2022-11-01 DMRWG Meeting Notes

Meeting Date

01 Nov 2022 The DMRWG meets bi-weekly on Tuesdays at 12:00-13:00 PT / 16:00-17:00 UTC. Check the ToIP Calendar for meeting dates.

Zoom Meeting Link / Recording

- Recording

Attendees

- Neil Thomson
- Steven Milstein
- sankarshan
- Kevin Griffin
- Carly Huitema
- James Schoening (IEEE Ontology Standards)
- Ken Woodruff
- Christoph Fabianek
- David Hoffman
- Burak Serdar
- Zaïda Rivai

Main Goal of this Meeting

Introduction of the Data Modeling & Representation Working Group

Presentation/Discussion - **Authentic Data - Simple in Principle, ...**

Agenda Items and Notes (including all relevant links)

Time	Agenda Item	Lead	Notes
5 min	- Start recording - Welcome & antitrust notice - Introduction of new members - Agenda review	Chairs	**Antitrust Policy Notice:** *Attendees are reminded to adhere to the meeting agenda and not participate in activities prohibited under antitrust and competition laws. Only members of ToIP who have signed the necessary agreements are permitted to participate in this activity beyond an observer role.* - Meeting Slides - New Members:
5 mins	Review of action items from previous meeting	Chairs	

5 mins	Announcements	TFLeads	News or events of interest to Governance Stack WG members:
			
			• Future Topics
			  ○ Determine generalized data support in the ToIP dual stack including data structures and exchange mechanisms at the application level
			    ■ Point - VCs contain data for a specific purpose (claims & attestation to support entity/entity trust). They
			  ○ Requirements on Authentic Data:
			    ■ Use of identifiers for Data and Authorities used to sign the data. Is this KERI or something simpler
			    ■ A general model for packaging data, its identifier(s), etc. Model for
			    ■ The interplay between Data and Governance management and authority (who signs?)
			  ○ Requirements on "root of trust" on general data. How is this different from VCs, and Identifiers?
			    ■ Does Verifiable Credential data provenance go deep enough?
			  ○ Detailing the (Authentic) Data Lifecycle and how it is different from current practice
			    ■ Detailing iterative development of a dataset (infrequently a "linear" process) (*presentation?*)
			  ○ Transformation & translation - mapping data from one schema to another. Requirements for an SSI Trust/Authentic model. <u>Area of immediate concern - VC interoperability</u>
			    ■ Role of Ontologies (*presentation*)
			    ■ Dealing with evolving & competing standards
			    ■ Experience with layered/OCA schemas (*presentation*)
			  ○ Authentic requirements on data advances happening in parallel with SSI/DID technology
			    ■ Data structures (relational, graph, RDF and newer variations, ...)
			    ■ Communication/exchange mechanisms (publish/subscribe, containers, data pipes (e.g., Kafka))
			    ■ Data agreements, Consent

20 mins	Authentic Data	Chairs	**Authentic Data** - A Published Dataset - available for use/consumption by 3rd parties is built on data and data governance used to design and build a re-usable dataset from "first principles"
			
			Presentation - Authentic Data - Simple in Principle - Nov 1 DRMWG
			
			  Summary
			
			  ■ Authentic Data is data that has been crypto-signed by an "authority" (role) using their private key for which users of the data can verify using the "authority"s public key.
			  ■ Creating publishable/sharable data is via a Data Lifecycle where the data is initially captured/collected, then checked for input and consistency errors, cleaned of outliers, duplicates and checked for overall correctness. Each of those stages needs to be persisted and linked to the dataset that is published for 3rd party use that are part of the data provenance (trust) chain
			  ■ The Data needs to be designed with respect to structure, metadata, and "fitness for purpose". Governance needs to be designed to ensure accuracy, consistency and correctness
			  ■ Governance drives requirements for error, consistency and accuracy as an active part of the data lifecycle. Data Governance is the strategy, Data Stewardship is oversight the data lifecycle
			  ■ Discussion at:
			    ■ Kevin: Need a definition of "authentic"
			      ■ Authentic: having an origin supported by unquestionable evidence; authenticated; verified: (in other words, securely attributable to a cryptonymous identifier)
			      ■ Neil: Working with Henk van Cann on shared terminology with ACDC (Authentic Chained Data Container) Task Force
			    ■ Burak, Neil, James: Data Transformation (source to target) is going to be core. The ideal solution would be a common model
			      ■ Neil: Added to future topics
			      ■ James: Ontologies are heading to a hub model of data
			      ■ James offered to present his perspective (presentation)
			      ■ Burak offered to present his perspective (presentation)
			    ■ Carly: while you the data lifecycle is laid out as a logical, linear set of stages/steps, our researchers see this as (very) convoluted process of data collection, cleanup, combination and re-combination with other datasets (presentation?)
			      ■ The collection, transformation, and recombining of datasets is claimed to be a feature of Ocean Protocol - where they track attribution throughout (at least that was my understanding).
			      ■ Ocean Protocol links:
			        • Ocean Protocol site
			        • Ocean Protocol Blog
			        • What is the Ocean Protocol
			    ■ Kevin: Root of trust is the authentic part of ACDC for more details ACDC (proposed IETF) specification
			    ■ Burak, Kevin: Data/metadata (incl. semantics) are independent of containers or exchange "channels". Need to consider Identity /Identifiers for Data vs. the SSI Entities (e.g., Issuer, Verifier, Holder)
			      ■ Neil: (post-meeting) this suggests that the ToIP tech/gov dual stack is two (dual) stacks
			        ■ Identity tech/gov
			        ■ Data tech/gov

5 mins	• Review decisions /action items • Planning for next meeting	Chairs	As the next meeting falls on the opening data of IIW (15 Nov), it is suggested to postpone this meeting until the next meeting date - Nov 29. Proposal - to have one or two presentation on data transformation and discussion on Nov 29

Screenshots/Diagrams (numbered for reference in notes above)

#1

Decisions

- Sample Decision Item

Action Items

- [ ] Sample Action Item