IEEE P1752 Minimum Metadata subgroup
Minutes of conference call held on January 14, 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)

Attendance:

<table>
<thead>
<tr>
<th>First Name</th>
<th>Last Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simona</td>
<td>Carini</td>
<td>UCSF / OmH</td>
</tr>
<tr>
<td>Anandatirtha</td>
<td>Nandugudi</td>
<td>University of Memphis</td>
</tr>
<tr>
<td>Ida</td>
<td>Sim</td>
<td>UCSF / OmH</td>
</tr>
<tr>
<td>Paul</td>
<td>Harris</td>
<td>University of Melbourne</td>
</tr>
<tr>
<td>Anna</td>
<td>T</td>
<td>Consultant.</td>
</tr>
<tr>
<td>Paul</td>
<td>Petronelli</td>
<td>PALM Associates</td>
</tr>
<tr>
<td>Christina</td>
<td>Caraballo</td>
<td>Audacious Inquiry</td>
</tr>
<tr>
<td>Shivayogi</td>
<td>Hiremath</td>
<td>Temple University</td>
</tr>
</tbody>
</table>

Agenda:

- Last call’s action items
  - Still working with AMA for the BP example
- Header.json
  - Write about source creation date time in the schema principles document
  - Difference between recorded time and reported time – is it still useful?
  - Suggestion - Use first created at the initial source
  - Do we need to make a distinction between sample data and processed data
    - Effective time frame captures the time at which the data point is valid
    - Source creation time reflects the time the data point was generated.
  - Remove the word ‘first’ from the description.
  - Acquisition rate needs description
- Schema-id.json
- References to datasheets
  - Multiple datasheet references could be in an array
  - Arrays makes it easy to extend to many datasheets
  - Personal datasheets could also be multiple
• URI vs IRI
  o Restrict to only references
  o Unrestricted strings could have PII
  o references should be resolvable
    ▪ Should be able to find it
  o Unique, identifiable and resolvable
  o Decision made to not support unrestricted strings.
  o Use IRI as the preferred reference as IRI is safer since IRI is a superset of URI and is less restrictive

• Other comments
  o Need more real-life sample data to stress test the schema
  o Use cases
    o Metadata after metadata aggregation

• Outstanding issues
  o UUID
  o sample metadata
  o Filtering
    ▪ Just set to Yes/NO or does it need additional information
    ▪ Algorithm information could be captured in data sheet
    ▪ Is it needed?
    ▪ Decision reached to not add filtering flag in the header
  o Flag identifying raw data
    ▪ Boolean flag
    ▪ Use principles from Digital medical association
    ▪ Is it needed?
      ▪ Could it capture some notion of ground truth.
      ▪ Decision reached to not add raw flag in the header

 Comments in the chat:
 None

Action items:
  • Ida: General principles for generating Datapoint ID
    o Consult with Brian
  • Simona Carini, Paul Harris: update header.json
  • Use cases

Conference call adjourned 10:00 am Pacific time
Next call: Tuesday Feb 4 at 9:00 am Pacific Time

Minutes taken by Subgroup Secretary Anand Nandugudi, University of Memphis