

IEEE P1752 Minimum Metadata subgroup  
Minutes of conference call held on October 15, 2019 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)

Attendance:

First Name	Last Name	Affiliation
Simona	Carini	UCSF / OmH
Anandatirtha	Nandugudi	University of Memphis
Paul	Petronelli	PALM Associates, Inc.
Ida	Sim	UCSF / OmH
Sean	McConnell	AMA
Jakob	Bardram	Technical University of Denmark
Shivayogi	Hiremath	Temple University
Paul	Harris	University of Melbourne
Anna	T	Consultant

Agenda:

- Last call's action items
- CARP overview
  - Use case for upload time?
    - Yet to find a specific use case for upload time
- Datapoint life
  - Not much use for a developer
  - Conclusion - Not really required
- Datapoint ID
  - Globally unique id
  - RFC 4122
  - Action item – make a specific proposal for defining Datapoint ID
- Data absence
  - Having error data points
  - Sampling rate vs effective time frame
  - State the provenance of the data
  - Offset are complicated, may not be necessary
    - Could lead to loss of precision

- Should every element in the series have its own ID?
  - Action item

Comments in the chat:

RFC4122

<https://tools.ietf.org/html/rfc4122>

"id":"59715530-c95a-11e9-ff41-e5771fd30900"

Example <https://developer.gnome.org/glib/stable/glib-GUuid.html>

More details on UUID hexadecimal format (including how one might generate a UUID that is compliant) <https://www.cryptosys.net/pki/uuid-rfc4122.html> and wiki has a good discussion on 'collisions' or non-unique identifiers

[https://en.wikipedia.org/wiki/Universally\\_unique\\_identifier](https://en.wikipedia.org/wiki/Universally_unique_identifier)

Meaning, using this UUID approach we would need to tolerate such error rate as 1 in a billion non-unique ids. So to defeat this 'bug', perhaps some error-handling would be required for larger datasets, such as searching to see whether UUID has already been used before assigning

My point is – if we're to talk about "sampling" we should use another example.

"sampling" is related to sensors and the stress example is not a sensor location is better as example

Another issue – should each stress reading has it own data point id?

We do this

Action items:

- Sean McConnel, Jakob Bardram: Investigate further into DatapointID wrt datapoint series
- Anand Nandugudi/ Ida: Data Absence, refine based on today's discussion

Conference call adjourned 10:00 am Pacific time

Next call: Tuesday November 5 at 9:00 am Pacific Time

Minutes taken by Subgroup Secretary Anand Nandugudi, University of Memphis