

IEEE P1752 Minimum Metadata subgroup
Minutes of conference call held on February 4, 2020 at 9 am Pacific Daylight Time

Conference call started at 9:00 am Pacific Daylight Time on zoom

Slide deck presented by Chair Dr. Ida Sim (available on WG website)
<https://site.ieee.org/sagroups-1752/files/>

Attendance:

First Name	Last Name	Affiliation
Simona	Carini	UCSF / OmH
Anandatirtha	Nandugudi	University of Memphis
Ida	Sim	UCSF / OmH
Paul	Harris	University of Melbourne
Anna	T	Consultant.
Paul	Petronelli	PALM Associates
Christina	Caraballo	Audacious Inquiry
Shivayogi	Hiremath	Temple University
Brian	Page	Ribbon Communications

Agenda:

- Last call's action items
- Header.json
- Schema-id.json
- Datapoint ID
 - Are datapoints Private or public?
 - General consensus public
 - Device/process should be able to generate the UUID
 - Can a device be identifiable with UUID
 - Metadata has the datasheet information of the device.
 - Issues/usecases relating to datapoints being public
 - Correlation of IDs for data analysis
 - Correlation of IDs to reveal PII
 - Back to the basic question
 - Do we need datapoints?
 - Is it needed for data analysis
 - Important to go back in the processing pipeline to identify the source.
 - Yes
 - Having datapoints increases the storage requirements.

- Lossless compression
 - Decision made on having Datapoints ids
- Are datapoints required or optional
 - Decision made to make datapoints – required.
- Generation of UUIDs
 - Which scheme to recommend
 - Are there any royalties involved with different schemes?
 - Restrict the options to generate UUIDs to keep it consistent.
 - Schema cannot generate UUIDS, can only validate the format of the ID.
 - We may need another field in the metadata to identify which standard/scheme was used to generate UUID.
 - This cannot be mandated as long as the UUID complies with the RFC?
 - Decision made to use RFC 4122 compliant UUID
 - Separate field to identify the method used in RFC 4122
 - Hashing (used in v5 of RFC 4122) makes it difficult to identify PII
 - Privacy issue
 - UUID/ MAC address could be reverse generated backwards as the metadata has the time stamp even if the MAC address is hashed.
 - It's better to use best practices upstream for storage and transportation
 - Encryption
 - The principle document should contain recommendations to use to best practices to protect PII
 - Datapoint assigned to each instance in a datapoint series?
 - Depends on how it is used for analysis (Action items, naming issues)
 - Discussion postponed to the next meeting due to lack of time.
- ULID
 - Higher probability of clashes.
 - Decision made to not use ULID scheme.

Comments in the chat:

None

Action items:

- Ida, Brian: General principles for generating Datapoint ID
 - Datapoint series
- Simona, Paul: update header schema

Conference call adjourned 10:00 am Pacific time
Next call: Tuesday Feb 18 at 9:00 am Pacific Time

Minutes taken by Subgroup Secretary Anand Nandugudi, University of Memphis