IEEE P1752 Minimum Metadata subgroup
Minutes of conference call held on July 9, 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)
http://sites.ieee.org/sagroups-1752/files/2019/07/2019-07-09-1752-Metadata-slides.pdf

Attendance:

First Name	Last Name	Affiliation
Simona	Carini	UCSF / OmH
Shivayogi	Hiremath	Temple University
Sean	McConnell	American Medical Association
Henry	Ogoe	DoD/VA Interagency Program Office
Paul	Petronelli	PALM Associates, Inc.
Ida	Sim	UCSF / OmH
Anna	Tee	Consultant
Jakob	Bardram	Technical University of Denmark
Pradeep	Balachandran	Consultant, e-health domain

Review of action items from previous call and of subgroup members' contribution via email

By changing the namespace (e.g., FHIR) we can add support for other standards, so metadata work can apply also outside Open mHealth work

Information relevant to the study should not be stored in each datapoint
Need pointers

Discussion question: see slide 10 of http://sites.ieee.org/sagroups-1752/files/2019/07/2019-07-09-1752-Metadata-slides.pdf

Open mHealth does not impose constraints on how data is represented at the sensor level; the header comes into play when data is shared so answer to question above is no

It may be good to calculate how much is the overhead

Guiding pronciple is to allow people to say what they want to say and not overcommit →
hence, keep metadata at a minimum
E.g., if sensor location is an important property, it may be collected in the datapoint

We need a definition of datapoint…
- sensor-acquired info on a single person
- instance of a schema

and of datastream = sequence of datapoints with the same metadata

It is possible to extend model to support datastream
Either you define a new schema or in the metadata you state this is a stream

If datastream is an ordered sequence of datapoints, then a datapoint can be defined as a stream with one measure

Open mHealth was not developed with IoT in mind, so need to work on scope

List of proposals:
- use sequence number for datapoint that is not tied to a timestamp: this can be in the data (rather than header)
- call datastream "datapoint series" (to avoid confusion with streaming of data)
- attach DOI to a file containing constant information and datapoint has pointer to it
- sampling rate in the header + timestamp in the datapoint

Sampling rate vs. sequence number:
Sequence number is representative of sampling time period
However, the actual timestamps reveal when data was collected
If the sampling rate is uniform, maybe the timestamp is not needed for every data point in the datastream (datapoint series), and if a missing data sample is not a concern the overhead from having a time stamp for every data point can be avoided
Sampling rate would be part of the overall description of a study

Action items:
- Jakob Bardram to present on August 27
- presentation of examples as agreed upon previously
- put reworked slides on iMeet

Conference call adjourned
Next call: Tuesday August 27 at 9 am Pacific Time

Minutes taken by WG Secretary Simona Carini, UCSF
(Subgroup Secretary Anand Nandugudi, University of Memphis, is OOO today)