While UIDAI has not released any details about their de-duplication process, given the requirements and capabilities that are claimed, the best way to share this data with the ABIS providers would be using a Multidimensional Fuzzy Vault Scheme (MDFVS), where the biometric feature data is converted to a vector [41].

When a person enrols, the features extracted from their biometric information are stored into the database. During deduplication, different biometric templates are compared with each other; however, to prevent the ABIS providers from having the original biometric data of the citizens MDFVS can be used. This way the ABIS providers only get extracted features from the templates and no other information about the templates.

The biometric templates need to be protected and that can be done in two different ways [42].

1) **Feature transform** Using a random key, biometric templates are transformed via a transformation function. The matching and storing is done over the transformed templates. The feature transform approach can be further categorized as (i) salting - the transform is invertible, this is key based, and (ii) non-invertible transform - using a one way function.
2) **Biometric Cryptosystem** Biometric template is stored along with "helper data". Using a binding key K with the template we get the helper data. Hence, such systems are called key-binding biometric cryptosystems. To match, we recover the key from helper data using the biometric features.

From Supreme Court of India Judgement [43], following details about information exchange with the ABIS providers are known:

1) The data sent to ABIS is completely anonymized. ABIS systems do not have access to any resident's demographic information as they are only sent biometric information of a particular resident with a reference number and asked to de-duplicate. The result with the reference number is mapped back to the correct enrolment number by the Authorities' own enrolment server.
2) The ABIS providers only provide their software and services. The data is stored in UIDAI storage and it never leaves the secure premises.
3) The ABIS providers do not store the biometric images, only the template.
4) The encrypted enrolment packet sent by the enrolment client software to the CIDR is decrypted by the enrolment server, but this is never stored.
5) The original biometric images of fingerprints, iris, and face are archived and stored offline. They **cannot** be accessed online.
6) The biometric system provides high accuracy of over 99.86%. The mixed biometric have been adopted only to enhance the accuracy and to reduce the errors which may arise on account of some residents either not having biometrics or not having some particular biometric.

Given that the biometric templates are stored in two ways, it is safe to assume that biometric templates are transformed using the feature transformation approach using salting so that later the reference numbers sent to the ABIS providers can be re-matched to the applicants. For ensuring a feasible de-duplication efficiency the ABIS providers must be using specialized feature selection. This would make sure that the vector distance comparison between two template vectors happens over the most distinctive features. Without more information it is not possible to get a better idea of the de-duplication process.

4. Defining Security and Privacy Breaches in Aadhaar

To identify and evaluate security and privacy concerns, we first define "security" and "privacy" within the context of Aadhaar. These definitions must be inclusive to large sections of marginalized people in India — precisely those whom the scheme was intended to help.

4.1. Security Breach

Security refers to how one's personal information is protected. This can be done via the existing CIA (Confidentiality, Integrity and Availability) global standard for information security. Specifically, this includes –

- **Confidentiality** – Access to a resident's personal data (demographic or biometric) collected at the time of enrollment or updation is granted only to authorized individuals within UIDAI and its partner organizations.
- **Integrity** – A resident's personal information within the CIDR or during transmission is not modified or lost in an unauthorized manner.
- **Availability** – A resident's personal data is available to authorized entities within UIDAI and its partner organizations when required.

Any violation of one or more of the above is considered a "Security Breach".

4.2. Privacy Breach

We define privacy in the Aadhaar infrastructure as follows: The Aadhaar database should not reveal any information beyond verification. If an entity has knowledge of

$a_1, a_2, ...a_k$ columns of a person's Aadhaar information, they should not be able to gain knowledge of the a^{th}_{k+1} column. This includes brute-forcing by checking against the same column multiple times. For example, if an entity knows the name and phone number of a person, they should not be able to query the database multiple times with a different date of birth. Any violation of the above can be considered a "Privacy Breach".

Aadhaar is meant to ensure targeted delivery of benefits and services to Indian citizens. Verification of the existence of a resident to receive a benefit/service must not lead to their identification (i.e. unearthing of their personal information). For services using aggregated Aadhaar data, no individual must be identified from within the aggregated data. Although data sharing must exist within the Aadhaar infrastructure, it must be done in a privacy preserving manner.

5. Media Allegations Analysis

5.1. Filtering Legitimate Breaches

Our primary database[2] of media allegations consists of 36 reports from various news outlets. From this, we filter out Security and Privacy Breaches that are legitimate based on our knowledge of the Aadhaar infrastructure and our definitions of Security and Privacy. Our reasoning for classifying a breach as legitimate or not is also available on the Google Sheets link provided in the previous footnote. This filtering yielded 17 legitimate Security Breaches and 10 Privacy Breaches, which were further analyzed. It must be noted that the list of Security and Privacy Breaches are not mutually exclusive.

Additionally, for each legitimate Security Breach, we ascertain whether or not there was a breach of Confidentiality, Integrity or Availability of data in the Aadhaar infrastructure. An example of this analysis is shown in table 1 and the entire analysis is available in the footnote. We now show the results of our analysis.

According to the breakdown in figure 4, the prevalent breach in security standards is that of Confidentiality. The most common occurrence of this breach entails a subset of Aadhaar data being made public. This includes a large scale Security Breach where thousands of private data points are leaked. Prevention of this type of breach goes back to the mitigation strategy wherein the data is secured in encrypted data vaults and access is limited.

Breach of Integrity is also a common occurrence. It compromises the quality of the central database. Fortunately, these are usually detected by UIDAI. They usually occur at an individual level which either involve a small set of rogue insider-agents or the hacking of individual accounts. If performed repeatedly, the compromised breaches are detected while for a specific usecase like introducing certain individuals into the database, the breach is virtually undetectable. This can be protected by having stronger security measures that require OTPs. People often forget their username and passwords; therefore, OTPs can be used to add another layer of security. Moreover, better quality and technically secure hardware would further reduce the chance of this breach since illegal biometric authentication is not easy to achieve against well-manufactured devices. Finally, standardized monetary penalty on enrolment agents who are caught (and lose their credentials) could disincentivize malicious behavior.

We see that breach of Availability is rare and occurs only in cases of insider attacks. The CIDR repository itself is fairly secure and removing/editing information is hard to do illegally. Internal attacks can be mitigated by using a decentralized system of checks and balances where no individual can commit edits [44]. For example, all operations by high level employees could require approval by randomly chosen officers (anonymously) ensuring a relatively secure system against internal breaches.

5.2. Security Breach Analysis

We define three broad classes of Security Breaches:

1) Server Hacking: Hacking of the UIDAI or Partner organization software/database.
2) Infrastructural Loophole: Access to private data via legitimate UIDAI channels.
3) Sub-par Hardware: UIDAI hardware tricked into approving false biometrics as genuine.

After categorizing the breaches, we analyse their Feasibility on the basis of the "Cost" of the breach and the "Level of Mitigation" against it. The Cost is defined as the time and resources it takes to produce the breach while Level of mitigation is determined by the documented Aadhaar security, technical or otherwise. After analysing the Aadhaar protocols and the type of breach, we suggest mitigation strategies to ensure robust security.

A detailed breakdown of our examination is provided in Fig. 4 Aadhaar is predominantly vulnerable to "Infrastructure Loopholes". These breaches exploit the general negligence to set or adhere to security protocols. As discussed, agents of Aadhaar such as especially EOs can effectively be a threat to the security of the database if their credentials are not stored properly (multiple instances of this have occurred). This is a breach that is detected often but measures taken to curb it are seemingly nonexistent. Complimentary and robust security standards like OTPs and Iris scans for these Aadhaar agents can be effective to ensure accountability.

Server Hacking is another common breach of the Aadhaar infrastructure. The CIDR Database in itself is secure and no reports of it being hacked have been received in popular media, but data in UIDAI's partner organizations is regularly being stored in insecure databases. UIDAI needs to set stricter standards and maintain them across the board. Essentially, no one should be allowed to store any Aadhaar data, but the CIDR. Any queries to the database should go through the CIDR and local copies shouldn't be stored since securing many local servers is a complex and costly task that arguably does not benefit the system all that much. The more direct or indirect connections to the database, the harder it is to secure the entire system; the infrastructure is only as strong as its weakest link.

The issue of Sub-par Hardware can be remedied by improving the quality and security requirements for the necessary equipment. This is also a question of time since

2. The entire analysis can be found here: https://bit.ly/aadhaar-media-reports

TABLE 1. EXAMPLE OF THE AADHAAR SECURITY BREACH ANALYSIS.

Each article was analyzed with regards to the type of security breach (Confidentiality, Integrity or Availability) with our reasoning. This table consists of a few examples. The entire list along with our table for privacy breach analysis can be found in our primary database here.

Allegation	**Confidentiality?**	**Integrity?**	**Availability?**	**Legitimacy? (Yes, No, Unsure)**	**Reason**
UIDAI reveals 210 govt websites made Aadhaar details public, did not specify when breach took place	Yes	No	No	Yes	We know that this is a legitimate breach since the UIDAI has confirmed this breach of data confidentiality through an RTI (Right to Information) request. It is important to note that the UIDAI itself did not leak this data. It was posted on the websites of over 200 central government, state government and educational instututes. It is still unclear as to how they had accessed the Aadhaar data.
Three Gujarat websites including government portal made Aadhaar details public	Yes	No	No	Yes	The Ministry of Electronics and Information Technology confirmed the breach. Even though this is not a direct breach of UIDAI data, this would not have happened had Aadhaar not existed.
UIDAI Blacklists Centre That Leaked Aadhaar Details of M S Dhoni for 10 Years	Yes	No	No	Yes	Human error with regards to the enrollment officer/agency
Rs 500, 10 minutes, and you have access to billion Aadhaar details	Yes	No	No	Yes	UIDAI sued The Tribune for exposing this breach and getting access to this leaked data. This suggests that it is a real breach on the database.
Indane Leaked Millions of Aadhaar Numbers, Claims French Researcher	Yes	No	No	Yes	Same as above
How a SIM Card Operator in Hyderabad Apparently Created His Own Aadhaar Database	Yes	No	No	Yes	He got the information from other sources but the Aadhaar infrastructure should be strong enough to not accept biometrics that are not live.
Aadhaar data: French hacker exposes flaws in its Android app, asks people not to use it	Yes	No	No	Yes	The video shows how this is possible.
Aadhaar's Dirty Secret Is Out, Anyone Can Be Added as a Data Admin	Yes	Yes	Yes	Yes	Access to Aadhaar details after enrollment or updation requires access to the CIDR. Therefore, this alleged breach points to an insider incident. In their review, Agarwal et. al. emphasised the Inadequate protection against insider breach of CIDR data. Thus making this breach a possibility.
Aadhaar card tampering racket busted in Surat	Yes	Yes	No	Yes	Since the enrollment officer is the focal point of information entry/updation, it is possible for this alleged breach to take place if the login details (incuding fingerprint) of the officer is forged.

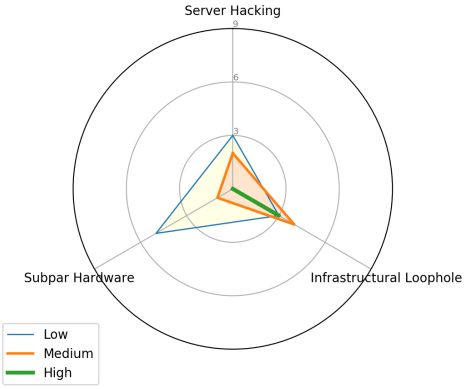
new and improved (more secure) hardware for biometric capture are introduced in the market every year. UIDAI should periodically update their hardware requirements such that none of their equipment is old and insecure.

5.3. Threat Actor Analysis

With regard to Security Breaches, we analyse the threat actors based on their capability, motivation and damage caused and give a value to each feature—low, medium and high. Our analysis, summarised from extensive analysis of allegations and confirmed hacks alike, is detailed in Table 2.

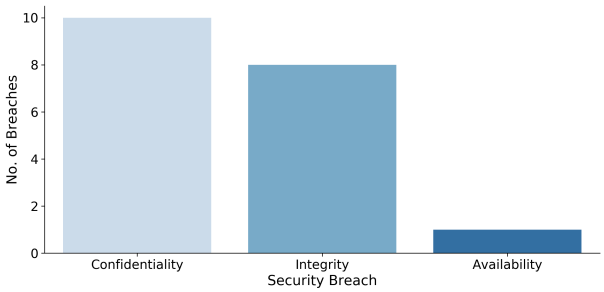
5.4. Privacy Breach Analysis

Listing the various allegations of Aadhaar privacy breaches, we find that limited access to the database and illegal and insecure storage of Aadhaar information to be the prevalent vulnerabilities. These are primarily due to improper or inefficient handling of data by UIDAIs partner organisations.

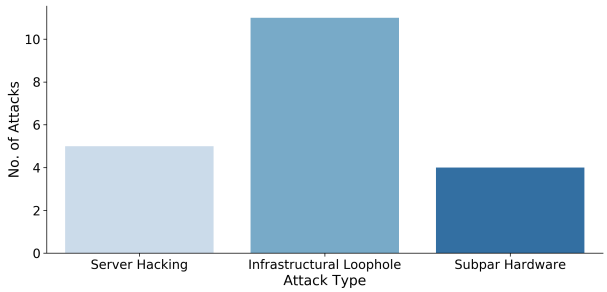


(a) Number of Breaches vs Feasibility

(b) Privacy Breaches



(c) Security Breaches



(d) Types of Security Breach

Figure 4. Analysis of Media Allegations

TABLE 2. THREAT ACTOR ANALYSIS **SECURITY BREACH**

Threat Actors	**Capability**	**Motivation**	**Damage Caused**	**Comments**
Enrolling Operator	Medium	Low	Medium	Damage caused is medium since rogue individuals can be entered into the database by the EO, thus hampering the **integrity** of the data.
AUA	Medium	Low	Low	AUA has the ability to store Aadhaar data provided by the resident when they come to get authenticated (Aadhaar welfare etc.), which is a threat to **confidentiality**.
ASA	Low	Low	Low	The ASA infrastructure has sound measures against Security Breaches.
EA	Low	Medium	Low	EA has medium motivation since the agencies have the ability to make fake Aadhaar cards, even though we are unsure of the number of cards they can generate. This is a threat to the **integrity**.
Registrars	Medium	Medium	Medium	Registrars control multiple agencies. Therefore, they can carry out the same type of breach as an EA on a larger public sphere.
UIDAI official (high ranking)	Medium	Low	High	A high ranking UIDAI official can potentially gain access to the database through internal connections, threatening data **availability**. They can also generate voter IDs via Aadhaar numbers, threatening data **integrity**. As discussed, Aadhaar is more vulnerable to insider attacks than from the outside.
External Governments*	High	Low	High	The damage caused is high due to the sheer scale of the breach given the high capabilities of the threat actor. This could compromise all 3—**Confidentiality, Integrity** and **Availability** of data.
Tech Companies*	High	Medium	High	Same as above along with the fact that Aadhaar data might be extremely useful for the working of numerous big data companies and hence motivation is kept as medium.
Digitally Literate Citizens	Low	Low	Low	These breaches take a lot of effort and time, especially if information is collated from physical documents. This endangers the **Confidentiality** of data.

* **Potential** actors that could take advantage of vulnerabilities

The breakdown detailing the number and type of Privacy Breach is specified in Fig. 4. In either case, the

pivotal issue is that an individual can be identified which could result in the misuse of their data by malicious actors. This can include surveillance, profiling and/or creation of new services without consent for the individual by the state or other private actors.

An interesting point to note is that although most security breaches happen within the Enrolment Ecosystem, privacy breaches are largely witnessed in the Authentication Ecosystem. For Aadhaar to be effective in targeted delivery of subsidies it needs to ensure that resident data is private beyond enrolment.

If organisations require Aadhaar data to analyse aggregated trends, we strongly recommend differential privacy algorithms be used. For other cases that require individual instances of data, it must be constitutionally valid and must be complemented by effective protocols to ensure that the individual is not identified after processing. We have mentioned mitigation strategies for individual breaches in the table. We are hopeful to see more privacy preserving protocols if and when the Personal Data Protection Bill is passed and implemented in India.

6. Conclusion

Aadhaar is the world's largest digital biometric identification system. In this paper, we perform the first comprehensive security and privacy analysis of the Aadhaar ecosystem. In contrast, previous work [7] tackles only breaches on the CIDR (the main Aadhaar backend database). In our paper, we:

- Exhaustively examine the technical and structural infrastructure of Aadhaar across all published documentation to compile a comprehensive overview of the complete Aadhaar security and privacy infrastructure.
- Analyse breaches on the Aadhaar ecosystem reported by the media by filtering (for) legitimate Security Breaches.
- Evaluate legitimate breaches on the basis of threat actors involved, time and resources required (cost), and the level of (feasible) mitigation available.
- Suggest mitigation strategies for the legitimate Security Breaches.
- Perform an analysis of potential threat actors on the basis of their capability, motivation, and the damage they could cause.

From our analysis of security breaches, we find that the Aadhaar ecosystem is most prone to breaches in Confidentiality, followed by breaches in Integrity. Our threat actor analysis reveals that Aadhaar is more susceptible to insider-attacks rather than those by outsiders. This is especially true with regard to negligence in the Enrolment Ecosystem. Furthermore, we find *infrastructural loopholes, breaches in UIDAI partner-institutions, and sub-par hardware* to be the main security vulnerabilities in the Aadhaar infrastructure. Although the CIDR is equipped with the latest security measures, the same cannot be said for non-UIDAI organisations that are part of the Aadhaar ecosystem. Specific technical and legal protocols need to be developed for these organisations along with regular checks by UIDAI and external auditors.

Similarly, using our definition of privacy, we examine media reports alleging privacy concerns within the Aadhaar infrastructure and provide mitigation strategies for the same. We find illegal data storage and data access to be the prevalent concern when it comes to privacy. Our study highlights the need for robust privacy protocols, especially for UIDAI's partner organisations within the Authentication Ecosystem, both, public and private. It must be noted that for Aadhaar to be efficient, it needs to engage and share information with other organisations, but it must be done alongside structural (such as penalties) and technical protocols (such as differential privacy) that ensure resident privacy.

Biometric deduplication forms a major chunk of the Aadhaar ecosystem, however there is not enough information about its security protocols available in the public domain. Future work can seek this information from the UIDAI and build upon our work to perform an analysis of Aadhaar's biometric deduplication system. Since India's Personal Data Protection Bill is still being discussed and debated in the Parliament, future work can employ context-specific definitions of privacy derived from future legislation.

References

[1] Z. Whittaker, "A huge database of facebook users' phone numbers found online," Sep 2019. [Online]. Available: https://techcrunch.com/2019/09/04/facebook-phone-numbers-exposed/

[2] D. McCandless, "World's biggest data breaches & hacks," May 2019. [Online]. Available: https://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/

[3] European Parliament and Council of the European Union, "Gdpr key changes," 2018, https://eugdpr.org/the-regulation/.

[4] J. Panday, "India's Supreme Court Upholds Right to Privacy as a Fundamental Right-and It's About Time," Oct 2017, https://www.eff.org/deeplinks/2017/08/indias-supreme-court-upholds-right-privacy-fundamental-right-and-its-about-time.

[5] C. of Experts under the Chairmanship of Justice B.N. Srikrishna, *A Free and Fair Digital Economy Protecting Privacy, Empowering Indians*. Supreme Court of India.

[6] Ministry of Health and Family Welfare, "Digital Information Security in Healthcare Act," Mar 2018, https://www.nhp.gov.in/NHPfiles/R_4179_1521627488625_0.pdf.

[7] S. Agrawal, S. Banerjee, and S. Sharma, "Privacy and security of aadhaar: A computer science perspective," *Economic and Political Weekly*, vol. 52, pp. 93–102, 09 2017.

[8] Tech2News Staff, "Aadhaar Security Breaches: Here are the major untoward incidents that have happened with Aadhaar and what was actually affected," Sep 2018, https://www.firstpost.com/tech/news-analysis/aadhaar-security-breaches-here-are-the-major-untoward-incidents-that-have.-happened-with-aadhaar-and-what-was-actually-affected-4300349.html.

[9] UIDAI, "Aadhaar Dashboard," Live web page, https://uidai.gov.in/aadhaar_dashboard/index.php.

[10] S. Aiyar, *Aadhaar: A Biometric History of India's 12-Digit Revolution*. Delhi, India: Westland, 2017, vol. 2.

[11] UIDAI, "Vision & Mission," Live web page, https://uidai.gov.in/about-uidai/unique-identification-authority-of-india/vision-mission.html.

[12] V. Goel, "'Big Brother' in India Requires Fingerprint Scans for Food, Phones and Finances," Apr 2018, https://www.nytimes.com/2018/04/07/technology/india-id-aadhaar.html.

[13] UIDAI, "Aadhaar FAQ," Live web page, https://www.uidai.gov.in/298-faqs/enrolment-update/enrolment-partners-ecosystem-partners/2014-what-are-the-fifteen-commandments-that-an-operator-must-remember.-during-resident-enrolment.html.

[14] ——, "Enrolment Agencies," Live web page, https://uidai.gov.in/ecosystem/enrolment-ecosystem/enrolment-agencies.html.

[15] ——, "Setting up and Managing an Enrolment Centre," 2018, http://www.nictcsc.com/images/AadhaarProjectTrainingModule/EnglishTrainingModule/module_3a_settingup_managing_enrolment_centre17122012.pdf.

[16] ——, "Roles and Responsibilities of Verifier and Introducer," Live web page, https://www.uidai.gov.in/images/training_nov_17/Roles_Responsibility_Verifier_Introducer_05122017.pdf.

[17] UIDAI, *Aadhaar Registered Devices - Technical Specification*, 1st ed. New Delhi, Delhi: MeitY, 2017, vol. 2.0, https://uidai.gov.in/images/resource/aadhaar_registered_devices_2_0_09112016.pdf.

[18] Ministry of Electronics and Information Technology, "Aadhaar technology & architecture," 2014, https://archive.org/details/Aadhaar-Technology-Architecture/page/n2.

[19] UIDAI, *Request for Empanelment of Enrolment Agencies*, ser. Empanelment Of Enrolling Agencies. New Delhi, India: MeitY, 2017, https://uidai.gov.in/images/RFE_SEPT_Final_11092017.pdf.

[20] ——, "Questionnaire - UIDAI Operators," 2011, https://uidai.gov.in/images/training-2019/QuestionBank-Operator-510/English_510QB_24012019.pdf.

[21] UIDAI and MeitY, "Training, Testing and Certification," 2019, https://uidai.gov.in/aadhaar-eco-system/training-testing-certification-ecosystem.html.

[22] UIDAI, IDBI Bank, "Memorandum of Understanding - UIDAI and IDBI Bank," 2011, https://uidai.gov.in/images/mou/partners/mou_idbi.pdf.

[23] UIDAI, "Registrars - Enrolment Ecosystem," Live web page, https://uidai.gov.in/ecosystem/enrolment-ecosystem/registrars.html.

[24] UIDAI, "Demographic Data Standards and Verification procedure (DDSVP) Committee Report," 2009, https://uidai.gov.in/images/UID_DDSVP_Committee_Report_v1.0.pdf.

[25] A. Srivas, "Millions of Rural Indians May be Hit as UIDAI Ends Contract With CSC Network For Aadhaar Enrolment," Feb 2018, https://thewire.in/tech/millions-may-affected-uidai-centres-csc-network-clash-renewal-aadhaar-services-contract.

[26] Ministry of Law and Justice and Government of India, "The Aadhaar (Targeted Delivery of Financial and Other Subsidies, Benefits and Services) Act, 2016," p. 3–4, 2016, https://uidai.gov.in/images/targeted_delivery_of_financial_and_other_subsidies_benefits_and_services_13072016.pdf.

[27] UIDAI, "List of Live Authentication User Agencies (AUAs)," Aug 2018, https://uidai.gov.in/images/list_of_live_aua.pdf.

[28] ——, "List of Live KUAs," Aug 2018, https://uidai.gov.in/images/list_of_live_kua.pdf.

[29] ——, "Authentication Requesting Agency," Live, https://uidai.gov.in/ecosystem/authentication-ecosystem/authentication-requesting-agency.html.

[30] ——, "Operation Model," Live, https://uidai.gov.in/ecosystem/authentication-ecosystem/operation-model.html.

[31] UIDAI and MeitY, "Circular No. 1 of 2018: Enhancing Privacy of Aadhaar holders - Implementation of Virtual ID, UID Token and Limited KYC," 2018, https://uidai.gov.in/images/resource/UIDAI_Circular_11012018.pdf.

[32] MeitY and UIDAI, "Compendium Of Regulations, Circulars & Guidelines For ASA and AUA," 2018, https://uidai.gov.in/images/resource/compendium_auth_19042018.pdf.

[33] Compliance Uncovered, "Aadhaar Data Vault - To whom it applies," Sep 2018, https://complianceuncovered.com/2018/09/03/aadhar-data-vault-to-whom-it-applies/.

[34] UIDAI, "Aadhaar Authentication API Specification – Version 2.0," 2017, https://uidai.gov.in/images/FrontPageUpdates/aadhaar_authentication_api_2_0.pdf.

[35] JISA Softech Pvt Ltd, "Aadhaar Data Vault," 2018, https://www.jisasoftech.com/aadhaar-data-vault/.

[36] Thales, "Complying with UIDAI's AADHAAR Number Regulations," 2018, http://go.thalesesecurity.com/rs/480-LWA-970/images/Thales-UIDAI-AADHAAR-cb.pdf.

[37] CERT-In, "Empanelled Information Security Auditing Organizations," 2018, https://www.cert-in.org.in/PDF/Empanel_org.pdf.

[38] UIDAI, "List of Live Authentication Service Agencies (ASAs)," 2019, https://uidai.gov.in/images/list_of_live_asa.pdf.

[39] ——, "Aadhaar E-KYC Specification - Version 2.0," 2016, https://uidai.gov.in/images/aadhaar_ekyc_api_2_0.pdf.

[40] Idemia, "Aadhar-based digital ID customer registration: a new era for the Indian telecommunications industry," May 2017, https://www.idemia.com/news/aadhar-based-digital-id-customer-registration-new-era-indian-telecommunications-industry-2017-05-17.

[41] A. Juels and M. Sudan, "A fuzzy vault scheme," in *Proceedings IEEE International Symposium on Information Theory,*. Lausanne, Switzerland: International Symposium on Information Theory, June 2002, pp. 408–.

[42] H. Liu, D. Sun, K. Xiong, and Z. Qiu, "Selecting distinctive features to improve performances of multidimensional fuzzy vault scheme," in *Biometric Recognition*, Z. Sun, J. Lai, X. Chen, and T. Tan, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 229–236.

[43] Supreme Court of India, "Writ petition (civil) no. 494 of 2012," Proceedings of Judgement, 9 2018.

[44] Cybersecurity and I. S. Agency, "Insider threat mitigation," 2019. [Online]. Available: https://www.dhs.gov/cisa/insider-threat-mitigation

Appendix A.
Categorization and Requirements for AUAs and KUAs

See figures 5 and 6.

S. No.	Organisation Category
Category 1	**Government Organisation**
1.1	A Central/ State Government Ministry/Department and their attached or sub-ordinate offices.
1.2	An undertaking owned and managed by Central / State Government (PSU)
1.3	An Authority constituted under the Central / State Act/Special Purpose Organisation constituted by Central/State govt.
Category 2	**Regulated Service Providers**
2.1	Regulated / Licensed by RBI – Banks and Payment & Settlement System
	2.1.1 Public Sector Banks (PSB)
	2.1.2 Private Banks, Foreign Banks Licensed by RBI to operate in India, Payment Banks, Small Finance Banks
	2.1.3 Regional Rural Banks
	2.1.4 Co-operative Banks 1. State Co-operative Banks 2. District Co-operative Banks 3. Scheduled Urban Cop-operatives Banks 4. Non Scheduled Urban Co-operative Banks
	2.1.5 Payment& Settlement System Network 1. Financial market infrastructure 2. Retails payments Organisation 3. Cards payment network 4. ATM networks 5. Pre-paid payment instruments 6. White label ATM operators 7. Instant Money Transfer
	2.1.6 Non-Banking Financial Company
2.2	Regulated by IRDA/PFRDA - Financial Institutions
2.3	Regulated by TRAI – Telecom
2.4	Regulated by CCA – Certifying Authority, Digital Locker providers, e-Sign providers
2.5	Regulated by SEBI – KYC Registration Agency (KRA),Depository Participant (DP), Asset Management Company (AMC), Trading Exchanges, Registrar and Transfer Agents
2.6	Regulated by National Housing Bank
2.7	Regulated by DGCA/AAI(AAI Act)- Duly licensed- 1. Airport operators having scheduled civil aviation operations, and 2. Scheduled Airline operators.
S. No.	**Organisation Category**
Category 3	**Other Entities**
3.1	3.1.1 Company registered in India under the Companies Act 1956 / The companies Act 2013 (Company under group of companies has to apply individually)
	3.1.2 Partnership registered under the India Partnership Act 1932 or under the Limited Liability Partnership Act, 2008
	3.1.3 Proprietorship firm
	3.1.4 Not-for-profit Organisations (under section 25 under The Companies Act 1956)
	3.1.5 Academic Institutions / Research and Development Organisations
	3.1.6 Societies registered under Indian Societies Registration Act, 1860or The Indian Trust Act, 1882 or The companies Act, 2013 (Sec 8) / Co-operative Society Act 1912
	3.1.7 Any entity other than above mentioned categories

Figure 5. Categorization of AUAs for Appointment by UIDAI [32]

S. No	Authentication User Agency (AUA)		Additional requirements for eKYC User Agency (KUA)
	Technical Requirements	**Financial Requirements**	
Category 1	1. Backend infrastructure, such as servers, databases etc. of the entity, required specifically for the purpose of Aadhaar authentication, should be located within the territory of India. 2. Entity should have IT Infrastructure owned or outsourced capable of carrying out minimum 1 Lakh Authentication transactions per month. 3. Organisation should have a prescribed Data Privacy policy to protect beneficiary privacy. 4. Organisation should have adopted data security requirements as per the IT Act 2000	No financial requirement	No additional requirement for KUA
Category 2		No financial requirement	No additional requirement for KUA
Category 3	1. Backend infrastructure, such as servers, databases etc. of the entity, required specifically for the purpose of Aadhaar authentication, should be located with in the territory of India. 2. Entity should have IT Infrastructure owned or outsourced capable of carrying out minimum 1 Lakh Authentication transaction per month. 3. Organisation should have a prescribed Data Privacy policy to protect beneficiary privacy. 4. Organisation should have adopted Data security requirements as per the IT Act2000. 5. Entity should be in business for minimum of 1 year from date of commencement of Business.	1. Paid up capital of Minimum₹1 (one) Crore. OR Annual turnover of Minimum₹5 (Five) Crore during the last Financial year.	Entity should meet Authentication Transaction Criteria as laid down by the Authority from time to time.

Figure 6. Financial and Technical Requirements for Entities seeking to become AUAs [32]

Appendix B.
Categorization and Requirements for ASAs

See figures 7 and 8.

S. No	Organisation Category
Category 1	A Central/ State Government Ministry / Department or an undertaking owned and managed by Central / State Government
Category 2	An Authority constituted under the Central / State Act
Category 3	Any other entity of national importance as determined by the Authority
Category 4	A company registered in India under the Indian Companies Act 1956
Category 5	Any AUA or KUA meeting authentication transaction criteria as laid down by the Authority from time to time

Figure 7. Categorization of ASAs for Appointment by UIDAI [32]

Category	Financial Requirement	Technical Requirement
Category 1, 2 and 3	No financial requirements	No technical requirements
Category 4	An annual turnover of at least Rs. 100 crores in last three financial years	A Telecom Service Provider (TSP) including All Unified Licensees (having Access Service Authorization) / Unified Licensees (AS) / Unified Access Services Licensees / Cellular Mobile Telephone Service Licensees operating pan India fiber optics network and should have a minimum of 100 MPLS Points of Presence (PoP) across all states OR Should be a Network Service Provider (NSP) or System Integrator having pan-India network connectivity for data transmission and should have 100 MPLS PoPs in India,
Category 5	No Financial requirements	Any AUA or KUA meeting authentication transaction criteria as laid down by the Authority from time to time

Figure 8. Financial and Technical Requirements for Entities seeking to become ASAs [32]