WHITE PAPER (draft - v1.0)

Abstract— Since the internet became a widely adopted platform for information exchange, internet companies, service providers and government agencies have been gathering data from every user reaching their servers. In addition to this many of them require their users to create accounts and enter personal information to use their services, encouraging people to give away more and more details in order to improve the benefits of their service.

This has become the norm when using platforms like Facebook, Twitter, Instagram, LinkedIn and Amazon, to name a few. The personal information they collect about their users is either provided directly by the users themselves (i.e. names, address, date and place of birth, credit card details, relationship status, friends' details, employment details, etc.) or inferred from their use of the platform (i.e. access frequency and location, search queries, interests, behaviour, personality, etc.). The latter usually even without the user's knowledge or hidden in lengthy terms and conditions.

The problems caused by this are threefold:
- **The data is siloed**— Each platform stores these data in their own servers and in their own custom format.
- **Users have no control over it**— It is hard for users to access, read and delete it when needed.
- **Lack of privacy**— A user's privacy is always at risk. Internal data leaks, hacks and software bugs have been increasing over the past few years making the users vulnerable to exploitation.

Our goal is to create a technology that gives users complete ownership and control of their data, whether it is created by the user himself/herself or generated by 3rd party platforms on their behalf.

This technology will give users:
- **Full ownership** of their data.
- **The power to grant and revoke access** to their data to any 3rd party requesting it.
- A **single source of truth** for their data.
- The **power to monetise** their data.
- A decentralised interface that **provides trust** to all or parts of their data.
- An interface for 3rd party application to interact with the data.

The above will be described in detail in the following sections.

INTRODUCTION

This white paper will provide an in-depth description of '*Identity*', its architecture, implementation, use cases and roadmap. Please note that '*Identity*' architecture, as described in this paper is holistic, all parts are intimately interconnected and explicable only by reference to the whole, which from now on we will refer to as '*Identity*'.

In future white papers we will provide details on how 3rd party applications could take advantage of the technology along with examples and sample code.

THE CORE

Identity's main goal is to build a standard format for the storage of a person's data, whether it is private or public, as well as a set of rules to access it and manipulate it while giving the owner full control over it. In addition to this, the data should be stored on a medium of a person's choosing (offline, online, private server, S3, etc.).

One question that arises regarding a person's data is in respects of trust. How can someone's data be trusted and/or be accurate at all? The answer, we believe, lies in peer reviews. The trust of a person's data (or small fractions of it) will increase based on the number of peers who trusted it, as well as the reputation of the peers. This can be computed with the *PeerRank* algorithm, listed below:

$$PR(p_i) = \frac{1-d}{N} + d \sum_{p_j \in M(p_i)} \frac{PR(p_j)}{L(p_j)}$$

A person's data could be divided into 3 categories:
1. **Self provided data**— A person may input their own data just as they would on a social media platform, bank or online store. The idea is that the person will have a single, private source of storage and when a 3rd party asks for some of it, the person will be able to grant or deny access to it.
2. **Peer generated data**— In order to add an extra layer of trust to a person's data (or a subset of it) it can be peer-reviewed. This means a single piece of data can be abstracted into a Smart Contract in the Blockchain and linked to the person's data via its contract and wallet addresses. This will allow known peers to review the data for accuracy and leave their mark, validating it.
   Anyone could potentially validate a person's data, however due to the nature of the Blockchain, storing any data involves a small gas fee which consequently helps to avoid spamming, abuse and over-validation of data.
   A 3rd party wanting to check the validity of a person's data can do so with our PeerRank algorithm which scores the quality of the data by scoring the quality of the peers who have reviewed it.
3. **3rd party service data**— a person may signup to 3rd party services by providing some of their data,

this 3rd party service may record data on the person's behalf (usage, behaviour, stats, etc.), a person may request this data for their personal records and for sharing with other services.
Due to current regulation, should a person request it, a 3rd party service should always be able to:

1. Provide a person with all data captured and associated to them.
2. Delete all data associated to them.

By providing the above mechanisms a person will always have full control and ownership of their data and be able to store it online or offline, delete it or share it with any other person or company on demand.

THE STANDARD

A person's data should follow a certain standard so that it can be understood by any 3rd party service without the need for building any custom adapters.
Identity has set JSON as the data-interchange format of choice due to its versatility and widespread usage.
The data should follow a modular approach where a person's general data (name, address, date of birth, etc.) should be under a 'general' key and 3rd party app data should be a unique identifier for the 3rd party app.
Data duplication should be avoided whenever possible by referencing the key of original path of the data.

An example is illustrated below.

```
{
  "general": {
    "firstName": "John",
    "lastName": "Doe"
  },
  "someApp": {
    "username": "john.doe",
    "firstName": {
      "$ref": "#/general/firstName"
    }
  }
}
```

A person's *Identity* may consist on a variable amount of data, it is up to the person and the 3rd party services they interact with to build these data. Due to this, any service can be plugged into a person's *Identity* and interact with other services and 3rd parties, making a person's *Identity* highly modular.

ENCRYPTION

A person's data should be encrypted when:

- It is stored
- It is being transferred

This means it is only decrypted when being read or written, reducing the probability of misuse or data theft considerably.

The method of encryption is by using a person's public key from their account's address. Decryption can be performed in the same manner using the private key, thus ensuring only the person who owns the account can read the data.

When granting access to a 3rd party service to a certain part of a person's data, the person will encrypt these data with the public key of the receiving 3rd party service. This guarantees that nobody else will be able to read the data other than the receiving person.

When a 3rd party service wants to write data to a person's *Identity* they will encrypt the data with the person's public key, and when the person receives it they will be able on decrypt it with their private key and add the data to their *Identity* if they so choose.

PRIVACY

One of our main concerns is privacy and compliance with EU legislation. A person should be able to delete his/her data at any time, which leads us to create two types of data: volatile and semi-volatile.

Volatile data— This data is fully controlled by the person. It can be stored locally on the person's device or on the cloud upon a person's request. The person may delete it partially or completely at any time. Any 3rd party services using these data will no longer have access to it.

Semi-volatile data— This data is controlled partially by both, the user who created it and the 3rd parties it had interacted with. It can can only be deleted upon mutual agreement and after paying the gas fee required to update the Blockchain thus avoiding abuse and spam. This data should not be of a sensitive nature and should only consist of wallet addresses and untraceable ID's.

DATA SHARING

A person is fully responsible for their data, its accuracy and how, when and who it is shared with.
There a two different approaches for a person to share their data with a 3rd party:

1. **Publicly**— Although it is not recommended, a person is able to publish their data publicly, should they choose to.
2. **Privately**— A 3rd party may request access to all or parts of a person's data. The person may or may not grant access to it. Should access be granted, data

must be shared as described in the 'ENCRYPTION' section.

SIGNATURES AND VALIDATION

Occasionally a 3rd party might require a certain entity (bank, insurance company, local authority, etc.) to validate a piece of data on behalf of a person. This can be achieved using any of the approaches listed below:
The parties involved are: Identity holder, authorised entity, 3rd party service.

1. **Acting as intermediary**— The *Identity* holder requests a signature from the authorised entity providing the public key of the 3rd party service. If granted, the *Identity* holder is then responsible for sending this signature to the 3rd party.
2. **Acting as a customer**— The *Identity* holder instructs the authorised entity to send a signature to the 3rd party directly.
3. **Bypassing the Identity Holder**— The 3rd party may request the authorised entity a signature on behalf of an *Identity* holder, without the *Identity* holder's involvement.

The signature mentioned above may consist on any type of data, as requested by any of the parties.

REPUTATION

A person's *Identity* is never complete until 3rd parties are able to interact with it via Smart Contracts. These interactions increase a person's reputation in the network which consequently affects the reputation of the network itself.
Reputation cannot be bought, it has to be earned via these interactions.

An interaction increases a person's reputation in a certain area of their data, based on the reputation of the 3rd party who interacted with the person's data (and this can be either positive or negative, increasing or decreasing a person's reputation).

As an example:
- Person A has a reputation of 5
- Person B has a reputation of 10
- If Person A interacts with data from Person B, Person B's reputation will increase by 1
- On the other hand, if Person B interacts with Person A, Person A's reputation will increase by 2, since Person B's reputation is double that of Person A.
- However if at some point in the future Person B's reputation decreases (due to lost trust or misbehaving in the network) all the interactions it performed will have a negative effect on all persons it interacted with in the past, thus punishing misconduct at a network level.

SCALABILITY

The modular nature of *Identity* allows the technology to, theoretically, scale indefinitely. This holds true when assuming no storage or computational constraints.
In practice, a person's *Identity* is bound to get larger over time (depending on usage and varying between persons). In most cases this size won't be greater than a few kilobytes but for some persons linking their *Identity* with various social networks and storing every possible piece of data they gather can make the size considerably larger.

The data is kept encrypted at all times except when it is read or written, this means it needs to be encrypted and decrypted every time, incurring a memory overhead for such actions. The memory required is determined by the system it is running on and the size of the data. In most cases this should have no UX impact for the end user since it will not take longer than a few milliseconds. However as the data size increases so will the memory required to encrypt/decrypt the data in a reasonable time.
In order to keep the size of the data small and maintain a more constant computational expense, the data should be encrypted at a 'value' level instead of at a 'key/value' level. Meaning that decrypting a person's Identity won't necessarily decrypt the whole *Identity* but just the parts that are needed to complete a task.

SUBSCRIPTIONS & EVENTS

A powerful feature of *Identity* is the ability to implement subscriptions. A 3rd party may request a person to subscribe to part of their data. If the subscription is accepted, whenever the data changes, the 3rd party will be notified of the change and get instant access to the updated data until revoked.
At any time, a person may view his/her subscriptions and subscribers and unsubscribe or revoke access with immediate effect.

IDENTITY APP & WALLET

This is the source of a persons's *Identity*, it contains all the data that a user generates over time, by default it is stored locally on the device that was generated, however the person has a number of options on where to keep it:
- Offline— on the device that was generated; can't be used on any other device
- Online— Stored:
  - Privately: The person takes care of storing the data online
  - On *Identity*'s server

- On a 3rd party service

The data will never leave the person's device unencrypted, it will always be encrypted using the person's public key and decrypted only via the person's private key.

Upon registration the person is provided with the option of creating a new wallet account or importing an existing one. This wallet address will be the unique identifier for the *Identity* and will be part of the person's *Identity* for interacting with any other online service or platform.

IDENTITY TOKENS

Identity tokens are the main currency throughout the system. They can be exchanged for services that range from implementing 3rd party services to centralised data storage. In addition to that they can be bought and sold in the marketplace and used to create and manage *Identity* accounts for businesses.
These tokens are also awarded to *Identity* account holders on a monthly basis based on their reputation in the network as well as participating 3rd party services. *Identity* tokens has a limited supply thus increasing in value as their supply scarce.

IDENTITY ACCOUNTS

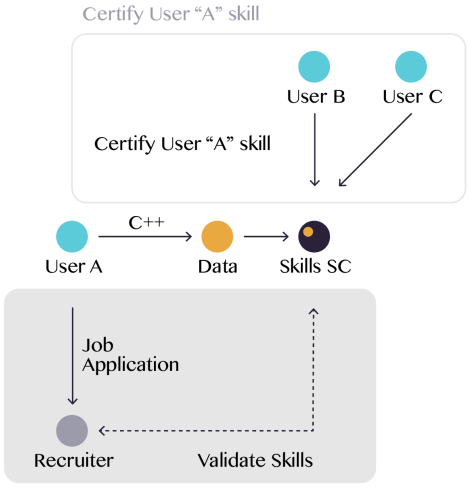
There are two main types of accounts:

- **Individual**— Any person is entitled to an *Individual Identity* account, they are free and can take advantage of any available 3rd party services. These accounts get awarded a variable amount of *Identity tokens* on a monthly basis determined by the accounts reputation increase on each given month.
- **Business**— Businesses can also open *Identity* Accounts, these accounts are managed by one or more *Identity* holders, they are also free and can take advantage of 3rd party services including some specify business related services (business verification, finance, ...).

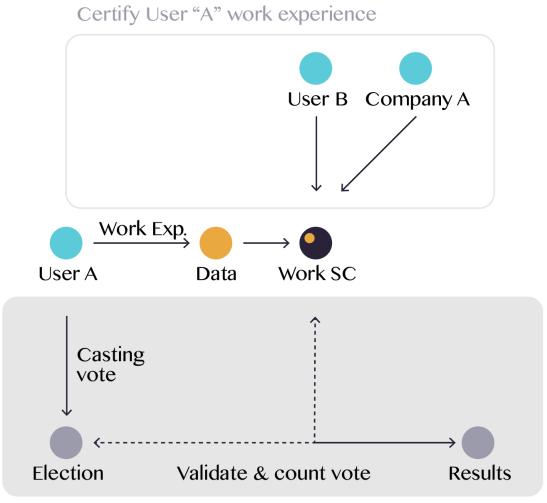
Opening an account can be achieved via any of the *Identity* apps.

USE CASES

Recruitment— Two smart contracts can be adopted by a person to list his/her skills and work history (company and job title).
Each skill is stored in the the Blockchain incurring a gas fee per entry, reducing the number of irrelevant skills submitted by the person. The person will then be able to share their skill's address with peers which will then be able to validate it. The more validations by high ranking peers a skill has the more confidence a possible recruiter will have of it being true, considerably reducing the time required to find, research and interview candidates.
A similar approach will be taken regarding a person's work history and work references.
The companies a person has worked for in the past (as well as colleagues working for those companies) will be able to validate a person's role at the company, giving a person's experience more credibility than traditional CV's and reference letters.



Voting— Voting has been a controversial subject lately, we believe that we can enhance the process in terms of trust and outcome considerably by using not only Blockchain technology and a person's *Identity* but also a *Weighted Voting* strategy.



Thanks to the input of related smart contracts (ie. skills, work history, interests, ...) an entity creating a voting event would be able to specify which traits of an individual are valuable and should bear a higher weight

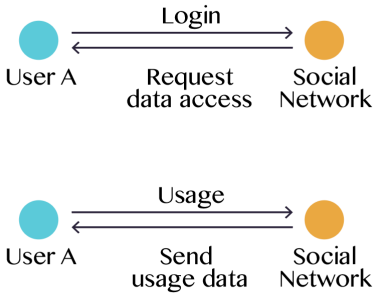
in the vote. Upon this being agreed, the voting platform will compute and take into consideration the weight of each individual and propose a result based on this.
By default all votes count as 1, but for highly skilled or relevant individuals, their vote can count as much as 50% more (1.5). Resulting in a fairer outcome.
Some possible use cases are:
- Elections
- Shareholder decisions (weighted based on stake held)
- Community decisions (weighted based on involvement, popularity, experience, ...)

Finance— The financial sector spends billions of dollars every year gathering financial information on individuals applying for loans, mortgages, credit cards...
Credit scores are biased, outdated and in many cases erroneous. These problem can be solved and costs can be significantly reduced by using certified data from persons as well as improving the processes of data sharing between any involved entities.
An example of this is a person requesting a loan from company A. Company A wants to know if the person will be able to repay the loan, a common way to determine this is by looking at bank statements from the person, which the person must submit to company A, leaving the door open for erroneous material.
Instead the proposed solution is for the person to request the last 3 bank statements from their bank providing it with the public key of company A, the bank will send the person these data encrypted with given key, this will assure the the user has no way to temper with the data and so it will arrive intact to Company A which then will be able to decrypt with their private key.

The same example can be applied to any other transaction which requires the transfer of sensitive information between two entities in the context of a requesting person.

Social— Having an *Identity* account removes the need for social media platforms to request and store more data than they need from their users. A much needed update to current social media networks needs to occur. The proposed update is to allow users to sign up, login and link their social media accounts with their *Identity* account and request access to the data needed to create the profiles. Later the social networks should no longer store any sensitive data on their servers, instead push it to the user's *Identity* thus maintaining a 2-way communication, giving the user a greater degree of trust towards the social media platform.

The benefits of this approach include:
- Users have full control over their data.
- Users decide what to share, when and who with.
- Data leaks and hacks will mostly be avoided.
- Advertising revenue could be potentially shared between the social network and the user.



Dating— Dating apps have made it simple for people to meet people alike. However there is still a large empty space between what a person says in a dating profile and what a person is really like.
By using a person's *Identity* on a dating profile certain aspects a the person's personality can be known and trusted beforehand, reducing the risk that meeting a stranger entails.

ROAD MAP

Phase 1—Working Identity app in Beta where users can create their identity and store it locally on their device. Along with this app there will be an extra 3rd party service created, for which, we have chosen a 'recruitment' service with the initial feature of 'skill' validation. Where users list their skills and have other users/colleagues validate them in order to create a decentralised reputation.

Phase 2—The Identity app will offer users the option to store their data in multiple devices and providing real time synchronisation between them. In addition to this there will be an option to store some or all of the data in the cloud, encrypted with each individual's private key. This will guarantee the Identity's availability and access from any 3rd party services which have been granted access.

Phase 3—More 3rd party services will be developed, these will include "Voting" and "Business validation" between other yet to be defined.
This phase will also see the development of the feature 'Sign in with Identity'. Which will allow any platform implementing it to request certain data from a user's Identity in order to instantly get access to read and write to it.
This will allow users to be in full control of the data captured by these platforms. ie. read, delete and remove access.