Mickey Mouse Model (MMM):
A global solution for safe and secure data sharing

Author: Paul Knowles, Head of the Advisory Board at The Human Colossus Foundation

In the midst of this COVID-19 pandemic, worker bees in the *Decentralized Identity* and *MyData* communities are frenetically throwing credential and schema constructs into the melting pot, not to mention UX [user experience] designs and dApps [decentralized apps], in a grand effort to quash the coronavirus. Although this boundless energy must be largely commended, community members must also take a step back in order to distinguish the wood from the trees and to better understand which part of a Decentralized Data Economy (DDE) their component resides and, in doing so, where to draw a line so as to not infringe on a different part of the model that may fall beyond the scope of their expertise. This scope creep can muddy the waters and become detrimental to the overall mission.

All components in a balanced DDE have male and female counterparts. The following component diagram indicates the "yin-yang" synergy of a balanced network.

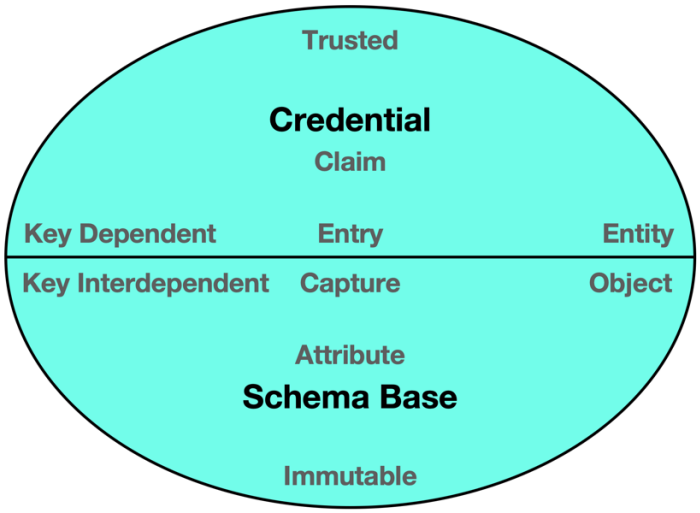

Figure 1. A component diagram showing male and female counterparts in a balanced network model.

In figure 1, DDE elements, components and characteristics in the top half of the model fall into the *Credential* space, the *data entry* domain. Everything south of the equator falls into the *Schema* space, the *data capture* domain.

The two fundamental kernel technologies on each side of the model are, in the case of *data entry*, Self-sovereign Identity (SSI), a term used to describe the digital movement that recognises that an individual should own and control their identity without the intervening administrative authorities and, in the case of *data capture*, Overlays Capture Architecture (OCA), a flexible architecture which represents a schema as a multi-dimensional object consisting of a stable schema base and linked task-oriented overlays. The key feature of SSI is that it enables private encrypted pairwise channel communication without centralised authority dependency or third-party interference. The key feature of OCA is that it facilitates

a unified data language so that harmonised data can be pooled into multi-source data lakes for improved data science, statistics, analytics and other meaningful services.

The root identifier types used by these two core technologies require different characteristics in each case. In the case of SSI, the pairwise endpoints are identified by *entity identifiers*, a type of identifier that is governed by an entity who controls the signing key. In the case of OCA, the data capture objects are identified by *object identifiers*, a type of identifier that contains a hash of the content of an object. The following hash grid table describing the different identifier states.

State	Key Dependent	Key Interdependent
Trusted	Entity identifiers	Object identifiers that include reference to an entity identifier
Immutable	Entity identifiers that include reference to an object identifier *(This state cannot exist)*	Object identifiers *(No entity identifier referenced)*

Figure 2. A hash grid table describing the different states of entity and object identifiers.

To help dissect the characteristics of the identifier hash grid table, the meanings of the associated characteristics need to be defined.

Key Dependent vs *Key Interdependent*

Key Dependent
The identifier is governed by an entity and therefore a signing key is required.

Key Interdependent
The identifier can either be governed by an entity (whereby a signing key is required) or not governed (no keys required).

Trusted vs *Immutable*

Trusted
The identifier is governed by an entity and therefore a signing key is required to establish trust.

Immutable
The identifier contains a hash of the content of an object which cannot be changed. If an object identifier is governed, the controller of the signing key has control over the content contained within an associated identity document and, as such, it can no longer be deemed immutable.

To scale this back to the *Credential* and *Schema* domains in the network component diagram shown in figure 1, this essentially means that all elements within the male side of the model should be signed by a private/public key pair as proof of governance to establish trust. All elements within the female side of the model, on the other hand, should contain a hash of

content to establish immutability. This is the crux of where scope creep can be detrimental to the overall mission.

In a harmonious DDE, SSI technology has given rise to the most compelling component of trust, a *Verifiable Credential* (VC) – a digital representation of a physical credential but more tamper-evident and more trustworthy than their physical counterparts. An example of a VC might be a digital representation of a driving license. An interesting characteristic of a VC is that, due to structural governance built into the component, with some additional cryptography, certain information within a VC can be masked to protect personally identifiable information (PII) where necessary. This is where zero-knowledge proof (ZKP) plays a valuable role in human-to-human interaction within the digital realm. For example, say you are celebrating your 21st birthday in the US with a gathering of friends at a local bar. In the physical world, the bartender might ask to see your driving license in order to prove that you are of legal drinking age prior to serving you. The driving license contains your date of birth, a PII element. What cryptographic ZKP enables is the ability to mask that element with proof that you are over the age threshold to be served a drink without having to display your date of birth. That makes VC characteristics compelling for human-to-human communication. All of this functionality resides in the *Credential* space.

However, to enable a safe and secure data sharing economy, the female side of the model takes the lead. A phrase coined by Philippe Page, President of The Human Colossus Foundation, for decentralized data capture components (including all related extensions, coloration and capture functionality) is "*Decentralized Semantics*". In the *Schema* space, to enable multiple parties to interact with a common schema, OCA represents a schema as a multi-dimensional object consisting of a stable "schema base" and linked "overlays". Overlays are task-oriented layered data objects that provide coloration to the schema base object. This degree of object separation enables issuers to make custom edits to the overlays rather than to the schema base itself. With schema base definitions remaining stable and in their purest form, a common immutable base object is maintained throughout the capture process which enables data standardisation. In the case of global adoption of the architecture, OCA facilitates a unified data language so that harmonised data can be pooled into multi-source data lakes for improved data science, statistics, analytics and other meaningful services.

It is in the synergistic combination of the trust characteristics of a VC and the immutable, yet interoperable, components contained in OCA schema that privacy compliant data sharing can be facilitated in a DDE.

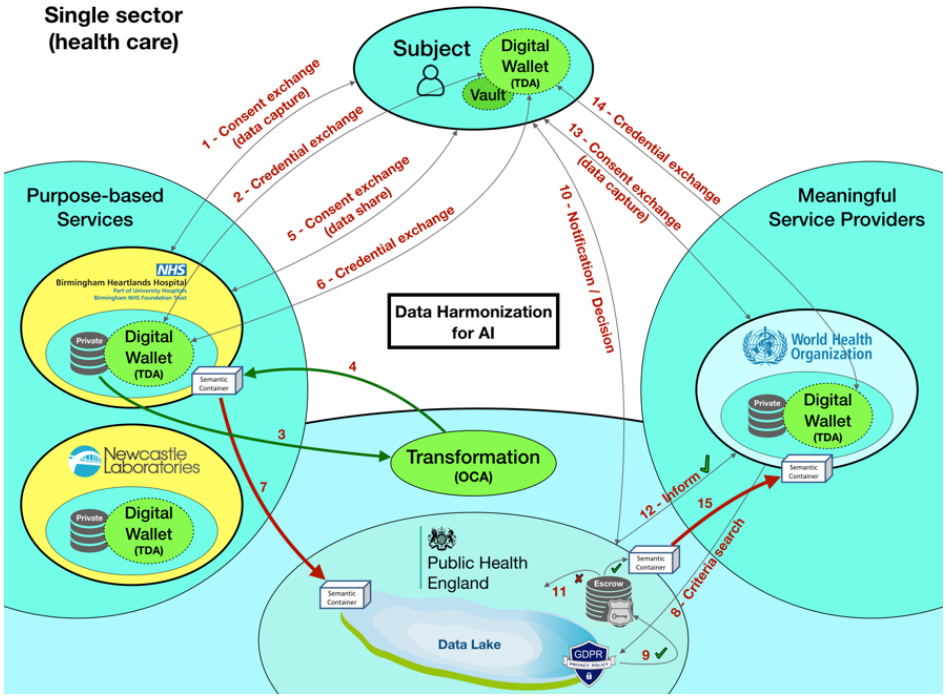

Figure 3a. A high-level representation of privacy compliant data sharing within the health sector.

The model represented in figure 3a is affectionately known as the *Mickey Mouse Model* (MMM). The left "ear" contains *purpose-based services* (PBS) - entities that issue schema for initial data capture into a DDE. The right "ear" contains *meaningful service providers* (MSP) – entities that enrich data by combining criteria searched data from harmonized data lakes. Consent is always obtained by the subject before data portability is actioned. VCs provide the mechanics to verify that the data flow stems from a trusted source. At the heart of MMM is the OCA transformation tool required for data harmonization prior to data porting. In this single sector use case, the processing for consented *data capture* and associated credential exchange is treated as a separate process to the equivalent *data sharing* flows. In the case of an emergency response situation, such as the COVID-19 pandemic, those two processes are combined to speed up urgent data portability to any specialised agencies responsible for national and international public health (see Figure 3b).

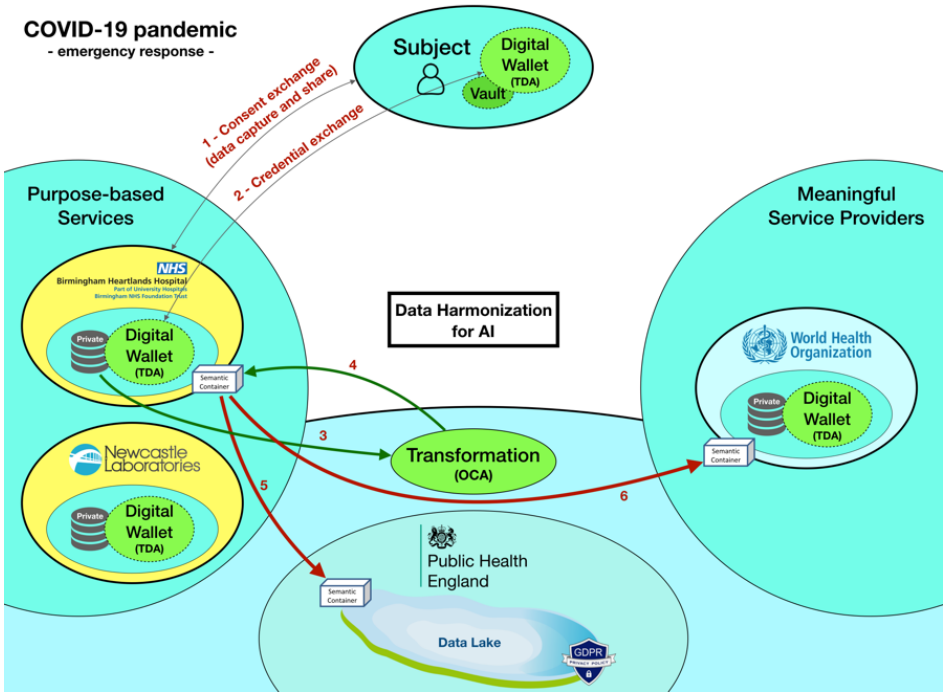

Figure 3b. A high-level representation of privacy compliant data sharing during an emergency response situation.

The level of object interoperability and subsequent language unification offered by OCA will enable more accurate trend analysis to be performed on richer sets of harmonized data leveraging AI to automate decision making and to engage machine learning for ongoing efficiency and effectiveness regardless of industry sector or societal situation. The power of the "*Mickey Mouse Model*" unveiled but don't let scope creep muddy the waters.

Note: A "decentralized" version of the Internet is currently under construction and The Human Colossus Foundation will be introducing the core technology components of the build via a series of blog posts throughout the course of 2020.