Decentralized Semantics Working Group

- Mission and Scope (as defined in the WG charter)
- Conveners
- Chairs
- Mailing List
- Meetings
- Description
- Core components
  - Overlays Capture Architecture (OCA)
  - Blinding Identity Taxonomy (BIT)
- Deliverables
- Shared documents for member contribution

Data semantics is the study of the meaning and use of specific pieces of data in computer programming and other areas that employ data. When studying a language, semantics refers to what individual words mean and what they mean when put together to form phrases or sentences. In data semantics, the focus is on how a data object represents a concept or object in the real world. The term “Decentralized Semantics” is a representation of “data semantics” within an age of distributed ledger technology (DLT) solutions which continue to mould the current decentralization movement.

Mission and Scope (as defined in the WG charter)

The mission of the WG is to define a data capture architecture consisting of immutable schema bases and interoperable overlays for Internet-scale deployment. The scope of the WG is to define specifications and best practices that bring cohesion to data capture processes and other Semantic Standards throughout the ToIP stack, whether these standards are hosted at the Linux Foundation or external to it. Other WG activities will include creating template Requests for Proposal (RFPs), and additional guidance to utility and service providers regarding implementations in this domain. This WG may also organise Task Forces to escalate the development of certain components if deemed appropriate by the majority of the WG members and in line with the overall mission of the ToIP Foundation.

Conveners

- Paul Knowles
- Robert Mitwicki

Chairs

To be elected

Mailing List

decentralized-semantics-wg@lists.trustoverip.org (subscribe by going to lists.trustoverip.org)

Meetings

This WG currently meets weekly on Tuesdays. See the Meeting Page for the meeting schedule, agenda, and meeting notes. For a calendar invite with complete Zoom information, please send email to the mailing list above.

Description

The post millennial generation has witnessed an explosion of captured data points which has sparked profound possibilities in both Artificial Intelligence (AI) and Internet of Things (IoT) solutions. This has spawned the collective realization that society’s current technological infrastructure is simply not equipped to fully support de-identification or to entice corporations to break down internal data silos, streamline data harmonization processes and ultimately resolve worldwide data duplication and storage resource issues. Developing and deploying the right data capture architecture will improve the quality of externally pooled data for future AI and IoT solutions.

Core components

**Overlays Capture Architecture (OCA)**

(Presentation and live demo / Tools tutorial / CSV parsing tutorial)

OCA is an architecture that presents a schema as a multi-dimensional object consisting of a stable schema base and interoperable overlays. Overlays are task-oriented linked data objects that provide additional extensions, coloration, and functionality to the schema base. This degree of object separation enables issuers to make custom edits to the overlays rather than to the schema base itself. In other words, multiple parties can interact with and contribute to the schema structure without having to change the schema base definition. 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 standardization.
OCA facilitates a unified data language so that harmonized data can be pooled for improved data science, statistics, analytics and other meaningful services.

- OCA Editor - https://editor.oca.argo.colossi.network
- OCA Repository - https://repository.oca.argo.colossi.network
- OCA source code - https://github.com/thclab

**Blinding Identity Taxonomy (BIT)**

The Blinding Identity Taxonomy (BIT) is a defensive tool created for the purpose of reducing the risk of identifying data subjects within blinded datasets. BIT contains a list of elements to be referred to by schema issuers and data controllers for flagging attributes which may contain identifying information about governing entities. Once attributes have been flagged, any marked data can be removed or encrypted during the data lifecycle.

- BIT in HTML format: https://docs.kantarainitiative.org/Blinding-Identity-Taxonomy-Report-Version-1.0.html

**Deliverables**

- Technical specifications for all core components required by "Decentralized Semantics" as defined by the Mission and Scope statement above.
- Also check out the ToIP Deliverables document for high-level deliverables of the Trust over IP Foundation.

**Shared documents for member contribution**

- Functionality requirements and desired capabilities that a fully interoperable data capture architecture should offer
- Frequently asked questions