Digital Identity Manifest

Context

The surprising upcoming of computers within only a few decades is an evolutionary event, instantaneous compared to the eons living creatures have spent on earth. The discovery of the nuclear fire and insight into the inner workings of particles have empowered man to shape material into myriads of tiny calculators, connected together by a network covering the whole planet. At the same time, human population has grown almost exponentially to an amount where side-effects of human lifestyle become more and more visible and questions of resource management start to become more pressing. The computer network could play a helpful role in the answer and that is what this text is about.

Scope

This document outlines the concepts and structures used to model what is called the digital Identity, without going into the any detail of mathematical formulation. Proposals to that will be published in separate documents. We are not going to consider any technical aspects of handling cryptographic secrets in this writing. While this is a very important problem, its solution cannot be purely technical. This information is addressing a general audience, avoiding technical notation and discussion wherever possible.

Vision

It has been stipulated that the computer network could be used establish a new level of democracy. This is certainly true and some advance was made, yet an acceptable solution is outstanding, as it comes to electronic votes and other civil rights. The people are wise not to urge the matter prior understanding, which can take time. Before accepting any computer based system as a replacement for whatever is in place now, it should be field tested in other applications which are less susceptible to mass manipulation.

Digital identity starts with commerce.

The advantage of commercial application as a starting point is mainly that it's optional. Gold, cash and credit-cards still coexist as payment methods and there is little need to enforce their replacement. If digital identity can be used to perform standard commercial transactions, such as payments, it is much more useful than any sort of Citizen Identity limited to voting every couple of years.

Once, however, digital identity is established in a society to a certain degree, for commercial use only, it can readily be used for political purposes as well, such as collecting votes for a referendum. The questions that arise from such undertaking are part of the evolutionary frame for the development of digital identity towards an instrument of democracy and that will take time.

It is tempting to imagine all the positive effects of the idea, including solutions to complex topics such as resource allocation, employment and ecology. It is far out of scope of this writing but the reader is encouraged to employ his fantasy. We restrict ourselves to the question of what the fundamental structures of such a digital identity might be and we'll try to formulate some requirements.

Claim

The ubiquitous presence of network enabled computers in peoples everyday life demands regulations on how much personal information these devices shall collect, process and disclose to the network. It is necessary to define control instruments and structures that ensure that devices act as intended by their owners. While these mere facts are well understood and the underlying principles were already put into the legal framework of many countries, the result remains very opaque to the individual.

To be of practical use regulations must be stated in a machine readable way

This also means, that inner workings of the devices and services must be disclosed to their users to an extent that allows them to verify every operation performed by the device or service was executed in a way conforming to what they agreed to. This is not equivalent to force device manufacturers to open all their secrets, not at all indeed.

> To illuminate the question, consider first the case of a *computer game*. The extent to which the manufacturer of the game needs to explain the inner workings of the game is fully under his control. There is no need a priory external rules, at last it's only a game. Should the players be unsatisfied, they can simply stop playing it.
>
> In contrast to that, in case of a *medical consultation* for example, we would expect much better assertions to what exactly who is doing in the process.

It is claimed that a *mathematical framework* describing computer based interactions is the first and most fundamental prerequisite for what shall be called the *Digital Identity* **DIGID** in the following. Based on that, individuals and organizations can independently verify the correctness of claims formulated in the language of the framework. This is much more rigid than any existing method and by choosing a suitable language, does not impose limits on how complex and specific such claims might be.

Basics

The mathematical framework for DIGID is based on calculations on finite sets of integers which are typical operations on a computer. A specific set of functions (cryptographic algorithms) is chosen so that a *digital signature* can be created on an arbitrary input using a *secret key*. If the secret key is considered to be accessible only to its owner, this already constitutes a simple form of DIGID which can be used to *prove knowledge* and to *authenticate* the owner.

Based on this we can already formulate an application if we allow a set of owners, called participants, to exchange signed documents, without going into any detail of what the content of these documents looks like or how they are exchanged. Instead we just assume that these are basically the same documents as are used with conventional signatures. Obviously, if digital signature is equivalent with conventional signature (a legal question) it can serve the same purpose. This is important because it provides a smooth transition between the old and the new system. But there is much more now.

Every interaction between participants creates a certain chain of (possibly signed and linked) documents. This may be to the advantage or disadvantage of participants, depending on the circumstances.

> If you signed a sales confirmation and you fail to deliver, the customer can use it not only to demand the money back from you. In addition he might also ask for compensation, especially if such compensation has been stated in the purchase order underlying the conformation.

As the number of documents involved can become large, one of the most basic structures of DIGID will be a method to store those documents, for the lifetime of the owner and well beyond this, if desired.

Naturally we would expect that some of the documents used in interactions are machine readable, so that computers can do the work of analyzing the content to provide useful information to the owner. The result of this analysis should be prepared so that it can answer particular questions, namely those that ask for trustworthiness or credibility of another participant.

It's left to mention here, that cryptographic algorithms allow for far more than just a simulation of our traditional ink signature. Besides for signatures, secret keys can be used for authentication and privacy, when accessing some resource, for example. Very interesting applications arise from group operations, where multiple participants act together to create a proof of their collective decision.

Architecture

The inner structures of DIGID are all graphs of different kinds. Graphs can readily be represented by computers and a rich set of tools is available to cope with them, we will not go to that level here. It should be clear that many different ways to implement DIGID exist, we don't want to create any prejudice.

The requirement underlying this architecture description here can be summarized as follows:

- An instance of DIGID is a collection of keys, corresponding certificates, security related data and other data.
- The access to keys that the DIGID uses to interact with others is controlled by a special process called **agent**.
- The agent maintains a history of changes and all data associated with the DIGID.
- The agent can issue certificates to be used for authentication with third parties.
- The agent can interact with other agents to fulfill tasks demanded by the owner or others, if allowed to do so.
- Except for private keys, the agent operates on data that is structured in a publicly accessible format. This implies that, for example, another agent implementation could be launched on some data created by the first implementation.
- The agent must provide an interface to the authorized user to request the full history of the system, including all data, except the private keys.

Security and Cryptography

From recent considerations it has been decided to choose Elliptic Curve cryptography exclusively. This allows, in contrast to RSA, richer schemes and protocols to be used. Besides that, it consumes less space and resources and could be considered more secure.

The security component consist of services operating on the secret values associated with the DIGID. It is therefore the most sensitive component in the system. Other components relay on the security to operate correctly.

Namespace

The namespace graph is a tree and resembles what we know from file systems FS, including directories, files and possibly attributes. There is one distinction, however: while most FS represent a single state (the current state), the namespace shall be able to represent more, namely not only the current state, but also all the states preceding it (the history), down to the initial state, which is represented by an empty namespace.

> Consider a situation where one makes a decision based on the content of some of the files in the namespace. One file could, for example, hold information about a price of something. Now the outcome of your decision is an action, possibly a purchase or sales order. Prices can changed, so you would not expect the same outcome of your decision if you asked at another time. The namespace structure is designed to answer the question given a parameter to specify the time explicitly.

The namespace graph comes with a serialization format that allows to extract sub trees and histories in a form that can easily be imported into another instance of a namespace. This allows to effectively distribute the physical materialization of a namespace over a heterogeneous infrastructure by means of replication.

Trust Graph

Given the vast amount of documents, including those resulting from interactions as well as others, originating from the owner directly, we need some additional structure to represent the implications of these documents in a suitable way. This applies to documents that are machine readable.

We assume that some machine readable documents contain information that expresses how much a certain property applies to a certain participant. Such documents can be extracted and the resulting structure can be represented by a directed, possibly cyclic graph, called the Trust Graph subsequently. Now queries on this graph can be used to ask if a property, or even a more complex expression involving multiple properties applies to a participant. By adding thresholds, the computer can handle obvious cases silently, while presenting us the rest for manual decision.

This can be extended to other objects, even simply all documents giving us the possibility to qualify these objects. Such a qualification being an object itself gives use the option to talk about the qualification itself and so on.

Agent

The agent is the very heart of every DIGID instance. The agent uses the key management, namespace and trust graph components to provide a service that acts for the benefit of his owner. This involves routine validation of certificates and other security related data. But it also includes keeping track of all decisions the owner has chosen to remember.

In most cases, individuals will not be able to operate their own agents, due to high costs related with running a server. Instead agents will be run by trusted third parties. It is desirable to make agents free to move amount different providers, at this will and cost of their owners. This shall guarantee, that the owner has the option to use multiple, different providers and to change the provider whenever he wants.

Providers operating agents should adhere to standards on how this is to be done. These standards must be available for the public.