Q&A #3: RFA #2020-036 OpenHIE COVID-19 Standards & Data Exchange

Carl Leitner: Hi everybody. Thanks for joining the Q&A on the RFA for the OpenHIE COVID-19 Standards and Data Exchange Round 3. This is an opportunity for you to ask questions to help clarify what we're interested in this RFA, as well as ask any questions about some of the programmatic objectives, etc. We want to start with a short presentation and overview of the work to date and where we're going and then turn it over to Q&A. To start, this RFA is centered around the outputs of the OpenHIE COVID-19 Task Force, which is chaired by Digital Square, Regenstrief Institute, and WHO. The mandate of the COVID-19 Task Force is to identify some of the key gaps and standards for data exchange as part of the COVID-19 solution. We have weekly calls, a wiki, and a number of threads on our community forum for discourse. We invite you to join that conversation as you consider this RFA and go forward. Some of the use cases that we're looking at specifically are around case reporting. So we have been developing a FHIR implementation guide that takes the who COVID-19 case reporting form and maps that onto the FHIR data model as both the clinical aspects, the observation, the encounters, as well as a questionnaire. And this Implementation Guide or the data model has also been mapped to a number of popular open source digital health tools. There will be plans to support that for a number of the case reporting solutions and contact tracing solutions such as DHIS2 tracker and SORMAS and Go.Data that are being employed as from the COVID-19 response. Related to that work is the weekly WHO Situational Report which has aggregate data and there's currently a work item in the integrating the help enterprise working group on specifying exactly how to submit that aggregate caseload data using ADX as well as the FHIR equivalent of mADX. That's work that is currently getting underway as well. Future work include potentially looking at screening and triage workflows and adding those into the FHIR Implementation Guide based on the synthesis of the screening and triage workflows from a number of countries. As well as looking at the child case report form for the multi-system Inflammatory Syndrome. For reference, you can find the WHO case reports forms here, the implementation guide on GitHub, and within the previous rounds of funding, there are two projects that were selected for funding. One is for Jembi and Intellisoft linking OpenHIE and DHIS2 components with the case reporting form and you can see some of their presentations on that work. The second is IntraHealth that's linking mHero which is connects Rapid Pro and other SMS relay and IVR apps to an OpenHIE infrastructure and health worker registries to support data collection and alerting as part of the COVID-19 response. As you're considering responding to this RFA, please take into note some of the other work that is going on and if you are interested in collaborating, we're happy to provide points of contact to the to these organizations. What we're looking for specifically in the RFA is a global good that is actively engaged in the COVID-19 pandemic response. Identify some of the key use cases such as case reporting, contact tracing, coordinating care management for which that global good is being deployed or is planning to be deployed as part of that the COVID-19 pandemic response. Please also provide information on the current deployment status, we're looking specifically to understand that this could be used in multiple countries or is being used in multiple countries ideally at a national scale. Next is to describe the exact data exchange and interoperability challenges the global goods are facing. So this is not funding to build a just a generic COVID-19 app for a specific use case but really, we do want to focus in on the standards and interoperability challenges as patients move through the health system and all aspects of their encounters, whether it's the contact tracing, triage, medical care. We could also consider support in areas that are less about the clinical care but related to, for example, the supply chain or vaccines that we hope to have soon. Please provide a clear description of the intended scope and development approach and how you will engage with the Task Force and contribute back to the community. The timeline for this third round of applications: today is our Q&A teleconference so thank you for joining. The letter of intent to apply is due October 2nd. That's how we know that you are intending to apply and we can direct communications to you through that letter of intent. And then the applications are due on

October 9th and we aim to notify you of decisions by October 16th. So with that I'll turn it over to on any questions that you might have. Thank you.

Q: You mentioned one of the responses should be the identification of a global good that is actively engage. Is it possible that you could just extrapolate and expand on that please?

A: Yeah so, what we are looking at doing is leveraging as much as possible existing digital health infrastructure, so less about developing a new app to be deployed but reviewing the digital health tools and infrastructure that are already being used that could be a part of the COVID-19 response.

Q: That tells me that you already have some candidates in mind. Can you expand on that?

A: To be honest, we don't have candidates in mind. I think there are a number of natural ones to consider just based on their deployment, but there isn't a specific tool that we're targeting. We do have some sense of what's being deployed on the ground, but there's a lot that we don't know. There are things like DHIS2 and CommCare and SORMAS that have high levels of footprint that can be natural things to consider. It could be looking at some aspects of the OpenHIE architecture that can be built out to support this since that's less on the front end tool or for the frontline health worker, but more in the back end supply chain systems such as OpenLMIS or others could also play a role. So sorry I am not being very directive here but we do want to pull in ideas that we haven't necessarily seen before and want to give you the opportunity to do that.

Q: You've got the WHO form that has got all the pertinent questions on it, you need to capture this data, aggregated it, and somehow get it back up to the WHO. Do you see the capture of this form as separate from the health system that is already in the institution/hospital or do you see it as integrated into the admission process?

A: I think it depends on the deployment scenario, so yes to both perhaps. For example, on the former, I know a number of solutions were developed just to focus in port of entry so this is not at the EMR level to capture some of the data, maybe not all of it that's represented in that case reporting form. We could also see if there's a screening process that's happening at the facility, perhaps looking at something like OpenMRS that's deployed at many health facilities, can we generate case reports from there and sort of what that that flow is. Both types of scenarios are something that we would consider.

Q: I would also imagine you try and make reuse of data as much as possible; the data collected should be repurposed which makes sense to me because if a patient is presenting for the first time with severe symptoms and you capture all the details into your admission process, you don't want to immediately afterwards go and recapture everything for the for the WHO form so I would imagine you use something like FHIR to plug on to their system or their data source and siphon the relevant data off and populate a database or structure and in FHIR format that contains just the information that you're looking for.

A: Yeah, I agree that that would be a good way to do it. We don't want to burden health workers with double data entry.

Q: You say that case reporting across countries is very unspecified so I would imagine a country on one continent has got a very different hospital system to someone on a different continent or even different country. So you want to unify this as much as possible as soon as possible in the in the whole chain of events rather than later on and then try and work out how do you get all the data speaking to each

other. So I would imagine you plug FHIR in as soon as possible in the whole chain of events and that just makes sense to me. What do you think?

A: Yeah, it might be worthwhile to walk through some of the slides. I know they were shared on the COVID-19 Task Force call but we didn't go into too much detail, but in terms of case reporting, you can see down at the bottom we have multiple potential points of service that we want to submit a case report, or two into the case report repository, and we have some data workflows here. And to get into some sort of specifics to break that down a little bit more and that lower level is we may have some legacy systems that have very little support for FHIR or even resources to upgrade and provide that support. We have the middle of the road where we have systems that can do some basic FHIR manipulations and API calls as well as on the far right, the more mature from this perspective, fully FHIR capable, can submit FHIR resources and work with them sort of natively. So what we want to do support a lot of these different scenarios for submitting data. The green boxes, these mediators, are being built. The upper two will be a part of that Jembi work that was mentioned previously, and the patient level monitoring mediator is being built through PEPFAR funding, but the intent is to quickly get as you're saying these point of system submitting data. If not environment quickly transformed into FHIR into a case report repository. And then one question is what do we do afterwards; how do we actually use the data? I think there's also interesting things here to explore that hasn't been covered yet in terms of aggregating the data or creating dashboards with something like Kibana or creating more of a clinical research data pipeline from the FHIR based case report repository.

Q: This looks like a data warehouse with data mocks to me?

A: Yep, but very narrowly constraint. We've got the FHIR questionnaires and their clinical representations in the composition of the bundle and the questionnaire response.

Q: The Jembi team's strength is in collecting the data, exposing those API endpoints for the outside world to connect to that will then collect it. And then from our side, we make use of it through dimensional models and then start reporting off of it. That's where I see we could add value to this whole process.

A: I think it's stated in RFA somewhere, but we're looking at open source tools and open source stacks on this so I think those types of solutions could be of interest. One thing that was not a requirement, but sort of a recommendation is that we're looking as much as possible to the Instant OpenHIE platform that brings together a lot of the components in the OpenHIE architecture already and using that as a basis for the health information exchange, but a gap there is on the dashboards visualization aspects specifically around case reporting. But more generally just having a good data visualization and data analysis pipeline are gaps in the current architecture.

Q: Is this repository still to be built or will there be multiple repositories provided by the entities in question per country?

A: It will vary country to country. I don't believe that there is any solution that's adopted in multiple countries as a FHIR case report repository. Some countries will use SORMAS like Ghana, Nigeria, Fiji. Other countries are using DHIS2 tracker but that's not yet aligned with FHIR. The Jembi team will set up a generic FHIR server, such as a HAPI FHIR server as the case report repository and I think that's where the good intermediary between the point of service systems and whatever the emergency operation

center or whoever's actually managing the COVID-19 response is using that the FHIR server could be that intermediary for the data exchange and normalization.

Q: Don't you see the HAPI FHIR server as limiting rather than enabling in this case? In yesterday's session, you said that it would be a fairly tall order to implement the FHIR structure map. I would see this solution as requiring you know quite a lot of the FHIR components to operate efficiently and successfully. If we are limited with the server that does not support the hierarchical map, I think we are against be limiting ourselves in how successful this could be.

A: There are three ways to map a questionnaire, which is essentially an unstructured or semi structured data collection form onto the clinical data that's in the bundle. There's observation, definition, and structure map-based. The observation is only focused in on observations and it is quite limited. The definition-based is fairly powerful and we've used it in a several other projects and it's really shown that it meets a minimum set of data transformation requirements. The structure map has nothing wrong with it, it's just more complicated and not necessarily every tool would support that. That being said, we also want to promote the adoption of the FHIR standards as well as adaptation of them for specific country use. This aligns with a number of global initiatives, for example, they WHO digital accelerator kits and smart guidelines are being expressed in FHIR and this will be part of their smart guidelines and digital accelerator kit work or COVID-19.

Q: I was just asking whether the HAPI FHIR servers not being able to support the structure map is going to be a limitation?

A: Yeah it theoretically is, I don't know if it's practically in the short term, and we're not at all saying it must be HAPI FHIR server. That was just an example.

Q: So a HAPI FHIR is not concrete in the stage?

A: No specific software tool is set in stone. We do encourage you to look at the previous investments and the global goods that we've made and that includes DHIS2, Instant OpenHIE, and the components there, but there is no requirement so to say.

Q: This mediator or this abstraction layer that we just spoke about, where we will shield and basically enforce the five protocols, it's important, considering the volumes of data which has millions upon millions of FHIR resources. We will have to do some kind of legwork and homework to ascertain whether the existing offerings are up to speed. The reason I say this because we had an experience with OpenHIM. Because of the volumes involved, it wasn't able to manage the volumes. So, I'm just concerned that we are going to choose a solution that come time to delivery and implementation that, it's just not going to cut it. I'm basically asking for your guidance here. What do you suggest we do, how do we connect with the teams who are directly involved with these mediation systems so that they can answer these questions?

A: Yeah, I'm happy to get you connected with the Jembi team that's working on this and you can ask some question there. In terms of scaling issues, I think there's definitely a concern on scale and there's a lot of approaches that can be taken and one could be what are the bottlenecks and OpenHIM. Let's try to look at what those bottlenecks are and address them, or at least define the constraints under which we would hit these bottlenecks, so we know at what stage a purely OpenHIM solution can work and what stage isn't in terms of scale and volume. Another potential is looking at the Instant OpenHIE, which

has OpenHIM in there, but since it's with Kubernetes you could probably proxy through a specific high-volume tuned service rather than fully routing through the OpenHIM. So, I think there are a number of potential options to explore on the scale issue.

Q: If we also look at the two scenarios, one is priming your warehouse and remember this is country specific. It's not going to be an international solution. If we look at the total global number of COVID-19 cases, we're not talking about the volumes of HIV records that we have to process, so you might need to prime your warehouse upfront and you use a different mechanism for that and then going forward, once it's live, to keep it up to date to trickle feed it with real time updates when it comes, so from that point of view, that's one way of looking at it and resolving that problem. It's priming it and then keeping up to date. And then, we're looking at thousands of records a day, which is totally manageable with the API top interface.

A: I think just in terms of scale, in the US we have 50,000 cases a day, that gives you some sense of transaction.

Q: On this current slide that's on the screen at the moment, under the composition in that light green box. What is bundle and questionnaire response and what's the difference between them?

A: Yeah, so let me just back up to this to the slides here. So ignore the HAPI FHIR capable point of service, but the questionnaire resource is generally a lot easier for submitting data systems to submit their data and for engineers and software developers, they have to know very little about the FHIR data models. And so as the easy onboarding route that we're focused in on that questionnaire which models the WHO case report form. Just as an aside, this could also be the HIV case report form or other health areas. That questionnaire is sort of an entry attribute value table and doesn't really have any clinical meaning behind it. So what we do is use that questionnaire response (the questionnaire is the metadata of what's in the form and the responses and instance of that form filled out) and package it together with this bundle of clinical information as a FHIR composition and that is the case report. This bundle of clinical information is the specific observations and counter data, the patient, the things that map easily from the questionnaire on to the FHIR data model. We'll use definition-based extracts, so basically, each item in your questionnaire is tagged with a FHIR path that indicates which data field of the questionnaire gets mapped to the specific clinical data model.

Q: Almost seems to me like you're splitting the question into two parts. One that maps to FHIR and the other part the questions that don't map to FHIR and then trying to transport them together. Is that correct or not?

A: The questionnaire or the questionnaire response has both data that easily maps into FHIR and other data that doesn't. When we go through this questionnaire mediator, we want to extract the stuff that easily maps into FHIR into that bundle so that that can be reused for example in a shared health record for clinical care or for further data analysis. We are preserving the questionnaire response that was submitted, both for the audit record keeping of, this was our case report questionnaire that was submitted as well as to retain the data that does not easily map onto the FHIR data model.

Q: But if you use extensions you can map everything to FHIR.

A: Yes we can, but we didn't want to get into a proliferation of extensions that can cause some issues in terms of interoperability, so in particular, what we want is an infrastructure that can support case

reporting across multiple health areas. And if we have to start doing extensions straight away just to deal with COVID-19 and then you have to do extensions to deal with HIV case reporting etc. that quickly becomes unmanageable at scale. Doing it in this way, helps separate more cleanly the content versus the structure if that makes sense.

Q: We don't just want to create extensions, just to get the data through. But at the same time, you are leaving it for another day to be resolved eventually. There might be some useful information in there but at least you are keeping it on record so you could deal with it at a later stage. Do you have location information coming in with the data? Do you know where the facility is that's capturing it as it relates? Is information available that could come with the data that we are capturing or that these APIs would be capturing?

A: I believe so; I'm 95% sure but I'd have to pull up the ID itself, which I can do but I believe there is a location resource in the case report form. Let me double check that but once we have that there is an IHE profile called mCSD for mobile care services discovery which details how we use location, how to use the FHIR location for mapping the facilities that have the geographic coordinates, whether it's a point or a boundary, and how to make use of that. I don't think it's explicit yet in the implementation guide, but we would recommend the mCSD as a way to manage the facility metadata.

Q: If you want to stay very specifically with FHIR without extensions, how are you going to track comorbidities, which is quite crucial to COVID-19?

A: In the questionnaire, there is a section for comorbidities.

Q: How you going to map it without FHIR extensions?

A: I guess there are two issues here. One, how to represent the FHIR resources. These would be generally mapped to observations and I'd want to check with Elliot on that but I believe that's the right resource for that. The second aspect of this is the coding and the terminology related to this. For something like diabetes, there's multiple codes. One thing that we're trying to resolve is to what level of detail do you need to know that somebody has diabetes is just yes or no, that they have diabetes or do you need to know, specific type one, type two or more specific details on that. Hopefully next week, we will have the Go.Data team to talk through those questions. Does that address your question?

Q: Yes, to a degree, and because I see comorbidities as either a yes or no, but also see it as time based because you might have a patient to present in September with a comorbidity of diabetes and diabetes might be something they've had for the last 10 years. So it's a yes or no question as the patient presented with diabetes as a comorbidity. However, the patient might present today and is quite clear of everything and within a month presents again and is showing signs of severe TB because that is also comorbidity, but that that is a temporary situation. It's not a permanent situation and same with asthma and even cancer. How are you going to cater for that?

A: My understanding for the general case reporting workflow is that you would record the comorbidities and the symptoms at present that those observations would be a part of an encounter and that encounter would have a timestamp associated to it so you'd be able to say when they presented with a comorbidity. I guess it depends on whether you're talking about clinical care management aspects, in which case you would want to see all of the comorbidity information in their shared health record and access to that. There might be conflicting information over time. Right, so you don't have cancer, then

you have cancer. It's up to the clinician to determine what to do with that information, we're not dictating the clinical care management aspects of that.

Q: So this is COVID-19 specific and are we talking about capturing and reporting COVID-19 specific data only so like you said at that point in time, you register a COVID-19 case for a person for a patient and whatever relevant comorbidity information you have at that point in time. But wouldn't be tracking it over time, you wouldn't be tracing it or following up over time. Is that just a response to the COVID-19 case specifically in capturing it at that time and not as it is or was with the case with HIV the way you track a patient over the patient's lifetime because COVID-19 is a curable disease?

A: This is less like the HIV use case, although we are seeing some instances of people having gone through symptoms but yes, I agree it looks the majority of the patients are time-bound and so hopefully there isn't a need for longitudinal clinical care management and really if you want to make the differentiation in case reporting, you can think of public health case reporting versus clinical case reporting, where the public health case reporting is focused in on understanding the size and scope of the problem to help make informed decisions on what are the best public health interventions. Whereas the clinical case reporting, which requires more detailed information than what's in this case reporting form that would be for the clinical care management.

Q: You may need to identify that a patient had COVID-19 four times instead of 200 million COVID-19 records. It won't make sense if you can't link the various episodes of one patient together, will it?

A: Yes, I agree that we need something in the longer term that's more comprehensive and that can take care of that. This one green box on deduplication mediator. That is where we would create a record linkage between different case reports assuming that we have enough identifiable information on the patient to make those linkages. Presumably, that interacts with a client registry that has deduplication service built into it and we can leverage that. But that is not work that's been done yet but I agree that that is an issue that we will need to deal with.

Q: But is that relevant to this proposal at all or is that for a lighthouse later stage?

A: It's hard to say. We definitely need that for HIV. This infrastructure is supposed to support HIV, COVID-19, and other disease areas. I think it's unclear. There are people that are getting reinfected and it's unclear what the implications of that are for the public health response. So whether we actually need to build this now or not, we definitely need to plan for it for that potential eventuality. We haven't seen a pandemic in the digital age. It's not like Ebola, where it was relatively contained and relatively easy to do contact tracing. This is a whole new beast and everybody's learning as we go.

Q: Now the solution I see to this whole problem is you want to plug FHIR onto everything as early in the whole process as you can and you want to plug as much FHIR on as possible?

A: Yes, and build your analytic and data visualization tooling on top of that, as much as possible.

Q: The point of service tool doesn't have to be FHIR-capable, right, just the layer above it needs to be?

A: Yes, that's essentially what these mediators are doing. They're a facade on top of those points of service that makes it look like they're submitting FHIR data.