OIX Workshop #1 - Standards

Held at 1-4pm on 05/02/2020 in AgFe Offices

**Agenda**

* (Re)introductions              13.00 – 13.05
* Recap actions and notes from Kick Off      13.05 – 13.30
* Semantic Web standards and tools      13.30 – 14.30
* A worked example              14.30 – 15.30
* Open knowledge base (GitHub)      15.30 – 15.45
* Look forward to Workshop #2          15.45 – 16.00

**Attendees**

* In person
  + Ben Helps - Factern
  + David Rennie - Idemia
  + Sid Kalitha - FSA
  + Ross McDonald - Future Borders
  + Wayne Robinson - HMRC
* On the phone
  + Richard Thomson - Idemia
  + Sahy ? - Idemia
  + Stephen Cowx - Future Borders (until 14.30)
* Apologies
  + Chris Allgrove – GDS
  + Nick Davies - HMRC
  + Robbie Fraser - Factern

**Meeting notes**

1. ***Introductions***

Sid

* Innovation Lead at the Food Standards Agency
* Has been leading on the FSA blockchain Proof of Concepts
* Focus has been on two use cases: export of pork from UK to China and Cattle traceability
* Opportunities to use DLT to make process improvements
* The pilots highlighted the importance of data standards
* DLT could provide an opportunity to look beyond the one up and one down on the supply chain by including multiple actors but there are challenges.

1. ***Recap actions and notes from Kick Off***

Ross

* Need to recognise that there will be many ecosystems emerging over time
* **Rule of engagement:** publish your ontology

1. ***Semantic Web standards and tools***

*Slide 8 - “Semantic Web” refers to W3C’s vision of a Web of linked data*

Ross

* Point to point APIs are brittle
  + As producer of a data set, you need to produce a different API for each different configuration or view that accesses it
  + As consumer of a data set, you need to form a query **and** have knowledge of the domain / data set which the query is pointed at
* SPARQL / Semantic Web
  + Advantage: it allows for a hugely wide range of queries on a single data set
  + Advantage: it is easy for event producers to deploy
  + Disadvantage: it is hard to model

David

* How well used is the Semantic Web?
  + Our model needs to be realisable

*Slide 11 - what is a Semantic Model?*

Stephen

* Need to put some thinking into use of DIDs (a recent concept) vs. URIs (foundational concept of the Semantic Web) as means of identifying ‘things’
  + What are the choices and trade-offs?

*Slide 13 - Semantic Web application architecture*

Ross

* Proposal is to use Apache Jena for delivering the application architecture to support the Alpha

1. ***Look forward to Workshop #2 [brought forward to capture input from S. Cowx]***

*Slide 25 - What kind of expertise might be available?*

Sid

* There is an opportunity to engage multiple actors across the supply chain to ensure greater assurance of products

Sid

* Provenance is important and can highlight the underlying value of a product.
* Food suppliers will do DNA samples on products. It would be good to have access to that data
* Port Health Authorities are usually from the UK local authority where the port is located. The Food Standards Agency along with DEFRA are responsible for the overall policy on Products of Animal original and Products of non-animal origin.

Ross

* We are looking to take the **assurance** piece of this and make it machine readable and executable

Stephen

* Suggested set of events that might be relevant
  + The consignment originates from the country which ot states it does
  + This consignment complies with regulatory requirements
  + This consignment contains what is stated on the manifest
  + This freight forwarder/trader/importer/carrier has a history of not being stopped
  + This freight forwarder/trader/importer/carrier has a significant number of imports which have been verifiably "good"
  + Has this freight forwarder/trader/importer/carrier been seen before
  + Is this freight forwarder/trader/importer/carrierbehaving in a way which is unusual
  + Is this freight forwarder/trader/importer/carrier linked to other traders or trades which have resulted in a successful intervention

Ben

* Stephen’s examples are more focused on the actor than the event

 Wayne

* Adding “weight of experience”
* Assessments and data can change over time

 Ross

* Longitudinal view – **time** is a key dimension to capture
  + Entity
  + Attribute
  + Value
  + **Time**
* Attestations themselves can be **time bound**

***Conclusions***

* Three parts to the metadata model
  + Event typologies, e.g.
    - Scalar values
    - Delta values
    - Attestations
    - Units of computation
  + Schema of required data for each, e.g.
    - Time stamp
  + Example vocabulary
* Dimensions of challenge to test whether Semantic Web is suitable
  + Security: this is a different layer, but ask for NCSC view on security concerns given
    - Anonymity of event data layer
    - Corroboration between independent sources difficult for bad actors to fake
    - Not creating a central data lake
  + Skills gap / difficulty of modelling: are we creating a soft barrier to entry?
    - “Trusted Service Provider” to support less capable entities
  + Ability to scale: tertiary concern? Implementations will adapt…
  + Too academic / not used?
  + How to fund development of the ecosystem?
    - This leads to need for ‘value exchange’ mechanism to reflect ‘work done’ – part of **rules of engagement**
* Rules of engagement
  + Must be recommended in + opt in to join
  + Sign up to rules of engagement that include
    - Production of instance data (held on a distributed basis by participants)
    - Publication of ontologies (accessible to ecosystem)
    - Registry of actors (on DLT)
    - Reasoning engines (bespoke and federated)
    - Value exchange (see above)
  + Consent / revocation / tracking is key
    - Crux of ability to scale legal frameworks that are so problematic today
    - Just give consent to expand access to other data points / tokens of value
    - Government may have a role to police this

***Agreements***

* To use GitHub to host collateral and artifacts
  + Meeting notes need to be vetted first to ensure not publishing proprietary / strategic IP