DIDs and Personal Data Storage for Children

Jonathan Endersby, Shaun Conway, Lohan Spies
Consent https://consent.global @globalconsent

Submitted to the 3rd Rebooting the Web of Trust Technical Workshop as a discussion paper.
October 2016

Introduction

The South African government runs a subsidy scheme for Early Childhood Development Centres (Creches for 0-6 year old children) to ensure that children from poor families get access to quality education, care and nutrition.
This program is currently operated mostly as a manual system with paper forms. However, this pays out many millions (USD) funding each month and the government wishes to increase funding in future years as a national priority. Large sums of money, remote communities, poor oversight, and a lack of infrastructure have inevitably created a situation where it is difficult to know how effectively the current system is performing, identify service gaps and prevent fraud, which is currently rife.

There is clearly a great opportunity to leverage technology in ways that can improve the efficiency, accountability and effectiveness of social development and benefit programs such as this.

Consent is implementing a Proof of Concept to create decentralized digital registries of children, their caregivers, service providers and the service centers they attend, using Decentralized Digital Identity standards. This will provide the basis for tracking the delivery of early learning services and administering the government daily attendance subsidy over future years. Verifications will be performed using mobile applications as oracles of the children's attendance.

Goals

Our primary goals are to:

- Bootstrap digital identities for more than 600,000 children.
- Streamline processes and reduce administrative overhead (for service providers and government administration).
- Increase funding efficiency.
- Reduce fraud.
- Provide richer and reliable strategic data.

Challenges & Constraints

Socioeconomic Challenges

1. **Complex Identity**. Many of the children in our target demographic may not be registered with the national registration authority. Some families have lost their paperwork or were simply unable to register the child due to a host of complex reasons. (eg. Living with family members who are not legal guardians, legal status in country etc).

2. **Cost of Connectivity**. While most ECD teachers will own low-cost Android smartphone, the cost of internet connectivity is such that many will not have adequate data on their phones to enable the use of an app that requires internet access.

3. **Guardian Availability.** Many parents work long hours and use public transportation, as a result many children are dropped off at school by a relative or community member. It is not unlikely that the person with the legal authority might not even know where the school is. This is the reality of single parents working 12 hours shifts with additional long commutes.

4. **Education, Literacy and Communication.** It is often difficult to explain what we are doing (both to service providers and parents). This slows down the process or registration significantly.

5. **Fraud and Corruption.** There is a financial incentive to misrepresent the truth. This muddies the water and complicates user experiences.

Legal Constraints

1. **Data Ownership.** Who does this data belong to? What data should belong to the child, parent/legal guardian and service providers?

2. **Rights and permissions**. In the complicated socio economic environment, can we legally create records for an individual when we might not have permission from their guardians?

Technical Challenges

1. **Time.** We are creating identities and user data that needs to be around in 10 years time when the child is old enough to responsibly use the information. Until this time, we will only have proxies or agents.

2. **Biometrics are tricky, expensive and cumbersome**. Children's fingerprints and eyes change, biometrics scanners (dedicated or in smartphones) are very expensive and getting children to go through the process of biometrics is complicated (enrollment and subsequent verification).

3. **Our User is not the User.** Service providers register identity accounts on behalf children of who they are not the legal guardians and who they might only have limited interactions with.

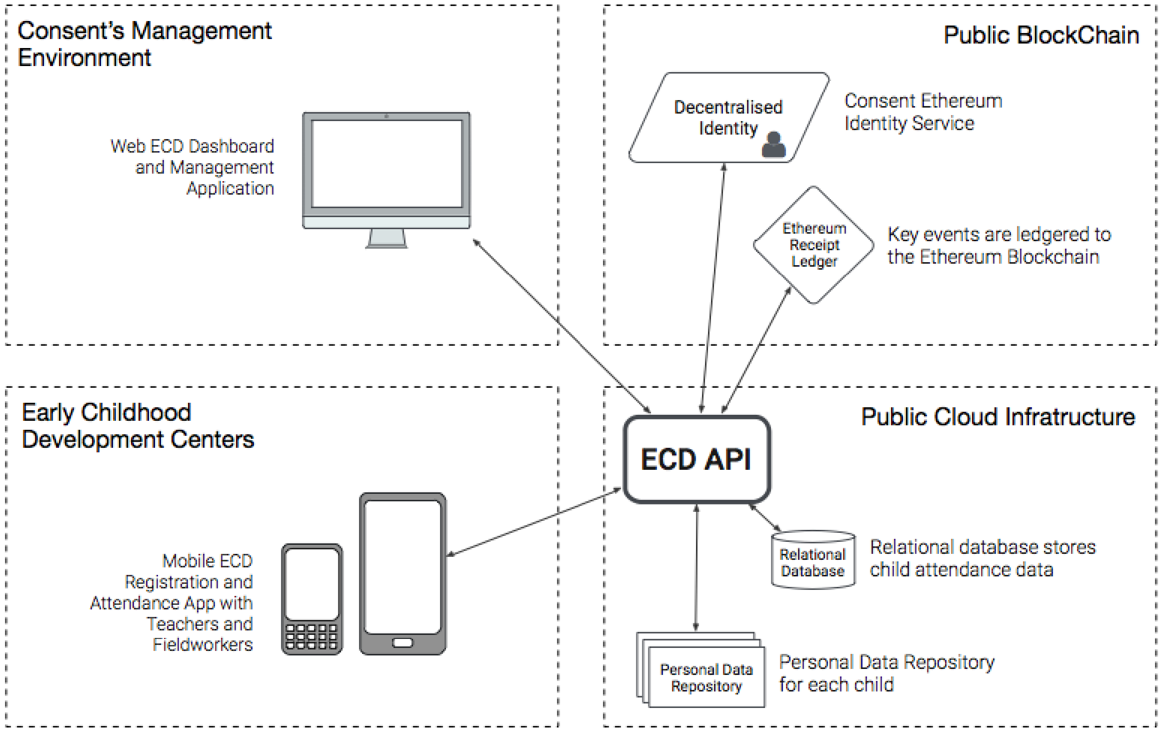
4. **Key Management.** Cryptography needs keys and those keys need to be kept securely, but we cannot expect service providers to take responsibility for securing a child's private keys.

5. **An effectively "informal" data entry environment.** Preschools are chaotic at the best of times, not least when their teacher/child ratio is very high.

The Opportunity

Build a system that efficiently registers 600,000 children living everywhere from cities to rural villages. Track their attendance, providing meaningful strategic planning data and prevent fraud. Do this with limited connectivity and on a decentralised system.

Architecture


Consent's Management Environment
Web ECD Dashboard and Management Application
Public BlockChain
Decentralised Identity
Consent Ethereum Identity Service
Ethereum Receipt Ledger
Key events are ledgered to the Ethereum Blockchain
Early Childhood Development Centers
Mobile ECD Registration and Attendance App with Teachers and Fieldworkers
ECD API
Public Cloud Infratructure
Relational Database
Relational database stores child attendance data
Personal Data Repository
Personal Data Repository for each child

While many of the components above should be clear to the reader, key components are expanded below.

Ethereum Identity Service (EIS)

A smart contract that allows for storing and querying DIDs and associated metadata. The public key of the registered party is stored so that only they can modify their record, enabling an individual to change the referenced public key for a given DID.

Personal Data Repository (PDR)

Personal Data Repository (PDR) is a hosted, secure (encrypted with off-platform keys) container for storing an individual's data. The PDR has a convenient mechanic for securely sharing with, and securely receiving data from, 3rd parties. PDRs can be hosted anywhere on the internet, from inside a secure data center with thousands of PDRs, to running a single PDR on a Raspberry Pi in your home.

Ethereum Receipt Ledger

The receipt ledger is a smart contract that enables the simple receipting of a piece of data, usually a hash of some cleartext that can be referenced later. This provides a ledger of events that can be referenced later by either party.

How does it work?

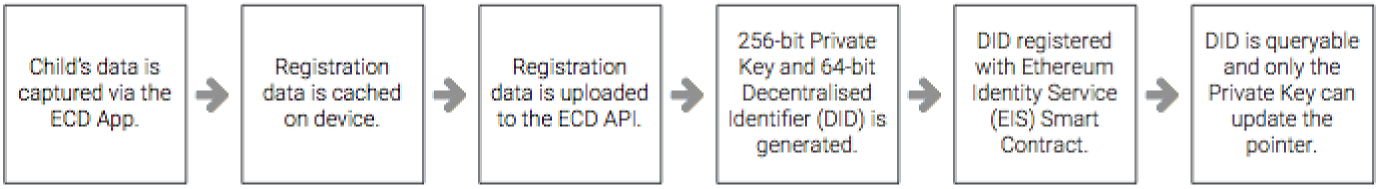
Digital Identity Guardianship

In an ideal world Children would have their digital identities managed and controlled by their parents or legal guardians. This would include hosting their Personal Data Repository, managing backups of their data and ensuring that keys are kept safe.

In practice this is simply not realistic, certainly not in South Africa. To address this challenge we have opted to take on the responsibility of acting as the Digital Identity Guardian for the individuals who enter the Consent system via this Early Childhood Development programme.

Creation of a Decentralised Identifier

The first step is to create and register a unique Decentralised Identity (DID) for each child against which future actions can reference as additional data and verifications are required.


Child's data is captured via the ECD App.
Registration data is cached on device.
Registration data is uploaded to the ECD API.
256-bit Private Key and 64-bit Decentralised Identifier (DID) is generated.
DID registered with Ethereum Identity Service (EIS) Smart Contract.
DID is queryable and only the Private Key can update the pointer.

Once this is done we are able to easily query the EIS service for a given DID and discover metadata for them.

Creation of a Personal Data Repository

The ECD system has a minimum requirements for an identity to be considered complete enough to warrant PDR creation. Once this requirements are met we create a PDR for the child on one of our shared PDR servers. The URL for this PDR is returned and the EIS is updated accordingly.

Attendance Data Tracked

Service Providers using the ECD App (Android only) can register learners, update their data and take a daily digital attendance registry. These data are submitted to our system as soon as the service provider has data connectivity.

Key Data Shared with the Child's PDR

As the ECD system generates data it shares key data with each learner's PDR. Examples of this would be:

- Copies of registration documents
- Registration Photographs
- Aggregated Attendance Data

Eventual Identity Takeover

When a child is old enough to take over their identity we enable them to do so by proving their identity to a nominated agency (likely a retail banking partner). This process will require the individual to prove that they are the person whose identity is being referenced by the DID and Personal Data Repository. This process is likely to be tied closely to the national identity number.

Once they have proved their authenticity we take a number of steps:

1. The user is asked to generate a new private key (likely with the Consent LifeKey App)
2. The public key referenced in their EIS record is updated to reflect their new private key.
3. Their PDR data is re-encrypted with their new private key and the old data is deleted.

Challenges

While we believe we have designed a robust system, we still have a number of concerns and are looking for solutions and would love to chat with anyone who has clever ideas.

1. The system’s reliance on Consent (or another party) to act as the Digital Guardian and host the PDR and keys is counter to true decentralisation.
2. A concern around teenagers taking control of their identity and PDRs but then acting irresponsibly and losing their identity and personal data. We have plans for private key recovery, but these need to be set up properly by the user. Luckily we’re roughly 10 years away from our first child turning 16, so we’ve got time to perfect the process.
3. Establishing a reliable mechanism for a child going through the takeover process to prove their identity.