

2021-04-06 Standard Data Models and Elements Drafting Group Meeting Notes

Attendees

- Brian Plew
- David
- John Walker
- Marie Wallace
- Paul Knowles
- Paul Murdock
- Rajesh Pillai
- Ramesh Raskar
- Sandeep Jain
- Steven Milstein
- Tony Little
- Vitor Pamplona

Agenda Items

Time	Item	Who
2 min	Welcome & Antitrust Policy Notice	Rebecca Distler & Brian Plew
5 min	Introducing Paul Knowles (new co-chair)	Paul Knowles
5 min	Data Schema	Brian Plew
10 min	Methods of Evaluating Data Models	Brian Plew
5 min	30 / 90 / 180 Day Framework	Brian Plew
30 min	CCI Data Schema Overview	Paul Knowles

Presentations

Overview of CCI Schema (Paul Knowles) - [Google Drive Folders](#)

Recording

Topic: Good Health Pass - Standard Data Models and Elements
Start Time : Apr 6, 2021 12:00 PM

Meeting Recording:

https://zoom.us/rec/share/B3G4euPq1C2-biS96uQTKVx_xEgVUVJmMMIBkergu3UL_FPxGGe2TNjNsuNqXt4m.TaFI_u2v-3mC5nPR

Notes

1. Welcome and Linux Foundation antitrust policy

2. Data Schema

- Q: Do we really need to go beyond the minimum? What would we do if we were trying to describe more than minimum? What would it be used for? Do we invest extra effort into it?
 - Huge amounts of data captured by legacy systems (e.g., physical address); this information isn't necessarily required as long as you have certain active identifiers you can authenticate
 - Cross-reference with CDC to have some mapping; EU has largely disregarded legacy attributes
 - With eHealth group, come up with 3 specs for minimum viable capture (testing, vax, recovery)
 - Same data capture used across jurisdictions
- Q: Do optional attributes become the "superset"?
 - Optional = not imposed on all jurisdictions
 - E.g., CVX code used in North America, but not required in Europe (done with market authorization holder)

3. Methods of Evaluating Data Models

- Q: If it's a small number of jurisdictions, does it become optional? Or is it based on the element itself?
 - If it's a totally new attribute for a jurisdiction, put as optional in superset
- Q: Are we doing a data dictionary (schema, layout)?
 - We can't define schemas used by each country / dictate what schemas are used (use any ontology or data model)

- But when it comes into capture space, can rebuild semantics so it can deal with multiple languages and decide on data model to use on data exchange side
- Immediate reality is that people won't agree on a single schema - have to identify data dictionary - compromise position to interoperate between schemas
- Q: Which data exchange models?
 - Discussion on using FHIR exchange (use data capture structure to come up with all the human readable labels)
 - Discussion on using JSON-LD (FHIR might be too healthcare focused)
 - FHIR not most consumable structure for end users to understand data - and majority of consumers are not going to be healthcare entities
 - Serialization shouldn't matter at architecture level; need to fit in payload off of a W3C cred
- Discussion on differences between data management; data collection and data exchange; collecting data from multiple sources (rebuild so you can exchange it for machine readability - can define an architecture without defining attribute names)
- Should explore a few options and weigh pros and cons against them

4. 30 / 90 / 180 Day Framework

- Objective: Large orgs can make announcement at 30 day point on "intent to implement;" smaller orgs may be able to move quicker
 - Provide guidance to enable development of credentials and passes by 90 days
 - Implementation / rollout = credentials and passes first; explore end-to-end interoperability with full health records later on
 - Still helpful to have an understanding of the "health record" first (e.g., superset) and then work backwards into credentials and passes
- Delineate difference between lab results vs. medical interpretation vs. passes

5. CCI Data Schema Overview

- EU documents
 - EU is first document defining recovery minimum dataset - super helpful because we can use this as a template to start with, and compare this with CDC, and others
 - EU also has vaccination and testing documents as well
 - Lists things like COVID tests authorized across all of Europe (capturing all of this in a global spec), including some more granularity into testing devices
 - Goes into minimum data capture for testing process
 - Cross referenced with CDC (vax reporting); for example, has not included address in spec
- Global reference docs in folders
 - Include formats / labels / some pre-defined entries for fields (might change for jurisdictions); some text to help users understand fields
 - Note unique numeric ID (e.g., iRespond use case); there may be a few use cases using biometrics. Rather than leave it out, suggests that we are aware that there is space for biometric ID if we need it
- Recovery = 3 fields (cryptographically link recovery to test record; link to positive test sample)
 - European, US, Canada specs will reference codes (e.g., ICB10)
- Can push into an OCA form; generates JSON-LD form
- Questions & discussion
 - Could use as a stand-in for SNOMED; difficult to tell countries what do about encoding data (or common set of codes)
 - International health standards org competes with WHO standard (?)
 - Different standards from FDA
 - Medical codes vs. ontology
 - A lot of sensitive data should not be held in the credential; should be held in a token (a report kept separately from a credential and certificate)
 - Explore use of 10 attributes to use as the credential; other sensitive information hidden
 - Specs - linking identifiers to capture the testing record, but also using a hash (e.g., cryptographic link between testing certificate and testing record)
 - Will need to differentiate COVID-19 from SARS-CoV-2
 - Challenges with existing ontologies is that they're not really human readable; need a short layman readable set of labels (is this going to be added?)
 - Unified medical language system allows mapping specific codes to consumer-centric terms (consumer health thesaurus)
- Yellow Fever solution example (Paul Knowles)
 - All the pre-defined entries change based on language selected; when you want to store this info to aggregate it - might want to store it in one language but capture it in another
 - One schema, and everything else is human readable language overlays
 - Having stable schema base important
 - Multiple layers of certificates, with translations just one of them
 - Are we going to be working with this?
 - Data governance?
 - Translations political - can't use the word pass, want to call it something different
 - Introduces questions around trustworthiness of translation - translating labels is one thing, but translating the results, can you trust what's been translated? Could risk immutability of the underlying credential?
- Need to think about what model to use on the data exchange side
- Note that with all datasets, high degree of convergence

6. Wrap up

Action Items

1. Need to discuss how to operationalize CCI Schema mapping into GHPC recommendations (including use of template + annex)
2. Further discussion on FHIR vs. JSON-LD data exchange model
3. Move working group meetings to Mondays