Global Alliance
for Genomics & Health

ARCHIVED - 2023/2024 minutes can be found here

GA4GH Clinical & Phenotypic Data Capture & Exchange Work Stream
Minutes and Actions 2021-2022

These are the minutes for the GA4GH **Clinical & Phenotypic Data Capture Work Stream**. For further information, please visit the GA4GH Work Stream page at www.ga4gh.org, or contact the Work Stream Manager.

- Clin/Pheno Landing Page
- Phenopackets Subgroup - Agenda & Minutes (ARCHIVED)
- Phenopackets Cancer Task Team Notes
- Cohort Representation Minutes
- GA4GH 2020 Overall Strategic Roadmap
- Clin/Pheno 2020 Strategic Roadmap
- 2019-2020 Meeting Minutes
- 2017-2018 Meeting Minutes
- Recordings

GA4GH Slack Workspace

- Clin/Pheno channels: #clin-pheno_ws, #cp-phenopackets, #cp-pedigree, #cp-cohort_rep

Table of Contents:

Meeting Protocols

List of Meetings

- 2022-12-07 - Clin/Pheno Year End Roundtable
- 2022-11-22 - Beacon v2 and Phenopackets v2 Integration
- 2022-11-21 - Pedigree Subgroup
- 2022-11-03 - Pedigree Subgroup
- 2022-10-20 - Pedigree Subgroup
- 2022-08-23 - Computable Cohort Representation - Summary Statistics
- 2022-08-11: Pedigree Subgroup
- 2022-08-04: Pedigree Subgroup
- 2022-06-23: Pedigree Subgroup
- 2022-06-09: Pedigree Subgroup
- 2022-05-26: Pedigree Subgroup
- 2022-05-25: Clin/Pheno Full WS
- 2022-05-xx - Computable Cohort Representation Hackathon
- 2022-04-28: Pedigree Subgroup
- 2022-04-20 - GA4GH Montreal Connect
- 2022-03-24 - Computable Cohort Representation Meeting

2022-03-23: Clin/Pheno Full WS
2022-03-17: Pedigree Subgroup
2022-03-15 - Computable Cohort Representation Meeting
2022-03-03: Pedigree Subgroup
2022-02-23: Clin/Pheno Full WS
2022-02-17: Pedigree Subgroup
2022-02-03: Pedigree Subgroup
2022-01-20: Pedigree Subgroup
2021-12-14 - Cohort Rep Report Back Meeting
2021-12-09: Pedigree Subgroup
2021-12-02: Computable Cohort Representation - PHENO Team
2021-12-02: Computable Cohort Representation - MIACC Team
2021-11-11: Pedigree Subgroup
2021-11-02: Computable Cohort Representation - PHENO Team
2021-10-28: Pedigree Subgroup
2021-10-19: Computable Cohort Representation
2021-10-12: GA4GH Connect
2021-10-05: Computable Cohort Representation
2021-09-30: Pedigree Subgroup
2021-09-21: Computable Cohort Representation
2021-09-16: Pedigree Subgroup
2021-08-19: Pedigree Subgroup
2021-08-05: Pedigree Subgroup
2021-07-28: Clin/Pheno Full WS
2021-07-27: Computable Cohort Representation
2021-07-22: Pedigree Subgroup
2021-07-13: Computable Cohort Representation
2021-07-08: Pedigree Subgroup
2021-06-29: Computable Cohort Representation
2021-06-15: Computable Cohort Representation
2021-06-10: Pedigree Subgroup
2021-05-27: Pedigree Subgroup
2021-05-26: Full Clin/Pheno WS
2021-05-13: Pedigree Subgroup
2021-05-13: Computable Cohort Representation
2021-04-29: Pedigree Subgroup
2021-04-22: Phenopackets Subgroup
2021-04-15: Pedigree Subgroup
2021-04-08: Phenopackets Subgroup
2021-04-01: Pedigree Connectathon
2021-03-25: Phenopackets Subgroup
2021-03-24: Clin/Pheno Full
2021-03-18: Pedigree Subgroup
2021-03-11: Phenopackets Subgroup
2021-03-01: Virtual GA4GH Connect 2021

Global Alliance for Genomics & Health

2021-02-25: Phenopackets Subgroup
2021-02-18: Pedigree Connectathon
2021-02-11: Phenopackets Subgroup
2021-02-04: Pedigree Subgroup
2021-01-28: Phenopackets Subgroup
2021-01-21: Pedigree Subgroup

Meeting Protocols

- Please note that by participating in meetings, attendees agree to adhere to the GA4GH Standards for Professional Conduct.
- Meetings may be recorded for note-taking purposes. Recordings will be deleted within three months of the meeting taking place.

List of Meetings

2022-12-07 - Clin/Pheno Year End Roundtable

Chair: Melanie Courtot, David Hansen
Attendees "Name (Project Affiliation)": Moni Munoz-Torres, Oscar Pastor, Marion Shadbolt, Ian Green, Grant Wood, Joe Cody, Chen Chen, Vanessa Vogel-Farley, Kathy Reinold, Soichi Ogishima, Jeffery Shu, Victor Lin, Francis Jeanson, Aarthi Krishnan, Coline Thomas, Nebiyou petros, Jules Jacobsen, Zornitza Stark, Nicky Mulder, Ayhan Serkan SIK, Orion Buske, Adriana Malheiro, Mogomotsi Matshaba

	Actions Arising	Assigned To	Deadline
1.	Review the Pedigree documentation	Vanessa (and anyone else interested!)	
2.	Possible use case/pilot with Australia Biocommons for Phenopackets and Beacon integration?		
3.	Follow up offline about modifying the terms in the KIN ontology to be more inclusive for reproductive technologies	Lindsay, Grant, Orion	
4.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH Connect: 2023 Strategic Refresh is now	Lindsay	5 min

	open for public comment! Deadline is Dec. 22nd
	- Call for New Driver Projects opening in the new year
	- GA4GH Connect 2023: April 19-21st in London, UK
	  - Please let lindsay.smith@ga4gh.org know if you have a workshop in mind!
	- Next full WS meeting January 25, 2023
	  - Proposal for a pharmacogenomics standard (TBC)
	- Student Projects		

2.	Phenopackets Year End Roundup - Slides (bit.ly/cp-pfx-fall-22)	Moni	5 min
3.	Computable Cohort Representation Roundup - Minutes (within this you will find linked the landscape analysis spreadsheet & use case collection doc) - Background & Context Slides - Hackathon Summary/Minutes	Lindsay	5 min
4.	Pedigree Year End Roundup - Github Space - Documentation	Grant, Orion	5 min
5.	Community Roundtable - Slides - Rare-X - Biocommons - Enabling Clinical and Phenotypic Data Sharing between Australian Research institutes - Improving IVF/third party reproduction language in the KIN Ontology	- Vanessa Vogel-Farley - Marion Shadbolt and Victor Lin - Lori Metz and Joe Cody	5 min presentation + 5 min for Qs EACH (30 min total)
6.	A.O.B.	All	5 min

Minutes:

A transcript has been captured using Otter.Ai.

Phenopackets Update:
MMT: We've created a standard within the GA4GH and this is a standard that literally allows you to pick up a packet of genotype data that can be used anywhere by anyone. And these packets are encoded using ontologies and can be used to develop a patient's phenotypic profile or one for model organisms. And it helps to improve diagnostics mechanism discovery and an integration with the environmental health data. Phenopackets allow us to exchange information about the patient, including him about identifiers, information about age of collection, as well as details about each of the absurd phenotypes. And this includes whether or not the phenotype was observed when they were first observed, how they were first observed in how severe a

phenotype is, and how each of the phenotypes is linked to a patient through their genomic information through the samples that were collected in when available even to the patient's parents and siblings. And so in the latest version of the standard, the latest updates include improvements for the representation of cancer in COVID-19 phenotypes. Phenopackets are going to benefit research biologists, physicians and patients alike and they can support a diverse diversity of use and use cases listed on the left in the context of biomedical ecosystem, right. So there are multiple providers of phenotypic data. This includes you as a patient you as a clinician, as well as a data that comes through a variety of mechanisms that include mobile health gadgets in the electronic health record, and the Phenopckets schema acts as a common model that can capture data from all of these different sources with a unified software representation. And in turn, it can be used by multiple receivers of the phenotypic information. And this includes journals, databases, registries, clinical laboratories, we're hoping that its implementation can help us to speed up the accuracy of the diagnosis, as well as the treatment effectiveness, and that it can facilitate cohort identification and patient matchmaking during the clinical trials, and just in general, improve the data landscape and facilitate the creation of all these tools and algorithms but as they find their way into databases in website and web based tools.
There's a lot of really good exciting things that have happened over the past few months. We have now a standard that it's both approved with the GA4GH as well as with the International Standards Organization. And there have been a few joint news releases with the standard Council of Canada that issued this celebration of being able to bring the first of the standards from our group. We have a paper that was published in June. Not quite the fall but we're still very excited about it. And the paper we just published in August, which is a practical introduction, basically a tutorial. We have work pending, we're working with generating a set of a validating tools, just working on an adjacent a validation tool that would allow you to basically recognize whether or not you have the correct information in the phenopacket as well as trying to do generate tools for a building where we have a new website that is not yet deployed but will be able for you to be able to transform your information into a phenopacket. And there's also of course, exciting work happening cross workstream with our partners in genomic knowledge standards as well as the ability to support the second version of Beacon so that we can accept queries with phenopackets and return a list of phenopackets as well. Of course, we're working with the rest of the Clin/pheno workstream to support pedigree.We basically started building an OMOP phenopacket exporter. This runs as a web server on top of OMOP and exports data into a phenopacket format for an individual. They started as a biohackathon work last year. We are also working with the HL seven organization. They work on interoperability for electronic health records and Phenopackets is one of the very select group of Vulcan accelerator projects, which are projects to foster growth and in the application of research standards.

Computation cohort still study group status. Spent lots of time exploring and defining queries for cohorts and phenotype for individuals in cohorts. 5 different breakouts during hackathon, eg Beacon, data access requests, FIHR & HL7…Overlap with other GA4GH groups - eg DURI group looking at standardizing Data Access Requests. Overlap with Beacon and PXF.
Hard to focus on one direction - too many interesting ideas!

Pedigree
LS: Working on resolving the last remaining comments for PRC approval, aiming for SC approval in January. In the agenda, you'll see I've linked the GitHub space in the documentation, and those two links will take you to everything you need to know about the pedigree standard. Through the PRC process, there has been quite a bit of changes. So if you haven't looked at the documentation in a while, I'd really recommend you take a look. One of the other big updates I'd like to share as we're working on, hopefully migrating our kin ontology to the OBO relations ontology.

Rare-X
VF: Starting coming into this space an build on what was out there. We are a little bit different in that we are patient-driven, spin out from Global Genes. We have a platform for collection patient and health centered data.

We guide with advocacy partners and share equity across all diseases. Create collaborations with researchers and biopharma studies in the rare disease space.
There is so much overlap in disease, but we also see domain specific areas. Wanted to bring something together that can show cross domain findings. This overlap drives the use of standards and working with GA4GH. We want to approach infrastructure in a way that is patient-centric. The patient is reading a question that makes sense to them, but coded in a way that is structured and machine readable. Focus in on an integrated or federated data model, connect where it makes sense and not where it doesn't. Everyone comes in the same door with data collection. We have a layered model that can be continually built upon. Mapping domains back to the standards to make them usable and identifiable. Make sure to do the landscape work and bring experts together.
Case study with Colorado Children's genetics. Clinician reported data collected in REDCap.
Industry partnerships with Illumina, and Q-state biosciences.

Australia BioCommons
MS: We don't have a national solution for data sharing and archiving. We needed something to bring everyone together to be able to query across them. Looked at Beacon, having systems where each institute can transform their data into the common Beacon model, and bring everything together in a Beacon network. Want to make this process as easy as possible. Within the Beacon network, looking at developing a better UI. Have developed a serverless Beacon (sBeacon). We have no active backend running 24/7, it's also auto scaled. No data transformation needed, uses VCF as is. Still moving parts.

Improving language in the KIN Ontology
JC: Collecting family history that is accurate is essential. Knowing whether it's donor, sperm, eggs, can change care. Want to change family history capture process through point of care or imputation. Want to work with stakeholders and promote conversations between patients and clinicians. Some of the comments we submitted had to do with the IVF example that was in the Pedigree documentation. Recommended changing that category to accommodate other forms of assistive technology. We had recommended some basic terminology to add to the ontology.
OB: We have some representation of this in the owl file.
VVF: How to describe self-reported things to be patient centric.

2022-11-22 - Beacon v2 and Phenopackets v2 Integration

The minutes/agenda for the Phenopackets / Beacon Integration meeting can be found here.

2022-11-21 - Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Alex Tsai (GA4GH)

	Actions Arising	Assigned To	Deadline
1.	Follow up with Chris about updates/timelines,	Lindsay	

	clarify namespacing		
2.	Follow up with Jules about Phenopackets content	Lindsay	
3.			
4.			
5.			
6.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping	Orion, Grant, Lindsay	5 min
2.	Reviewing Past Actions: ● Check in with RO contributors about timelines (ie. does it take weeks or months to merge things in?) - Lindsay ○ Likely will not be ready for January? ● Create issues for the outstanding KIN Ontology fixes, cross-post in RO - Lindsay / Orion ○ Done, Chris created an issue in RO ● Create PR in RTD for Phenopackets content - Jules ○ Not done? ● Fix up the ecosystem diagram and add to RTD - Lindsay ○ See here ● FHIR IG Updates (KaryotypicSex, Population Descriptors) - Meen ● Share text about FHIR IG Balloting to add to RTD -Grant ○ See PR here	All	10 min
3.	Remaining To-Dos - See PRC Comment Tracker - We'll walk through the Reds and discuss. Outstanding items are summarized below: - Documentation Updates, including: - Revised examples with FHIR modelling (nice to have) - Clarify Phenopackets & FHIR linkages		10 min

	- KIN Ontology - add content about moving to RO and making issues - Class Diagram on the Concepts page needs to be updated (still says REA instead of populationDescriptors) - KIN Ontology updates (comments from Open for Public Comment) - Two different valuesets to incorporate - GEL and FHIR - IVF Examples - FHIR IG Updates - Karyotypic sex, population descriptors - Phenopackets integration - Need a set of instructions for using the standard in the context of phenopackets - Jules is on it - Outreach to implementers - draft email her - Pedigree-tools - Need to tweak the standard first		
4.	A.O.B.	All	5 min

Minutes:

Walk through of actions and remaining to-dos.
Timelines: Send documentation to PRC either Dec.1st or Dec. 5th. Ask for thumbs-up before the holidays. Documentation to SC by Jan. 12th.

Consider “implementathon” for post-approval - bringing together implementers to Connect

2022-11-03 - Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees “Name (Project Affiliation)”: Meen Chul Kim (CHOP/KF), Zornita Stark (AGHA), Arthur Hermann (Kaiser)

	Actions Arising	Assigned To	Deadline
1.	Check in with RO contributors about timelines (ie. does it take weeks or months to merge things in?	Lindsay	Done
2.	Create issues for the outstanding KIN Ontology fixes, cross-post in RO	Lindsay / Orion	
3.	Create PR in RTD for Phenopackets content	Jules	

4.	Fix up the ecosystem diagram and add to RTD	Lindsay	
5.	FHIR IG Updates (KaryotypicSex, Population Descriptors)	Meen	
6.	Share text about FHIR IG Balloting to add to RTD	Grant	

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH Connect - The Variant Summary Statistics Format v1.0 is now open for comment	Orion, Grant, Lindsay	5 min

2.	Updates:
	- Steering Committee timing confirmed! Thursday January 26, 2023 at 20:00 - 23:00 UTC
	- Pedigree Ecosystem Diagram
	  - New version, suggested revisions welcome
	- Moving KIN to the Relations Ontology
	  - Action: Cross posting issues to the RO Github?
	- Phenopackets integration documentation
	  - Jules will write some text, I advised he make a PR to the RTD
		Lindsay	5 min

3.	Remaining To-Dos
	- See PRC Comment Tracker
	  - Note: Lindsay has changed the colour-coding: Green = PRC Approved/Resolved, Yellow = We have completed but not yet PRC approved, Red = To Resolve
	- We'll walk through the Reds and discuss. Outstanding items are summarized below:
	  - Documentation Updates, including:
	  - Revised examples
	  - Addition of content around scope, rationale, use cases, etc.
	  - Clarify Phenopackets & FHIR linkages
	  - Add karyotypicSex
	  - *Class Diagram on the Concepts page need to be updated (still says REA instead of populationDescriptors)*
	  - *Are we still adding a free text field alongside populationDescriptors?*
	  - KIN Ontology updates (comments from Open for Public Comment)
		All	40 min

	- Two different valuesets to incorporate - GEL and FHIR - *See comment on RO and changing the relationship with the standard* - FHIR IG Updates - Karyotypic sex - Phenopackets integration - ~~Delete phenopacket-pedigree repo~~ (DONE) - Need a set of instructions for using the standard in the context of phenopackets - Jules is on it - Outreach to implementers - draft email here - Send this later, set deadline for ourselves first - Pedigree-tools - Need to tweak the standard first - IVF Examples - For KIN Ontology integration, list of suggestions for RO		
6.	A.O.B.	All	3 min

Minutes:

Worked through the PRC comment tracker and updated text and color-coding as appropriate. Main outstanding items are KIN Ontology Fixes, FHIR IG Updates to reflect a few changes to the model, and Phenopackets integration text. Next meeting will be a quick check-in scheduled for Monday Nov. 21st @ 4:30pm ET / 9:30pm GMT / 8:30am AEDT.

Link to class diagram:
https://drive.google.com/file/d/15zESxFd6z17iKB2e68sRYbX1iDreWjTn/view?usp=sharing
XML export is on the Slack channel.

NOTE: The below transcript was auto-generated.
GA4GH Connect is coming up in just over a week from now.
There is as of right now. Put potentially one meeting coming from Clin/Pheno on related to Beacon and Phenopckets integration which couldn't be somewhat related to us if you know we get the the pedigree integration with Phenopckets going as well. Timing of that hasn't been scheduled, but I'll let you know when it has been.
And there's a couple of other sessions that might be of interest. Now. There's one on data access. There will be a data model and schema consensus group meeting on November 30 A little later this month was the other one.
It's not coming to me right now, but there's a couple of different vegan meetings and such as well.
And then other note, and I've linked this in the agenda as well. The variant summary statistics format is now open for comment. You can find that right on the GA4GH dot org main page. And so essentially this is a format for GWAS summary statistics originally developed outside of GA4GH, the version one and now they're they're hoping to bring it in and develop it further.
Are their own work stream or is that another work stream that has done that? No. So this is a new standard that's going to be put through in LSG large scale genomics, but the standard itself was developed kind of like outside of GA4GH but like Yasi if you know y'all see, he it was him and his team so he's he's kind of part of GA4GH already, but it's not a official standard of ours that that we had created. So they're hoping to bring it in and have us kind of adopt it.

So yeah, anyway, I think that open for comment is there until the end of the month, so you've got a bit of time to take a look at it. It's really short documentation. It's only about two pages. So should be should be a quick quick review.
Okay, and then I'm just going to go into a couple of updates from our meeting last week.
Big One is that the steering committee timing for January is now confirmed. So it's going to be Thursday, January 26 at 120 UTC, which I think is this time.
So pretty much at the timing of one of our regular meetings is the steering committee meeting in January.
So knowing that we have to have all documentation submitted to steering committee two weeks before that, so that would be January 12.
But you know, if we stick to our original plan of kind of having that final PRC checkmark in December we'll be in really good shape and have lots of buffer time there. So so just to make note of that, that that has been confirmed. And then typically, we do have to do a presentation at that meeting. So no grant or Orion if you wanted to take that on but we will have to make a presentation for that assuming the the PRC I'm assuming when you point that way, it doesn't go to Orion for me but that's not pointing at me. Yeah, there you go down. Okay. Anyway, yeah. Orion is this direction on my screen.
So So yeah, so we will have to make a presentation for that. But I'll get you the details for that. Once I have more. Now, Lindsay, is that is that the very final step for us? For Approval? Yes. Yeah. Once we kind of get all this other stuff done. Then we do that presentation, the steering committee so when they receive the materials two weeks in advance, that will be their opportunity to kind of read it and decide, you know, how they might want to vote or any other comments they might have.
Then they'll vote at that steering committee. And so what we'll know by the end of that meeting, whether we get thumbs up or thumbs down hopefully it's a thumbs up. And then after that point, then there'll be like a little bit of implementation, outreach sort of stuff, like as your comms team will write some sort of article. And at that time, we'll also have that student who we are planning to have do a lot of the outreach and writing for us which is awesome. So we'll get lots of lots of communication out there.
Any other questions about that process?
Oh, good. Pretty straightforward.
All right.
One other update I had and I'll share my screen here.
So I've been working on a revision of the ecosystem diagram and I wanted to check with everyone that I've kind of got it. So are you seeing this slide here? Yes. Okay. So based on the conversations we had at our last meeting, I have moved the export import ingest arrows up to the implementation box.
I've split the family health history, data collection tools and pedigree management tools into two things, but I don't think these are split properly. Stephanie had done this originally added another box for risk analysis tools. renamed this to the pedigree conceptual model.
Did I miss anything?
Yeah, because that's not a pedigree management tool. I think other than sorting out some of these things in the top left, everything else looks great. I like it. I think this is a huge improvement. Lindsey, and I think it'll be okay. Perfect, are really, really useful. Or and maybe I opened up a can of worms by suggesting that we split those two and maybe we should just group them back together again, depending on open to whatever people think but I think all of the other changes that we that, that you've gone ahead and made are great. And I think this will clear up so many of the confusions that people had and it'll, it'll work very well on that top. Okay, awesome. Yeah, and I mean, any. Oh, Grant, go ahead. Yeah, just a couple of questions. To help help me further understand what you've done. You're especially on the on that top left there. So you have the three data collection tools, and I don't have an opinion on this or not. I was just curious, did the three that you pick there or was that because they're most well known in the research world? No, this is what you're seeing on this screen are ones that either have implemented or will potentially implement. So it's an example of kind of our Oh, good ones in the room. Yeah.
Yeah, no. So hopefully we'll be adding more, you know, logos to that box.

scan that the GC chat box. Can you tell me a little bit about that? Because, yeah, so that was I know, maybe Zornitza. You know more about this one than I do, but I understand. So Zerah developed this kind of counseling chat box that can be used to kind of, you know, a patient or someone could talk with this chat box and it'll collect, you know, some of the information you share with it.
And they missed, but that's about as much as that's it. Yeah.
Has a name to it.
Did Alejandro start working on this? I can't, couldn't remember. I don't remember who did the work on it either. I feel like Ella hundra at least did something with it because I think that's how we originally heard about it. But okay, Edna Chatbot. That's good.
Great. Okay. Thank you. So yeah, um, so again, you know, I'll leave this up. I think it's open for editing. So feel free to play around with it and, you know, edit as needed, or let me know if something can be edited and I can, I can play around with it.
But again, yeah, if I missed anything, please do. Do let me know.
Okay.
Other updates.
Okay, so we heard back from Chris Mangal. about moving kin into the relations ontology. And so he suggested as a first step, either creating an issue in their ontology, kind of to get them start looking at all the relations ontology folks to start looking at kin.
Although Melanie had suggested, alternatively, to cross post some of the issues we had received in our repo over to aro as well.
There was never an immediate response of which one we should do.
But maybe, maybe both create an issue to say, hey, we're gonna move this thing over and then also cross post some of the other ones that we've received. So it's all in one spot.
I don't know if anyone has a preference of how that's moved forward or not. But or we can wait and see if Chris has any preferences that come back.
But that's good. We've all been moving forward on that.
I just want to make sure because I don't know if I heard a Ryan's direct opinion on this. I I know that Lindsey you originally sent out emails about this question. Orion, are you happy with this move or was there any concerns that you have?
I don't have any concerns about this one. Grant. I think it I guess. Yeah, I think we talked about that on the week when you weren't able to join a vote. I think my my general feedback on this was is that I think it from a sustainability and having having the ontology people invested in keeping it up to snuff over time. perspective, it seems like a really useful home for it, as well as from the perspective of in order to keep the ontology is on our side. Generally, the rule is don't create new ontologies if you don't need to, and so we want a new ontology, but we can get keep them happy by giving it a home. So I think that those all points in one direction for me and there aren't any particular concerns that I have about it. Are there any that you have, or thoughts that you want to share? Right? No, I didn't have any either. I, it sounded to me like it was the best idea that was on the table. And I just wanted to make sure that there wasn't some other issue that I wasn't aware of right about doing this. And so that's why I just want to ask you, but no, I'm happy to. Well, just I guess we'll just have some ongoing relationship with this group then. At the very least, it might sound like we want to help promote them in some way, because that would benefit us right if we did yeah, like by promote you mean relations, ontology, or well, both both their group and our kin ontology. But and when I say promote, I just mean more like get the word out. Let the world knows that this group exists. And that's where our kin ontology will live. Yeah, yeah. And I think we can, you know, once it's in place, we can update that in the documentation as well. Because we have that that kind of chunk of information about the kin ontology, so we'll just have to update where it can be found, essentially, right.
So yeah, yeah, I agree.
And Great, okay. The one last update I had is I did reach out to Jules about getting some instructions on how to use pedigree in the context of Phenopckets.
He said he'll make a PR for in the read the docs.

For what he's going to write. I asked if he could do it within the next two weeks at the latest when he seems to be no problem. So so we'll just wait until the PRS ready there and we can get that in. So that's fantastic.
Great fielding his questions. Yeah, no problem. It did. I answer, okay. Or did you? You're good. You're good. I fell in between the one the one day each week where I let all of the GA4GH emails queue up before I can deal with them.
Okay. Any other updates? And we wanted to bring up before we dive into the two dues, or maybe they'll come up as part of the to dues? Yeah. Okay. All right.
So I'll take us over to the tracker here.
So I mentioned in the email previously, I've changed all the color coding in the track tracker. So green means it's resolved. It's done. We don't have to look at it. Yellow means we've resolved it. It's just PRC. This was one of the things PRC requested, so they'll just need to look at it. So from our point of view, we don't have to worry about yellows, and then reds are things that are to do although a lot of these I think we can change today.
And a lot of them are things that just we just need to resolve or I've kept it read because I have a question about it now and we can talk about those today as well.
So, before we get into things, one, I guess the main update that I'll give is that I've put forward a PR a pull request to try to handle some some subset of the to dues that were supposed to be done. Would you like me to merge it right now? And then maybe people can see it reflected on read the docs when we're going through these things?
Otherwise, we can just reference kind of the pull request and I can just use words to describe what was done.
I mean, I'm happy for you to merge it now. Does anyone else object?
No. Alright. Okay.
I need to Oh, no, I'm not allowed to purge my own pull request. Oh, really? Oh, it will need to be one of you. Let me approve, or approve it. Yeah. Okay. Oh, no, you can't even review your own pull request and say that it's okay. And then merge it, which is really the is that a setting that I have on the repo or is that just a it's a repo setting? Isn't it isn't my done? You've turned on the requirement for review, and require at least one review before merging which will make sense.
And yeah, okay. You know how I'll look at that later. But anyway, if approved me and approves it, then me.
Good to have friends.
in high places. Yes, I placed this Exactly.
Right. Okay. So that'll theoretically be reflected soon, thanks to means GitHub action. Thank you.
Okay, um, so why don't we just kind of start at the top and we'll roll down.
So the first unresolved item here was, again, just related to the documentation. So this was about
oh, so including some of the, you know, rationales behind the certain decisions we made. So, you know, why did we create a kin ontology instead of using you know, the fire set?
What else was there? Why didn't we use the family member history genetics profile? Why did we make a new IG?
So in terms of those rationale type questions, has that been added to the documentation yet?
There is some discussion of the alternatives that is now broken out into some of the different value sets and profiles and the limitations of them. It is
Present. Oh, yeah, you've got the updated one there. Lovely.
So yeah, so it talks about those existing standards for pedigree for family history, some of the rationale behind it.
And then what our goals were. So I guess this, this intro probably had the most textual changes out of other sections in terms of what was changed. And there was a fair bit of reorganizing at least that I did to try to make the documentation more coherent and kind of flow better with I think how people will come in and need want things presented to them. Perfect. No, that looks great. Okay, so that sounds like we can change that too. I hope so. Yeah. So there can always there can always be more and so what we'll probably want to do is make sure at least at a high level, that we've attempted to check these boxes, and then what we'll probably want to do is as we were going to send it out to stakeholders to read the actual human like the actual pros and see if it

all makes sense still. And then see if it covers there. Okay. So I'll turn it to yellow for now. And then I'm also just going to flag that box in blue, so I have more text.
Great. That's good. That's good.
Oh, okay. So this was this one, I changed to red because I had a question about it.
So I was thinking about the move to the relations ontology. And so there was this comment from the PRC about, you know, adopting kin as part of the standard and CO evolving it. So now that if we're moving it to relations, ontology, like do we need to say anything about that relationship? Is that relationship going to change from what we had written earlier?
Like, I think we could still frame it as you know, we'll co evolve it and, you know, bring those issues to aro when we need to.
But I just wanted to double check that there wasn't anything I'm missing. There.
That's fine. I think that's okay. Yeah. The
Is there anything else there that we should probably be sure to do? I think that's okay. For now. Yeah. Okay, we're just saying and, and for sustainability sake, it's been incorporated into blah, blah, blah. And it will be
Yeah, we'll take a request there.
So the the kinship ontology documentation page will need to be updated with all of this information. And these changes before we before we submit that good point, okay. So maybe, then maybe I will not change this to that and just add here
do see
okay, in that case, I'm just gonna turn that back to red for now. So remember to do that.
Okay, so this one is one that I don't think we can really deal with right now. But So expanding the kin ontology with terms to maps, those additional terms that were that had been listed by a gusto. I think this will kind of fall under the stuff to do for kin with aro so I don't know if there's anything we can really do about this right? Now. But so the plan last time, I think, was that we were going to kindly ask our O to do this. Is that right? Like the Chris Chris and the ARO maintainers to do this, if they're if they're so willing?
Yeah, yeah, so I think I mean, I'll leave it as read just so we still see it as a thing here, but I guess we can not not no action we can take right now on it. Okay, but we're gonna do it beforehand. And I did see that they had started making some pull requests to the terminology.
repo, so I don't know when that switch over, that they're planning on doing is they're probably just making sure that what is there is clean before pulling it in? Yeah, I think so. Those issues that were made, I think were the ones Melanie suggested cross posting over tomorrow.
So we could we could potentially do that.
Okay, so they already submitted pull requests that fixed though Oh, six. So therefore, I guess what they're, how many issues are there? There? are 22 2345677 are from the open for comment period. And of those seven, I think at least two two were fixed and then they made two other changes. Okay, whether that's great for context.
Awesome. Thank you.
Um, okay, actually. So one more question I had as well. So, I suppose there is the potential that you know, moving the ontology to aro could take longer.
You know, I mean, I guess I don't know how long it could potentially take but we're kind of dependent on this. You know, external group to help move these things along. So I was thinking about the timelines to approval. And so I guess the question I have, can the conceptual like the conceptual model, theoretically can be used without the kin ontology, right.
Yes. Okay. Because if that's the case, then I think it's fine for us to still go to January without necessarily having kin ready.
Because essentially, so I was having this conversation with Angela, but like, what do we do if the ontology is not ready?
And she said, you know, if the if we consider the ontology to kind of be the package piece to the model, then they should probably go together. However, if, you know if we just bring forward the model and the other associated pieces we have, it can be used without kin, than that is okay. And can can be added later.
Does that make sense?

Does the
the terminal like the conceptual model itself, theoretically works without like it?
The work is not meaningless without kin. But as a standard, it kind of doesn't make sense without having that reference to the ontology. Because even the basics of like, can you represent a trio requires kin? Yeah, so the alternative here if we wanted to unblock it on the external party, would be that we make the improvements to kin. Like, at the very least we make it so that terms exist for the things that we want to collect that we need to collect for coverages sake, and then when we hand it off to the ontology lists, they actually make sure that everything makes sense.
Right? Okay. Um, for her for lack of a better word, other than just saying like, yeah, yeah, we can define terms and, and give them identifiers. And then the ontologies can come in and say that all of the markup and the way that you've structured this is all wrong and it needs to be done differently, and then they they can do that. Yeah, I don't think that is at all at odds with the way that we anticipate this, you know, feedback and improvement process going. So that's a way of checking goals with well keeping to the timeline. So all it means is that we shouldn't depend. We shouldn't wait on aro to incorporate those. We should just mass merge them and ourselves, which is work, but not Not, not really not huge work. Okay. Yeah, no, no, because I was just thinking about that. And I was like, I would hate to like that. One reason we might not make it is because we're dependent on that external piece. So this is how we do work around it, or we just wouldn't even incorporate those and it's not the end of the world. Yeah, okay. So, okay, so what do you think should we add
so, you know, we had said that this one we were going to put to our row for example, should we make this ours? Like our issue to solve or not?
If Auro is actively working on this, then let's have them do it. But if we're so So I guess the question is, is how quickly do things get incorporated? Yeah, that I don't know. I don't know how to first entirely Yeah, so we should just have even just a an email check in with. Okay, Chris, just to make sure that we're not that we're aligned with what their expectations are because they're in the middle of understanding incorporating reviewing Kenya, okay, are going to start incorporating it in the next couple of weeks. It would be it doesn't make sense for us to start the day the experts in for this and we should just create the issue in cross posted in both probably, and then they'll get to it in short order. It sounds like we should, at the very least posted in our own is the answer. So no matter what. So the first thing to do would be to create those issues. On our side. And then when, when it is incorporated in our row, then we cross post any that haven't been resolved, is probably the right way to do it. So we're just going to track them in our repo. We'll keep making changes in our repo if we need to. And Bill keep in you know, fixing, fixing, fixing issues with Well, we've done and then when they pull everything over then we won't be making any more changes on our side. So we'll post over everything that's still open. Okay, so just to confirm that I've got this correct. We're gonna post these issues in our repo.
Then I will also check in with Chris and even potentially, Melanie might know more about how all this works too. So email check in and then I guess once and just get a sense for timelines like when when would they when would they expect to pull it into aro like, Is this in a matter of weeks, or is this in a matter of months, and that'll really dictate what we do.
Okay, you can let them know about our steering committee deadline. Yeah. Oh, yeah, I definitely will. I guess I'll give them a sense of kind of where we're at and where we're trying to get to, and hopefully something will align in there. If that doesn't work for them, then I guess we can go back to the original planners who just didn't write. Okay, so it's just a question for them. Like, do if they're moving quick. They can incorporate these. You if they're not, then we'll do it and that'll be more stuff that they have to clean up before they pull it into arrow. Sorry.
Okay, that sounds good. All righty. Um, great. Moving on.
Oh, okay. Use cases into the documentation.
I did a pass on cleaning up the use cases. When people have the chance they can go through and read it. I hope it is. So there are two things one, in the introduction, I reworked to kind of example use cases and I put tried to articulate the primary use case of both here and actually at the very, very beginning. And then two I pulled out the the well yeah, no, I guess this is really the main the main part of the point. I mean, they're basically the same as what was there. It's just cleaning them up and trying to make sure that they're consistent

with what we've actually done. Okay, no, that's perfect. I will change that one to yellow for now. Great. And if upon reading it, anyone feels like it doesn't do a good job that please just let me know or make the change. Make a PR.
No, that's good. And then there was also the comment about kind of having that example continuity throughout the documentation. That's not there. Done.
That was a nice to have.
Okay, yeah, true. Um, so I'll leave the to do piece there then and just put continuity question if we get to it great. If not, at least the use cases are very good there.
This one I have is read but it's not really a huge thing. It was just please provide hyperlinks so I figured when the documentation is all done, we can just check everything that's linked should be linked.
So that's not a big one. I even I could do that one at the end.
That was good.
Oh, so improve the documentation. So it's clear how to link to Phenopckets and fire. Jewels is taking care of that Phenopckets piece so that's good for fire what do we have in there right now?
Um, I don't know. Is there anything additional we need to add here about kind of this HL seven fire implementation guide piece? I
mean, what else would someone a developer minded person or or otherwise coming in? Who wants to understand what this looks like? Like how to actually use it in the context of fire? Is there is there more more stuff here that they would want?
So I am assuming that person is already experienced in fire or implementation guide, How to Read implementation guide at least so I think it is pretty sufficient.
Okay, and we do say You know, if you want to learn more, yeah, Firebase Yes.
Okay, okay, that's the final to do. We'll be making sure that any of these last round of changes for additional fields that we've added here or there are fully reflected in that implementation guide.
Is it I don't think so, which they haven't been yet. Right. I think this draft this, this IG is from a brutal end of April. And I think there have been a few small changes that we've made to a few of the fields, your current like the the population descriptors and Yes, exactly, yes. Unless you are very, very quick on updating the chromosomal carrier.
Can I ask a question about that, please? Yeah.
Has actually are you could you go back and click on that link that was there for this
video, yeah. Okay. So this has not been validated. Is that correct? No, not yet. So I think that it's important for you to say that in your text, because it's not. It'd be like throwing something onto the GA4GH website that really nobody in G GA4GH has looked at or approved.
You guys have built this.
It's sitting in its own world. I'm not aware of it. It doesn't follow any of the processes that we use to get official HL seven fire artifacts built.
Okay.
Okay, yeah, we can certainly make that clear. It's not sitting on any HL seven space, though, right like this.
Well, it's masquerading I don't mean this in a nasty way. But it's using the standard.
It looks like a fire implementation guide because somebody created it using the standard template. So it looks like an official fire implementation guide. And, you know, I forgot when it went to the top, which is where it would say anything about its maturity.
Okay, it's a local development bill, which is appropriate, meaning it's nowhere right? It's not really a valid HL seven artifact yet.
Okay, that should be easy enough to do to add, you know, a sentence or two that says it's not hasn't been validated yet. That's yeah. And then I got other questions about that at another time, but about how to proceed with this. And grant on. I would expect that something I would hope that something you've been thinking about or looking at, but definitely when we got to that point, yes. Yeah.

So this is still under development. So Right. And that's my point. I mean, I think we probably should say under development or something like that in your in your words. Okay, next. Okay. Thank you. So I just I changed the to do on this row to clarify the text that is not been validated.
Grant Oh, that's about the classes given that you have more context on the HL seven validation slash development process. Would you be up for putting together a couple of sentences that we can plopped into the docs that have the appropriate amount of hedging or caveats with the right language? If you like, what Yeah, I can certainly do that there. Is any specific, you know, HL seven required language here, you know, we have to follow but yeah, I'd be happy to do that. Okay. That'd be great. Thank you. I want to make sure that it's, it's clear to all communities right. Okay. Um, and we will we probably should put some description next to the link also, right.
onto this page. You mean? No, the the the page that has the link to oh, that's, yeah, that's what it was. here and we would put the sentences probably Yeah, right. Yeah. Right. Okay. Yeah, I'll do that.
All right.
Okay, skipping on down to Okay. So the key topic sex field. So we talked last meeting about adding it to the common data set, which I think we've done updating the IG in the documentation accordingly, and then provide implementation guidance or documentation on how to represent that. So this is another documentation piece. Have we tackled that yet?
I added it here. I didn't go out and add it to any of the other link situations. And I don't see. And so it's it's not in the it's there. It isn't there or in the IG or in the UK stock yet. Oh, actually, I think it is in the use case, but it's not in the IP or in that diagram. I mean, do you have a nice way of regenerating that diagram? Oh, yeah, sure. Or do you is do you have like this training manual process, but I can do that. Okay, that would be do you have like a source file that's downloaded and uploaded somewhere that we could save to make it simpler? Yeah. Let me share is draw that IO on Oh, yeah. So if you, that'd be great. So if you can even just export the draw.io.
Then we can commit it to source code. And then we can kind of iterate on it.
Sure.
Okay. Oh, that'd be great. Thank you this. Yeah, because I actually on that topic, because I also noticed, although we've got population descriptors down here, it still says race, race, ethnicity ancestry in the diagram. So that'll need to be updated as well.
Okay, so Okay, so we've got it in common. Jade says in the documentation, but not the IG Correct.
To IG last diagram, those are the only two things correct to the IG and the class diagram. Those are the two spots it's missing.
And population descriptor probably.
Yes. Okay. So that's the next row here. Oh, yeah, update the concept map. Because although it's in the text, it's still says RA and that and those are that's not in the IG either. Right?
Not yet. Not yet. Okay.
And then, oh, so, this was a question I had. We had talked about adding, like a, potentially a second free text option to that are we still planning on doing that or have we abandoned that idea?
Because I remember when we yep, that's a question for the IGs because it the conceptual model is a little is generally pretty light on how defined they are.
We I think we say that it should be a free text option. But what that functionally means is that we are going to have to figure out how to support that in fire in the fire ag and then in the unit packet. And make sure that they connect.
Okay.
I mean, like, I don't think it's, I don't think we have to do it. To be honest, it's more of whether we want to or not, because I mean, now that we've changed the field to population descriptor.
I guess it would be if there was, you know, someone's trying to use I don't put an ontology in there or something and it doesn't have a term they want to use. So they want you maybe use the free text to fill in more detailed information. I don't know.
So yeah, are we adding that option or No? Should I delete this sentence is basically what I'm wondering.
The make sure there's a free text option.

I would just delete that. We, the conceptual model, has actually does say that free text must be supported. So we've got checked on that. And so now we have to make sure that the fire IG and the SR packet implementation guides both support that.
Okay, that's perfect.
Okay, I've updated that. Great. And I think the last thing we have, and this was again, just a question from me. So it was this was a comment based on the common data set.
And, you know, if you were looking down the columns we had, you know, the pedigree IG and I think we had fire genetic pedigree in there as well. And then across the rows, we had the rows for race, ethnicity and ancestry, we have put not applicable, and I think one of the PRC members interpreted as like, we were saying that that couldn't be represented fire, and I think can just not in that specific profile that we have listed, I think so basically, I wanted to confirm that the text I have written there is correct.
It can be represented in fire however the common data set tables only representing the fields in fire genetic pedigree not.
Does that make sense? What I was trying to explain there
though those specific fields can't natively be represented in the fire genetic pedigree itself.
Yes, and I think we updated I mean
it's a bit of a because on one hand
what we're really saying is fire the fire core, and basically, only fire Core Plus genetic pedigree, right?
Because I think other sections in there do reference things in the fire core, like outside of just that one profile. Like patient dot gender, for example, or patient.id. But what we're saying here is that these are not also not natively. Yes, they're not natively supported. There. Okay. Everything can be represented in fire. So it's a little bit of a delicate Okay, so if they're happy with it, then I can just change it to yellow we're done. And Arthur, oh, Arthur's gone now. Ah, okay.
I was gonna put him on the spot and be like either. Does this make sense given the I'll tell you why. When we're done here, I'll tell you why he was had joined for a few minutes, but this person Yeah, well, but I guess for you, do you have any comments about the right way to toe the line here with what we're trying to say? Because you're in the best position, I think to speak to well, it's what we have here is true.
You know, it's not represented in phylogenetic pedigree.
Okay. All right.
That's good. Well, I'll mark that as result.
So that is that's everything.
So going back to our agenda here, so I guess the my follow up question is we had talked about you're kind of using Nov to do some outreach to, you know, some potential implementers or current implementers. And, you know, have them check the documentation. Make sure everything makes sense. You know, if they've never seen the documentation before, maybe get another set of eyes on it.
When would we like to start that? Because it sounds like there's still a couple of little things here or there. We need to fix in the documentation.
So what what would be a reasonable time to start that then? If we still want to make those couple of changes? Do we want to wait for the thinner packet stuff that Jules is doing before sending it around? Or?
Yeah.
I mean, I, it's a good question. I lean towards maybe waiting those two weeks and then we have a more complete package to send, rather than them coming back and be like this. There's not enough information or whatever it might be represented or whatever else. Yeah. Yeah. Okay. Then okay. We should use them. Let's wait those two weeks and use that to make sure that the fire fire piece is IG is updated and referenced properly and things like that before we Okay, perfect.
So for do we want to meet in a couple of weeks, then where are we now? 31 to 24th is Thanksgiving that's not ideal.
On the 17th
Well, so, potentially, I gave I asked Jules to hand it to us by the 18th. So that's why there's a potential if just in case you waited till the very end of the deadline there. We might miss it.

Actually, for the day anyway, we don't have to do it Thursday. No, we don't. I guess it depends on you know, if we change the day will that you know, disrupt anyone's calendars like if we did earlier that week following Well, it seems from the people that have joined today that this final detailed stuff is not a big exciting thrill ride for the rest of the group, which is fair.
Okay, so why don't I mean, there's a Clin/Pheno leadership meeting scheduled.
On Monday, the 21st at our usual time, we could use like, I can even schedule the half hour after that if people are free, since we've kind of already got part of that block scheduled.
Yeah.
What works for me and would you have that half hour there? So 430 to five on Monday, the 21st. Monday, the 21st. Want me?
One second? Sure. Yeah. And then it's a two I don't know if you were planning on coming to that as well.
That half hour block right there.
The alternate should be available. We can just send it around that people are happy for it. If we can if we don't.
Yeah, I mean, I can put it in the calendar. If we decide. We don't want to use that half hour. That's totally fine, but at least we've got something in there. I agree.
Okay, I'm just gonna, I'll get that done after.
Okay. So, with a couple of minutes we have left grant, did you want to You said you want to talk? I think I know what you're going to talk about a little bit because Bob had sent me a message as well. But take it away. Well, yeah, I hate to to do this because my brain is just thinking of different ideas of where what we might do next, and one of them is with genomics.
And you were thinking of something else. I was thinking of something else. But yeah, so So I was gonna have Arthur, you know, if there's going to be more people on the call that weren't aware of genome x, just do a very quick description of what it is but then I was going to then present some of my ideas and that would be that genome X could put as part of the different projects that they are lining up to do. I could envision a project that combines genomics and family health history we have in our common data set document in one in the appendix, where we have the HL seven variant reporting, pharmacogenomic reporting, somatic reporting, etc. That's in there because there's this idea of a you know, a future use case where we take the pedigree model that we're working on and people create a family health history tool. And not only do they want to collect that data for for one particular proband and send it somewhere. They also want to attach some variant data or you know, some pharmaco genomic data or data to that family health history report that we're building. And so I wanted to talk with not only our group but then also Arthur, as that might be of interest for genome x because what we want to do is genome X is supposed to bring in people who are not working in HL seven who are not working at GA4GH who are not doing standards. There really are people who are building products right and they want the industry to adopt quickly adopt differences standards, including fire and GA4GH. And under g genome access where this could happen, that's the umbrella where this can happen. And so I was just gonna raise the idea that how about under genome x, we create some sort of a project, you know, get a lot of people interested, including those who are actually building products on the commercial side, where we can we can start adding genetic and genomic data to family health history reports. Anyway, so that was the reason why I invited Arthur. I think he just had to leave quicker than then he had time. But anyway.

2022-10-20 - Pedigree Subgroup

Chair: Orion Buske, Grant Wood

Global Alliance for Genomics & Health

Attendees "Name (Project Affiliation)": Meen Chul Kim (CHOP/KF), Chen Chen (GA4GH), David Hansen (CSIRO/AGHA), Alex Tsai (GA4GH), Zornitza Stark (AGHA)

	Actions Arising	Assigned To	Deadline
1.	Aim for the green highlighted to-dos below to completed **by the end of October**	ALL	Oct. 31
2.	Schedule a follow up meeting for Nov. 3rd	Lindsay	DONE
3.	Follow up with Jules re: instructions for Phenopackets implementation	Lindsay	
4.	Revise the ecosystem diagram	Lindsay	
5.	Implementer Outreach - aim for November	Lindsay	
6.	Update Pedigree tracking spreadsheet	Lindsay	

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH Connect: Call for Workshops closes Oct. 21 (tomorrow)	Orion, Lindsay	5 min
2.	GA4GH Website Question - do we want the pedigree ecosystem diagram (see here) to be our main graphic?	Lindsay	5 min
3.	Review of Timelines to Approval - The next GA4GH SC meeting is TBD but will be in January. This gives us ~ 2.5 months.	Lindsay	2 min
4.	Update - Moving KIN to the Relations Ontology	Lindsay, Orion	5 min
5.	Remaining To-Dos - See PRC Comment Tracker - Documentation Updates, including: - Revised examples - Addition of content around scope, rationale, use cases, etc. - Clarify Phenopackets & FHIR linkages - ~~Harmonized classes and model page~~ - ~~Remove concept of minimum form~~	All	40 min

	- Add karyotypicSex, ~~rename REA to "population descriptor"~~
	  - Documentation complete for REA
	- KIN Ontology updates (comments from Open for Public Comment)
	  - Two different valuesets to incorporate - GEL and FHIR
	- FHIR IG Updates
	  - Karyotypic sex
	- Phenopackets integration
	  - Delete phenopacket-pedigree repo
	  - Need a set of instructions for using the standard in the context of phenopackets - Lindsay to reach out to Jules
	- Outreach to implementers - draft email here
	  - Send this later, set deadline for ourselves first
	- Pedigree-tools
	  - Need to tweak the standard first
	- IVF Examples
	  - For KIN Ontology integration, list of suggestions for RO		

6.	A.O.B.	All	3 min

Minutes:

Pedigree Ecosystem Diagram
Include risk models in the ecosystem diagram.
How to represent the GA4GH Pedigree Format circle, when we're moving away from the standalone model. Soften to 'conceptual model'.

AT: Having a core Pedigree model for the tools would be the best way moving forward. Wouldn't be able to expand in the future otherwise.
OB: Do we allow exporting in the format other than those two.
Is it useful to distinguish the types tools - patient/FHH collection, vs management. Document could be improved by distinguishing these.
Move the arrows to the implementations.
Emphasize the analysis tool side to highlight genomic use cases.

Most important is the documentation. Tooling needs to be consistent and run locally. Phenopackets implementation in the docs. **Aim for end of October with the green highlights - solicit feedback throughout November.** While we solicit feedback on the other pieces, like FHIR IG and KIN.

Next meeting November 3rd.

Transcript (generated from Otter.Ai):

LS: The main thing to do today is regroup from our last meeting and confirm all of the main things we need to do over the next month. Few months or so until the next GA4GH steering committee approval. So what I wanted to first talk about in terms of housekeeping, GA4GH Connect, which as a reminder is our virtual working session meeting, is coming up on ONovember 14 to 16th as well as November 30. That extra date was just added because for a few contibutors, the 14-16 didn't work. The call for workshops closes tomorrow. I wanted to get a sense of whether we wanted to submit anything for workshop. In my opinion, I think right now we're really going to be focused on getting things done that we need to do for approval. So I don't know if there's anything sort of connect cross work stream stuff we need to do at this time, but I wanted to throw it out there in case anyone had any strong opinions or thoughts on this.
OB: I don't have strong opinions. I'm fine to skip Connect. Unless someone feels like there's something that we should do during it.
LS: Perfect. Okay.

LS: All right. Next thing before we dive into some of the approval timelines is related to the new GA4GH website. I had received a question from our comms team about our ecosystem diagram that's currently living on the introduction page of our documentation. They were wondering if we wanted this to be the main image for the pedigree page. Eventually, when we're approved, they will create something more kind of like the cartoony style looking ones that they do for phenopackets. I wanted to, again, kind of open to the group one, are you okay with that? And two, is there anything we need to modify on this before it goes live?
OB: I have three questions for people about it. One is, so I was digging into that earlier today. To one would be does it make sense for us to add any other analysis tools that aren't yet part of the GA4GH quote unquote ecosystem but might make sense or do we ask people whether they want to be looped in like, some of the obvious ones? More like hereditary cancer risk side, like can risk or something like that Boadicea, one of those tools that I think would be useful to show here from the perspective of the user in terms of thinking about the ecosystem.
LS: I wouldn't want to, I guess specifically name any group or thing if we haven't reached out or brought them in? If we could just put something kind of generically labeled in there, like risk models, that would apply to them.
OB: I think risk models might be a good way of dropping the hint of what those capabilities might be.
LS: Will forward edits to comms.
OB: The second question is, we highlight the pedigree format, pretty front and center. How do we feel about this? Given that we then mostly in the text hedge away from that as much as humanly possible?
LS: relabel it from pedigree format to conceptual model, and we can make the circle small. I don't know what do you think?
OB: I don't know at this point. The most important, from my perspective, the most important things that I think that we need to do before we submit are going to be getting the tooling. Like there's some documentation stuff, and then there's getting the tooling up to date with an appropriate version of the fire and fina pack and implementations in order to do that. If I recall, right, that that when we get into some of the as basically if we're going to be building an adapter between these two different formats, a phenopacket version and a fire version, the code for the pedigree tooling is going to need its own internal format in order to support the switch between the two, at the end of the day, right? Or it's just going to use one of the two, it's just going to use fina packet and then we run with it, what is your thought on that? Because it to some extent, does impact this because then we could just export from that internal format to and then when we talk about it, downplay that component, but it's at least not inaccurate to say that there is this capability that makes sense.

AT: Having a pedigree own structure to store certain fields. I think it would be the best way going forward. That way, like if we want to incorporate something else we could do that. Whereas if we were only to support the specific phenopacket/FHIR stuff, we wouldn't be able to expand in the future.
OB: yeah, I guess the question is do we allow exporting these data in any format other than those two? Is there an intro or are we just opening up this whole can of worms.

Unknown Speaker 13:40
so. Yeah. What the tooling would need to be able to do is basically say, you the tooling would need to be able to say import fire export fina packet, or import fino packet export fire or import from Boadicea ped for gedcom. Output, fire or import one of those three and outputs, output SR packet. I think those are the set of conversions that are needed, and it can either be done as importing any of the five and outputting, a common format and then inputting that common format and outputting whichever you know, at least the fire in the pheno packet, but potentially some of the other formats as well. So it can go through a central place or it can that doesn't need to be exposed in the way that we are, but it may make some things simpler. It's just about right now that whole right edge of ingesting it for analysis tools and things. It's what it really is, is those all get ingested from the implementations above. Right. Yeah. And that's really, so if that, especially the circle and the arrow to the right, is is where I think the most confusion is happening. Or would happen from a user's perspective, because that isn't really what is happening in the way that we've set things up. Right? Now. It's,

In theory the problem could be solved by having if that tooling that they have is able to support the what we're proposing in terms of the implementation guide in terms of fire and the improvements to the pedigree representation in the context of Fina packet, which I'm not sure it would, but if it did support that, then they would have that conversion already there. If there, if the scope of what they're talking about include the pedigree data in in this relational way then that would solve that problem. We wouldn't really, but if what we're thinking about is saying there is pedigree data that's coming in from these different formats. I guess maybe then the most important part of the tooling is supporting the conversion from existing structures like gedcom and ped and Boadicea into a GA4GH compliant version implementation like the phenol packet one that could be enough and then there's this other tooling that can get you from packet to fire. I knew that the lift of sort of natively supporting fire might be a little bit high, or to support that in terms of the tooling.

OB: 3rd question. This one's easier. I wonder whether it is useful to distinguish the two different types of tools, buckets of tools that we have on the left, especially if as we add more, which is that you have fundamentally like patient family health history, data collection, versus pedigree management. I think one of the ways that the documentation could probably be improved is by clarifying the distinction between what we talked about when we say a family history, or a family health history and what we talk about what we're saying when we talk about a pedigree. And that in the context of this standard, we're actually trying part of the challenges that we're trying to be ambitious and cover both of those. If we were just covering your regular genomic pedigree you might be doing then it for some use cases of ped is enough. But if you're incorporating family history and the consumers of family health history information, in a pedigree sense, which is where risk models and things come in, then you need a more comprehensive format. So that's the other idea that I had for this, but those are the three.
DH: we don't want to just make it look like an exchange pedigree exchange format, you know what I mean? So emphasizing maybe more on the analysis tool side, or analysis side. expanding it into being a little bit stronger in terms of highlighting the genomic use cases.

Lindsay to make the diagram edits.

LS: I just wanted to do a quick recap of the timelines that we have just so everyone's on the same page. So the next GA4GH steering committee hasn't been scheduled in the calendar yet, but it will be January of that we know. But the exact date, it's kind of up in the air. But roughly this gives us all of November, all of December. A little bit left of October and presumably a little bit of January at the end there. However, I realized this gets into

holiday season near the end, so let's assume we've only got about the two months up until Christmas and we can try to finish by then. So roughly seven weeks or so.

LS: And then the one last thing, I wanted to give a quick update for anyone who hasn't heard of this discussion yet about moving the kin ontology into the relations ontology. Had an offer that we could potentially integrate the kin ontology into the relations ontology, they would support resolution of the open for public comment feedback. And from what I understand, it would still be appropriately attributed to GA4GH the work we're doing here. It would essentially almost be like a subcategory in RO. The added bonus of this is that they will maintain it so we don't have to take too much of the maintenance on ourselves. So since that point, there has been an email outreach to determine the next steps of how we might take this forward. I just sent a another ping earlier today about it.
OB: I think the only other change that we need to do is what is the list of critical to dues before we submit, and there are probably a you know, a few things on the ontology side that are on that list, but we'll get into that later.

LS: What I've listed here, I just quickly pulled from that comment tracker, just kind of in a more shorter, digestible way. So if you want to see more detail about all of these, again, just click this link here. So one of the big things I think we need to do and this comes largely from the PRC is all of the outstanding documentation updates we need to do. So this includes: we talked about doing some revised examples that kind of follow that go through the whole documentation. Related to that, we talked about addition of some content around the scope, the rationale and the use cases. This was suggested from the PRC, when they were kind of trying to understand the differences of, you know, like, what does the kin ontology do that the FHIR family member history value set cannot. What was the rationale of creating some of these things, which I think will also be really helpful for users and implementers when they're, you know, learning about the standard and wanting to understand some of the choices we made at the modeling. There was also a request just to clarify some of the fina packets and fire linkages. So I think this this is referring to right now in the documentation. I think we've only got maybe like a paragraph that kind of explains the Phenopckets integration right now. So I think we just maybe need to expand that a little bit more.

OB: Minimum form? We tweaked it. There's like, whatever the minimum is, and then there's probably more that can be done. I had gone ahead and adjusted the way in which it was talked about in order to meet the minimum needs of that PRC comment. And I'm sure that as we work and rework the documentation that will further adapt, but at least at a high level, that should be covered. Okay, excellent. So talks about kind of the choices that a data input that an implementer will have to make about how they want to serialize it and some guidelines for it, but it's no longer called a minimum form and it's now you know, softened significantly.

LS: Perfect. Okay. I'll add that response in the spreadsheet as well.

KaryotypicSex/REA Edirs
OB: We did make the change, at least in terms of the documentation, I think of updating it to be called Population descriptor.

How strong was the comments? I'm just noting, because I know there were some great updated guidelines around pedigree nomenclature and sex and gender representation and things like that. We've been digging into this a whole bunch on the phenotype side recently, too. And it's, I mean, it could be part of this

Zornitza - Can I put you on the spot for a quick sec? Do you have any thoughts about the karyotypic sex comment, Have you ever encountered this? We're talking about the karyotype. XX and how much how often that is necessary from a pedigree representation perspective explicitly, as distinct from sex assigned at birth and gender.

ZS: Very occasionally. And, and I guess, we often find that out from the genomic data itself. Sometimes the referrer does know that there's a discrepancy. We do have a pathway for disorders of six differentiation. But it's a rare event. It's a small subset of patients that we test, but it's definitely a group of patients that we test.

OB: In terms of the representation of it. I could imagine there being two different approaches. For saying that we represent this. One would be saying this is less a property of the individual and therefore doesn't belong there, but is more something that would be represented through a disorder coding system like yeah, that's one of the which is already supported by fire and, and fina packet and we would just say, in these cases, here's how you represent it, which we could do. The alternative would be to say, we'll just make an explicit karyotype xx field it's a string, just text. And unless there's a more structured way which packet might actually have and could point us to, do you have a sense of which of those two routes might make the most sense to people who are looking at this?
ZS: I guess it is helpful for it to fade into genomic testing system as an explicit thing, because that it's an unpleasant surprise for us to find that the karyotypic sex is different to what the phenotypic sex was stated as usually causes a lot of panic.
OB: Okay, so in other words, having it be explicit as part of the pedigree individual, if known, may be a better way of making it apparent rather than just having it like nested within some clinical diagnosis somewhere.

Meen any concerns with us adding that and implemented in that way (same as Phenopackets)?
MCK: Not on my end.

KIN Ontology
Include Super set of the fire value set and the GEL list. And there's not really a reason it shouldn't be included based on the way that things have evolved and we made some original design decisions early on, that would have been a justification for it not to cover everything in the fire value set, but I don't think that applies anymore.
** possibly pass on to RO folks

Phenopackest
So there we need to delete the phenopacket pedigree repo that isn't needed anymore. And then basically I think we just need a set of instructions at the end of the day. For “I am a developer who is trying to use this. And I would like to use it in the context of a Phenopacket”. Explain it to me like on five sort of an approach.

And the problem like it will change as it gets incorporated or not, or like the relationship between what we've done and phenopacket evolves over time, but it has this certain relationship at this point in time which is I think that it is implemented on a branch and in a way that can be used in the context of phenopacket which is fine. We just need to tell a human who is trying to use it, what that means about how to actually use it. And I have a very vague, cloudy brain foggy version of a sense of what that might look like. But I think Jules is the person who really practically knows like, this is what you would do to do it.

LS: I'll make that an action item. For myself to reach out to Jules.

2022-08-23 - Computable Cohort Representation - Summary Statistics

The minutes for Computable Cohort Representation Summary Statistics meeting can be found here.

Global Alliance for Genomics & Health

2022-08-11: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Meen Chul Kim (CHOP/KF), Melissa Cook (NCI), Grant Wood, Amanda Spurdle (BRCA), Alex Tsai (GA4GH), Hannah Calkins, Chen Chen (GA4GH), David Hansen (CSIRO/AGHA)

	Actions Arising	Assigned To	Deadline
1.	See Comment Tracker for the most up-to-date listing of To Dos	ALL	
2.	Reach out to Hoa re: updating KIN Ontology	Lindsay	
3.			
4.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH 10th Plenary - GA4GH Code of Conduct - Open for Public Comment until September 8	Orion, Grant Lindsay	5 min

2.	Continued PRC Comment Review and Planning
	- PRC Comment Tracker
	- Documentation in Google Doc Form
	- Update on REWS/Security Questionnaire
	- REA Feedback (see below)
	- Asking about a standalone serialization in the GA4GH DaMaSC Group
	  - Recording (starts at roughly 23:00)
	  - Passcode: LFK9bM.M
		ALL	50 min

3.	A.O.B.	All	5 min

REA Feedback from REWS contributors:
This is a super important question. The answer is that this data structure should be able to hold as much data as is available for the individual when it comes to population descriptors.

The simplest would be a field "Population descriptor(s)". And then explain that this is to capture the info available in as much detail as possible about ethnicity, ancestry, race, tribe etc..., e.g. Don't put "Asian" put "Indian from Delhi" etc.

There are many other things that could be potentially useful, e.g. language spoken, genetically inferred ancestry, where the sample was taken.

[Do you think it would be important to distinguish whether the information was self-reported (vs. clinician or someone else reporting it)?
Yes, ideally! Maybe this could come in brackets. i.e. the description could read something like: "Capture **as much detail as possible** about the individual's ethnicity, ancestry, race, tribe etc, **as well as indicating what information that is based on** (self report, genetic inference, administratively assigned, language spoken, geographical area) e.g. Don't put "Asian" put "Jat Sikh (self report)" etc.

If the field needed to include as much information as possible, would you agree that it would be best to make the field free text rather than trying to adhere to ontology terms or other structured format? The downside to this is that it reduces computability if users can throw in anything they'd like. I suppose later down the road we could update the field constraints if someone decides to take on the herculean task of creating a truly globally representative ancestry/ethnicity ontology (Hancestro is currently abysmal for African populations).
Definitely free text, for two main reasons. a) If it becomes important to process this data later, useful information will have been lost, and b) if we can avoid putting people into boxes, we should -- humans don't come in distinct types.

I've just been taking another look at the paper that intros Hancestro. They go to quite extreme measures to go from starting data to one of their categories, pasted below. I hope this illustrates that getting to a nicely harmonized dataset is a huge undertaking, and that people who are accessioning samples should not be expected to do this. What enables others to do this at a later point is information that is as detailed as possible (now, whether others should even attempt to do this harmonization is another question...). So I would vote not to even mention Hancestro, and perhaps give more examples in the definition.

FYI, When it comes to variant interpretation, I've heard that a main use case to get population descriptor data is to identify whether two reports of a variant actually come from one and the same individual or not.

Hancestro excerpt:
"We relied heavily on data stated by authors in the GWAS publication, giving precedence to information inferred using genomic methods, such as principal component analysis (PCA; see Additional file 1: Box S1 for a list of methods commonly used to ascertain ancestry). In some cases, we considered other sources, but only when the information provided by authors was limited or ambiguous. We consulted peer-reviewed population genetics publications to obtain additional information on lesser-known groups that were not adequately characterized by authors or when samples were described using ethno-cultural terms (for example, "Punjabi Sikh"). When the only information provided in the publication was the location of recruitment, we consulted The United Nations M49 Standard of Geographic Regions [12] and The World Factbook [13]. The latter is a regularly updated compendium of worldwide demographic data, covering all countries and territories of the world. Additional file 2: Table S1 provides a list of countries of recruitment in the Catalog, together with the sources that were consulted and the inferred categories.

In rare instances, the ancestry information provided by authors was not detailed enough to allow the resolution of samples into ancestrally distinct sets. For these samples, we created complex, multi-ancestry detailed descriptions and selected multiple categories (for example, Catalog entry for Jiang R et al. [14, 15]). For admixed samples, we selected either one of the categories that includes individuals with well-defined admixture ("African American or Afro-Caribbean" and "Hispanic or Latin American") or the category "other admixed ancestry". We also captured additional information to describe the ancestral backgrounds that

contribute to the admixture. No ancestry-informative detailed descriptions were generated in the absence of ancestry or recruitment data; for those samples, the category "Not Reported" was selected.

Where possible we also curated country of recruitment (Fig. 1; Additional file 1: Figure S2b) and country of origin as this provides additional and complementary demographic information. Country of origin was extracted when the country of origin of the study participant's grandparents was stated or when the genealogy of the sample could be traced to a particular country.

The detailed extraction guidelines utilized by Catalog curators are included in Additional file 1: Supplementary Methods. A full list of Catalog detailed descriptions and categories is provided in Additional file 3: Table S2. Examples that illustrate application to specific samples can be found in Additional file 4: Table S3. All curated ancestry data are available from the GWAS Catalog website [4] (Fig. 1) and via download [16]."

Minutes:
The transcript of the meeting can be viewed here.

REA Field:
GW: Should we give this to the implementer, they can decide what ontology they want to use?
AS: Good idea. If people don't know what to use we can provide recommendations.
GW: We don't decide the ontology or term.
OB: Be clear in the implementation that when something is codable concept or ontology, there is some fall back, like having free text. Question of whether we need to include provenance, like self-reported information. Make sure the standard reports free text. In Phenopackets, there is no fall back to free text, in FHIR there is.
To re-name field to "Population descriptor" and make sure there's a free text option.

New KIN Ontology Terms:
Lindsay to email Hoa re: term additions.

Examples & Use Case
OB: Would it be useful to choose some, represent in Phenopackets and FHIR. Or choose one and go into depth.
Say we choose the first one, let's come up with a pedigree for a trio or quad and demonstrate how that works in the context of Phenopacket.
OB: I think we should choose a story that highlights a few, like a patient is seen, information is sent for testing, then later a second family member comes in, etc.
GW: Choose an example that shows more than one condition.

Karyotypic sex
Add field to the common dataset, update the IG, and the documentation

2022-08-04: Pedigree Subgroup

Chair: Orion Buske, Lindsay Smith
Attendees "Name (Project Affiliation)": Zornitza Stark, Meen Chul Kim, David Hansen, Chen Chen, Amanda Spurdle

Global Alliance
for Genomics & Health

	Actions Arising	Assigned To	Deadline
1.	Check Grant's availability for next meeting	Lindsay	
2.	Move the readthedocs content into a Google doc	Lindsay	
3.	See Comment Tracker for actions!	ALL	
4.	Start to harmonize the PRC responses and tag people where needed for comments	Lindsay	

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH 10th Plenary - GA4GH Code of Conduct - Open for Public Comment until September 8	Orion, Lindsay	5 min
2.	PRC Comment Review and Planning - PRC Comment Tracker	ALL	50 min
3.	A.O.B.	All	5 min

Minutes:

LS: I think one of the challenges and I think one of the reasons we're also going to be updating the GA4GH approval process, is that a lot of the questions they had were the kind of things that we've already discussed and made decisions on and now we're having to reexplain those decisions that we've made. Because we've had these discussions they should hopefully be easy comments for us to answer. We have to give the PRC a little bit more context and background to help them with some of their review. One of the big questions that we had and I highlighted it in red, was just about two things. One was this question of the standalone-ness of it,. And then from that there was also the question about FHIR and the family member history resource. I think they wanted a sense of why the new implementation guide what does this do that the other things can't. I think we can pull a lot of that from our use cases and even some of the folks that are using the FHIR implementation guide to help us explain this. There was a related question about the kin ontology about the FHIR valueset - what can KIN do that the valueset cannot. So a lot of it was just kind of wanting to know wanting to better understand some of the choices we made.
The sooner we can get back answers to some of the questions like the rationale of why we made certain decisions, I think the sooner we can get the PRC on the same page. There was a question from the PRC about the strategy for GA4GH and HL7 interactions. I flagged that as something that I'll just take care of, we'll work on this, you know, kind of more from a GA4GH high level point, so don't worry about that for us.

LS: I think it's going to be part of my job to remind the PRC that that's not really the type of questions they should be asking, because that should already be something we've solved - is this needed? Why did we do this? We've done that work already. And that question should really be answered when we first proposed it and

showed all these groups that want to use it right. It's my job to show the PRC they should be asking: does this do what it needs to do? Not necessarily the why.

So first comment here about how the pedigree is computable and how comprehensively it is able to represent any pedigree, and to what extent it's able to represent various disease situations such as risk factors, incomplete penetrance, multiple noncommissioned officers. So I think if there's something in the documentation we're missing that would clarify this. To me this isn't a model issue. This is just a communication issue.
OB: This is a taking of the examples and those example use cases that are outlined at the beginning and following them through to explain how it would be this to maybe give tangible examples that are specific to those use cases later on in the document.
LS: How do we want to address that? Can we work some of these into the examples we have later on?
OB: I think that's probably what we need to do.
LS: The priority question they had - is the pedigree standard expected as a standalone or will it need to go with a further standard that describes clinical and other individual level data? So of course, as we know, we've got FHIR and phenopackets, so maybe we just needed to make that a little bit more clear in the documentation. There was additional comments on the call itself about thinking of GA4GH's goal of interoperability and what it means for someone to might have to switch from one implementation to another to a standalone. There's a question also here about the a family member value set in FHIR. And so what can it or what can't do that our kin and ontology can do? And so a suggested action here is just to explain this a little bit more. If the problem is tooling? That's a separate conversation. So I think again, this is just something that we need to explain.
OB: I'm curious what other people's thoughts are. I mean, I think yes, we've gone down this path of state of not having a separate standalone standard. If they feel strongly that that is not the right choice. It will be helpful to get that feedback. But I mean, ultimately, the choice of presenting it in this way, and maybe that's just what we have to mitigate. People are focused on the translation between a phenopacket versus a FHIR implementation of this rather than in reality if it were a standalone standard, then people are going to be focused on the translation from FHIR to it and from it to phenopacket and from phenopacket to it, etc, - having a third standard doesn't necessarily help either. Or introduces similar challenges around conversion, like having a totally separate standalone standard doesn't mean that there isn't conversion that's necessary then to get into and out of it. But it's a fair point. I mean, if there were four serializations that we were talking about, then it might have made sense at that point for us to have it be standardized separately, and then do the individual mappings to and from it. But I think that's sort of where we landed in terms of it. So it sounds like we just need to do a better job of explaining, clarifying that and the rationale and making sure that there is clear interoperability between those two implementations, which is hard. But not any harder than the interoperability between those two implementations and some third independent standard.

So that's sort of my sense in terms of the first one. In terms of the second, and related to the goal in terms of the FHIR family member history value set. I think there are a few things that we can talk about. One was part of the rationale for starting a new one, there was a subset and actually a different set of terms that were relevant to this use case originally. In terms of having it be, let's say, gender, or sex agnostic, and a few other attributes for the very first version. That philosophy then sort of changed as we have gone through this design implementation process. So we've adopted, I think, a much more inclusive approach in terms of adding terms that are, you know, adding additional terms and then linking them in logically so in the context of the there should probably not be things that are in the FHIR value set and not in ours. It shouldn't be a superset of it. There are already things within our terminology that are important to represent in this case. Around in particular, I think donors that were not at least at the time represented in that value set. The
DH: For me, just looking at the family member value set is that it's a value set, not an ontology. So what I think we set out and you can correct me is that it is to have an ontology with relationships so that we can do that reasoning across it. And that's not what the family value set does. And value sets in FHIR, necessarily part of the FHIR standard because they point to codes from other code systems. You can specify you have a value set

which contains codes from SNOMED. I think we didn't find anything that met that requirement of reasoning across. We did have a demonstration with that, from Alejandro? Yeah. I think that's a really strong argument.
OB: Thank you for bringing that up David. I think the value set is a taxonomy but not so so terms in it have parents and they can have multiple parents. The thing that I found is that it is quite Euro centric, in or at least Anglo, Anglo in its structuring rather than semantic in its structuring. And so for instance, there's a term for aunt and a term for uncle and not one for what that relationship regardless of sex, because English doesn't have a common word for it. And so it's not in the value set. Yeah, from any conceptual model, and you would have a term there.
DH: And I think that's a strong point too, but we have to be careful how we make it, because we could always have asked or written an extension to the family data set to add more codes or put in a request to HL seven to add more codes. Semantically it didn't give us what we needed. And so an example is the aunt and uncle, gender, those lack of relationship so you know what I mean? You should probably point out, Lindsay that, and I don't know how you do this, but, you know, Grant and Alejandro, or Grant for sure, were involved in developing this system.
LS: what about the implementation guide itself versus the family member resource? I mean, I think it might be a similar argument.
OB: Yeah, the critical reason around the family member history is it's all siloed in the context of a proband and like the entire use case for family member history in the context of this is that it's siloed. There, if we're talking about the family member history genetics profile. flashing back to what were the rationales there. Meen is that more top of mind for you? In terms of the reason for some of the gaps around the higher family member history genetics profile,
AS: Are you talking about the fact that if we wanted to do a genetic analysis, we would say the first person identified to have a variant might be a pro band? I think it's that's one of the fundamental limitations of the family member history like a family member history, pro, resource and the way that they behave in the context of fire is that they're reported. They're basically rows in a family member history statement from a person.
OB: and that makes sense but it is also restricted always in the scope of a single patient and their family member history. Yep, it also does not allow you to recenter or basically to do extended relationships. It's always radiating out from one person.

The transcript below is voice to text and is unedited. Some of the discussions may not have translated perfectly.
It comes back to what this is set up to be in health in the in the clinical versus what we're doing in research. This is set out to be if you have a person, what is their family history, and then you have two people who are related. That's a separate question. In terms of genetics, as I said, right from the beginning, the only way to do it properly is to collect a proper pedigree. And then you can change who the proband is, or decide what the ascertainment criteria were. But you were the ones who pulled me up right in the beginning to say the purpose of this was only family history. So I think it's fine. Yeah,
absolutely. And so Thanks, Amanda. And I think I'll, oh,
you can quote me as being a person who was dissatisfied and had to be re satisfied what the purpose of this code was for. For the one in 100 time I join these calls.
Every time you join, it's perfect. So So that's for the so I think that is also true of the family member history genetics profile. That genetics profile allows you to, if I can try to explain my understanding of it in, in non technical languages, with with a family member history, resource, it's all radiating out from one person. And so the farther away you go, the more you lose words to describe the relationship. Like just like as if you were describing relatives of a person. The family member history genetics resource allows you to kind of recenter and describe relationships relative to other people within that family which is great because it allows you to get a larger, it allows you to collect a larger family history for one person. What it again does not allow you to do is to have a joint pedigree for a family
that maybe you should just just use those terms family history and pedigree because that's really the difference family it is, is a description in words. Yes. Whereas a pedigree is describing the relationship of each member

relative to the people around them, because every person has a mother and father. And in fact, if you don't have a mother and father at the top of the pedigree, even if they're not provided to you pedigree analysis doesn't work. You have to make a mother and father you know how often a person will pitch on their own. And they and they will describe them. Well, I don't know if you know that. But we have this problem where we have to add people into false people into pedigree provided for us through analysis to work because everyone has to have a mother and a father with an ID. I.
Think the distinction between family history and pedigree as words is a very important. Thank you for nailing that down as a clear, clear point and we need to drive that home front and center beginning
well as long as you can start working on a real pedigree standard. Maybe that's the problem is that the term is called pedigree standard, not family history standard.
Sorry, I've just worked it up myself because when I came in here, I thought you mean pedigree,
right from the beginning, and I was getting well it's not good enough for me, because I talked about pedigrees. And I think what
we've ended up what we tried to end up with is something that can robustly be used to describe a pedigree, pedigree pedigree, but also in the context of fire or in the context of you know, packet, but also can be used to represent that family to to take the existing family history. Data and get it into that format.
So maybe that maybe the thing to say is remind people that this is if you want to stick with the word pedigree, because I think we'll probably be too late to change it. Now the family history standard is that it's version one and didn't we always say that it would be extended? The plan would be that we would extend it.
Yeah, I mean, I think every GA4GH version one typically is intended to be you know, the minimum viable product, right and then it can be built on to version two, version three, whatever from
and I think and I think it's fair to say my knowledge of what gets collected in the clinical setting, even for me where I demand that a pedigree is the perfect thing a real pedigree, not a family history report, by the way. Is that's how it's given. It's written as aren't affected, age 30, blah, blah, blah. And the other thing about a pedigree and that's an important thing is people only for family history. They generally only tell you about the affected. Family members are first and second degree. They might talk about the kids who are unaffected because they the people at risk, but it's not a full pedigree. It is just what people remember to tell you. And that's important because if for us if you're trying to see the veteran really is causal, if they've got 10 females, we've all reached 60. And they don't have breast cancer and they carry the variant that tells us something that they may not be in the family history report.
So I'm just watching the time to I want to make sure we have time to go through through all the comments we have. Are we happy with kind of at least the direction of response I think we want to give here and what we need to explain with regards to you know why we developed the new implementation guide and why we have the new ontology.
Yep, I think there's enough there. I think to some extent, we have to show that we've really thought about it not just we've done it
without Exactly, exactly, yeah. And this what I've written here doesn't have to be our response. These are just me scribbling notes. Of course, we can obviously make it much more comprehensive. Okay, so let's keep on rolling down here. So the next one was about the direction of relationships. So I'm wondering if the minimum standard form is sufficient to represent a pedigree, excuse me. I would like to think that the condition five and I wasn't sure what that was, but is not able to uniquely draw a pedigree with the kin ontology used. So comment more extensively about what the minimum standard is for. Any thoughts reactions to that one?
We need to know what condition five is.
Yeah, I wasn't sure what that was either.
So that one of the examples, it's under direction of relationships. There's a little bit of a note about some concept of minimum standard form. So what the basic gist here is that and this gets a little bit towards what you were I mean, what we were just talking about, it looks like there was a bit of a difference between restructured text and marked down. So it all got collapsed onto one line. Sorry about that. Um, the the way that we've set up this standard is pedigree slash family history standard is such that it can represent a family history, and but in a computable way that allows you to resolve it into a pedigree and can also describe a

pedigree computationally in the context of because we're describing it as a graph, structure with nodes and edges that are directed edges that are relationships between people that say, this is this person is the mother of this person. There is a unique ID a unique sort of fully fleshed out representation in which every possible relationship between every pair of people is, is is shown as an edge or is communicated as an edge but that's an awful lot and not necessary. Because one person being someone's mother implies that the other is the first one's child. So we only needed to choose one of those two, and the other is logically implied. And so there's a notion that you could have a set of rules that would, you know, guide people towards sort of defining what the minimum standard is here kind of what the minimum amount of information that they need to have to have it be a connected, connected tree or connected structure is connected graph. In this context, there is still I would say, not a unique minimum form. I don't know whether or not that's a problem. I don't think of it as a problem. It would be a problem if we wanted to hash that structure in order to generate a unique ID or something like that, like happens in verse but I don't know if that's a problem in our case. That is a structural problem, like it would it would require a structural change to fix it. In other words, to make it so that there was only a single unique solution and the reason is because of relationships at the same level. So in particular siblings, twins, partners, partners are like, which of the two partners do you put the relation the edge from or two? And because right now, all of those edges are directed you have to choose one. So unless we added to the structure undirected edges, then partnerships are there's always an you there's not a unique option. Sorry if this is way too technical for folks.
But I understand it because I learn every time I come on this call so you talking about say, if it says subs child has cancer age, whatever which said, because the subs Oh no,
it's more like if we say this person is if we say person one is the partner of person two. We could have equally said person two is the partner of person one, but we probably don't specify both. So both of them.
Yeah, well, but if everything's in relation to the program, then surely that's, that's okay. Because then if the proband is the pro band and they are person one, by definition I mean, that's not what we do. We call someone a probe and they have a one next to them. And so everything that would be the probands partner, the probands child, we don't even call about Person Number Two's child. It's all probands child. If they have two relationships, then it would then then you would have to show try and say that it was with you know, marriage one on marriage to I don't know how you the hell you
were. So So I think so. So then An example would be, let's say the proband has some siblings. And let's say those siblings, they're twins. They're a couple of a couple of twins. In the current way that this is structured, it's a graph model. So it's the twin is not a property of the person. It's that it's the relationship that those are huge twins have to each other. And there is an ambiguity about which twin has the relationship to which one, where do you define that in the way that it's structured right now? If it were important, important, and like, the basic answer is that it wasn't. I don't think we thought of it necessarily, or we didn't in terms of the implementation. I don't think have a hard requirement that there be a single unique representation of a given pedigree. If that is the case, then we should make some structural changes to support that, in particular in terms of in terms of relationships, that our
view ambiguous, ambiguous, yeah,
or or collective. Sorry. I don't think it's not symmetric.
I guess it comes back to back to the fact that this is called a pedigree standard and that if the idea was that the family history should be sufficient to make a mini pedigree, either you need to say, make those changes so that you really could be maybe not 100% Sure, but 99% sure you've got the pedigree, right. Or you have to just say this is deficiency.
I think this one, I hear what you're saying. Amanda, I think this is that's part of almost validating, pedigree representation and saying like, turn this into a pedophile or something right like you and we could do that. The what what I think this is getting at is if you take a given pedigree, you can represent it in this pedigree standard in a few equivalent ways. Yeah. And that's where they're saying like, I'm wondering if the minimum standard form is sufficient to represent a pedigree uniquely, and the answer is it isn't. It's enough to represent it more compactly but it

doesn't matter if the information that you would draw from nope, it doesn't it like that's definitely comment to me.
So so as long as we're not saying that it is required to be and then there's the question of what is why even talk about a minimum standard? And so maybe the or a minimum forum and so the answer might be we just removed that, because it's confusing the issue. I think I added it in because I felt like there was it might be useful for some guidance there for tools who are implementing this. Um, I'm in terms of saying well how of the many different ways that they might represent a pedigree using this structure, how to do that, but maybe it's just confusing the the issue.
Yeah, and I guess the only the only the only con of that is that if two people have the same pedigree, by chance and the same bloody program they'd have to have as well. It could look like two different pets like if you matched it on the basis of the family history report. It could look different ways it was the same.
It could look different in terms of the actual
the word, describe it.
Yes, there are lots of different equivalent family histories. And that's the problem. Like or that's the complexity here, right that you be honest
I think that's a minor problem compared to just knowing that it's the same one because you're more likely to have two different probands from two different centers representing the same family. That's that's a, that's a bigger problem that you will not solve in this FH. You'd have to have a pedigree to do that.
And even with the pedigree, you still need identifiers or
we are more Yeah, we did ourselves support
name, you know, enough enough demographic information in order to know
this old variant type. You know, that's what we use. Yeah.
Okay, so I'm just starting to color code some of these comments because they're very overlapping, because I was just looking down here on this one. Yeah, given the lack of unique tools will struggle implementing the standard in the same way. How do you envision that happening, which I think is very much related to what we're talking about now?
Okay, copies are there.
So there's roughly six to 10 plus we have about 11. Through the public comment, although I've realized this one here is the same as one of the ones the PRC gave us. So there are some overlapping pieces here. Which will hopefully make it a little bit easier.. So okay, let me have time we got left about 15 minutes. So do Sorry, I don't know if I captured it. correctly. But are we going to be removing this minimum form from the documentation? Did we come to an agreement on that? And does it sufficiently answered this one as well about implementation?
It's a minimum, it's a smaller form. It's a form that says, you don't need to include all directions of all relationships. You don't need to say that person one is the mother of person two and person two is the child of person one, what is
the term you would use to describe it? If it's not going to be a minimum standard? It's like, you know, it's a reduced form..
It's recommendations for a reduced form is really what it is and we can decide whether or not but the reality is from an implementation perspective, you're not going to want to include the full graph of relationships that doesn't make sense. You're going to want to basically have a pedigree, like what you're saying, Amanda, you're going to want parent child relationships, right? Like you're just that's and the intent of that section was to say basically Hey, you know, when communicating, if you're communicating a pedigree rather than a family history, you should probably communicate it as parent child relationships, if you can reduce it into that for him. That's that's all that section says. But they can because our use cases included family history use cases like risk, cancer risk assessments that are not based on the pedigree but are based on knowing that the proband has an aunt some aunt who tested positive for BRCA that that needs to be representable that was the notion okay.
So again, this is mostly a documentation clarification. Again, I think

this is documentation verification, if we're okay with it not being unique, which unless someone else has something to say about that, like adding that as a constraint will involve, I would say considerable changes at this point. If that's a requirement, then we would want to make a whole bunch of changes, but it doesn't sound like they're asking for that to be a requirement. It just sounds like they're asking for guidance given that and so we can provide that. Okay, um, it's just about adding a little bit to the tooling. To help with reducing that form maybe. Or,
okay. All right, we have a 10 minutes left. So really, briefly to go through some of these last ones here.. And so this one, number six here, actually talked about with the PRC for a little bit, because the suggestion was to adopt kin as part of the standard and CO evolve. I think what was happening was, because kin is in its own separate repo, they saw it as like a separate standard or a separate thing. But I think this is mostly resolved because I just kind of explained like, we do intend, it's part of the pedigree ecosystem, and it's kind of a piece of our standard. So I mean, correct me if I'm wrong, but I think it would co evolve and it's something we would maintain as part of the standard. So I think that was resolved. There's nothing we need to do there. So continuing with the ontology, so this one augusto provided a list of some of the enumerations they have used. So this gets back to the uniqueness point again. So and I think also some of the earlier points about why the ontology and not the the code set. So quite complicated but they need to be specified to have uniqueness here as examples we use either add or abandon expectations of uniqueness and reflect. So I think we abandoned unique this based on our last discussion of Yeah, okay.
So I would say that, given this feedback, what we might want to do is update and expand the kin ontology with either terms that map to these, you know, add these as, let's say synonyms for additional terms that are logically defined in that space.
Yep, that makes sense. Okay, um, okay. Please provide hyperlinks to the terminologies used across the document. Where do we reference the terminologies if any. Because when I first read that I couldn't immediately
we referenced the kinship terminology, I'll be working with the pedigree model, oh, there. If there are other places, we also reference it for the actual property that it's defined under, which is the relation property under the relationship class.
Okay, hancestro. We did leave that one there.
Right, if there other Yeah. So usually, we're doing it in the classes for the property that it was scoped under, but it sounds like they would like it maybe at the top level.
Yeah, I think so. That's what I'm getting
at is provide hyperlinks to the terminologies used across the
whole document. So I'm thinking like, you know, if we look at the pedigree model page, we we say ancestor here and it's certainly do
Oh, yeah, we should probably the pedigree model page probably should be lightened and instead we should we, I thought I had a PR that did that. Um, what what triggers the rebuild of the the republication of the documentation, does that happen automatically,
ah, mean has to rebuild it right.
So, in order to set the automatic build, you should meet it you should add me as an administrator to the repo. So,
the action the pages build deployment action does not auto push to GitHub Pages.
No, so I have to rebuild by clicking a button because I did it that way because I am not an administrator of the repo so I don't have the access to the web hook.
Okay, I can probably do that right now.
Yeah, great. If we could just set it up as an action so that then we can make sure that nothing is
then I can change the activity, so that whatever we merged into master, then we will be rebuilt. Right? Okay. Okay.
Yeah, I will. I'll make an action for myself to add you as a as an admin.
Yeah, sounds good. Okay. And I think yeah, we should just replace on the pedigree model page. Basically just replace things with a definition of that concept, and none of the information about the individual and instead

just point to the classes. So you have to now that the I think we have stuff a little out of sync in two places, so we don't need it written classes is a better representation. Than
Yeah, which I think gets out this next comment pedigree model does not agree with classes,
right? Yeah, yeah, I think
okay, um, so classes I'm struggling with how to link to affected status, particularly for more complicated situations, such as potentially incomplete penetrance. CO segregating conditions affected is zero to one and there is not a disease description class, you clarify. Take us back to the classes page here.
Yeah, so we only have in terms of effected we only can do a single condition that's sort of built in everything else would be expected to be done under fino packet or fire, because
I don't understand how you can how you can represent incomplete penetrance someone's either affected at the time you see them or not the only way to do that would be to build in an age related thing. Because the only way you can see that someone something's you know, there are two things you need to have. If you're talking about penetrance the relationship is to the age at presentation and the genotype.
So, I think what, what the comment is getting at, if I can is that they looked at the model and they see from the explicit part of the model that only a single condition can be defined when what they care about is the fact that there are multiple conditions going on in this family at once. And maybe multiple feature clinical, you know, phenotypic features or clinical signs that they want to track across the family at once. And that individuals might be affected or not for any one of those things
that I can get, but I don't understand how you can how you can ever capture potentially incomplete penetrance without knowing a genotype for the disease and the age of presentation. I think this is really coming at it from a you know, like, it sounds like this person's coming in from a rare disease model,
a rare disease model perspective where they're looking at having features HPO terms across multiple individuals
and they expect that person to be have a nasty disease and they don't and they're surprised by it. We are never surprised by incomplete penetrance that's what we deal with. You know, if you have a 50 to 80% risk, only half of the people are going to get the disease. Right. That's not complete penetrance
but they're thinking about in terms of certain signs or symptoms potentially manifesting and some individuals
Yeah, this is the quick answer to this is that descriptive power that Augusto is looking for is when the standard is used in the context of FHIR or phenopackets.
Yes, I agree. Yeah. Basically, you need to have multiple conditions
Exactly. Which, which both fire and fino packet have natively and we just we don't we didn't do it in this because we wanted to keep it as bare bones as possible and that supported by the carrying standards. We only included one in order to make sure that we were covering basically a peds sort of standalone use case. So that would be an extension of this where you allow for more than one phenotype and it could be more than one phenotype related to the genetic condition, actually. And it could be more of more than one phenotype. unrelated to that genetic, so you could have two diseases one was explained by the genetic variant and another that's not or you could have a range of presentations for that genotype, like polyposis. And cancer, for instance, is a good example. If you have an APC variant, you have the precondition and the condition that's a I would call that
there's not a disease description plus so well. Could be so isn't that what Bing was saying? Is that you you describe your disease in a different way. And then you decide on what's not one later. Maybe? Almost because what's not in one will change over time as you get more information about the gene disease relationships.
I think it's that we didn't say what terminology we give no pointers in terms of the terminology that a disease would be recorded. To the extent that we even allow recording that which is slightly, you know, at the edge, I would say of the scope of what we've been doing because that's supported by feedback and
fire. Yeah, so it all comes back to you know what people see this as the term minimum standard implies that it will carry on working in the future you can add on to it and if not one right now. doesn't do it. What you need to have is a normal one. And then allow for a column that says what the disease description classes
and how much are we focused on I mean, the unique challenge that we're in is if we're not a separate standard, really like we are but it's both fina packet and fire are I would say very much at the edge have a

format and into a modeling language. And so they're both very, very happy to expand. You know, sort of as as the need and use cases arise. So unclear whether or not that's
Yeah, but I think I think at the moment if this person's right that if you're saying node one, you're not actually aligned for node 12345. For different for different disease descriptions, and less you, you make it clear that node one is just the way let's go back to the word minimum. And you'll have another column that allows you to fill in different versions of one. They're wanting this to be more than more, more electable
even exactly, they're wanting us to include capabilities into it directly that are that are supported by fina packet fire. And how would you do that? Why Why would we do that? Because at the moment, it would not be this person is saying it wouldn't meet. What is the minimum standard?
That's because they don't understand they're not seeing it in the context of pheno packet or fire?
Yeah, they're seeing it as a standalone,
I think. Exactly. And very much doesn't support what they're looking for. As
I just I guess I'm trying to say is a right now, I wouldn't know how you would link that to Phenopckets where there will be multiple conditions when you've only got an alternate one.
Okay, then it sounds like we could improve our comp are documentation so that that's really clear to to initially you Amanda, not clear to you that you know, our, our PRC doesn't stand a chance.
So cool. I do see we're at time here. Does anyone have like five minutes to stay on and we'll quickly burn through these last? I think there's mostly two or three. I just wanted to quickly check with you guys. Sure. Okay, let's go really quick. So, I want to talk about this one because it came up in the PRC. And through public comment as well. So the question of differentiating between sex at birth carrier topic, sex and gender. So right now our model includes gender and sex at birth, but not carrier topic sets. And through the public comment. Someone had provided an example where they have for example, sex reversal cases where the phenotypic sex was different from the XY karyotype. And so they explain how an adjustment of wording to account for molecular sex or carrier type or adding they said adding a field like gender I did respond saying we do have a gender field. And then, so right now they're adjusting sex in the pedophile to resolve XY karyotype mismatch warning. So my take on this is, I think from both of these comments is maybe we should think about adding a carrier topic sex feels like an optional one if that's possible. I don't know what it was the reactions to that.
Lindy Let me chime in. Okay, I know we should decouple our pedigree tender from fire but in terms of Fire, fire actually, highly recommends to represent karyotype effects in observation as an observation because that requires clinical observation for sure. So if we include coyote presets in our individual profile then there are some sad when someone submits the pedigree information that one should know the carrier to pick sacks which will not be required of course, in our class, but I will say it is not very likely for data submitted centimeters to submit carrier to pick sacks on their own. So I am not very down for the adding chaotropic sec store profiles class.
So I wonder if those because I'm pretty sure. So fire can represent it as an observation, I think helps to double check. I think Phenopckets you can represent carrier to big sex directly represent a yes, yeah. So I don't know. Do we just say this, depending on your implementation context. represent those things,
and we just give them guidance on it. Yeah, and
actually, if I remember correctly, including the inclusion okay with typic sets was version one. Because I remember show him the version two revamping and then they actually made it as an observation, not as a part of individual like a person.
Okay. So so it sounds like what we need to provide here is the guidance implementation guide guidance or at least like documentation, guidance about the fact that we're not really helping, let's say around the interoperability then of it. We're just saying well, we're saying if you have this information, here's how you should represent it in fire and here's how you should represent it, you know, packet and then maybe it's on us in terms of the tooling to make sure that those are interoperable? Yeah, maybe the reality though, is that and okay, I can keep talking about this at length, but we're over time so carry on, Lindsay.
Okay, we can continue on this further. No worries. Okay, so there was a one here. I've assigned this to myself to connect with REWSt, this but you're just kind of a inquiry about for race, ethnicity, ancestry, we're still using

the term race. So they suggested just checking with rules about any best practices here so I can do that. And they might have some some guidance for us.
Yeah, we weren't sure about whether to split them out or only use ancestry because that's the most relevant one, even though existing data is going to be from all three categories. Yeah.
So I can I'll ask them about that. The next one is just consider how the tools we expect that they evolve with the standard. Oops, I put that in the wrong one. And I put Yes, we do. So they can until it we've already answered this one. This is a related question to kind of that race, ethnicity ancestry one, and there's just a, there's a general conversation on the PRC about, you know, are we representing some of these things as best as we can in terms of gender? Some of this I think, went a little bit off track. In the sense that one of the PRC members thought that this standard would have a role in the actual like drawing of pedigrees and how you might represent, say, certain gender transgender, for example, in a pedigree, which we that's not that's out of our scope. So I included this year anyway, just to say just for us to maybe think about if there's, you know, any way we could improve or provide guidance on some of these things, we could think about putting that into the documentation. So that's that and then lastly, there was just the public comments this last one we've already addressed. This one here, I put a suggestion, and there was a group from Yeah, from Argentina, that just kind of they weren't techy people in terms of software devs. And so we're having trouble trying to understand aspects of the documentation. So I thought, you know, if there's a way we could think about adding some sort of more laid description, or even a thing that says this is geared towards mostly software developers who are implementing that just might help some of the confusion people are having there. Then we have a whole bunch of great comments that Chris had caught about our ontology. So subtle issues, like the direction is reversed, or you might need to distinguish things better. I've linked to all of the GitHub issues he raised about them where he describes these in a little more detail. So I think what we can do here is, is maybe comment rate on those threads rather than going through this spreadsheet. And then lastly, there was the one question comment here from Tony. With regards to the again this Do we want a standalone or this conceptual version that we've created? And so I just copied his sort of comment here thinking that we need we need both we need a standalone in addition to the conceptual that we have, but in my opinion, you know, I guess Yeah, something to think about if we decide we want to do something standalone and correct me if I'm wrong, I don't think they're mutually exclusive. Right? We could create a series usually, if we wanted to. So yeah,
the nasty detail is that we may end up having done so implicitly, in order to actually have the tooling fully convert between, you know, packet and fire, you have to have a common way to move them between them. So you may end up with it. But it's about whether or not we want to publicize and standardize that
or not. Yeah, okay. Something to think about. Anyway, we are way off time and I want to be respectful of everyone's everyone's time. So I'll open the floor quickly. Any last comments, thoughts about all of this? What I can do after this call is start to organize some of our responses here maybe even get some up on up as issues. I can start tagging people where, you know, we need attention on things. We can start working through these in terms of the timeline for resolving some of these comments. So the next PRC meeting isn't until the end of August, I think it's August 28. So there's quite a couple of weeks here. However, I think the sooner if we could set a priority particularly to just some of this description. I think that would be good. We can get that back to the PRC sooner. And I think that will then help with working through the resolution of some of these other comments as well if they have that additional context.
So, about Angela Solanas comments. I know Angela. Oh, yeah. And so I think even for someone like me, I struggled to understand what the purpose of this was. And so if you want to write something for Angela, and I can understand it might be good, right. And if you go back to the version two comments for the PRC it's it's almost like you need that general introduction. The same lay speak would actually be the thing that you would include it the introduction as to what, what, what the purpose of this model was be and how and also some future proofing like it. The point is that it will interact with Phenopckets To do this, or it will interact with whatever
what's the best way to get feedback on these things. So one of the things that I'm so I don't know, I mean, like, can we have a staging site where the the changes to the readme like we can merge them in to a dev version, or is it should we be doing this in Google Docs or something so that humans can read it before it goes off on

the dev side? I'm just trying to think about what's the right way. Lindsay? Have you? What have you seen in other working groups where the documentation is in GitHub, but you want feedback from non technical people? Their markdown, because then you can just pre pretty preview. Yeah, we're in restructured text land with read the docs. And so you can
Yeah, I mean, I've seen it in both ways. We have had some GA4GH public comments where the documentation is in GitHub. And what we've done is provided like pictures and comments of how someone could add comments and error create issues. I think for simplicity's sake, we could just do this in a Google doc because everyone knows how to do it, and we wouldn't have to kind of write an explainer. And that's typically much more easier to share doesn't require someone to sign up for GitHub if they want to get into the stuff. So I think it'd be easier if we just did a google doc okay. There are some groups that do that have used, like docs or some version like that, for these sorts of things.
So then once we're happy with it, then we can copy we can throw it
into the docs. Yeah, yeah. Cool. That's fine. Let's do that. Thanks. Perfect. Do you want me to like guess, copy everything we have in the read the docs into a Google doc first or it's not
too much. It's not too much to ask Lindsey. No, no, not, not at all. Anything at least that we're going to need to change? Yeah.
Yeah. And I can do that. Thank you. Yeah, I can definitely do that. Perfect. Okay, um, so I guess one last question for everyone. Do we want to check in again, in like, two weeks or so because I think their meeting is on.
It'd be good to have make sure we can get granted the next meeting. I think it's the only thing I'd say about that. Right? Yeah. Two weeks is fine. Be good. If Grant was there?
Okay. Yeah, they're the next PRC is on the 26th.

2022-06-23: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Melissa Cook (NCI), Meen Chul Kim (KF/CHOP), Alex Tsai (GA4GH), Chen Chen (GA4GH), Moni Munoz-Torres (Monarch)

	Actions Arising	Assigned To	Deadline
1.	(see list in agenda below!)		
2.			
3.			
4.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping	Orion, Grant, Lindsay	5 min
2.	Outstanding Tasks ☑ ~~Update the Tooling for BODICEA export? Check in with Tech team (Lindsay)~~	ALL	20 min

	☑ ~~Check in with Alex re: BODICEA output (Orion)~~ ☑ ~~Pedigree tools - still referencing the old ontology and conceptual model~~ ☐ **Approval Docs:** Note to implementers - we will need a description of your pedigree implementation (please keep an eye out for the notifications!) ☑ ~~Follow up with Jules (Lindsay)~~ ☑ ~~Adding column to the Common dataset~~ ☑ ~~Phenopackets implementation - ok to live in Phenopackets repo?~~ ☑ ~~Meen - add link to public server when ready~~ ☑ ~~Add detail to the implementation interop notes (eg. Implementation Demo) - DONE but needs review~~ ☑ ~~PhenoTips link? (says "url to be added by submission")~~ ☑ ~~Review TrakGene's JSON Export (Orion)~~ ☑ ~~Update examples (Orion)~~ ☐ Help with the FHIR examples (Meen) ☑ ~~Check if FHIR IG needs updates (Meen)~~ ☐ Representing pregnancies, will need to add an extension to individual? ☐ Confirm last PRC member (WSL) ☑ ~~Readthedocs - note we still have example text on the front page~~		
3.	A.O.B.	All	5 min

Minutes:

AT: So you mentioned that it was the terminology that was updated. I remember it was like two weeks ago or something like that. And there were there were a lot of lines that got updated so I wasn't too sure what mattered or if I was supposed to manually map out everything ontology.js
OB: I think what's going on here is that the code in it, it has a map for just convenience purposes. Right? It's just so that then we can refer in the code to terms like you can say is biological mother in the actual code rather than having to say like KIN 27. And so I think the way it was implemented originally is that it throws an error if one isn't there. But I think that's kind of just optional. In reality, I think we could just only list the ones that are actually used by the converter tooling, which I think is just those two at this point right now in terms of the

structure because it's all just dealing with direct family trees. So I have a commit I can either push it or I can turn it into a branch and I can push that even right now which gets rid of that error. If that wouldn be helpful.
AT: Yeah, as a branch, it would be good. And there is this other issue where the tool itself wasn't storing, like every single thing. If you were to import a Boadicea, do you prefer to have those Boadicea inputs stored in the properties?
OB: It's a great question. It is one that I think is the tip of the iceberg for the conversation that we should probably have today. If that's okay, can I build on your question and then loop in everyone else on the call to try to get this feedback for that?
So the basic challenge we're running into is in a few different flavors all right here. I think this is a critical question for the people on the call, which is the pedigree standard as we have it is a conceptual model, and it is distilled down to only those kinds of those pieces that are the most essential. And we haven't even really structured those pieces, but instead have provided guidance for now what that looks like when implemented within a phenopacket model or within a FHIR model. And then there's sort of those that are connected into all of the different data that can be collected there. One of the challenges that we've run into is maybe lack of consistency between all of those and in terms of that standard conceptual model. We can't have full alignment between the property names in the document. And so because we've gone down the route of not having a separate, actual standard that you can actually put actual data into. Instead you have FHIR and what it looks like in FHIR for a Phenopacket, what it looks like in your packet, you can put data into one of those two things. I think we're running into some meshing difficulties. And those are coming to bear I guess with the TrakGene implementation, which was right now, I think based on some of the early documentation that we've had for what it would look like if there were a separate standard.
If we're truly going to implement a converter to link , the question for the inverter to link basically is if we want to use risk models. If we want to run a risk model, a risk model depends on it an awful lot more than what's at the core of the conceptual model. It requires some of this richer data that we're leaning on phenopackets, or we're leaning on FHIR for, like cancer diagnosis. We haven't given any specification for the structure in which a cancer diagnosis details would be specified. But that is implemented in one way and in a packet that's implemented in a different way and FHIR and in theory that can be gathered from either of those two models. Our question then, I guess the question for us now is in terms of the converter tooling, to talk specifically about Alex's situation, or for someone else who's coming in and trying to implement this standard. How far into those weeds do we want? Now it’s not just the Pedigree standard, but also into FHIR land and other fields.
MC: Is there a framework you’re working with? Guidelines on number of transformations? Be crisp about the scope and also minimize scope.
OB: Tried to minimize this by talking about data elements in a theoretical sense and leaving the specific implementations up to the user. Where those disagree is where there is friction.
MC: Even just being clear about the content specification definition before you get a string, middle ground between the implementation and then there is the instantiation of it in someone’s system. Where those systems don’t align, there needs to be some acknowledgement. Where does this end and Phenopackets pick up, same with FHIR.
OB: Led us down to the path that we're on and where we're at in it, I guess which is where we're saying, this is a statement of data elements. Here is a specification for or here's a description of what that looks like. Here's an implementation that meets certain criteria, perhaps within these two different standards. But, then the question is, how deep into the interoperability then do we go? Which is sort of related to what Alex is ultimately asking because right now there are a whole bunch of things that is part of like supporting a certain amount of pedigree interoperability would make sense to have be there in the tooling, but then that should be there in that data model, too.

OB: Even if we ignore BODICEA, the tooling will allow you to convert with Phenopackets and FHIR. So much overlap with all of these things.
MC: Why have two, why not align them.

MMT: Trying to come into things that help us integrate more. If, because we're doing this off representing the pedigree inside of the phenopacket, then the fact that there exists an implementation guide for the Phenopackets itself, are the people who worked on that implementation guide able to incorporate the element of the pedigree in that so that there doesn't need to be an additional transformation.
OB: The things leveraged in those two systems are very different. Physical structure looks different in these two worlds. Can convert them, but where it gets hard - looking at the common dataset, we have elements that include rich data models. Unbounded if we want to include all of these.
MC: It sounds like there's overlap but different for those use case. Think about when those wrappers are different. Containers are still holding a core of the same information even though some of the other bits around need to be different.
GW: When Melissa said we need to unify everything, in my mind, I thought our goal was to create a unified standard that everyone would follow and that standard needs to be somewhat universal. Pick and choose what they need to do for their use case. Whatever people read that we publish, it should argue for what exactly that solution is. Should be clear that it is our universally recommended standard. Are we getting to that goal.
OB: I felt like the goal would be - here are the data elements we think are important, here is what is or isn't support in ped or FHIR, and here is a much richer set of things that are supported if you use the GA4GH Pedigree. In other words, I didn't think we were trying to use 1.
MC: Confusing as an end user, which to use. Too many choices and too many paths doesn't help the community come together.
MMT: I wonder if one of the thoughts that led to this point - what each of those platforms allow you to do? When you are thinking about conducting, just exchanging information about a pedigree, you can think about a FHIR server allowing you to exchange that and maybe that's how that conclusion was reached. It would be good to represent it in a FHIR format so that it can be utilized and exchanged. But when you're thinking about conducting analysis over the information, then we're not talking about FHIR because FHIR is not going to allow you to do that. If you are thinking about the analytics that you can do by utilizing the information about your pedigree would have to be done in a different environment, which is I believe what all of this documentation that you've produced about what ped can and cannot do and how feedback and support in a more robust representation of the pedigree is also allowing you for that part of the of the empowering the analytics. I don't know how we arrive at the fork. But I wonder if those are some of the thoughts that led to people believing that this was an option to be at this work and to have both options. And then if you choose to represent both of them, then I think if there's a case to be noted for the user - if this is about exchanging information, here it is, if this is about analytics, over the pedigree, here it is.
MC: Brilliant to clarify it for these different use cases because collecting and exchanging data is one type of use case and then taking that collected stored data and analyzing it as a different use case.
GW: Right, FHIR is just to transmit one to another. But can transmit to an analytics tool that can read a FHIR format. So if your data in FHIR, it doesn't preclude you from using an analytics tools.
MC: Clarifying each one of those things is whats needed here.
OB: Real world feedback on the lack of direction, and we haven't made the data more directly interoperable. We gave a structure to it, but not really connecting the ecosystems.

What's next?
MC: Be explicit and document, which hopefully shouldn't be a huge lift.
OB: Clarity is not quite there in terms of implementations. Question is, can we wrap this up with a bow. How to resolve the FHIR Phenopacket friction.
MC: Document the friction points to help get them resolved. What to accomplish now and get on the roadmap for later.
GW: Do have the phenopackets column added to the common dataset column, is that helpful?
OB: want to make sure we're clear to implementers what they want to do, last things I want is to make a problem worse or have it not be used. Need to tighten this up.
GW: Still go forward, helpful to get the feedback and fresh eyes.

MC: Public review will be a great way to hone this deliverable.

OB: Back to Alex's question - yes, the internal model needs to support those things for the common dataset document. Your placeholders are probably good enough. What we can say is even import from PED for Bodicea into a common model, you can then take that and export into ped, Bodicea or others out of that common model, and as long as that kind of common model is compliant, which might require a few tweaks here or there based on some of the changes recently, then we should be good.

2022-06-09: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Bing Feng, David Hansen, Meen Chul Kim

	Actions Arising	Assigned To	Deadline
1.	(see task list below!!)		
2.			
3.			
4.			
5.			
6.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH 10th Plenary Event Page - Meeting schedule during product review process - Google Summer of Code Projects - Note Tim's project on using Pedigree! "Developing user-customisable forms on pedigree data"	Orion, Grant, Lindsay	5 min
2.	Final Actions Review: ☑ ~~Create PR for outstanding changes in the Classes section of the documentation (Meen Chul Kim) - DONE~~ ~~Issue #25~~ ☑ ~~Add maternal/paternal terms to the KIN ontology - David's team member Hoa has made some~~	ALL	15 min

- ~~edits, will attach in next post. Please review! (Orion + All)~~
  - [x] ~~Also see Issue #24 re: sibling ambiguity~~
- [x] ~~Add updated pedigree ecosystem diagram to the docs (Lindsay)~~
  - [x] ~~Check in with Stephanie about "beautifying" it~~
- [x] ~~Add PRC member suggestions to the timeline document (EVERYONE)~~
- [x] ~~Update the Tooling for BODICEA export? Check in with Tech team (Lindsay)~~
  - [ ] Check in with Alex re: BODICEA output (Orion)
- [x] ~~Work on first draft of submission documentation (Lindsay)~~
- [ ] Review TrakGene's JSON Export (Orion)
- [x] ~~Merge new KIN owl file with the old (Lindsay)~~
  - [x] ~~Replace purl in the readme and readthedocs~~
- [x] ~~Take a first stab at the documentation for REWS/Data Security, than tag people as needed (Lindsay)~~
- [ ] Approval Docs: Note to implementers - we will need a description of your pedigree implementation (please keep an eye out for the notifications!)
  - [x] ~~Follow up with Jules (Lindsay)~~
    - [ ] Adding column to the Common dataset
- [x] ~~Reach out to potential PRC members (Lindsay)~~
  - First outreach sent! Augusto & Lori confirmed
- [x] ~~Check in with Meen about Product Maintainer status (Lindsay)~~
- [x] ~~Merge PR #24 - Lindsay's attempt at adding the GA4GH logo to the docs~~
- [x] ~~Update the Pedigree repo readme (Lindsay)~~
- [ ] Add detail to the implementation interop notes (eg. Implementation Demo)
- [ ] Update examples (Orion)
- [ ] Check if FHIR IG needs updates (Meen)

3.	Product Approval Documentation Review
	- Timeline
	- Approval Documentation
	  - Implementer notes
	- REWS/Security Documentation (also linked in the approval documentation above)
	  - There are a few questions that need group feedback
		Lindsay	20 min

4.	Next Steps / Remaining Tasks - Ontology Tasks	ALL	15 min
5.	A.O.B.	All	5 min

Minutes:

Decision on meeting scheduling during PRC process:
Keep the upcoming 23rd meeting just in case we need a last minute check in, cancel the rest. We can schedule meetings during the PRC process ad hoc.

[From Otter.Ai]
LS: First, what we could do is just go through all of these actions. Some of these can actually be crossed off, which is great. And then I've got some questions on the documentation to go through and then if anyone else has something they'd like to chat about, feel free to jump in. So starting up first, Meen had an action to create a PR for the outstanding changes for the classes section. He has done that today. So I'm going to cross that one off but the issue is not yet merged. Orion - Meen has tagged us both as reviewers on this. We can come back to this once he's here. Next, maternal/paternal terms being added. That's done. And then Orion today you had updated to move those classes into properties, correct?
OB: Yep. There are a few oddities, sort of the way that it's set up right now. There are a few more things that we might want to do as part of this first version, but that was just digging into the sort of current state of that when merging those in but yes, that part is probably done now.
LS: Is that something we'll want to come back to in this call like ontology tasks?
OB: Maybe if there's time it's not a high priority.
LS: Orion reviewing track genes JSON export?
OB: So I followed up with Tim, I might have been too long to follow up with him, I did follow up with him again today.
LS: Okay, sounds good. Some tasks that were in the approval Docs. So this mostly has to do with the implementation details. So David, I saw you had had other David add in some details. That is fantastic. Thank you for that. And then Orion you've got the not yet fully implemented details in. I did reach out to Tim, I'll have to follow up with him as well. Meen, we can ask him again when he gets here, but I know he said he was re formatting. Jules, I'll need to follow up with again. I did briefly talk to Soichi this week at an ISO meeting. He's seen my email and he says he'll get back to me as soon as he can. Bing's implementation is currently blocked by the tech team. One thing I did want to note there is this new section here called interoperability notes. It hasn't been in past versions when I've taken products through. And so I was looking at the Beacon documentation to see how they filled it in and they talked about a validator tool that they had put together and they run centrally. So I figured we could use this as a place to share, for example, some detail on the ontology and our validator tool, unless you think there's something else we should be doing here.

OB: It seems like this would be a situation where we would want to demonstrate a producer consumer case, validated on fire, and on Phenopackets like looking at it from sort of a high level as I think what's the project we're trying to achieve here? That would be what I would think would be kind of a minimum that we would want to have going into the approval process. Okay. Yeah. That would be kind of ideal, unless we want to, I mean, we can tie the tooling in either the validator or the converter tooling into it as part of that story in order to weave a little bit more in terms of connections. But ideally, it would be communication between more than direct pairs.

OB: Are we able to consider doing ingestion from one source? Might be able to demonstrate Meen's that way. The other ways would be between either a pair of different or sets of different pedigree tools. So like, if we were to take a pedigree from the REDCap plugin, and show that, you know, you can serialize it and pull it into a different pedigree tool, whether it's tracking or phenotypes is something that would be compelling for the team, the other any other pairs, or taking a pedigree, or pedigree data from any of the tools that can export it, and pulling it into, like exomiser, or for the or being coseg analysis tool like from I feel like that that would be the most compelling if we can, I don't know how close to that we can get in this timeframe, but that would be ideal.
LS: I know we kind of touched on some of this with the implementation demo we had done while ago. Right? So I did link in the slides, but we might have to fill out. Maybe what I could do is find the old recording for this if it's still existing somewhere and even add some screenshots or something to help fill it out.

OB: I think I was supposed to update the example in the documentation to cover explicitly the phenopacket and the FHIR representation. And I don't think I've done that yet. So right now they're basically in a lightweight, just Yaml format we need to give people actual tangible JSON representation. I think we need to connect with Jules and if he can tell us what we should do in order to make that be part of phenopackets, either this or the next version or whatever.
LS: Yeah, okay, I will. I'll start a new email thread on it. So a bit of an update on our PRC. We have two or three members already confirmed, which is great. We are just waiting to hear back from Bob, I'll follow up with him. I did ask Meen about being added as a product maintainer and he agreed. So that has now been added to the docs. We added a GA4GH logo to the docks.
I added this action about the readme for myself today. So back in our repo, we have this readme page that was kind of our holder before the docs. What I might do is kind of shave some of this down and basically point it to the read the docs.
OB: Agreed. I think that readme should have very little for even just technical information. And then the link at the top right, where you configure the actual link for that is what people will go to, so we're good.
LS: Anything else task list wise that we want to talk about.
GW: I just wanted to confirm Orion when you felt like you were ready to go back to the the document and fill in the column where we have the GA4GH pedigree standard, fill in information for the new data elements that you're adding to the model
LS: You mean for the common data set?
GW: Yes, yeah.
OB: So Meen updated the classes to accommodate the common dataset document [lost connection]
I don't think anything needs to be updated in the documentation more is that wrong?

Meen is everything in this document now represented in the classes? Or is it just harmonized with it?
Are there rows in a table that are not supported? By the classes, the updated classes?
MCK: So I will say most all of them are reflected in the updated class documentation.
OB: That's awesome. Perfect. Because so I think that is the big improvement that I think we were hoping for. So I don't think we need, I mean, we can discuss whether or not we need a column here to represent the specific property within the standard. It isn't clear to me that we do, grant if you feel otherwise then we could certainly add it. The piece that might be missing is having phenopacket as a column.

It's only for all of the stuff that is covered by things that are already there. Take gender as an example. This is already supported within fire. And this is also already supported within phenopacket and so it just needs to be the equivalent phenopacket of family.pedigree or is it .GA4GHpedigree.
GW: Let's keep it simple at least in this first draft.

OB: Meen the question was just whether there are specific things that you would like for us to review in the issue.
MCK: Yeah, since the deadline is approaching, I guess we can do a very simple quick review and then merge because what I did was revamping the unclassified section against the common data set using the field name cardinality and simple description, so that's it.
LS: Okay, action on myself and Orion to do a quick review and then just merge this.
LS: I had some questions about the REWS security documentation. I took a first stab at filling this out for us over the past two weeks or so. There's a lot of questions in here about data use, you know, have we thought about security risks, what regulatory ethics concerns? A lot of my answers, I just kind of put, you know, we're a model so we don't store data. Uses are dependent on the implementer and their use case. So a lot of these answers are kind of along that theme. If anyone has time and can go through these answers, and make sure that I didn't say anything totally incorrect, please do.
There's a question here about listing any security standards, papers, threat model, attack papers or best practice documents specific to the domain of your product. Does anyone know off the top of your head? So I guess, maybe just as a bit of context, a lot of these apply to API's not models, so they don't apply to us. But I still kind of took a stab at filling some of this out anyway.
GW: I haven't heard of an attack paper on a data model.
LS: We'll just say this one is not applicable. Second question lists the risks and threats that apply to your product.
OB: I think even that is inviting. I hear what you're saying. I think it's a great way of trying to give them something. A stronger statement and one that we can make here, where it's like it's the security risks of this are the exact same as the security risks of a text file.I think we could be similar to Phenopackets I would frame it as the information that we're representing is identifiable sensitive information. And therefore, like the way that people handle that information, they need to handle that information appropriately. And they need to treat data represented in this format, as though it is containing sensitive, identifiable patient information. It's a higher risk version of what phenopacket is doing because being a packet is designed to be de identified by nature. And ours has name and date of birth. We just need to build on it. And maybe tweak it so that it's less about privacy, per se and more like just the risk is people put data in this format, and then there is a leak or something like that around it. Are they mishandling.
LS: Okay, the questions following ask you to estimate like risk likelihood. And I really struggled with this one. Because, you know, it's asking us what is the skill level? Is it a good attacker, a medium attacker, I have no gauge.
GW: Attackers aren't relevant in this particular case,
OB: Accidental disclosure is the real risk.
LS: if we feel the question doesn't apply to us, and we can't really answer it, that's okay. We can just kind of say that. You know, as you saw above, I had answered a lot of them being like, this doesn't really apply to us. And that's totally okay.
I think I'll have to rewrite this one now, but this kind of again, goes off the same theme.
This is one I had a little bit of a struggle with. So you know, what risks are there that uncertainty or misunderstanding of results could be exacerbated due to the product, have you referred to the return of results policy? And so I took a stab at kind of saying, you know, if the information contained in is fed directly back to a patient with any interpretation or counseling, then it's possible the patient could mistinterpret.
GW: Our model does not produce an interpretation or a risk assessment. implementers will do that. And so also here it says, Have you referred to the return results policy? We're not in that field. We're not in the field of doing genomic research, where you have a study participant who you might have a return results.

The last section here is, this is a new section, it's on equity, diversity inclusion.
I added a brief blurb here. But I think this is actually no longer true because we can include race, ethnicity, ancestry now, right? Yeah. Okay, I need to update that.

So the question is, what assumptions about resource availability have been made in the product's design and development? So computing power students and technologies time and money? How are these weighed against competing concerns such as moving the field forward responding to cutting edge cutting edge use cases, etc?
GW: Well, that's actually an interesting question, but I'm not sure that that should be asked here. You know what I mean? I don't think it should be asked of our subgroup, it's a question that lives higher than our subgroup, I guess I could try to throw something in there. There are groups that are thinking about that stuff, and are doing pilots and we're learning but I don't think a data model affects any of those issues.
The last two Qs here, do it require significant time resources or deep GA4GH knowledge that could present a barrier to lower resource projects. So I'd written because it's intended to be used in the context of Phenopackets or FHIR, it does require knowledge of those. And then I was gonna write something like you know, is it it's feasible for one software developer to do it? I don't know. That'd be a question for you guys. I don't know.
GW: I hope the answer to that is yes.
MCK: I'll say yes

OB: I have one quick question. Meen were there any additions that you made that should be reflected in the FHIR IG?
MCK: I don't think so, but let me check again
OB: Good. I wasn't sure if like the race, ethnicity ancestry or something like that we need to make sure are covered. I can send you a Slack message and we can chat a little more about it.

2022-05-26: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Zornitza Stark (AGHA), Bing Feng (UUtah), Chen Chen (GA4GH), David Hansen (AGHA/CSIRO), Melissa Cook (NIH), Alex Tsai (GA4GH)

	Actions Arising	Assigned To	Deadline
1.	Merge new KIN owl file with the old	Lindsay	
2.	Take a first stab at the documentation for REWS/Data Security, than tag people as needed - Note to implementers: we will need a description of your pedigree implementation (please keep an eye out for the notifications!)	Lindsay	
3.	Reach out to potential PRC members	Lindsay	

4.	Review TrakGene's JSON Export	Orion	
5.	Check in with Meen about Product Maintainer status	Lindsay	
6.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - Montreal Connect Meeting Report - **GA4GH 10th Plenary**, Sept. 22 & 23, Barcelona (Hybrid) - Save the Date - **GA4GH Hackathon** (July): The GA4GH Hackathon will be a 1-2 week-long event focusing on API improvements of existing GA4GH specifications and implementing GA4GH specifications in new environments. We are looking for coordinators to lead topic areas! To learn more or volunteer to be a topic coordinator, reach out to neerjah.skantharjah@ga4gh.org. The overall process can be viewed here.	Orion, Grant, Lindsay	5 min
2.	Reviewing the approval task list: ☐ Create PR for outstanding changes in the Classes section of the documentation (Meen Chul Kim) ☑ ~~Add maternal/paternal terms to the KIN ontology - David's team member Hoa has made some edits, will attach in next post. Please review! (Orion + All)~~ ☑ ~~Also see Issue #24 re: sibling ambiguity~~ ☑ ~~Add updated pedigree ecosystem diagram to the docs (Lindsay)~~ ☑ ~~Check in with Stephanie about "beautifying" it~~ ☐ Add PRC member suggestions to the timeline document (EVERYONE) ☑ ~~Update the Tooling for BODICEA export? Check in with Tech team (Lindsay)~~ ☑ ~~Work on first draft of submission documentation (Lindsay)~~	ALL	20 min

	☐ Review TrakGene's JSON Export (Orion)		
3.	Product Approval Documentation - Timeline - PRC Members? - Approval Documentation - REWS/Security Documentation (also linked in the approval documentation above)	Lindsay	20 min
4.	Next Steps / Remaining Tasks	ALL	10 min
5.	A.O.B.	All	5 min

Minutes:
(from Otter.Ai)

LSHello, everyone and welcome. This is our second last meeting before approval! Just a few really quick things before we get into the main agenda. I've added some housekeeping items to the agenda. I won't go over them all right now because I did talk about all of these things yesterday in the full meeting. But very, very briefly, if you haven't seen it, the Montreal Connect meeting report is linked, also linked to save the date for the upcoming GA4GH puneri in September. And lastly, I've also put in a bit of information about the GA4GH hackathon that will be coming up in July. So these are also known as the FASP hackathon for anyone who's familiar with FASP where essentially you can propose some ideas for interoperability or standards improvements via this hackathon structure that we've been working on. And so if you want to learn more about that you can email Neerjah. Anything anyone else has housekeeping wise before we dive in?

I thought the first thing we could do is just go through some of the tasks that we have assigned ourselves from the last meeting. I'll share my screen. So I'll just walk through these and have a couple of comments about some of these things as well. The first was about some of the outstanding changes in the documentation for Meen, he did send me a note yesterday. He's still working on it will send it to us when it's ready. Related to the kin ontology. Thank you David for connecting me with Hoa. He's made some wonderful edits to the ontology, the new version is now up in our repo. The one thing I wasn't sure how to do was generate the Purl link for it. If someone knows how to do that. That would be fantastic.

OB: Should it just be a matter of updating the main owl file, so that then the Perl link keeps going rather than adding an additional file alongside? Does it contain just the differences or does it contain everything but does it just contain everything? Everything?
LS: I think it has everything.
OB: So then it's just merging.
LS: We could also just delete this old one, and add this one and rename it.
OB: Just rename it Yeah, either way.

LS: So there's an action item for myself to update the pedigree ecosystem diagram we have that has been done however, I did also send it to our comms team to see if it can be beautified. Stephanie knows of our June 21 deadline and she'll get a version back which is also fantastic. There was an item for everyone to add PRC member suggestions to the timeline document. I'll come back to this because it's a little bit later in our agenda. I did reach out to our tech team and I know Chen's here and Alex as well. About whether we can get that Bodicea export into our tooling. They do have some priorities with an upcoming ISMB conference. However,

they have offered to try to fit this in around some of that work for our June deadline, which is amazing. Thank you so much, Chen. I also added this in a comment here just so everyone knows. And then lastly, for myself to work on the first draft of the submission documentation. I'll get to that again in our next item. And then this one I actually just added today for you Orion on if you saw the email from Tim he forwarded TrakGene's JSON export for you to take a look at
OB: I will look into that. Thanks.

LS: Okay, perfect. So, what I'm hoping we can do a little bit today is work through some of this documentation because I have some questions for everyone here. And maybe the first place to start off is our related back to that action item about PRC members. As a brief recap, we had put together this list of eligible PRC members here in this document was added to the chat here and so, what I would like to do is ideally, before we even submit our documentation, have started the outreach on this because depending on whether people are available or not, especially because we're getting into the summer, the sooner we can, you know, confirm some of our PRC members the smoother it will be. So I know we had highlighted a couple of name. Any preference on order of who I could reach out to First or are there any thoughts on any other potential leads we'd prefer to have? Otherwise, I'll just start all over and work down.
GW: I just would like to consider Bob because Bob does work with HL seven HL seven, you know, has had pedigree standards in the past. They're going to be really interested in what we're doing here.
LS: Second would be our driver project representative. And so we had highlighted again, a couple of different initiatives here. In my opinion, I would propose to prioritize genomics England because once James left we've kind of lost that genomics England representation there and I would I would love to get them to take a look at the standard
GW: Well, I wondered about BRCA challenge, you know. In the world of breast cancer, family health history is huge. It's, it's critical. It's a critical tool that they use in the world of well all cancers, but specifically definitely breast cancer. And I don't know if that's an argument that means that somebody should be from BRCA challenge should be invited or not. Just throwing that out there.
LS: Very true. I think the thought process behind not including them as a representative is because they have been involved throughout the development. So it might be to our benefit to have someone who could adopt the standard but isn't already part of this work to have a chance to review it and take a look at it. But otherwise, yeah either Mandy or Melissa would make great reviewers as well.
DH:I wondered about h3 Africa
LS: Yeah. Same thing as BRCA because we do have Katherine in the group
OB: Okay, so I guess Yeah, it's a question of how for the people who have been participating. Do we want to hold strong in terms of excluding them from this aspect of it? And we're really looking for people who aren't who haven't been, you know, relatively actively involved in the development process? Or can we have someone who's more, you know, who represents a very, very good use case for the adoption of this?
LS: Either way is fine. It's kind of up to us to make that choice. As long as it's a representative of a driver project, that's really the only criteria. We have had approval processes in the past where it has been kind of either case, and it's been fine.
DH: I think you're right. I think genomics England should be interested in what we're doing. How many do we need?
LS: So we just need one from each category. One work stream lead, one driver project representative and a third member, which can is very flexible can be pretty much anyone but we generally encourage it to be someone outside of GA4GH.
We don't have the third member yet. Could be anyone who's an active member of a work stream, not from Clin/Pheno, though, or you know, anyone with expertise in the field recommended this person is outside of GA4GH. And so grant I know you had suggested Lori in the past. And then there's also the suggestion of Raymond who lead OCHRN, which is a hereditary cancer network in Ontario, and has looked at our standard in the past so he's not part of GA4GH per se, but he's somewhat in the loop with what we've been doing. Are there any other suggestions or thoughts on a third member?

GW: , I don't know Raymond. But obviously I do know Laurie and she's an MD doctor, who is deeply involved in the world of family health history has published papers on the subject. So that's why she just she would just be excellent. She'd be perfect for this. I think now, obviously, I'm, I'm sharing for my candidate here, but but I think yeah, I think she'd be great.
LS: Are there any objections to if we reach out to Lori first? All right. So I'll make an action item for myself to start reaching out to these folks. So if they aren't available, this will kind of give us a nice window to do some follow up with other candidates. If needed. Fabulous. All right. Okay, so the approval documentation. So there was the action item for me to start putting this together for us. And this is linked from the agenda. So where are we here? I have started to link in our GitHub spaces, destination URLs, product backgrounds for the read the docs. I think the two main spots that we still need to fill out here are one the product maintainers - Orion, I'm guessing you might be one of those. But anyone else we can put up to three people. So these are just you know, who will be the point of contact for if there's maintenance issues or people have requests, comments, that sort of thing. And then the second thing we need to fill out is the implementation details. This would be all of the work that has been done setting up implementation so far. So like the track gene one for example, the work Alenjandro has done with like the red cap and open pedigree and such these would all need to be filled in here. It's not a ton of information, we need to put in here. Essentially, we just need to explain to the PRC these are the implementations we've set up and here it is working in the wild, and then should provide a link so they can see that it's functional. If you need an example of the type of information to put in there I did link the Phenopackets approval documentation at the top. What I'm going to propose and let me know if anyone objects to this. I'll make a section for each of the implementers and tag our point of contact there to add a couple sentences. I think that's probably the easiest way for us to go through it. But does anyone object to that? Okay, perfect. Okay. So I'll make that listing there. We can start filling it out. Product Review Committee, I'll fill that out once I've reached out.
The other kind of new piece of news is about this REWS / security documentation. So, normally, the REWS and security questionnaire is done at the beginning of the standards proposal process. So, you know, roughly two years ago is when we maybe should have done it. Long story short, there is a new rews and data security questionnaire. It's very similar to the past one; it just essentially merges it all into one document. And there's also a couple of extra questions they've added in about EDI. But I think that section is still actually under development, so we might actually not even need to fill it out. We will be the first group to test it out. Which sounds scary, but I promise it's a lot of the same questions that were in the previous one. As we work through this, if we can capture some feedback, you know, it takes too long to complete it or this question doesn't make sense or or anything like that. I.

So I had a couple questions as I started to work through this for us. There's a section here on product information. So this would have been stuff we would have filled out at the beginning. You know, what's the product name? What are our milestones? I just put in the ones we had written way back when product type, that sort of thing. Is this the name that we're gonna go with? Is it GA4GH pedigree? Is it just a pedigree standard? I don't know if we ever fully decided on the name we wanted to officially go with
OB: what is the common format? Is there a common convention within the GA4GH for this Lindsay?
LS:Some of our standards have GA4GH in front of it. Some of them don't
OB: It seems like it's the sort of thing where if you were to Google it, whether or not you would need to add GA4GH in front in order to find it.
GW: You Well, let me let me run through my previous thought once again and see what the group thinks I know I was trying to think in terms of how do we present this to the world in such a way that it's very marketable where people outside of GA4GH understand what we're talking about what this is about. People outside of the standards world, right. would see this name and know what we're about. And so that, you know, that's why I had thought about the idea of using the term that you know, common data set for family health history. Because I thought that was like a really good generic term that everybody could relate to. Now, if you feel like it's too generic, you know, I could understand that but you know, let me know what your thoughts
LS: So you're thinking of just the naming of the Common Data Set document itself?

GW: Well, you know, and that, that could be something that we need to distinguish I mean, does does our product with our product name benefit from saying the common data set? Or do you think just the document would benefit from just saying, saying common data set, you know, I mean, or do we use common data set for universally for
LS: apologies if I'm misunderstanding grant, but so I know in the past, when we talked about the naming of that document, we I remember, we went back and forth about minimal data set, common data set and so on. And I think we had agreed on common data set for that document. I'm happy with leaving that name as it is, if everyone else is. And so, to me that would if we think of GA4GH pedigree standard. I think this name would apply to, I guess, the conceptual model, and then then everything else would kind of be you know, deliverables almost under that umbrella in some way.
GW: Okay, then that that makes sense to me.
LS: Okay, I will resolve that then. What else did we have in here? It looks like a daunting sheet, I'll admit, but having worked through something very similar in the past, it doesn't take as long as you think it does. Usually, each question just needs a couple of sentences. There's some regulatory ethics questions, you know, how does it relate to these policies? There are a lot of questions here that are specific to API's. So a lot of those we can just write “not an API”.

Any questions about any of the approval process or documentation so far?
GW: So just so my mind understands the approval documentation and the Roos documentation are two different things.
LS: There are two different documents but part of the same process so if you were to click the approval documentation here and go all the way to the bottom there, foundations are workstream questionnaires. So there's the rews and data security one within that sort of overall. Yep, there's also a partner engagement questionnaire. We don't have to worry about this yet. That's just once approved and provides the communications team with some information to help formulate the outreach they're going to do. So they are separate documents, but still all part of the same process.
GW: So were you going to do the first draft of both of these then, or
LS: I can if that would be helpful. Well, most of the main approval one here is done except for the implementation details which we'll need our implementers to fill out.

Why don't we review it now? So we've got the work Alejandro did which includes got the REDCap open pedigree. David, do you know much about how the FHIR implementation guide is being used, if at all?
DH: I don't, but I can ask Dave Conlan who worked with alejandro on it.
LS: Okay, perfect. So I'll add that to the list. There. We've got Orion and PhenoTips. TrakGene. Bing’s tool but we need that bodicea output first, let's put it in progress for now. I don't know if h3 Africa has anything concrete yet.
OB: Other than potentially via the REDCap.
LS: Would we want to put about Phenopckets? Would we put that as a implementation? But it's not the same implementation as this listing? I'm going to list it here. And I'll reach out to Jules about maybe what we
OB: Exomiser, via phenopacket has been a thing for Jules. And that would be if it's supported by phenopacket, then in theory, it'll be and so I guess it's a question of timelines. Add GEM-Japan as well.
GW: Now, this list is just for the approval right? We will still working work on getting other implementers over time, when they need to be added to this list, or is this list just a snapshot in time?
LS: Just just a snapshot just for approval. Yeah. And we need two at a minimum, right. Yeah. Which we have which is great.
GW: Quick question to have this supported on a fire server. Is that important for what we want to accomplish here to have that as an implementation somewhere by somebody?
OB: So this is Meen, we should probably add that right. Kids First. Yep, that's another one.

LS: So we've got this listing here is essentially who would be point of contact for maintenance issues for question issues if someone you know will be requesting updates for example. Of course, there will be GA4GH staff who will kind of be maintainers in some way as well. But usually we try to keep at least one to two of our volunteer community that actually helped develop the standard be a point of contact after it's approved. So for example, for phenopackets, Peter and Jules are our maintainers. So we need at least two for pedigree as well.
OB: I'm happy to be one. There's anyone else who would be interested, I guess in terms of support - grant. I don't know if that's something that you would like to be
GW: My question is, am I the right person? Because you know, I'm not deep into the modeling part of what was done here. And is that type of maintaining that, that people will want to talk to people about or or is my knowledge my I'm you know, I'm just sort of like a high level, subject matter expert type person. Will that help people? If that will, I'm happy to do it. I just I'm trying to think of what what type of questions will come to the maintainers
OB: is it I guess, and Lindsay to the extent that you have visibility into that for other projects, how much is it? You know, questions like, why we you know, can it be can it do x rather than Y or can we have it also do Z or is it questions like, you know, we're interested in this as a thing, like, how do we get started with it? What is it? How does it fit in those sorts of questions, or is it both?
LS: It's both
GW: put me down. Okay. I would say,
OB: It'd be helpful to know whether or not Meen could be listed here, has some dedicated cycles through kids first, it would be helpful to know if he's interested in taking on that more technical
LS: Okay. Yeah, I'll reach out to hi.

GW: Lindsay, I see as you start going through the first draft. So these if there's something that you'd like me to think that I could do just as well, please just make the ad assignment. Just shoot me an email and say, Grant do this.
LS: For sure. Will do. I have a feeling once we get into some of the rews, actually more so the data security questions, I'll probably have to tag some people in it. Okay, if there is nothing else, so looking at our timeline here, we've got a meeting in two weeks, which is June 9. And then roughly, just under two weeks from that is our submission deadline. So our next meeting will be our last before submission. So I would really encourage everyone to take a look at the documentation. Get through any of those tasks you need to do prior to our next meeting so we can talk about it then while we're all together in a room for last time, if needed. And again, Slack channel, GitHub issues, whatever there is, please put them out on there and so we can follow them as well.

2022-05-25: Clin/Pheno Full WS

Chair: Melanie Courtot, David Hansen
Attendees "Name (Project Affiliation)": Ian Green (SNOMED), Philip van Damme (Amsterdam UMC), Chen Chen (GA4GH), Orion Buske (PhenoTips), Meen Chul Kim (KF, CHOP), Grant Wood, Brandon Wright, Francis Jeanson, Alex Tsai (GA4GH), Jules Jacobsen (QMUL/Monarch), Ayhan Serkan ŞIK, Steve Laurie (RD-Connect/Solve-RD)

	Actions Arising	Assigned To	Deadline
1.	Add comments/feedback to the draft 22/23 ClinPheno Roadmap	ALL	June 15

2.	Please add any comments to the Pedigree model, ontology, etc. via Github issues or email as soon as possible. V1.0 approval deadline is approaching!	ALL	ASAP
3.	Follow up with Peter R. about the supplement for Phenopackets implementation with Broad/Terra	Lindsay	
4.	Send Melanie C. the linking to Catalan biobank	Ian Green	DONE

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - Montreal Connect Meeting Report - **GA4GH 10th Plenary**, Sept. 22 & 23, Barcelona (Hybrid) - Save the Date - **GA4GH Hackathon** (July): The GA4GH Hackathon will be a 1-2 week-long event focusing on API improvements of existing GA4GH specifications and implementing GA4GH specifications in new environments. We are looking for coordinators to lead topic areas! To learn more or volunteer to be a topic coordinator, reach out to neerjah.skantharjah@ga4gh.org. The overall process can be viewed here. - New “Materials of Interest” section added to the agenda/minutes (see below)	Melanie, David, Lindsay	5 min
2.	Phenopackets Updates - Upcoming publication in Nature Biotechnology - "The GA4GH Phenopacket schema defines a computable representation of clinical data" - "GA4GH Phenopackets: A practical introduction" - submitted to Advanced Genetics - Finalizing software to work with Phenopackets - Supplement to work with Broad/Terra to implement Phenopackets as a part of GA4GH APIs	Peter Robinson	5 min
3.	Pedigree Updates - Preparing for Steering Committee approval submission on June 21st - Github Space - comments welcome on the model, KIN ontology and FHIR IG. Please create	Orion Buske, Grant Wood	2 min

	an issue in the appropriate repo.		
4.	Computable Cohort Representation Updates - Hackathon kicked off April 21 - minutes - Master Doc - Five topic meetings scheduled throughout May (see Master Doc). Last topic meeting is Wed. June 1st (CQL/Use of FHIR/Mappings).	Lindsay Smith	2 min
5.	Feedback on the 2022/2023 Clin/Pheno Roadmap - Draft Roadmap	Lindsay, ALL	15 min
6.	A.O.B.	All	1 min

Materials of Interest (Publications, Webinars, etc.):

1. (PRE PRINT) The GA4GH Phenopacket schema: A computable representation of clinical data for precision medicine
   Julius O. B. Jacobsen, Michael Baudis, Gareth S. Baynam, Jacques S. Beckmann, Sergi Beltran, Tiffany J. Callahan, Christopher G. Chute, Mélanie Courtot, Daniel Danis, Olivier Elemento, Robert R. Freimuth, Michael A. Gargano, Tudor Groza, Ada Hamosh, Nomi L. Harris, Rajaram Kaliyaperumal, Aly Khalifa, Peter M. Krawitz, Sebastian Köhler, Brian J. Laraway, Heikki Lehväslaiho, Kent C. Lloyd, Leslie Matalonga, Julie A. McMurry, Alejandro Metke-Jimenez, Christopher J. Mungall, Monica C. Munoz-Torres, Soichi Ogishima, Anastasios Papakonstantinou, Davide Piscia, Nikolas Pontikos, Núria Queralt-Rosinach, Marco Roos, Paul N. Schofield, Anastasios Siapos, Damian Smedley, Lindsay D. Smith, Robin Steinhaus, Jagadish Chandrabose Sundaramurthi, Emilia M. Swietlik, Sylvia Thun, Nicole A. Vasilevsky, Alex H. Wagner, Jeremy L. Warner, Claus Weiland, Melissa A. Haendel, Peter N. Robinson. medRxiv 2021.11.27.21266944; doi: https://doi.org/10.1101/2021.11.27.21266944

Minutes:

(Otter.ai transcript)
LS: Welcome everyone. Glad you could join today. So this will be a relatively quick meeting today. Just 30 minutes but we're gonna go through quite a number of topics today. So first up some housekeeping. Everything we're going to talk about right now is linked in the agenda. So first up, so we had the Montreal Connect meeting in April about roughly a month ago, we haven't had a full work workstream meeting since then. So if you haven't seen the meeting report, I've now linked that in the agenda. As a reminder, we just hosted one meeting which was on the computable cohort hackathon kickoff, which we'll get to a little bit later in the agenda. There's all the reports, recordings, minutes for all of the other meetings as well in case you haven't had a chance to see it. Second up. As a reminder, our GA4GH 10th Plenary is coming up September 22 and 23rd that will be in Barcelona. However it will be hybrid in case you can't make it. I've linked to save the date. One thing to note that's going to be a little bit different this year, we will be separating the plenary events from the working meetings. So normally our plenary events are you know, roughly four to five days long. There's the public facing events and then all the working meetings. Those working meetings are going to be virtually a month later. So we're going to be separating it a little bit. Plenary will be just those two days but should be a great program. I'm looking forward to seeing some of you in person again. Other events to be aware of coming up soon, the general GA4GH hackathon. And so for those who haven't heard of these events before, it's under the umbrella of FASP or the federated analysis systems project, which is essentially a project that works to enable interoperability of GA4GH standards or very specific, you know, working meetings on standards improvement. And so these hackathons are very self driven, it's our volunteers and contributors that come up

with the ideas and help organize these hackathons. But we as staff helped provide the structure to make these things happen. So right now there's an open call for volunteers or topic coordinators. And so you can reach out to Neerjah, heremail in the agenda. And these topics can be anything GA4GH related to say you wanted to hold a hackathon on I don't know connecting pedigree and beacon that could be a hackathon topic meeting for example. And I've also linked just a document there that kind of gives an overview of this, this whole process. And then lastly, I just wanted to add, I started our I'm going to start adding a new section to our agenda which I'm just calling right now materials of interest. So before all of our full workstream meetings, I'll send out a call for any, publications, meetings of interest, anything, you'd like to share about the CLIN pheno space with the work stream let me know and I can add that in the agenda so we can really use it as a way to build community within clin pheno. And so right now I just linked one of the preprints for the phenopackets schema publications. Alrighty, so that's all I had for housekeeping. Next up, we did have Peter who was going to talk about Phenopackets But I don't see him here. Jules, do you know much about the topics Peter was going to talk about? It was about the supplement to work with Broad and Terra, he asked for a couple minutes. Do you know much about that?
JJ: Oh, no, I don't. I'd like to know.
LS: no worries. We'll skip to the next one. I'll follow up and add some content maybe after the meetings done there. Okay. I guess the two things we can mention though is there are two Phenopckets publications that will be coming down the pipeline soon. One of them is essentially a Phenopackets ecosystem. Overview. So that one is titled, The GA4GH Phenopckets schema defines a computable representation of clinical data. I think many of you were authors on that. It has been accepted into Nature Biotechnology, and that will hopefully be coming up soon. And then there is a second one that was recently submitted to advanced genetics, which is an example case where you could take a say a rare disease patient and show how you would fit it into a pheno packet. So that one's been submitted. And I think we're working on some review comments on that now. Okay, pedigree updates. We've the pedigree subgroup is preparing to submit version one to steering committee on the 21st of June. So it's roughly a month from now. In the agenda. I've linked to the GitHub space. Here. Let me share my screen so everyone can see as I walk through this year. Okay, you're seeing the agenda?
Speaker: Yep, yes. Thank you.
LS: Okay, so in the in the agenda, I've linked the GitHub space here. So this first repo here pedigree is where you will find the conceptual model itself. The family history terminology, you'll find the kin ontology which is an ontology we developed to help support the pedigree. You also see a link to the FHIR Implementation Guide, as well as some tools here. So what I would really appreciate from everyone here is if you have a chance to look through add issues, add comments, you know, ideally, the sooner the better before we get to approval in June. That would be fantastic. We do also have some read the docs documentation. That is currently in progress. There's still a couple of pages we're finalizing. But again, if you want to take a look at that, let us know if there's anything that's unclear, maybe something we need to expand on please let us know that would be really, really helpful. And so we'll be working on finalizing some of these things, the documentation and then we have a series of your documentation we have to fill out for GA4GH for approval as well. Grant Orion anything you wanted to add to that?r
GW: Nothing for me. We're just really excited for this June 21 date we've been working on for a very long time. So we're getting very close.
OB: Second that yeah, fantastic
PvD: just a quick question. If we want to give feedback on this, do we do that through GitHub create an issue or
how?
LS: Yeah, so if you can create an issue just on the appropriate repo, that would probably be best. We'll see it there. However, if you prefer to send it through us, we have a Slack channel. It's in the GA4GH organization. It's hashtag CP dash pedigree. You could do there. Or also if you want to email it to me, and I can make sure it comes up in our next meeting, which is tomorrow, actually. So either route whatever works best for you. Okay, perfect. Thank you. Any other questions on that? Okay, I'll make sure I send an email to the full workstream once we kind of have everything together and it's ready for submission just so everyone can take a look again

as well. Okay, and the last update I had here before we dive into some of the roadmap stuff is just about the computable cohort representation group. As a bit of a bit of background, so computable court representation we're hoping is going to be you know, one of the new subgroups within clin/pheno, although it has not been formally proposed as, you know, a standard or we haven't officially defined that one output yet. So in at the GA4GH April connect, we had this computable court representation kickoff, which was kind of part planning of the meeting and deciding you know, how we're going to organize this hackathon. So we ended up coming up with five different topics, essentially one topic per week, over a couple of weeks. And you can find all of these listed in the master doc, which is also linked in the agenda here. Scroll down a little bit and you'll see the topic sessions here. So we had one about access and a darathon, and ones about beacon and data Connect. There was supposed to be one this week on phenotype matching tools, but we have postponed that one just because it's some conflicts that arose. And then next week, we're going to dive into CQL. And use of fire as well as terminology servers and mappings. And so if you haven't had a chance to attend any of these meetings, we have the action items, the recordings and the minutes all within this document. And then I do have an extra time for myself to put all of these action items in our Slack channel as well so that they'll all be kind of together in one spot. And so what we're hoping at the end of this is having explored all of these different tools and standards and things available to us, we can better define, what an output might look like. And yeah, what else did I want to say about that? Oh, if you haven't, you don't have immediate an invitation to one of these meetings, please let me know and I can forward that and I think the plan is after this last meeting next week, we'll get a report back or a retro that I'll invite everyone to we can explore some of the topics and findings from that and any questions about that before we dive into the roadmap? Good, okay. We are flying through here.

Okay, so the last thing I wanted to spend some time on today is about our draft roadmap. So I introduced this at our last meeting, but it's significantly more filled out than it was the last time. The structure is still essentially the same. You've got this motivation and mandate up at the front proposed solution and then these blue boxes which represent the different standards that we have in the work stream. We've got one for pedigree one for phenopackets, if you scroll down a little bit. And then a little bit further, we have the computable cohort representation one however, this one is still as you can see, got some gaps in it. And that's partially because as mentioned, we were still kind of working out what is exactly going to be. So right now I've kind of just put this in a to be iterated on section and not in the official roadmap per se that we will submit. These roadmaps are to be considered working documents. Just because we're putting in the two standards now does not mean we can't add anything at any time and in fact, I'd welcome anyone if you have ideas or projects you want to run, please don't hesitate to approach me or Melanie or David. If you can see if we can get something up and running in the work stream. I would say our biggest limiting factor in clin/pheno is just
having volunteer support to do some of these projects. But of course I'm happy to help with outreach and you know, any sort of project you're interested in. So, if you go right down to the bottom, I have a little spot for future roadmap items for discussion. Feel free to add ideas here or again to send me a note as well and after I run through this little bit, I can save some time if we want to have some open open floor space to talk about that. What I would really appreciate from everyone here is feedback on whether everyone agrees with this statement. So this is essentially our work stream statement - what we do this will be put on a lot of our you know, workstream pages and outreach documents. So if there is something you don't agree with in the statement, please let me know as soon as possible. Or add again, add your comments right in here. And then I don't want to go through all of the details of you know, every single standard here, but mostly to say, you know, if there's something you wanted the Phenopckets team to do, for example, we would love to know and maybe think about how we can incorporate that into our roadmap. Jules - about the journal interactions. Has there been any progress on that? I know we were talking about it being a potential goal, but I'm not sure where where that's gotten to.
JJ: We've had I think some sort of vague interest from one or two journals but nothing sort of concrete that they've actually said, oh, we'll definitely adopt these and start using them.
LS: Okay. So, do you think this is still a goal we would want to keep on the roadmap is something to pursue?

JJ: Yeah, I think so. Because it's really valuable part of you know, capturing the data that sort of that level but that that place but we didn't need sort of, I guess, I don't know, outreach, some some way of reaching the journals and convincing them that that's a good thing.
LS: Okay, so maybe we can talk about how we can plan for getting that to happen. Okay, that sounds great.
JJ: We'll have more things we need, you know, easier JavaScript tooling, maybe to embed into websites so that people can create feedback. It's easier than out of that rather than sort of expecting them to hack together a Yaml file.
LS: Yeah, that sounds great. I think we've got something about that web interface here. I got that from Peter. So I'm assuming that's probably what you're talking about. Yeah. Yeah. Okay. Great. So just as a bit of a heads up, I suppose. We will have to present this roadmap to the GA4GH steering committee by June 22 is the next steering committee. We've got just under a month to finalize our current roadmap. And so again, this is a roadmap that we will publish on our webpage do some outreach with so if there is anything, any comments, you know, put them in now. So I'm just going to pause here because I feel like I've been talking for a while. Are there any questions about the roadmap or expectations for it? Thoughts you have on other project directions we could take or if I hear silence, I'm gonna assume everyone is happy.
Speaker: I think we just want to sort of push adoption
LS: Whole section here for implementation support. Which I think is a theme we're gonna see with a lot of GA4GH deliverables that have already been published. It's, you know, how do we get people actually using them? What tools do they need to support them? Looking at interoperability with things like FHIR do we create fire implementation guides, that sort of thing. I think yeah, that's gonna be a common theme, as well as the more dissemination type stuff like doing publications doing webinars.
JJ: Yeah, those sorts of outreach materials. So I think definitely, we're just talking to people about it, because not everybody's heard of them. And, you know, there's there was what's interesting, the opposite. Pediatric clinical trials, I think they were interested in.

LS: I'm curious from everyone else on the call, you know, whether you're a Clin Pheno standards implementer or not, is there anything you think we need here that we haven't done yet? In terms of whether it's implementation support, you know, additions to standards, anything like that? If you don't have an answer now, that's totally okay. You can think about it and send us a note after mostly I just wanted to use this meeting to raise awareness that this is something that we have ongoing.
GW: Lindsay, one quick comment about the pedigrees section of the roadmap. Obviously, what we have in there right now, was something that we wrote a couple of years ago. Do you feel like we need to update this? So it also stalks talks about adoption?
LS: Yeah, this version that's here now does include some new content. It's not just a direct copy from the previous version that we had. So there is a little bit there is a section here about outreach implementation support, so things like a webinar from a publication, maybe a Connectathon. I've just put kind of guesses on dates here. So I'd welcome any sort of revision on that. But some of the proposed ideas are a webinar so once it's approved, introducing the broader genomics community to what it is and what it does. A publication I know in the group we've talked about turning that common dataset into a publication. So it could be something there as well as one on the standard itself. There was the idea about proposing a FHIR Connectathon to support the implementation guide, as well as just generally collecting use cases for a potential future version there as well.

OB: Lindsay, if there comments on this, or should people be making edits right here or in the working doc that's linked from the header of that?
LS: Yeah, good point. So what's in here now is a direct copy from that working doc. So maybe I would say just make it edits on this. I'll remove that link.
GW: I made that comment, Lindsay because I looked at the description there and the description looked like we had changed that much but that's no problem. The rest of it is updated and that looks great.
LS: If there is some of this description we need to change, please feel free to add the comments or suggestions right into the description. We've got for both phenopackets and pedigree here we've got listings of

proposed in the case of pedigree or current implementations. If you in your initiative or you know of an initiative that is going to be implementing or hoping to, please feel free to add them to this list here. So we can keep a good tracking where our standards are being adopted and where they might eventually be so we can provide that support if needed. So you'll see it here. Pretty much right before the milestone section. Okay.

IG: So we were doing some work with biobanks and certainly, the bio banks that I'm working with in Europe have started to expand their data collection into collecting phenotypes from directly from patient records. So one, they're a potential user for Phenopackets But I'm wondering with and isn't just one country, so this is all across Europe. I'm wondering whether on the roadmap, a focus on biobanks atleast in some form or another might be a good idea. But I'm throwing that out as a suggestion.
LS: Yeah. Are you thinking so this would be something specific to the Phenopckets roadmap or just kind of the work stream ingeneral?
IG: So I think it's the work stream in general, I think there is a link to Phenopckets so potential implementation sites for Phenopckets. Yes, but I'm thinking more in general terms about really about the way bio banks are using data potentially using GA4GH standards and having a big focus on where you're on phenotypes and on the clinical side and on the terminology that goes with it.
LS: Okay, I've added a comment here to our proposed solution. Do you have any comments about sort of the Phenopckets approach? I know we have biosamples, for example, that uses it as well as the gem Japan, biobank and it uses Phenopckets. Because I'm thinking, so the way the roadmap is structured is it's by deliverable. So we could put a general comment, you know, in the proposed solution, like what we'll focus on biobanks, or you know, clinical systems or whatever makes sense there. But I think where we can add, some of this is actually for in the specific deliverable sections. You know, whether it's a specific webinar or for biobank audiences, for example.
JJ: Yeah, we'd love to talk to some the UK Biobank people and it could be useful for them.
MC: We had talked a little bit with BBMRI Eric on the cohort front that maybe I'm always interested.
GW: Okay. Quickly, I just want to share quickly that since we're talking about this idea, which I think gain is a great idea. I know Intermountain Healthcare has a great biobank and they're actually moving to try to generate some some revenue from you know, selling access to that biobank from outside organizations. And so I would think, if they're gonna do that, to have the capability to attach fino packets to it would greatly enhance the value of their service. So please include me in any discussion. If you want me to help approach the inner mountain commercial biobank about fina packets and a pilot I'd be happy to do that.
IG: I was gonna say the the other group that I'm dealing with directly use the Catalan biobank network. So it's a network of a number of buybacks and they've set up a fire terminology server to serve the network. So they're building up connectivity between a number a large number of buybacks.
LS: That's great. And considering we've got our plenary in Spain, there's maybe some connection we can make there. Okay, thanks, everyone. So I'm just noting we're at our 30 minute time here. Any other comments questions from anyone otherwise, action item for everyone is to take a look at this roadmap and add your comments, suggestions and so on as well as to the pedigree standard. That's preparing for approval. So I'll pause here for any last questions comments.
MC: Quickly for Ian. I was just trying to find the Catalan biobank, do you have a direct linking
IG:I'll send you via email
JJ: If you want to put them right in my direction or ask if they want to have some kind of an introduction to Phenopckets as to what they are then. I'd be happy to do that.

Global Alliance
for Genomics & Health

2022-05-xx - Computable Cohort Representation Hackathon

Following on from the Computable Cohort Hackathon Kick-off at Connect, 5 topic meetings are taking place throughout May to dive deeper into the standards/tools/strategies that could support this effort. This master doc can be found here.
Please reach out to lindsay.smith@ga4gh.org to learn more.

2022-04-28: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Chen Chen (GA4GH), Meen Chul Kim (KF/CHOP), Bing Feng (UUtah/BRCA), Katherine Johnston (H3Africa), Steve Laurie (CNAG, EJP-RD)

	Actions Arising	Assigned To	Deadline
1.	Create PR for outstanding changes in the Classes section of the documentation	Meen Chul Kim	
2.	Add maternal/paternal terms to the KIN ontology - Sub-action: reach out to people potentially experienced with protege ontology editing - David/CSIRO, Melanie, Katherine's colleague	Orion - Lindsay	
3.	Add updated pedigree ecosystem diagram to the docs - Check in with Stephanie about "beautifying" it	Lindsay	
4.	Add PRC member suggestions to the timeline document	ALL	
5.	Cancel next meeting - Instead, request updates and check in on progress via Slack/email (apologies in advance for the hounding)	Lindsay	
6.	Update the Tooling for BODICEA export?		

7.	2022-2023 Roadmap	ALL (Lindsay to follow up on this)	

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - Montreal Connect Report Backs - GA4GH-wide feedback opportunities: Town Hall Meeting Registration, Community Survey (due May 13)	Orion, Grant, Lindsay	5 min
2.	Updates: Documentation - What sections are still outstanding? - Pedigree Ecosystem Diagram	All	5 min
3.	Updates: Extending the conceptual model with elements from common dataset - Progress?	All	5 min
4.	Next Steps/Outstanding Tasks	All	10 min
5.	Overview of product approval submission form - Proposed Timeline	Lindsay	10 min
6.	A.O.B. - Schedule of meetings moving forward	All	5 min

Minutes:

Documentation Updates
OB: Outstanding PR to merge and the docs will be more representative.
MCK to rebuild (done!)

Docs -Classes Page
MCK:Will create a PR for outstanding changes. At our last technical meeting we decided on a few changes
OB: Great feedback from TrakGene about adding some more detail and additional examples. The way the terminology is set up, we have terms for aunts and uncles, but not maternal or paternal aunts and uncles. They have questionnaires where it is possible to have information about the grandparents but not the parents. Questioning whether we're sure if we want these, whether to have a node. Small change either way. As a consumer, fewer ways of representing things.
GW: What we have done is optimized that design, removal of certain relationships where it's fully known how they would be related. Do we want to tell the world to go forward and follow what we've done as opposed to use being backward compatible?
OB: How much we meet people 100% where they are. We're somewhere in between the spectrum.
SL: In the example where you have the information with the child and the grandparent, do you make a ghost mother?
OB: Yes, create an individual there you would have not other info on

GW: What would be the most strategic decision here?
OB: It causes a bit of friction with people who have the data already. Creates a little less friction for those on the consuming side. I can see both sides really clearly. There's a rationale behind saying it's intended for this reason. We will be presenting in a world where it aligns with the FHIR valueset.
SL: What do you do when you don't know what side the grandmother is on?
OB: There are terms for ambiguity. It has grandparent, but not sex, or maternal.
GW: Is it possible for us to develop something that allows for both paths. Can we allow them to follow that maternal path?
OB: Lowest friction option is to just add terms, but it adds more options for the way things can be represented
SL: What about incestuous relationships?
OB: Can just add more edges with the graph structure.
GW: Easiest thing is to add the terms.
Lindsay to reach out to Melanie about protege ontology editing. Also check in with David re: folks at CSIRO.

Lindsay to re-add ecosystem diagram to docs.

Common Dataset Updates
OB: Relates to what Meen was talking about with updating classes,we discussed this at the technical meeting. Make sure everything is synced up.

Timelines for Submission
June 21st seems reasonable for submission. Need to have all docs, model and PRC ready for this date.
PRC member brainstorming in the timeline doc - please add any suggestions for members in the doc.

If we can reproduce the hounding we have with in-person meetings, let's cancel next meeting.

2022-04-20 - GA4GH Montreal Connect

Please find the meeting report here.

2022-03-24 - Computable Cohort Representation Meeting

The minutes for Computable Cohort Representation can be found here.

2022-03-23: Clin/Pheno Full WS

Chair: David Hansen, Melissa Haendel
Attendees "Name (Project Affiliation)": Grant Wood, Lyndon Zass (H3Africa), Marion Shadbolt (Australian BioCommons), Chen Chen (GA4GH), Orion Buske (PhenoTips), Alex Tsai (GA4GH), Peter Robinson (Jax/Monarch), Ian Green (SNOMED), Jordi Rambla (CRG-EGA), Ada Hamosh (MME), Ron Miller, Brandon Wright, Nicky Mulder (H3Africa), Melissa Cook (NIH), Shahim Essaid (Monarch), Katherine Johnston (H3Africa), José F. Reyes, Nebiyou Petros, Freddie, Mogomotsi Matshaba (H3Africa)

Global Alliance for Genomics & Health

	Actions Arising	Assigned To	Deadline
1.	Please review and add your comments/thoughts/likes/dislikes to the Clin/Pheno roadmap draft	EVERYONE	End of April
2.	Connect with Lindsay if you'd like to join Computable Cohort Rep Hackathon Planning	Everyone	First meeting is March 24, second meeting is April 5
3.	For a future meeting - presentation on Beacon v2.0 phenopackets integration		
4.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH-wide Strategic Roadmap feedback opportunities - Town Hall Meeting Registration - Community Survey - New Work Stream Leadership Announcement - Montreal Connect (April 20-22) - Clin/Pheno will be hosting a *Cohort Rep Hackathon* - please let Lindsay know if you'd like to help with planning - GHIF Newborn Seq Workshop (April 5)	Lindsay, David, Melissa	5 min
2.	Cohort Representation Presentations - ~~CINECA~~ (postponed) - H3Africa	- ~~Mamana Mbiyavanga~~ - Lyndon Zass	30 min
3.	Clin/Pheno Strategic Roadmap Refresh - Draft Roadmap	ALL	25 min
4.	A.O.B.	All	

Minutes:

H3Africa Cohort Representation

Lz: H3africa is one of GA4GH's driver projects. Several genomics projects running across H3Africa. Compelte cohort over 100k research participants. H3ABionet ensures African data is FAIR, implementing standards within the consortium. We're using GECKO to describe the depth and breadth of H3Africa cohort, make the data more findable. Aiming to map metadata (cohort descriptors). Mapped in two ways - automated and

manual. Successes: mapped all of the phenotype to GECKO, the standard CRF was used. GECKO mapping completed for half of the H3Africa projects. For the most part, automated mapping was successful. Challenges: some minor gaps identified in the ontology itself. Challenges with sharing of data dictionaries and case report forms across H3Africa projects, incomplete mapping for some projects. Next steps: get a broader view of where things have been done.

NM: H3Africa is reluctant to share a lot of their phenotypes. We'd like to demonstrate to the public that the cohorts are much richer than the datasets in EGA. Want to search the data in the EGA, and also have a broader view of what else is collected and sitting with the PIs. Trying to show in a summary way, this the number of projects HIV status, or other phenotypes.
GW: How are you collected the FHH information and what analysis do you use it for?
LZ: Being collecting by some projects, not all. DDD-Africa and a rare diseases group. Collected a bit differently, the rare disease group is limited in what they capture. Can't speak much to what they do in the analysis piece.
MS: The automated approach, are there's scripts available?
LZ: Yes, though I don't have the script, but will share that.
JR: How many terms have you been able to map?
LZ: Phenotype data is quite broad. Have been able to map most lifestyle factors. With project metadata its' more limited, but gecko allows us to map the type of experient that was conducted. Don't have specific numbers.
JR: How many rows in the master spreadsheet?
NM: GECKO doesn't have that many terms, so maybe under 100.
LZ: 143 rows, but only looked at a subset of codes. Will look at variables to group things together.

JJ: With the codes, did you link to the cohorts themselves?
LZ: Haven't done yet, will be looking at that next.
JJ: Might be useful to take aggregate codes at a certain level, know about useful groupings.
NM: With using an ontology, can go to the different levels in the branches. It's metadata, so it's less identifiable at an individual level rather than a cohort level. Don't have the values of each.
Cohort representation will help, because it will have those study levels and how to get to that next level of granularly.

Roadmap Planning
[Lindsay shared screen and walked through the roadmap]
Please review the roadmap document and add your thoughts, comments, things you like/dislike, etc. There is a section near the bottom to add in thoughts about future roadmap ideas.

OB: Does it make sense to make implementation more general, and where do we want to put actual implementations.
LS: Generally, implementations that are not done by the workstream (eg. an initiative's implementation) are not included in the roadmap. They are listed in the planned implementations section. We can add implementation support content like webinars/workshops, publications,etc.

PR: The IG itself is complete, and what we're doing now is writing HAPI FHIR Java code that provides a Java equivalent for the classes. The purpose is to use FHIR to transmit and translate that back into a native GA4GH phenopacket. Can leave this on the roadmap as something completed in Oct of this year.

Implementation support for Phenopackets?
PR: Had a great conversation with GA4GH staff on implementation support, waiting to hear back. Working on phenopacket creation tools that use text mining and quality control to create phenopackets. Have some good desktop tools and are now working on a REST API frontend that might be useful.

Have had a great conversation with the Beacon team. Idea to have Beacon v.2 accept queries that include phenopacket bits and optionally return a list of phenopackets that match that query or a yes/no. Looks like we will be getting a supplement to implement code that would support that API. Would be interesting to include Pedigree. Would be great to have a call with Beacon and Clin/Pheno.
JR: What Peter is referring to is Beacon query by example, where you provide an example of a record then you look for records that match these kinds of examples.
One thing I'm missing in this roadmap is where the implementations from Driver Projects go in the roadmap. Not a deliverable from the WS itself but it's very important. If not in the roadmap it should be somewhere.
LS: Agree, we need these listed somewhere. Some are listed in the implementation's section of the roadmap, and I think the tech team has been working on some kind of repository to track implementations.

JR: Gathering feedback from implementations so we can focus on the next version. Some kind of outreach or report.
JJ: Mostly have Github for feedback, might be the easiest thing
LS: I'll have more of a think about that. Github issues are great but there could be a more accessible way of doing things

2022-03-17: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Zornitza Stark (AGHA), Chen Chen (GA4GH), Alejandro Metke (CSIRO/AGHA), David Hansen (CSIRO/AGHA), Melissa Cook, Steve Laurie (late and lurking for RD-Connect), Katherine Johnston (H3Africa), Alex Tsai (GA4GH), Meen Chul Kim (CHOP)

	Actions Arising	Assigned To	Deadline
1.	Next meeting will be a working meeting. Set up reminder closer to the meeting date	Lindsay	
2.	Review the content in the Roadmap planning doc and revise as appropriate. - Detail about the standard needed in the description - Confirm source DPs and other contributing initiatives are correct	ALL	
3.	Add any additional content on the KIN ontology or FHIR IG to the docs	Alejandro	
4.	After next meeting (the working meeting), decide on which Steering Committee to aim for		
5.			
6.			

Global Alliance
for Genomics & Health

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - Montreal Connect (April 20-22) - GHIF Newborn Seq Workshop (April 5) - Lindsay is OOO next meeting (Mar 31st)	Orion, Grant, Lindsay	5 min
2.	Updates: Documentation - PRs #18 & 19 merged, new build of docs available: https://pedigree.readthedocs.io/en/latest/ - Technical Model pages	All	5 min
3.	Updates: Extending the conceptual model with elements from common dataset - Confirming that we will be using “common” dataset, doc titles have been updated - Revision of the introduction - REA progress - Steering Committee Update (June 22nd)	Orion, Alejandro, Grant, Lindsay	10 min
4.	2022/2023 Roadmap Refresh - Pedigree Standard Template - Would ideally like to have a draft ready for **March 23rd** (next full WS meeting) - Overall WS Document	All	35 min
5.	A.O.B.	All	5 min

Past Tasks:

- Merge PRs 18 and 19. Once complete, rebuild doc site
  - Lindsay, Meen
- Interested in participating in the proposed Testbed Session at Connect? Please let Lindsay know
  - All
- Confirm with Grant that we’ll use “recommended” instead of “common” moving forward (for this doc)
  - Lindsay
- Add additional field for relationships in the Phenopackets Family message
  - Jules
- Take a stab at including REA information in the conceptual model
  - Orion
- Circulate a comment box/table for roadmap updates prior to next meeting
  - Lindsay
- Share updates on developing an RDF, when ready
  - Soichi
- Plan for submission to GA4GH PRC by MAY
  - ALL

Minutes:

MCK: Going to update the classes page about updating the type and adding more elements on the common dataset
OB: Sounds great, happy to help coordinate on that. Just pulling from the doc and looking at how that might map
AM: +1

OB: Made a new PR for the docs (https://github.com/GA4GH-Pedigree-Standard/pedigree/pull/23). A lot of it is the minimum amount of information, getting to the root of what we're talking about. Summaries of key concepts, will replace with links to classes. A bit of a description of relationships. Alejandro, if you have more information on the IG you'd like to add, let me know.
AM: Sure, I'll have a look and add something
OB: Once we've updated the model, we can update this and potentially add the phenopacket and FHIR version. Does it make sense to add them all, tiptoeing the line between how concrete we are making it or not. There are a few pieces I'll want to coordinate with Meen and Alejandro on. Give people a realistic sense of what this look likes.
AM: For example, would be great to have something like those tutorials with tabs for multiple languages.
OB: Is that possible in readthedocs?
AM: Could imbed html. We can take a look.

GW: You can spend time creating the standard and then there's the marketing of getting that standard know. Common - was thinking about giving it that marketing name, this is a particular product. The recommendation part was to think that the outside world doesn't incorrectly interpret what the document is about. Meant for a wide range of stakeholders.
Because we went though the tables and pointed out that some elements were not part of the conceptual model, and we highlight ped, hl7, the text in the introduction need to be revised to fit that structure.

OB: Seemed like there were two different approaches to REA I wanted to run by everyone. Given there isn't an overall appropriate ontology, treat it as an ontology and recommend hancestro as a first step. There were a few competing options for REA ontologies, on that last call we decided it still makes sense to put a placeholder in there. As long as it encourages someone to take a term, or allow a free text to allow a variety of census data use cases.
GW: When we were working through this question, there were many out, and there were conversations that people preferred different ontologies. Have a code for the different ontologies?
OB: We're not going to force people to use Hancestro
KJ: the H3Africa Ethno-Linguistic Ontology is in the final process of stepping through the OWL requirements and have website (just struggling to find link at present) but it promises to be a vastly superior way to identify African ethnicities
AM: In the FHIR implementation, FHIR is very explicit with this through bindings. There are different categories of strength for bindings. Required closes the door to anything else. Start with a recommended.
OB: Is it possible to have multiple recommended?
AM: Haven't seen it. Could include a value from multiple codesystems, just not as clean. Hancestro already in FHIR.

KJ: can you set up front your choice of ontologies and that gets ported out as description of the flexible components you have used in your pedigree dataset? If there are these removable components such as choosing a different ontology, if you implement the pedigree common dataset, is there something that goes along with it like the metadata?

OB: When you define a term in the pedigree, it would say “this concept from this ontology as laid out by this identifier”.
KJ: Imagining implementing in redcap where you have an upfront setting. If you’re just going to associate with individual ontology code in every instance it doesn’t really matter
OB: functionality of the tool doing the collection. When the data is actually transferred, then just use the codes.

OB: Aim to have a work session prior to next meeting to see if bumping back or forward. Make a decision based on this.

2022-03-15 - Computable Cohort Representation Meeting

The minutes for Computable Cohort Representation can be found here.

2022-03-03: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees “Name (Project Affiliation)”: Chen Chen (GA4GH), Alex Tsai (GA4GH), Bing Feng (UUtah), Alejandro Metke (AGHA/CSIRO), Jules Jacobsen (QMUL/Monarch), Mamana Mbiyavanga (H3Africa), Soichi Ogishima (Tohoku, GEM-Japan)

	Actions Arising	Assigned To	Deadline
1.	Merge PRs 18 and 19 - Once complete, rebuild doc site	Lindsay - Meen	
2.	Interested in participating in the proposed Testbed Session at Connect? Please let Lindsay know	All	March 18th
3.	Confirm with Grant that we’ll use “recommended” instead of “common” moving forward (for this doc)	Lindsay	
4.	Add additional field for relationships in the Phenopackets Family message	Jules	
5.	Take a stab at including REA information in the conceptual model	Orion	
6.	Circulate a comment box/table for roadmap updates prior to next meeting	Lindsay	March 11th

7.	Share updates on developing an RDF, when ready	Soichi	
8.	**Plan for submission to GA4GH PRC by MAY**	ALL	

	Agenda Item	Speaker	Time

1.	Welcome, Housekeeping
	- Montreal Connect, April 20-22
	  - There will be an in-person component!
	  - Call for workshops
	  - Proposed Testbed Session
	- GHIF Newborn Seq Workshop (April 5)
		Orion, Grant, Lindsay	5 min

2.	Documentation Updates - Merging PR#18 - Merging PR#19 - Updates to classes section - Technical Model pages - Aligning with Phenopackets Docs	 - Lindsay - Meen - Orion	5 min
3.	Extending the conceptual model with elements from the common dataset	Orion, Alejandro	15 min
4.	RDF Model - From Soichi: We are starting development of RDF data of pedigree in the Tohoku Medical Megabank project (22,000 pregnant women/mothers, in total 73,000 family members in our Birth-and-three generation cohort study)	All, Soichi	10 min
5.	2022/2023 Roadmap Refresh - Pedigree Standard Template - Would ideally like to have a draft ready for **March 23rd** (next full WS meeting) - Overall WS Document	All	20 min
6.	A.O.B. - Redraft of pedigree ecosystem diagrams - Connect Workshops - Do we need another FHIR IG discussion?	All	5 min

Past Tasks:

- Send reminder to the Pedigree mailing list to review acknowledgements/affiliations
  - Lindsay

- Update the ‘Classes’ Section of the documentation and send a ping when ready for review
    - Meen
- Review PRs #18 and #19 and add any comments/suggestions prior to next meeting
    - ALL
- Follow up with Jules about Phenopackets implementation
    - Lindsay
- Extend the conceptual model with elements from the common dataset
    - Orion, Alejandro
- Revise the pedigree ecosystem diagrams to remove ambiguity about the “third model”

Minutes:

Doc Updates
OB: Started a skeleton of the pages I’ve been working on. Fleshed out the example in one of them. Start them as draft PRs, but as always, feel free to add comments. One placeholder is what makes sense to add around the use of this within Phenopackets.
JJ: For things like concepts and identifier, will just say its a thing to map to Phenopackets?
OB: Yes, use native phenopackets.
JJ: Some concepts in phenopackets - we have enumerations, that have been mapped to actual concepts in NCIt. FHIR I don’t think has the same mappings. We went with phenotypic sex description, gender was moer self-identification. Other difference was with life status, we’ve mapped that to an enum in a class as well.
OB: Question is there a more native way of representing in pregnancies or fetuses. We’re going to back out of a few specifics for Meen’s PR, the implementation of the concept would not be within the scope of the conceptual model.
JJ: Would be helpful for Pedigree model to say, when we mean male, this maps to this FHIR and this in Phenopackets.
OB: Put that as the description in Meen’s class file. Lighter on specific of string or int, and more on definition.
JJ: Still need to go to string to int
OB: As long as there’s another entity in phenopackets or FHIR that handles it.
JJ: FHIR doesn’t map sex as a concept?
OB: Either way, in the context of this, can’t assert representation in FHIR, but can we say they’re good for interoperability purposes.
JJ: If you look at the sex documentation (https://phenopacket-schema.readthedocs.io/en/latest/sex.html), that can be mapped to the NCIt ones.
I was thinking implementing with the org.ga4gh namespace for pedigree, so can include in the phenopackets schema with a separate namespace. Work with the objects as they come in. Should we be replacing or adding as an optional one or the other as a Family message.
OB: Ideal to be it have one or the other
JJ: No drawback really.
OB: In ideal world wouldn’t need a separate representation in the context of an individual. Individual model in phenopacket is quite robust, only thing is can you represent fetus or pregnancy through it.
JJ: yes we can, we have gestational ages
OB: Any placeholder for REA? Codeable concept but attached to an individual?
JJ: Didn’t end up including it. Could record in the pedigree field, but not individual. Would be a point update to take from pedigree model and add into individual.
OB: Replace ped row approach with a list of relationships and slightly more fields at the family level. Won’t work in the context of the pedigree person. Needs to replace existing pedigree model completely.
JJ: would then get cycles? Family has a pedigree and a list of phenopackets as individuals. Could still do this, they just might get out of sync

OB: As long as we don't mess with a Phenopacket. We need to make sure a sufficient amount of the pedigree as a whole is in the Family model. Really just replacing Family pedigree type with a different type. Gets us 95% there, for a first version that's great.
JJ: Nice way of keeping them together in a separate way. Don't need to replace, just have an additional field with relationships, alongside pedigree.

OB: Sounds like we need more discussion here. Two options: 1) making the family pedigree support an option for a non-ped pedigree, 2) adding something like family.relationships in which case the way in which that interacts with the pedigree field might need to be clarified somehow
JJ: Probably easiest to say GA4GH pedigree or ped pedigree. Affected is the only one that doesn't really map.
OB: Almost all use cases are being more explicit about this anyway. If there's a particular concept that is affected in some terminology, is there a parent one in ORDO or somewhere like that?
JJ: Not sure
OB: Where should this guidance live? What needs to happen?
JJ: I'm fairly certain I had a good play with this already. Sitting in a branch somewhere but have not merged it. Might be on a stale version.

JJ: Could create exactly what you've done in the conceptual model
OB: Probably adds enough additional complexity to what a phenopackest looks like because it already has a lot of that natively. May be a gap here or there that we can improve over time.
JJ: Interesting to see how that will work with the Phenopackets FHIR IG.
OB: Goal here is to use all of the native Phenopacket entities wherever possible and add value in bringing relationships in.
JJ: really only need to add one new class in that case

Aim for end of April early May for documentation/ready for approval.

Stick with 'recommended' for document naming - run by Grant.
OB: Taking a stab at updating this extensions. Trying to understand at what level of resolution would be helpful to share with people. Balancing between two pressure - one making it mean something beyond the document, but also not trying to give it the same level of specification as a format. Want that sway in terms of the implementation of that model. Get feedback on what level people think would be best?
JJ: The layout with field names was useful, but didn't know how conceptual this needed to be. Has to be some limit to these things.
OB: Having a minimum data about these entities and relationships between them. Would love to get clarity if REA is something to bring into the conceptual model now. I see options as either elaborating the description to justify it not being part of the model, but leaving it in the table. Or, removing it from the table but leave the paragraph that explains why we're not making recommendations at this time. Used by a lot of risk analysis tool. Probably something we should have. Important enough to take stab.
JJ: Hard part was defining all of the different things that can be put in there. For interop purposes, would have to do mappings with local REA fields.

RDF Model
SO: Work on the representation of our cohort family relationships in RDF. We need an ontology so would like to use the GA4GH terminology to represent family relationships.
AM: Almost done with all of your comments.
OB: Would love to hear feedback as you start using the ontology, if there are any other pieces of feedback that you end up with.
SO: I think we finish in 3 months or so
OB: Anything else you'd like to see from us to assist?

SO: Nothing at this time

OB: Would it be worth adding to the documentation the use with RDF? Can we highlight this project in any way?
SO: Would like to feedback our experiences
OB: It sounds like we could potentially include in implementations in documentation
SO: Yes, would like to contribute
AM: FHIR also has an RDF rendering, so would be interesting to compare with what Soichi develops.
OB: You are planning to use the kinship ontology in a larger RDF model. It would be interesting to see what that RDF model looks like. Is it based on FHIR RDF? Or separate mode?
SO: Making a separate model, but need to harmonize. I would like to share my ongoing RDF model development work with you to get feedback. It would be very helpful for us.
OB: Sounds like a third ecosystem to align pedigree information, in addition to FHIR and Phenopacket. Another place to guide people with interoperable alignment. Please share the models when they are developed, we would love to take a look.

2022-02-23: Clin/Pheno Full WS

Chair: David Hansen, Melissa Haendel
Attendees "Name (Project Affiliation)": Orion Buske (PhenoTips), Ian Fore (NCI, NIH), Jordi Rambla (CRG), Marion Shadbolt (Australian BioCommons), Grant Wood, Alex Tsai, Chen Chen (GA4GH), Alejandro Metke (AGHA/CSIRO), Steven Laurie (CNAG-CRG, RD-Connect), Mette Peters (Sage Bionetworks), Francesca Frexia (CRS4/BBMRI-ERIC), Melissa Cook (NCI), Brandon, Anita Rauch (MedGen, UZH), Mogomotsi Matshaba (H3Africa), Shahim Essaid (Monarch), Peter Robinson (Monarch/Jax), Aly Khalifa (Mayo), Jules Jacobsen (QMUL), Bimal Chaudhari, Jose Reyes, David Hansen (AGHA/CSIRO)

	Actions Arising	Assigned To	Deadline
1.	Organize a meeting to demo Beacon v2.0 to Phenopackets team	Jordi	
2.	Potentially plan for another IG discussion at April Connect	Lindsay	
3.	Connect beacon/search/phenopackets use cases in Cohort Rep hackathon		
4.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - Call for Google SoC Projects - Public Comment for GA4GH Product Approval Questionnaire - due Feb. 24th	David, Melissa, Lindsay	10 min

	- Clin/Pheno Lead Nominations
	- Montreal Connect, April 20-22
	  - There will be an in-person component!
	  - Call for workshops will be opening soon
	- Welcome new members!		

2.	FHIR Implementation Guides Alignment: Status Reports & Next Steps - Phenopackets IG - Pedigree IG - Vulcan & OHDSI Partnership	- Peter Robinson - Orion Buske & Alejandro Metke - Shahim Essaid	35 min
3.	2022/2023 Roadmap Refresh - Clin/Pheno Working Template	Lindsay	10 min
4.	A.O.B. - Cohort Representation & FHIR - Beacon v2 PRC alignment request - Connect Workshops (if time permits)	All - Susheel - Jordi - Lindsay	5 min

Minutes:

Phenopackets IG
PR: Phenopackets manuscript has been submitted to Nature Biotechnology - https://www.medrxiv.org/content/10.1101/2021.11.27.21266944v1, will be resubmitted as a Correspondence.
One of our goals is to integrate rare disease clinical care into EHR. The implementation guide of Phenopackets is not yet to cover all aspects of the Phenopacket but to cover phenotype driven diagnostic exome or genome sequencing. If ordering an exome can provide a description of phenotypic features, most people use HPO for that. Most programmes for that are manual so it would be great to have something to extract from the EHR, send that over to the lab in a structured form so they can run a program like Exomiser with a phenopacket. Exomiser and other similar tools can ingest phenopackets.
The phenopacket-schema repository in Github is the official repo for the shema. The schema is written in protobuf but think of it as a JSON schema. The core IG repo is the FHIR IG (http://phenopackets.org/core-ig). It's written using FHIR shorthand (FSH). The published version shows all of the artifacts. An IG is a way of specifying how to use FHIR to achieve a certain goal. In most cases, you want to restrict the elements used in an existing FHIR resource. For the GA4GH individual element, this matches well with the FHIR patient resource, but is missing KaryotypicSex.
Have developed some Java packets that will create a FHIR phenopacket, send it to a server, perform searches on the server, retrieve that back and reformulated as GA4GH phenopacket that can be used for research.

Pedigree IG
OB: https://github.com/GA4GH-Pedigree-Standard/. The documentation is not too far along yet. The pedigree standard is intended to build upon ped and other file formats in terms of exchanging family health history and pedigree information. Graph based model with relationships modeled independently. Works in the context of both FHIR and Phenopackets. Built tooling to support interoperability between pedigree and other formats.
AM: The IG is a mapping of the conceptual model in FHIR. FHIR already had some resources dedicated to representing family health history information but they didn't quite map the graph model we're proposing. In the

documentation you'll see a description of all of the resources we've profiled, as well as some simple examples using the Simpsons family. One of the key artifacts was how to represent relationships in the graph. Ended up repurposing the family member history resource. Key thing was using an existing extension to represent a patient. Finalizing a recommended data set. In the appendix there is a graphical representation of the existing family member history. After some discussion we decided it would be best to come up with our own ontology for relationships. Binding the relationships in the family member history profile. Idea is that we can use logic reasoners for validation and infer new knowledge. The ontology in OWL can be transformed into a FHIR codesystem and tie in with the IG.
AR: Wondering about the concept of AffectedStatus under investigation. Can only use for one clear disorder? Can there be different HPO terms?
AM: The reason is to have backwards compatibility of ped. The idea of having this in FHIR is that you can add any number of additional information but it's not covered in the implementation guide. Phenotypes would need to come from some other resource or IG (Phenopackets).
IF: As we put together a hackathon for cohort representation, I went in search of some asthma studies. At Least some datasets had pedigree information. Is there a potential here? Use as a test for whether this standard works with real data.
AM: That would be great, always trying to test out with real data.
IF: link here: https://github.com/ga4gh/fasp-scripts/blob/master/notebooks/search/Asthma%20tables.ipynb

Vulcan Project
SE: The Vulcan effort is about acceleration the use of FHIR in medical research and other new environments. Phenopackets was proposed as an use case two years. A year ago we participated in a connectathon where we demonstrated a simple use case where we took existing LOINC laboratory data, a LOINC-HPO library and a simple user interface to curate existing LOINC medical results into HPO base phenotypes. We sent them to a FHIR server and had a simple UI to read them back and review them. The representation of the LOINC information as HPO phenotypes were based on an earlier version of the Phenopackets IG, so we need to re-review. We're trying to implement a more realistic use case where we have existing clinical information already in the EHR, and EHR apps collecting data (like PhenoTips), have that information represented in FHIR, sent to a central FHIR server. A lab will be able to read the request along with the phenotypic features collected in order to perform an analysis, write whatever can be represented in FHIR back to the server, and then bioinformatics tools can have access to the information. TIMS (terminology service) is a new group trying to understand what additional requirements you need to support this efforts, you need HPO in a production manner for example.
The idea is to implement this in a future connectathon.
JJ: (Sarah): Trying to make it that phenotype data can travel test orders for genomic sequencing and be shared downstream in a repo for cohort analysis, etc. In the NHS there is a lot of concern over information governance. How do you see access and information control?
SE: Not addressing that right now, mostly focusing on how to take the information we're interested in and represent in FHIR. With the data, looking at a bottom up approach rather than start with an IG and decide what is the right profile. Just take the existing data and fit it in the base spec and write examples. Won't get validation without an IG, but starting to understand what data elements we have and where to place them. The hope with the FHIR server is that it will be a stable place to share/review JSON representation of data. Can think gradually about what is in PHI and where. That is likely well addressed in the clinical FHIR world and we need to learn from them. Security is a real concern in the clinical world.
SJ: A clinician is going to order a test, and the data is captured when the test is ordered. I see that it could be a possibility to pull it as late as possible, info is being added to that record continuously.
SE: Yes, all issues to address. I understand that there is a need from lab folks to collect info on their, but also for the clinician to assert the request. Everyone has to be independent and put a line around what they're responsible for, and that's where documentation and FHIR composition is important. Many aspects, and not any one is the only way of doing things. Need for a packet from whatever system or clinician. Use case is a

one direction - initial intake that gets collected and sent to a FHIR based repo, with the lab sending information back.
OB: Sarah, is the workflow you're talking about in the context of clinical care, like doing another analysis separate from the test order, or in the context of a research reanalysis?
SJ: mainly thinking of the former, a clinician orders a test, they don't get a result, time passes and maybe some research project has gone on and new variants of significance have been found. Don't want to re-send the sample for sequencing, just want to reanalyze the data already there, but also want to make sure patient data is up to date. Want a fresh look at the data, theoretically should be the same mechanism.
OB: depends on whether it's triggered by some upstream or downstream process. As long as it's being synced to a lab that has access to a repo it's okay.
SJ: My assumption is that a clinical thing will always be driven by a clinician, but a research thing may not be
AR: If the updated clinical information is then noted with a timestamp? Or is it all one input?
SE: In the clinical world, information doesn't get updated in that sense. A note is done during the visit and that is permanent. The information doesn't get revised, just additional information that accumulates.
AR: If there is a disorder developing over time
SE: The medical records will have snapshots
PR: The purpose of the phenopacket is really to be a bioinformatics representation. The raw FHIR is not arranged in a way to immediately use for analysis. Have to make some decision for analysis whether something is present or not if mentioned previously.
SE: What is the packet. It spans time and the individual in some cases. A layer that is somewhat abstracted. The packet has to be asserted by someone, that's a clinical decision if someone has a feature. Depends on how records are maintained, it's not straightforward. There are all of the modeling issues we need to try and understand. The OHDSI piece is a reference from the TIMS group.
AM: Question about terminology - you want to be able to access HPO via FHIR. What's the status of that? We've been using HPO in FHIR through some infrastructure that generically transform, but would prefer the official release.
PR: Working with Shahim's group to determine the best way of putting HPO in place. Also working on a HAPI FHIR model of the phenopacket. Essentially have to write a Java class that derives from the patient class. Going to need another month or two to finish it up, but welcome your involvement or comments.
SE: One piece I'm working on is trying to figure out how to use HAPI server. There is some coding and understanding to do there. Goal is to stand up a HAPI instances that behaves in a similar way to the default HL7 FHIR one.
IF: On that last point, is there a FHIR HL7 position on who else might stand up terminology servers and why? One thing we run into is when to extend the model. There is a reluctance to allow extension of the model, and does the same apply to terminologies.
SE: There is a need to stand up the server, and HAPI does not do it by default. In terms of thinking about new resources from scratch, that's a major effort and they don't want folks doing that. The TIMS group is trying to understand code mappings in FHIR. R5 has a new model called conceptmap2. At least two different ways to represent mapping.
DH: There are a bunch of FHIR servers are available -
https://confluence.ihtsdotools.org/display/FHIR/Features+of+Known+Servers
Tells you what features, APIs, etc. that they implement. R5 version almost ready.
IF: This is service software, not individual servers?
DH: There's links to where some of them stand up. HAPI doesn't really have a separate FHIR server, as Shahim mentioned.
SE: Needs some customization.

SE: How to start collaborating across these different efforts moving forward? How to regroup and have a bigger community around it? Make sure we're aware of everything out there.

PR: Be good to think about a common Github software implementation of a lot of these things.Each project has mixed rules and hard to see where they interface. Figure out core software artifacts that need to be developed, make is easier to do pull requests.
SE: In another call, talk about who is using what and where

JR: Coming to you with a request! Given that Beacon v2.0 is using standards from GA4GH, want confirmation that the developers of these standards are happy with their use. Tried to be agnostic to standards, and allow them to be plucked. In this case, involves Phenopackets. What are the considerations?
PR: Be good to have more practical use cases. Would love to hear more about that?
IF: Is there a timeline for this?
JR: We should have recommendations addressed by the end of March.
JJ: Could we have a meet up about the implementation and you can demonstrate? Probably the quickest way.
PR: +1. Would be good to discuss with Connect API.
IF: Could go into the hackathon!
JR: Shall I organize the hour meeting?
[yes!]
IF: And extend to Search.

2022-02-17: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Alejandro Metke (CSIRO/AGHA), Chen Chen (GA4GH), Meen Chul Kim (CHOP), Tim Jackson (TrakGene), Soichi Ogishima (GEM-Japan), Amanda Spurdle (BRCA), Bing Feng (UUTAH, BRCA), Zornitza Stark (AGHA), David Hansen (CSIRO/AGHA)

	Actions Arising	Assigned To	Deadline
1.	Send reminder to the Pedigree mailing list to review acknowledgements/affiliations	Lindsay	
2.	Update the 'Classes' Section of the documentation and send a ping when ready for review	Meen	
3.	Review PRs #18 and #19 and add any comments/suggestions prior to next meeting	ALL	
4.	Follow up with Jules about Phenopackets implementation	Lindsay	
5.	Extend the conceptual model with elements from the common dataset	Orion, Alejandro	
6.	Revise the pedigree ecosystem diagrams to remove ambiguity about the "third model"		

Global Alliance for Genomics & Health

	Agenda Item	Speaker	Time

1.	Welcome, Housekeeping
	- Call for Google SoC Projects
	- Montreal Connect, April 20-22
	  - There will be an in-person component!
	  - Call for workshops will be opening soon
	- GA4GH WS Video Explainers
		Grant Wood, Orion Buske, Lindsay Smith	5 min

2.	Documentation Updates
	- Intro and Acknowledgements (PR#18)
	  - Brief explainer of protobuf & FHIR added
	  - Reminder to please review and confirm affiliations for the acknowledgement page, and let me know if you'd like to acknowledge any funding.
	- Classes (PR#19)
	  - Adding ERD
	- Other sections outstanding: Pedigree Technical Model and Working with the Pedigree Model
	  - Documentation Skeleton
		- Lindsay
		- Meen
			15 min

3.	Recommended Dataset Document	Grant and Alejandro	10 min
4.	Other Updates? - Implementations, ontology, etc.	All	10 min
5.	Roadmap Planning - Draft Template	All	15 min
6.	A.O.B. - Full WS Meeting - FHIG IG - Connect Workshop?	All	5 min

Past Tasks:

- Confirm affiliations with pedigree contributors (send email)
  - Lindsay
- Add protobuf, Json, FHIR explainer to introduction.rst
  - Lindsay
- Work on the minimal dataset doc
  - Alejandro, Grant
- Put roadmap document together
  - Lindsay
- TrakGene proof of concept (present experience when ready)
  - Tim
- Add ERD to the Classes page
  - Meen
- For future agenda: outreach push with Redcap

Minutes:
Lindsay to send a reminder to review acknowledgements.

MCK: Will re-arrange and send a ping to merge section by the next meeting.
OB: Working on the two sections for the technical component and the pedigree model. Will flesh those out more. Alejandro, will need your help with the ontology section.
LS: I'll follow up with Jules about documentation.

Recommended Dataset Document
GW: AM went through this document and brought some great questions forward. The FHIR (GA4GH) column makes so much sense, we didn't have our own model in there. Going to remove the FHIR (plain). Few cases where we have an element on the left but not in the pedigree model. Might need to make a comment about putting those in in the future. I will add some text to the intro section about the appendix and new section of **
AM: Need to decide whether we be explicit about whether to extend or not.
GW: Would it be easy to extend our model to match these ones?
OB: I think so. Especially for the things that are optional. Goal was always to create the minimum model. Ensure we're being consistent in messaging, but the trade off is in complexity.
GW: Should we delete element rows? I don't think we should, it's more of an inspirational document to create a robust tool and repository. Have replaced "minimum" with "common".
AS: To me, age, DOB and age range are not even optional. Some element of age in the pedigree is important.
OB: In these contexts, we didn't include those because they can be represented in FHIR and Phenopacket environment. Comes back to the line about where we draw the boundary for information about people.
AM: Could call them out explicitly and say this is represented by 'x' class in FHIR.
OB: We have a conceptual model of an individual, and this is x /y in FHIR /z in phenopackets. Either we have a third to interoperate, or say we conceptually have this information. Or we leave it out and say it's handled by these other things. If we call it out, are we saying it's a separate class. If we were to say use the existing implementation, there's no mechanism to add additional properties that aren't handled by Phenopacket or FHIR. Concept of a node that points to an individual rather than an actual individual.
GW: With FHIR you can mix and match resources. Had always thought that's what we would do with our conceptual model.
OB: My understanding is that's correct, but the way in which amending or adding on to what's already there does still define another concept related to an individual, but doesn't have age because the main individual that would be able to be referenced could have that information in FHIR.
AM: We don't have an extra individual, just extend the patient. Highlights the complexity of different implementations with different strategies and having to map.
GW: Would suggest, why not come up with a new solution?
AM: Sure, and compare and contrast with Phenopackets.
AS: Getting stuff from one pedigree tool to another is the bane of my life. Anything that can help us is great.
GW: Don't want to add just another new standard, but a universally accepted one. Break those barriers and come to a technical solution that will share data easily.
OB: There are a few Qs here: how conceptual do we want the conceptual model to be. Do we want it to be usable on it's own. Hope was that there as a coherent core kernel independent of and could be used to define something that represents a ped file and could be implemented in FHIR or phenopackets. But could also work on its own to a lower degree. Anything that can be shared in Phenopackets and FHIR would not be shared in that core kernel. The other way is looking at it as a true conceptual model and not using it all, then we have an implementation guide for the recommended fields in FHIR and for Phenopackets. Bigger departure than directly implementing the implementation in either of those environments. Loosening the constraints on those two.

AM: I don't see us having a third implementation. I always thought the protobuf version would be the independent but it became part of Phenopackets.
OB: I think we had meant for it be separate and independent but not Phenopackest in the beginning, just protobuf. Is there a third is the question to some extent, and does it need to handle all of the other things we care about (like age), which is separately structured in FHIR and Phenopackets already. Let's use diagnosis as an example, do we want to define yet another standard for defining diagnoses so that we can have diagnoses for individuals in a pedigree. We can all agree that's important information to have in a pedigree. But, if you're using in FHIR or Phenopackets, that work has already been done. Either going to choose the FHIR or Phenopacket version, or create another that sits in the middle. Would have to do the work Phenopackets-FHIR did twice.
AM: I agree, but in some cases, if we keep at the conceptual level it might reveal gaps. Patient resource in FHIR doesn't have some of things called out in this document.
OB: So there isn't a third. Conceptual model is a conceptual model
AM: Would it be useful? What would be gained by the third implementation?
OB: Mixed feelings. Do we go beyond and say here is this file that works on its own, pedigree file that is not a ped file. Piece that you don't get is a full reduction in mapping complexity, because there's not one model per se, but you're always going to have to handle the mapping between a phenopacket and FHIR on the tooling side.
AM: Inherited this problem from discussions in Clin/Pheno when decision was made that the probuf would be a thing that lives outside of FHIR
OB: Back out of the conceptual model to be less technical, less grounded on implementation. Focus less on saying there is any particular implementation in that conceptual model.
AM: Doesn't mean we can't eventually work on a third. If there is a need it might help transformations to other models.
OB: Dramatically simplify the process of extending for all of these things in the common data model because we're not defining them.
GW: The more conceptual you keep it, the more adoption we'll find.
LS: Do we know of any DPs or initiatives that won't use in the context of FHIR or Phenopackets? Suggest we go by community needs - if no one needs that third implementation then we don't need to go there.
OB: No. Agree, I don't we planned to ever really implement. If we can back that conceptual model into anything that looks like protobuf will be helpful.
OB: There are a number of high level diagrams that we constructed historically and might be still be used that paint that concept in a way that is no longer the way we want to communicate that. Better ways of framing the blue sky architecture. Make the pedigree model less concrete, don't leave any ambiguity about whether there is another format or not.

Connect Breakout Ideas
OB: Whether or not having a conversation with Discovery/Search. Digging into retrieving more information. Priority of having a conversation with Jules and get his feedback down with incorporating into phenopackets.
GW: Could be something with GKS? Looking long term about connecting with variant data, having a "fun" conversation about the next phase.

2022-02-03: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Alex Tsai (GA4GH), Chen Chen (GA4GH), Tim Jackson (TrakGene), Alejandro Metke (CSIRO/AGHA), Melissa Cook (NCI), Amanda Spurdle (BRCA), Zornitza Stark (AGHA)

Global Alliance for Genomics & Health

	Actions Arising	Assigned To	Deadline
1.	Confirm affiliations with pedigree contributors (send email)	Lindsay	Mon Feb. 7
2.	Add protobuf, Json, FHIR explainer to introduction.rst	Lindsay	Next meeting (Feb 17)
3.	Work on the minimal dataset doc	Alejandro, Grant	
4.	Put roadmap document together	Lindsay	
5.	TrakGene proof of concept (present experience when ready)	Tim	
6.	Redcap Outreach (for future agenda)		

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - Public call for Clin/Pheno Work Stream Leads - Alex & Chen, GA4GH Staff Software Devs	Grant Wood, Orion Buske, Lindsay Smith	5 min

2.	Documentation Updates
	- Intro and Acknowledgements (PR#18)
	  - Revised Acknowledgements section
	  - Image for Intro page?
	- Classes (PR#19)
	  - Adding ERD
	- Other sections in progress?
		- Lindsay
		- Meen
			15 min

3.	Outstanding Tasks & Next Steps - Documentation - Ontology - Recommended Dataset Doc - FHIR IG / Phenopackets	All	10 min
4.	Implementation Updates?	All	10 min
5.	GA4GH Staff Dev Projects - Conversion Tooling	Orion	15 min
6.	A.O.B. - Roadmap Planning	All	5 min

Past Tasks:

- Introduction Page in Docs - add link outs to technical descriptions of protobuf, json, FHIR

    - Lindsay
- Revise acknowledgements page, prep an email to contributors to confirm affiliations and “opt-in”
    - Lindsay
- Work with Alejandro on completion of the recommended dataset doc
    - Grant
- Add ERD to the Classes page
    - Meen
- For future agenda: outreach push with Redcap

Minutes:

Documentation Updates
LS: Revised the acknowledgement section to include contributors who did a minimum of one “thing” (eg. attended a meeting, commented on a doc, raised an issue). Made a separate section for those who completed the pedigree driver project survey at the start of the project. I will send around an email next week to confirm affiliations, let me know which ones you want represented, and if you have any funding you want to acknowledge. I also need to add to the “brief explainer” section on the intro page.

General Updates
AM: Will take a look at the minimal dataset document today. Some outstanding tasks on the ontology work. The IG is in pretty good shape, not much to do for an initial release.

https://github.com/phenopackets/phenopacket-schema/blob/pedigree/src/main/proto/pedigree.proto → updated Phenopackets pedigree integration

Implementation Updates
OB: Saw update from David Conlan for open pedigree import/export. Issue now in open pedigree repo.
AM: Working on optional with some working examples. Fixing some functionality plugs.
TJ: Taken one of our families and have a test case. Run it through the conversion tool. Developed a family history questionnaire webapp, patients fill out survey before they come to clinic that auto generates a pedigree.
OB: Proof of concept would be awesome
GW: Share back with us the experience and any issues you come across.

OB: We wanted to talk about the conversion tooling and Bing’s analysis. There are few different options here but opportunity to get staff support around software dev needs. Where it seemed like their efforts might be most beneficial is around conversion tooling. Updating the tooling to allow export in BODICEA. Beyond that, updated any of the open source resources if valuable, like pedigree validation.
Some sort of CI automated testing would be valuable. Expand those formats to support FHIR and Phenopackets implementations of pedigree.
GW: Is there something we could consider that we can develop some simple tool that could take the ped apps that these Driver Projects use, and convert into our standard? Starting down the path of providing some solution.
OB: Question is what is the best way for the driver projects to deliver this. Goal of the tools repo. We could go further and offer a library but there are so many languages we would want to do this in. Could have a front end app to do conversions
GW: Both great places to start, library and web page.
OB: Chen and Alex could take what is currently there and set up to release as a node module. Now it’s just a repo with command line access. Provide value for web based pedigree solutions.

TJ: Second the idea of a web interface because you'll get more buy in. Follow the same train of thought of implementing a library you can imbed in something.
AS: From analysis POV, one has a bunch of data we can't access. In AUS, family cancer clinics are using a bunch of different tools. Often have to re-enter the pedigree. Sometimes the pedigree is set up for a specific tool.
OB: Benefit perhaps, if we do the client side.. If we leave it as node and release as a module, we could do it as a client side browser plug in. Batch analysis, having a web interface makes it a lot more tangible.
AM: Potential scope down the road for help with the validation stuff.

Roadmap Note
GW: Add note about doing an HL7 accelerator. Promote the standard outside of the GA4GH community.

2022-01-20: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Bing Feng (UUtah), Meen Chul Kim (CHOP), Melissa Cook, Zornitza Stark (AGHA), Amanda Spurdle (BRCA), David Hansen (AGHA/CSIRO)

	Actions Arising	Assigned To	Deadline
1.	Introduction Page in Docs - add link outs to technical descriptions of protobuf, json, FHIR	Lindsay	Feb. 3
2.	Revise acknowledgements page, prep an email to contributors to confirm affiliations and "opt-in"	Lindsay	Feb. 3
3.	Work with Alejandro on completion of the recommended dataset doc	Grant	
4.	Add ERD to the Classes page	Meen	
5.	Reach out to TrakGene re: implementations	Orion	DONE
6.	For future agenda: outreach push with Redcap		

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - Happy New Year! (is it too late to say this now?) - Public call for Clin/Pheno Work Stream Leads - Clin/Pheno Full WS meeting moved to Feb. 23	Grant Wood, Orion Buske, Lindsay Smith	5 min
2.	Documentation Updates - Intro and Acknowledgements (PR#18)	- Lindsay	20 min

	- Questions: how best to populate the "Brief Explainer of JSON, Protobuf, FHIR"? Any funding or external groups we should be including in Acknowledgements? - Classes (PR#19) - Other sections in progress?	- Meen	
3.	Outstanding Tasks & Next Steps - Documentation - Ontology - Recommended Dataset Doc - FHIR IG / Phenopackets	All	20 min
4.	Roadmap		
5.	Software Dev Projects		
6.	A.O.B.	All	5 min

Minutes:

GA4GH Software Dev Projects - please let Lindsay know if you have any other ideas!

Docs

- Acknowledgements: update list with names who have attended at least one call, commented on a document, etc. Lindsay to send an email to confirm affiliations
- Intro: For the intro to protobuf/json/fhir, add a very brief description of each, with a link out to the technical details

MCK: Created a PR with the classes section. Set it as a draft so I have your feedback and thoughts. Not a lot of descriptions per class.
Reflected was I did with the FHIR IG.
OB: Is there a standard or best practice when documenting conceptual models, whether to do it in a hierarchical view?
MCK: Why don't I add an ERD?
OB: I think that might help people see how they're related. Make explicit how it would be handled over the wire.
GW: At HL7, different groups would create different classes that would overlap. Had to do some sort of harmonization later.

GW: Concerning the recommended dataset doc, I need to work with AM to get that completed [Action]

Implementations?
OB: We targeted implementations involving tooling and supporting connections. Is that still reasonable? We'll have Redcap Plugin, a connection between one system to another that runs things off of Phenopackets or FHIR would be the second. If there was a way to loop in Bing's implementation either via Bodicea or not, than that is a well rounded set. Is there a group or other tool that people use yours in conjunction with?
BF: There is a group in France that uses my tool with BODICEA.

OB: The canrisk developers, I know they just came out with a new file format. Maybe looping them in might help.
BF: The BODICEA format has been changing, they have one format for each gene.
OB: On our side, can do a Phenotips to Exomiser push.

GW: For redcap, as a lot of research organizations use redcap. Can we consider doing some promotion where we approach research groups and ask if they want to do a pilot.
ZS: I think it's freely available. We've been approached by a lot of other research projects in Australia wanting it. H3Africa uses redcap heavily.
GW: Let's discuss this further. Put on future agendas [Action]
OB: Can also put pressure on TrakGene and us!

GW: Focus on implementation and adoption for 2022 roadmap.

2021-12-14 - Cohort Rep Report Back Meeting

The minutes for Computable Cohort Representation can be found here.

2021-12-09: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Alejandro Metke, Meen Chul Kim, Bing Feng, Melissa Cook

	Actions Arising	Assigned To	Deadline
1.	Start filling out the documentation - Intro Section and Acknowledgements - Technical Model - Classes - Working with the Model	- Lindsay - Orion - Meen - Alejandro	Jan. 20 (outline/draft)
2.	Reach out to Jules about the Phenopackes piece	Lindsay	
3.	Add link to new minimal dataset doc to the old one	Lindsay	
4.	Propose recommendations for which rows to remove from the minimal dataset table	Alejandro	
5.	Open a PR re: outstanding ontology issues	Alejandro	

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - Last meeting of the year! We will return on Jan. 20th (to confirm). Happy holidays! ✨	Grant Wood, Orion Buske, Lindsay Smith	5 min

2.	Documentation Updates - https://pedigree.readthedocs.io/en/latest/ - it's live, but not yet populated - PR #15, PR#17 (both now closed) - Next steps / Writing sections	Meen, Lindsay	10 min
3.	Updates on remaining Ontology issues/comments?	Alejandro	5 min
4.	Updates on minimal dataset mapping to conceptual model?	Grant, Alejandro	5 min
5.	Planning tasks to be completed by Jan. 20	All	10 min
6.	Updating or re-affirming the Pedigree Roadmap for 2022/2023 - Old (2020) roadmap	Lindsay	15 min
7.			
8.	A.O.B.	All	5 min

Minutes:

Filling out the docs
OB: Happy to do the overview and classes section. Meen and Alejandro, would appreciate your help with the tools and ontology sections
MCK: I volunteer for classes

OB: readthedocs has good support for versioning. Probably want to create a branch and have the community review.
GW: When to connect with HL7? Want to keep the group informed but don't want them to slow us down.
LS: Can share after GA4GH approval, or via the public open for comment process.
AM: Agreed, public consultation
OB: Getting phenopackets fleshed out. How to get that done. Might need to rope in Jules if he is willing. I could give it a shot.
LS: I can reach out to Jules and lay out a skeleton in terms of communicating that. Get an up to date sense with the recent conversations.

Minimal Dataset Mapping
AM: Made a copy of the doc, made some changes to the table and the appendix. The bit I found a bit confusing is that these data elements seem to be coming from the FHIR family member history resource, and another column about ped and FHIR. But FHIR has three pedigree models. Have a column for each FHIR model for how it will look like. In the FHIR spec, there isn't any models for reference so I thought it would be useful to compare the three models.
Nodes of the graph are the resources. This is different from the model we've proposed. Major differences is that the family member history is not allowed to contain patient information. Also have a composition resource. I think this is helpful to show the different ways to use FHIR.
In the tables, have added additional columns for FHIR, FHIR pedigree and FHIR GA4GH. Can have more meaningful conversations about the gaps.

There are some things here that are not in the conceptual model, so I'm wondering whether they should be here.
GW: Why would the source metadata not be in the conceptual model?
OB: Kept things as simple as possible, but no reason we can't add it now. Maybe the thing is to add and move on.
AM: Exactly
OB: The title is minimal dataset, but many things listed here are optional. Should we split these things out. Goal was to be permissive, whereas if we put things here, it could lead to questions about mismatch.
GW: Minimum or Minimal might need to be deprecated. Started with that original work that we called minimal, but what we've achieved here is a bit different. Maybe just call it recommended dataset.
AM: Also add a column for saying explicitly, this is part of version 1.0 of the model.
GW: My preference would be to add to the conceptual model. We want to add to the conceptual model to make it more robust and capture more use cases. If it's not a problem to add, let's do it.

AM: Could quickly go through the table. Some of these things appear to have come from very FHIR specific things. It's easy for me to implement in the FHIR IG, but will be more work for Jules.
GW: If you feel like it doesn't belong in the table part, we can move it. Thinking was to not only include the ped columns but also all of the FHIR resources that were linked. Need to decide what needs to be in the conceptual model. If a particular resources doesn't help, let's move to the appendix.
OB: There are a lot of different ways we can do this. The data depends on which specific use cases we're talking about. Happy to review your recommendations Alejandro.
AM: Happy to have a go and set some for review.
Most of the extra comments are around that.

Intention of the extensions was to connect each node to those closest.
OB: But can also re-base that. Can do the parent-child relationships and sibling-siblings.
AM: Still need to point at the proband. So if you go mult generation, still need that node with the patient. One hop type relationships. That's why I think our approach works better.
GW: As we talk to people like HL7, that will be a major point to share with them.

AM: ped had maternal and paternal id, but those don't fill well in the individual table because they're nodes in the graph
OB: have a relationship table?

AM: In the conceptual model, sex at birth and the HL7 defiintion is a bit wider. At this level, maybe adopt a higher level. Should match whatever is in the conceptual model. Most of the other comments are about stuff not in the conceptual model, but we've already discussed this.

Work off of the new doc.
Lindsay will add link to the new doc from the old one.
https://docs.google.com/document/d/1GQRd5jeZeB5qhHclLZxDe6kPD173bXWGYITsmCbTeuI/edit

AM: Made a pull request, Orion to review. Five issues outstanding. Might review offline. 1 PR, 2 comments, and two more to go.
Feedback from Soichi - feels uncomfortable with the term gestational carrier. I looked at the definitions. Really depends on the definition that we want to give to gestational carrier. Found one definition that implies IVF. The current hierarchy we have is wrong if we follow that definition.
OB: Was trying to get at the concept of being explicit about the ways people are biologically and environmentally related to offspring. Sounds like it needs to be pulled down and we need a different term for that parent.

AM: I'll have a look and open a PR. Other ones are pretty straightforward.
OB: Gestational parent?
AM: Yes, might just be avoiding a term with social implications.
GW: Maybe ask a genetic counselor.

2021-12-02: Computable Cohort Representation - PHENO Team

Minutes: The minutes for Computable Cohort Representation can be found here.

2021-12-02: Computable Cohort Representation - MIACC Team

Minutes: The minutes for Computable Cohort Representation can be found here.

2021-11-11: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Meen Chul Kim, Katherine Johnston, Zornitza Stark, Alejandro Metke, Bing Feng, David Hansen

	Actions Arising	Assigned To	Deadline
1.	Check in with TrakGene about any implementation support	Lindsay	
2.	Prep some slides and description of Pedigree subgroup efforts for Katherine/H3Africa	Lindsay	
3.	Work through the remaining ontology issues/comments	Alejandro	
4.	Set up the readthedocs in the Pedigree repo - Ticket: https://github.com/GA4GH-Pedigree-Standard/pedigree/issues/14	Meen	
5.	Map the minimal dataset doc to the conceptual model, ensure they align	Alejandro / Grant	

Global Alliance for Genomics & Health

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - No Pedigree meeting in two weeks (American Thanksgiving)	Grant Wood, Orion Buske, Lindsay Smith	5 min
2.	Reviewing Task List - GA4GH staff support for tooling - Check in with H3Africa & TrakGene with implementation support - Shrimp reference for KIN: https://ontoserver.csiro.au/shrimp/?system=http%3A%2F%2Fpurl.org%2Fga4gh%2Fkin.fhir&concept=KIN%3A001&version=20210805&valueset=http%3A%2F%2Fpurl.org%2Fga4gh%2Fkin.fhir%3Fvs - pedigree-validator repo	ALL	10 min
3.	Reviewing Soichi's Ontology Comments - KIN repo	Orion, Alejandro	5 min
4.	Pedigree Documentation - Draft Skeleton	Lindsay	20 min
5.	Minimal Dataset Document - Final review and vote	Grant	15 min
6.	A.O.B.	All	5 min

Minutes:

KJ: Dug into the project's collecting pedigree data, disappointed to find so little pedigree data in it. Haven't heard back from Guida yet. Next step is to communicate with all of the projects about the standards. In terms of implementation, play around with the redcap component and get the projects who are collecting to try it out. Mamana mentioned he heard about a group in Ethiopia interested in collecting pedigree data and were working on something.
GW: I'm aware of a project being done in Ethiopia with U of Michigan who have developed a mobile based family history collection app. Data gets stored on a server back in Michigan, but they provide the data back to doctors in Ethiopia.
KJ: Potentially, Manana mentioned a software group.
[Lindsay to check in with TrakGene]
[Lindsay to send slides about Pedigree to Katherine]

Validator Repo
AM: Command line to validate an instance of a pedigree. Java app using Hapi libraries. Read a pedigree JSON in FHIR and shows anything non compliant. Make sure codes align with the right ontology. Shows how to include validation from ontology reasoner. Instructions on how to run are included. Let me know about any feedback.
OB: Is the scope of the repo is that it will validate in multiple formats, or is it mostly focused on FHIR?

AM: No reason we couldn't include other formats in there. Original scope was just the IG. No need to create new repo.

Ontology Comments
AM: Got additional comments from Chris Mungall. Took a bit of time like adding natural language. Five things outstanding, should be able to quickly go through.

Documentation
OB: Maybe merge Model and Classes together? Or expand classes
MCK: Happy to help set up the skeleton.
Just created an empty ticket: https://github.com/GA4GH-Pedigree-Standard/pedigree/issues/14 will fill out descriptions.

Minimal Dataset Doc
GW: Want this group to happy with a good version 1.0. In the documentation we explain that it's a living doc and we'll review over time. Went through and cleaned up some sentences that weren't very clear. Made a few formatting changes as well. Is at the point where we'd like to subgroup to approve.

OB: Can we resolve the red text in one way or another?
GW: Yes, I don't describe that well in the document. My intend it to highlight the gaps.
LS: Move the red text to the notes?
OB: Depends on the framing as well. From the perspective of us putting this out, it isn't on us to make this change to FHIR, nor is it on use to say FHIR needs this change, but we can say that there is a gap.
AM: Descriptions are not aligned with choices made in the model. We should highlight gaps, and moving forward, we really can't modify the base FHIR resources.
GW: First point is critical, seeing things that aren't matching the model. Since you have that knowledge of the model, could you help with this?
AM: Can add comments to the doc.
GW: Go right in and make the correct change.

GW: If anyone wants to add in a paragraph that describes the best method please do that.

2021-11-02: Computable Cohort Representation - PHENO Team

Minutes: The minutes for Computable Cohort Representation can be found here.

2021-10-28: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Zornitza Stark (AGHA), Bing Feng (UUtah), Melissa Cook (NCI), Meen Chul Kim (CHOP)

Global Alliance
for Genomics & Health

	Actions Arising	Assigned To	Deadline
1.	Follow up with Yash and Jeremy about available cycles to help with the converter tooling	Lindsay	
2.	Check in with H3Africa and TrakGene re: implementation support		
3.	Draft documentation skeleton to review at the next meeting	Lindsay	Nov. 11
4.	Connect with Phenopackets IG team about the Pedigree IG, plan for IG updates session at the next full WS meeting	Lindsay	
5.	Move Clin/Pheno full WS meeting to Dec. 1st, cancel Pedigree meeting on US Thanksgiving	Lindsay	
6.	Follow up with Alejandro and review Soichi's ontology comments at the next meeting	Orion, Lindsay, Alejandro	Nov. 11
7.	Get the Shrimp reference for KIN - https://ontoserver.csiro.au/shrimp/?system=http%3A%2F%2Fpurl.org%2Fga4gh%2Fkin.fhir&concept=KIN%3A001&version=20210805&valueset=http%3A%2F%2Fpurl.org%2Fga4gh%2Fkin.fhir%3Fvs		DONE
8.	Complete one final read through of the minimal dataset document, resolve any outstanding gaps	Grant	Nov. 11

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - Connect meeting recordings are up on the Attendee Hub, as well as in a sharable Google folder. Please let Lindsay (lindsay.smith@ga4gh.org) know if you'd like a sharable Pedigree recording	Grant Wood, Orion Buske, Lindsay Smith	5 min
2.	Debrief: GA4GH Connect Pedigree Session - Meeting minutes - Thoughts? What went well? Key takeaways?	ALL	10 min
3.	To Do List: - Polish off implementations (these are needed for GA4GH Approval) - Finalize converter tooling - Finalize minimal dataset document	ALL	30 min

	- Fill out the documentation
	  - Do we want a readthedocs (like Phenopackets or VRS) or a detailed readme (like DUO or Data Connect)
	- Outstanding Ontology Tasks?
	  - Two pull requests to review (tagging Orion on this!)		

4.	A.O.B.	All	5 min

Minutes:

Updates from AM (via email):

- The REDCap pedigree plugin has been updated based on the FHIR IG, so it should be ready for the approval process. We are adding some more features at the moment and we will release a new version in the next couple of weeks.
- I am happy to contribute the command line validator tool that I showed during Connect. If you create a GitHub repo for me I can upload it there (you can call it pedigree-validator).
- I have no strong preference regarding the way we write documentation.
- I have been going through the different issues in the KIN ontology and solved most of them. There is a bit more work to do but I think it is stable enough to be usable.

Connect Feedback
GW: Everything went well! What do people think if we did a demo again in the near future? What would we want to add?
OB: Having more examples of direct support of the standard. Tooling that patched some of that. Would be compelling to have the standard used directly.
GW: What the group needs to focus on next.

OB: Redcap - getting cross talk between the different tools.
GW: G2MC - had a presentation from Brazil using the Redcap pedigree plugin. Would it be helpful to have these people talk about it?
ZS: Would be interested to hear how others are using it
GW: Use case was general practice/primary care type stuff. Had a focus on cancer. I think he was a physician from Sao Paulo.

OB: GA4GH Staff software dev support?
LS: Will follow up after this call with Yash and Jeremy

OB: BODICEA support, is there a way we can support you to import this data through Phenopacket.
BF: Yes, decided to use Phenopacket. Need guidance and documentation. Use the modules from Phenopackets to integrate, using C++. Haven't connected with Jules about C++.

MCK: When we first modelled in FHIR, discussion about family member history vs observation. We're now discussing implemented our model into the KF FHIR server. What we implemented before was just the relationships between patients. No deadline yet. We released our FHIR server as open access.

Need to check in with H3Africa & TrakGene on implementation/testing.
OB: Putting the documentation into place for how we see and expect people to use this with FHIR or Phenopackets. Can leverage Phenopackets docs. Align with Jules on

Doc preferences?
- Question of the extra work to implement in a form with usability on readthedocs
- Aim for more lay language, but making the techy information available

MCK: How to integrate with the Phenopackets FHIR IG?
LS: I will connect with the Phenopackets IG team, plan for an update session in November.

Lindsay to draft a documentation outline for next meeting.
MCK: I like to readthedocs website, can help with the deployment

OB: AM added definitions to all of the terms, incorporated additional terms based on the open pedigree.
Let's follow up with AM on Soichi's comments - go through them on the next call.

OB: Would it be helpful to give a visual summary of the ontology representation?
ZS: Visual always helps.
MC: Agree
OB: Get the Shrimp reference for KIN

Minimal Dataset Document
GW: Go through one more time, create a final version (formatting, final errors).
LS -> check process for minimal dataset approval
GW: Not a companion for the schema itself. Want people to know the document is separate. Talks about high level issues. I will finish this up.

Lindsay to move the full WS meeting to the following week - Dec 1st.

2021-10-19: Computable Cohort Representation

Minutes: The minutes for Computable Cohort Representation can be found here.

2021-10-12: GA4GH Connect

Minutes for each breakout can be found below:
Pedigree Implementation Demo
Computable Cohort Representation
Clin/Pheno Roundtable

Global Alliance
for Genomics & Health

2021-10-05: Computable Cohort Representation

Minutes: The minutes for Computable Cohort Representation can be found here.

2021-09-30: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Moni Munoz-Torres (Monarch), Meen Chul Kim (CHOP), Bing Feng (UUtah), Alejandro Metke (CSIRO/AGHA), David Hansen (CSIRO/AGHA), Guida Landoure (H3Africa)

	Actions Arising	Assigned To	Deadline
1.	Reach out to Jules about joining the demo + Exomiser	Moni	
2.	Check Connect registration status	Lindsay	
3.	Add content into the slide deck for the Pedigree Connect Demo	EVERYONE	
4.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - Plenary recordings are up on the attendee hub! - GA4GH Connect registration	Grant Wood, Orion Buske, Lindsay Smith	5 min
	Task List Review: ● Add ideas/notes/image into the slide deck for the Pedigree Connect Demo- EVERYONE ● Reach out to Soichi about slide 12 and where his work might fit in - Lindsay ● Reach out to Jules about phenopackets/exomiser piece of the demo - Lindsay ● Reach out to Driver Projects for additional use cases for the relationship terminology - Lindsay ● Create an example pedigree and send to Alejandro - Bing ● Check with Meen if there are any interesting	All	5 min

	cases where the pedigree was caught by QC - Lindsay ● Take a look through the original use case document and pull out any key pieces relevant to the terminology/reasoner - Grant + any other willing volunteers		
2.	Preparing for GA4GH Connect: - Agenda Template - Draft Slide Deck Status updates from each activity area in prep for the plenary and connect ● RedCap Plugin and OpenPedigree ● Trakgene ● FHIR Demo ● Cosegregation Analysis [the conversion tool likely won't be ready :(] ● Phenopackets	Grant, Orion, Lindsay	45 min
3.	A.O.B.	All	5 min

Minutes:

Planning for Connect:
OB: Draft agenda, feedback welcome. Kicking it off with Lindsay - welcome + intros (depending on the number of people the room. Orion to do the intro, Grant the talking about the minimal dataset.
Hand off to Alejandro to talk about the terminology and FHIR IG, and start off the demos with the data collection demo with the REDcap integration.
Then to data management and data analysis. DO we want more guidance on what we'd like to solicit from the crowd.
GW: Be prepared with some questions if people are shy. First thought is to allow everyone to see the whole picture, and then comment on the whole picture.
[Lindsay to check registration]

GW: Would be nice to learn if they are using pedigree and if anything is helpful to them.
LS: Community input, getting interest in engagement
MMT: Hopefully we'll talk to some people we don't usually interact with, do some use case collection from the crowd.
DH: Requirements as well. Also interested in analytics they want to do on top of pedigrees.

AM: Missing a section that might correspond to the native implementation of the pedigree.
OB: Split out the terminology from the IG. Branch in the phenopackets github that has pedigree implemented in it. Ask Jules if it makes sense to share this [Moni to reach out to Jules about this + Exomiser].
MMT: Important to highlight it's not implemented on the main branch. Important to get from the community what they would need from such a connection.

GW: Want to go through fairly quickly. 5 minutes max on the minimal dataset doc.

OB: How people can use the doc and any findings we had.

Presenting terminology - what to share?
AM: Could add slides or just use the github, have a live instance. Talk about the process and how the current output was organized.
OB: DO you feel comfortable for the motivation of it being an ontology?
AM: Yes

MCK: Given 10 min, I'm ready to go over demo, implementation and live query. Concerned about the selling point.
OB: Happy to talk through this now. One of the ways of talking about this is that it provides a way of recording robust FHH information in the context of FHIR, which ped does not have. That is valuable from a data management perspective that KF can talk to.
MCK: I agree with that. Will demo in FHIR. Based on my prior experience, wasn't able to come up with a good selling point.
OB: Might be that we're not quite there, and that's ok. Still in early days of some of this. When contextualizing with the KIN terminology and FHIR, that works.
GW: Test environment for the pedigree work we're doing. Could be a goal or next step. Not thinking as far as a starter kit.

AM: We have multiple FHIR servers we use - HAPI, our own version of HAPI, SMILE for some projects. If I have the time I might show how validation will work, but it might not be ready. Basis of some sort of test environment. Command-line application where you could consume the IG.

BF: Still don't see the export to BODICEA
OB: Will make sure we get that done. If you have any cycles, it's just about adding support for a new export type in the tool repo. I think it can import almost all of the columns. Some of the columns are outside of the existing conceptual model.
BF: Run with the example sent to Alejandro
OB: What's the best way to show what this looks like?
BF: Need to describe cosegregation analysis, how it is used, then run the analysis within the website.

MMT: I'm less familiar with the inner workings of EXOMISER, but could potentially present in Jules' place.

2021-09-21: Computable Cohort Representation

Minutes: The minutes for Computable Cohort Representation can be found here.

2021-09-16: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Bing Feng (UUtah), Alejandro Metke (CSIRO/AGHA), David Hansen (CSIRO/AHGA), Kathryn Van Diemen (TrakGene),

	Actions Arising	Assigned To	Deadline

Global Alliance
for Genomics & Health

1.	Add ideas/notes/image into the slide deck for the Pedigree Connect Demo	EVERYONE	Next meeting (Sept. 30)
2.	Reach out to Soichi about slide 12 and where his work might fit in	Lindsay	
3.	Reach out to Jules about phenopackets/exomiser piece of the demo	Lindsay	
4.	Reach out to Driver Projects for additional use cases for the relationship terminology	Lindsay	
5.	Create an example pedigree and send to Alejandro	Bing	
6.	Check with Meen if there are any interesting cases where the pedigree was caught by QC	Lindsay	
7.	Take a look through the original use case document and pull out any key pieces relevant to the terminology/reasoner	Grant + any other willing volunteers	

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH 9th Plenary registration - GA4GH Connect registration	Grant Wood, Orion Buske, Lindsay Smith	5 min
2.	Preparing for GA4GH Connect: - Agenda Template - Draft Slide Deck	Grant, Orion, Lindsay	25 min
3.	Status updates from each activity area in prep for the plenary and connect ● RedCap Plugin and OpenPedigree ● Trakgene ● FHIR Demo ● Cosegregation Analysis ● Phenopackets?	All	15 min
4.	Terminology Use Cases & Issues & Updates		10 min
6.	A.O.B.	All	5 min

Update from Meen:
RE FHIR, I found another KFDRC project (also released on dbGaP) having more complex pedigree examples. Details:

Study: GMKF: Congenital Cranial Dysinnervation Disorders (CCDD) and Related Conditions

PI: Elizabeth Engle
dbGaP URL:
https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001247.v1.p1
FHIR resources:
- Individuals (FHIR Patient resources): 899 (270 probands)
- Relationships (FHIR FamilyMemberHistory resources): 772 (uni-directional, 27 types)
  - Aunt-in-Law: 1 (REL:019)
  - Brother: 29 (REL:007)
  - Cousin: 1 (REL:014)
  - Daughter: 5 (REL:032)
  - Father: 327 (REL:028)
  - Husband: 1 (REL:026)
  - Maternal Aunt: 3 (REL:013)
  - Maternal Cousin: 10 (REL:015)
  - Maternal Grandfather: 4 (REL:017)
  - Maternal Grandmother: 6 (REL:017)
  - Maternal Granduncle: 3 (REL:002)
  - Maternal Great-Grandmother: 1 (REL:018)
  - Maternal Half-Sister: 2 (REL:013)
  - Maternal Uncle: 2 (REL:013)
  - Mother: 334 (REL:027)
  - Nephew: 2 (REL:002)
  - Other: 1
  - Paternal Aunt: 3 (REL:013)
  - Paternal Grandfather: 1 (REL:017)
  - Paternal Grandmother: 1 (REL:017)
  - Paternal Uncle: 1 (REL:013)
  - Sister: 24 (REL:007)
  - Son: 3 (REL:032)
  - Spouse: 4 (REL:026)
  - Twin Brother: 1 (REL:010)
  - Uncle-in-Law: 1 (REL:019)
  - Wife: 1 (REL:026)
- Diseases (FHIR Condition resources): 359 (5 kinds)
  - CFEOM
  - CFP
  - DRS-HGP
  - Moebius
  - Ptosis
- Pedigrees (FHIR Composition resources): 270 (23 families with more than trios)

Query results against one proband patient here:
Individual (Patient)
Relationship (FamilyMemberHistory)
Condition
Pedigree (Composition) resources

This family is one of the pedigrees having the most complex relationships:

- There are 8 family members (including the proband)

- They have 12 one-directional relationships between them (among them, the attached Relationship resource below contains 7 relationships regarding the proband)

Minutes:

Preparing the minimal dataset doc for the toolkit. Next steps involve:
- Finalizing the doc and making sure the Pedigree subgroup & Clin/Pheno WS overall are supportive
- Post it for public open comment
- Resolve any comments that come in. If there are significant changes, we may need a second Pedigree Subgroup/ full WS approva.
- Presentation at Steering Committee, with SC vote/approval

GW: Add in a slide for the minimal dataset document

How should we organize the demo?
AM: Spin up a redcap instance and show how to collect a pedigree. Would show the graphical representation, go back into the form. How the conversion tooling will interact with that component. The way that the plugin works, if you interact manually, should store the FHIR representation. Not sure how to show the demo because this happens on the backend.
OB: Manually download or export for visual purposes.
AM: Yes. Redcap has some licensing constraints, not sure if it's a problem and we should check. Make sure that we're not in violation of any constraints.
GW: The licensing shouldn't restrict anyone from doing a demo? Are you going to be the person to do this part of the demo?
AM: I'm happy to participate in the ingestion bit. Show how sample pedigrees are collected, then hand over to Orion and Grant.
OB: I don't there should much of the demo on conversion tooling. Bing - we'll be doing everything to ad BODICEA v4 export support between now and the next few weeks. Would be great if you could then take one of these files and run cosegregation analysis on this. Focus on look, we have pedigree data going from collection to interoperability to all of these different tools.
BF: Definitely.
OB: Trakgene has been poking at the pooling and found a bug already with import functionality.

AM: Need an example that makes sense for the analysis tool. Bing, if you can send a pedigree along I can get things pre-populated.
BF: Yes, I'll create an example pedigree and send out.
GW: How do we fit in Trakgene and FHIR in.
OB: Might be this bit of taking input to analysis. Could have research systems and their angle.
GW: I think that makes the most sense.
OB: Meen would be serving this data via FHIR, is there a consumer of that we could pair it up with.
AM: Exomiser could be one of the candidates? It would have to go back through ped.
OB: Weak flex of various functionality to take rich representation in FHIR and cut it down to ped. Anything in terms of terminology that you'd like to show off?
AM: I just made some changes, this prefix is now different (KIN). ** Would be good to try and do some validation. Not hard to implement.
OB: Something obvious like a logically inconsistent family.
AM: Only concern is that because if we're using OpenPedigree there are limitations on creating the families. Might have to show a drawn pedigree or some hypothetical.
OB: Maybe Meen has an example where his QC caught something.

AM: If Meen can provide an interesting example I can take a look. I can have a go at having a component that takes in a pedigree an outputs validation results.
OB: Yes, almost a proof of concept to get people thinking about capabilities.
Lindsay - Send Soichi a copy of slide 12 and see where he might fit in.

For Trakgene, is everything moving forward to be able to show something or fit into this story?
KVD: Realistically we might not get to the import side in time. Probably the bit we can show is similar to the redcap story. Perhaps have AM take the lead and note that there are other systems doing similar things.
OB: PhenoTips would be in a similar perspective as well. Happy to share a slide or add in some videos. Would prefer RedCap is demo'd, then talk about other support for other systems.

Lindsay to reach out to Jules about Phenopackets piece of demo.
OB: If we can do an import and either map to phenopacket or ped with the tooling. If there's something else he had in mind it would be great to get his feedback.

AM: Two use cases - inferring higher order relatedness, kind of what I'm already doing the test cases. Just need more examples. Want to follow up on the second use to see if we can get more details. Never thought we'd get into the problem of figuring out who someone is, might not be something we can solve with the reasoner. Get a concrete example of where that happens.

Worth having a slide on calling for use cases. Frame in a way that already have a core set of use cases.
OB: A slide about how to be engaged - show that it exists in the Github and where user stories are,
GW: Have we reached out to Driver Projects?
LS: Yes, I can reach out to the folks that originally filled out the survey

GW: I can take a look through the use case document and pull some pieces out.

AM: Initially I had a quick look and there are interesting changes but not too difficult. I think some of those are good for getting started with.
OB: Meen and KF might have some additional cycles as well.

2021-08-19: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Alice Mann (GA4GH), Elizabeth Janes (GA4GH) , Katherine Johnston (H3Africa/H3ABioNet), David Hansen, Moni Munoz-Torres (Monarch Initiative), Melissa Cook (NCI/CRDC), Meen Chul Kim, Bing-Jian Feng

	Actions Arising	Assigned To	Deadline
1.			
2.			
3.			

Global Alliance
for Genomics & Health

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH 9th Plenary registration - Submit your feedback on a GA4GH Schema Library by August 31	Grant Wood, Orion Buske	5 min
2.	Task List Reminders: - Add brief paragraph and link to the model in the minimal dataset document. We will need to have an open for public comment *prior to SC vote*. If we can have the document ready by Sept. 1st (at the very latest, the sooner the better!), we will be in good shape - Circulate the terminology github discussion page to the Subgroup for collection of use cases - **DONE**, but Lindsay can send a reminder - Update the tooling to support BODICEA v4 - Connect with Jeremy/ GA4GH dev team re: dev support task items	Grant, Orion	5 min
3.	FHIR Demo Update - Using one of the released Kids First studies titled *Kids First: Adolescent Idiopathic Scoliosis*. This study is about congenital birth defects having 651 participants (including family members) and 557 pedigree relationships between them.	Meen Chul Kim	15 min
4.	Status updates from each activity area in prep for the plenary and connect ● RedCap Plugin and OpenPedigree ● Trakgene	Orion	15 min
5.	Feedback on terminology use cases that were submitted (thank you Ronald and Bimal!)	Orion	15 min
6.	A.O.B.	All	5 min

Minutes:

- Task list: Orion will do this by the end of today, we need an open for public comment period. Having that ready before the end of the month is ideal
- Lindsay circulated a link for the terminology github, thank you for the use cases contributed. Please contribute any others if you have them
- Incredibly valuable for understanding the scope and reasoning behind the terminology. Any feedback in next week will be especially helpful and will be incorporated in next version of the terminology
- Connect with Jeremy re tooling → Orion

Global Alliance for Genomics & Health

Meen, Kids First Study

- Kids First: Adolescent Idiopathic Scoliosis
- - PI: Jonathan Rios
- - Individuals: 651 (168 probands = 168 families)
- - Affected Status: 380 (true: 282)
- - Relationship: 557
- - Condition: 286
- - Pedigree: 168 (TBD)
- This study is genetic analysis focusing on scoliosis, released on dbGAP. See above for statistics
- Clinical trial not based on full EHR so one limitation
- 282 affected participants, focused condition
- 557 relationships between participants, so some of participants consist of one family
- 168 families, in terms of our pedigree model, 168 pedigree resources in FHIR. for next meeting will do part 2 demo about pedigree and have been making some updates to the pedigree tool as well. This will be demoed next week
- Show you two-fold how to query FHIR API using postman and also show you FHIR data dashboard that Tim has been developing
- Here is one disclaimer when comes to data dashboard expect to see advanced analytics functionality but this dashboard is for FHIR so will add such functionalities but at the moment, querying FHIR API is not very visually appealing/comprehensible, returns JSON blocks. Wanted to make a similar visual prominence of this result as a form of dashboard. We have 651 patients, fetches the data in real time.
- Shows a bit of statistics but not very interesting
- Click into actual patient
- Observation shows infection status, annotation etc
- relationships : indicating extension relationship, one directional relationship explained.
- What does it mean when you hadn't pulled in the pedigrees? We have the relationships but the composition resource, grouping them together has not yet been pulled in.
- In implementation guide 3 pillars, individuals (patients) relationship (family) and composition (pedigree)
- Orion: Parent term in terminology for biological parent, you could query that through this right? Find anyone with any parent in the terminology, is that true?
- That is also possible in theory because in the data ingest module, map one term to one rel code. We can convert this dictionary to a list.
- Orion: does the FHIR end point you built have the full structure of the ontology? (yes) you could query rel2 to find anyone with any biological relationship, yes this is correct for the next demo.
- This is part of the dream to do those queries with that specificity. If you could import the full value set into the FHIR resource so that then you could do those queries rather than just exact matches of entries but instead to closure queries. Very powerful.
- Moni: what is the condition resource in this?
- Means abnormal phenotypic features
- In FHIR condition resource type normally used for diagnosis or relatively permanent phenotypic features.
- So i will do the demo a few more times and please let me know how we want to organise this demo contact for plenary.
- Orion: we will sort this out in the next couple of weeks
- How many of these families were more than trios?
- Meen: don't have a specific number at the moment, might be for an example to dig into more if we demonstrate some of the expressiveness of it.

Status updates

- Alejandro sent update on some of work on red cap plug in, has moved forward well and is set on track for being ready for the plenary for a demo
- Anyone to give an update on the trakgene side? There have been some convos for one of their developers to work on the open pedigree.
- No other updates to share here.

https://github.com/GA4GH-Pedigree-Standard/family_history_terminology/discussions/9

- Inferring two people are the actually the same person
- Use case where the definition of what it means to be related means something different in different cultures.
- Person being described differently on different sides of family when collecting family history
- Evaluation of whether or not if one person will be mentioned twice
- Any thoughts here or in general as to how this might inform the terminology?
- Think its family collection tools rather than terminology itself
- KJ: someone raised living in their homes called brother or sister but actually their cousin, some of the ways in which somebody could be duplicated in different situations. Really important when collecting family history that the participants have an understanding of what consists of a biological family member before you collect family history, so you know you are getting the right people.
- Orion: in terms of trying to make steps towards helping with those challenges, should we be explicit as possible in terms and leave to others to ensure interface uses those terms where appropriate. Should we have terms that culture those terms culturally better?
- KJ: challenging, taken approach that the terminology is clearly described and differentiating biological or nucleating family member. Really comes down to the people asking the questions
- Are there relationships that should be a term but there isn't one?
- Justs need to be clearly stated that definitions need to be clearly understood and explained upfront
- MC: because of all the cultural nuances, what is it that we are actually asking with collecting these data, what do we hope to understand and assess from it if there's so much variability in the meaning of the terms
- Orion: fair question and absolutely, how many layers of the onion do you want to peel away.
- MC: what do you hope to learn by asking the question and asking it that way? If so ambiguous that you can't group and interpret the data, what value is the data. Can we ask it differently to get a better position? Even biological mother is difficult, is it gestational carrier?
- Grant: We are going after learning familial risk, this conversation is out of scope for the activity of this group. This group needs to come up with a standardised recommendation on how to collect family history data and used in many different ways to do risk analysis. The out of scope part in my opinion because of the complexity of trying to manage how different cultures might define different terms and managing that is just out of scope for what we need to do here.
- Specific use cases are always helpful but not if use case is we need to come up with answers that satisfies different cultural definitions. That is out of scope
- Orion: we aren't going that far down this path, we need a set of terms that are useful for solving practical problems, which knowing who contributed an egg to a person's birth, gestational carrier are both important information, we need to explicit to defining these things but the added onion layer, given data may be coming from an EHR system that this is someone's mother, how useful is that? Other cultures where there is data collection, what is our relationship with possible ambiguity
- We want to have a bit of making sure we articulate why we made choices to maybe not have a term that reflects a particular familiar relationship
- Grant: but understand the work we are doing, especially standards work, actually can slow down the progress if you try to cover every single possible question as opposed to coming up with a solution that solves 95-98% of use cases. Instead of looking at all of these complexities, to really come up with initial solution that does the majority and then try to tackle other outliers in another phase.

- Moni: personally when to consider standards it is important to consider the social political climate. Take into consideration the ethical considerations.
- MC: echo both, have to be crisp about the question and the use case. These things are very complex. Use a different term that is really crisp if biological mother is not clear enough.

2021-08-05: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Alejandro Metke (AGHA/CSIRO), Bing Feng (UUtah), Melissa Cook (NCI), Elizabeth Janes (GA4GH), David Hansen (AGHA/CSIRO), Zornitza Stark (AGHA), Hannah Calkins

	Actions Arising	Assigned To	Deadline
1.	Add brief paragraph and link to the model in the minimal dataset	Orion	Next Meeting (Aug. 19)
2.	Once the above is complete, circulate the draft to the full WS for approval	Lindsay	ASAP following completion of the above item
3.	Talk to Bob about best method of engaging with the HL7 community	Grant	
4.	Circulate the terminology github discussion page to the Subgroup for collection of use cases	Lindsay	
5.	Update the tooling to support BODICEA v4	Orion + Dev Team	
6.	Connect with Jeremy re: dev support task items	Orion	

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH 9th Plenary registration - Submit your feedback on a GA4GH Schema Library by August 31	Grant Wood, Orion Buske, Lindsay Smith	5 min
2.	Reviewing previous action items: - Testing Kids First data (MCK) - Talk to Bob about best method of engaging with the HL7 community (GW)	All	5 min
3.	Pedigree Implementation Demo - Draft Diagram - Confirming tasks and responsibilities: does everyone know what they need to do?	All	30 min

	- Lindsay started a task board (needs filling out!) - From GA4GH Dev Team: "It would be great if you could sketch out what code needs to be written to make the demo possible, and then I can see where my team fits into things. For example, if Exomiser is a Java application, we could write an open source library for GA4GH Pedigree file management and put it up on Maven Central for Exomiser to import, that sort of thing. There's potentially some further cross-connections with other GA4GH work streams we could apply to this project, namely, if we wrap the necessary tools in a workflow language, we could apply the Cloud WS standards to execute the tool over REST API. We could see that as a second phase of the demo once we get the first phase up and running."		
4.	Publication of our minimal dataset?	All	10 min
5.	A.O.B.	All	5 min

Minutes:

From Meen:

1. pedigree-tools: I have local commits (will create a PR to master shortly)
    a. ~~Pulling relationships directly from the REL OWL (Done)~~
    b. Linting the JS scripts against ESLint/Airbnb JavaScript Style Guide for enhanced code readability (in Progress)
2. Pedigree demo in FHIR
    a. I got a confirmation that we can use released Kids First studies for the upcoming plenary meeting. One of the released studies has a very nice multi-generation structure. So, the steps I am thinking about include:
        i. ~~Create the complex pedigree using the pedigree tool (Done)~~
        ii. Write an ingest script in order to transform/load the Kids First family relationship data into a FHIR server (in Progress)
    b. I have two questions:
        i. By when do we want to get this ready?
        ii. Which FHIR server do we want to use? We (KFDRC) have multiple FHIR servers, but they have regulatory issues (like a study funded by an A fund should go to an X server and this data can be shown to the ones under the A IRB, etc.) and we may not be able to ingest this test pedigree data there. I can test-ingest into my local FHIR server and show you during our WG meeting, but the reporting purpose, it may not be ideal.

AM: The Demo with Jules was on a local machine, so depends on where we want to run it. Could be a HAPI FHIR sever
OB: He can run locally, so that sounds like it would be sufficient

AM: What are we planning to show in the demo. Draw pedigrees? Run exomiser?
OB: Are we visually demonstrating the flow, or here we have done those things. Better story if we take entry through to analysis.

Storyboard it. Here's a tool where someone is drawing a pedigree, have a tool where we show transferring to another format, putting into visualization tools and downstream applications - exomiser or Bing's cosegregation analysis tool.

Balancing which tools are used in the process, balance out screen time/visibility with each. Show each team that has gone through the effort of integration, have some opportunity to present that. Follow the form of the diagram.

BF: My tool support Bodicea v4, but the tool only supports version 3. In V3, the er+ columns have joined into one column for v4. Not sure whether there are any new version included.
OB: Should update the tooling so it can support export in v4.

AM: Seqr platform supports ped that's something that could be added to the legacy tools.
ZS: Trying to move data from my project from Redcap. Moving HPO terms, pedigrees. Only been exploring how to do it, so don't want to overcommit.
GW: Is seqr only going to support ped or will it move beyond that?
ZS: I could ask.

OB: In order to get here, assuming we're all in agreement. We would be making a call at beginning of October who has implemented what, and demonstrating it. Each of the tools have their work cut out of them in terms of import/export. Q around shared tool and terminology. Would be helpful for us to do a pass through and add comments.
AM: Yes, please have a look. Trying to address everything Soichi had commented.
OB: I'm happy to contribute on the tooling side. Sounds like we need import/export for pedigree in the context of phenopackets and in the context of FHIR, and export in BODICEA v4.

AM: Interested in understanding additional use cases where we can use the reasoner and do more sophisticated things. Will need some help with these. Discussed having a list of use cases or examples. Find some use cases that can be added into actual test cases. I have some minimal testing but would be good to have additional tests beyond that.
OB: Do we flag the discussion capability in Github user stories?
[Lindsay can re-circulate this]

AM: If you have limited information that was captured (in a chat box for example), might want to infer additional relationships that were not captured. Need more example of where missing relationships might be useful to fill out. Another would be how to deal with conflicting information, eg. multiple pedigrees referencing the same person and one might have a conflicting fact, wouldn't want to merge without checking.
OB: Merging on it's own. As part of merging, there is the related task that is more specific to the terminology. Resolving down to the lowest common denominator. Part of the way the terminology was originally set up, was that it only has ambiguous relationships aside from parent-child. But if we were to add things that weren't ambiguous, like maternal aunt, resolving down to explicit relationships with a reasoner.
AM: In the end, what you want is the core set of relationships you can use. Can also help you determine if you have enough information in the pedigree to get to that point.
OB: The extended family only being ambiguous is probably not serving us well anymore. Make sure we support all common terms from the FHIR family set as well, so that given a family history coming out of an EHR system, we can take it, reason on it, and map it down.
AM: One of Soichi's comments was around the absence of some of those.

OB: From a demo perspective, where are the sorts of places we'd want to connect in? Grant, given your connection to patient-facing side, any thoughts there with intermountain's effort?

GW: Possibly, but not sure about the timing. We are reinvigorating our family history tool and they are looking at adding new developer resources. That would be an opportunity to do something, but don't think it will be solid in the next month or two.
OB: If the only systems we have generating pedigree data are pedigree tools, then it will be reduced to parent-child relationships inherently.
AM: Could still simulate a conflicting pedigree.

AM: Mentioned we're testing the redcap plugin. They're happy with open pedigree, but the only bit they've commented on, is that whenever you create a new node, always have to create a couple (eg. parents). Couldn't find a way to create a single parent.
OB: Would love for that support to be added.

Combine the minimal dataset and standard into one publication?
Circulate to the wider Work Steam for approval -> aim for consent agenda item at the next steering committee meeting

2021-07-28: Clin/Pheno Full WS

Chair: David Hansen, Melissa Haendel
Attendees "Name (Project Affiliation)": Susan Fairley(GA4GH), Moni Munoz-Torres (Monarch), Arthur Hermann (Kaiser), Orion Buske (PhenoTips), Grant Wood (Intermountain), Ian Green (SNOMED), Nebiyou Petros, Hannah Calkins, Elizabeth Janes (GA4GH), Nicky Mulder (H3Africa), Davera Gabriel, Francesca Frexia (CSR), Alejandro Metke (CSIRO/AGHA), Robert Carroll (VMC/NCPI), Jules Jacobsen (Monarch/QMUL), Francis Jeanson, Heikki Lehväslaiho, David Hansen (CSIRO), Ada Hamosh, Adriana Malheiro (NIH), Bing Feng (UUtah), Meen Chul Kim (CHOP), Katherine Johnston (H3Africa), Mogomotsi Matshaba (H3Africa), Shubashini

	Actions Arising	Assigned To	Deadline
1.	Please review and add your thoughts the the GA4GH Schema Library requirements document	ALL	**August 31**
2.			
3.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH 9th Plenary registration - Driver Projects leads/representatives - please	David Hansen, Melissa Haendel, Lindsay Smith	5 min

	keep an eye out for an invitation to present at GA4GH Connect		
2.	Subgroup Updates/Resources: **Phenopackets** - v2.0 has been approved by the GA4GH Steering Committee! - Regular meetings have been cancelled - v2.0 has also been submitted to ISO for DIS (draft international standard) ballot - Latest documentation - Draft core implementation guide **Pedigree Standard** - Moving into milestone 3: implementation and testing. Aiming to have a demo together for GA4GH Connect (mid-October) - Draft pedigree model - Family history relations ontology - Draft FHIR implementation guide - Draft implementation demo diagram **Computable Cohort Representation** - New subgroup meeting Tuesdays (rotating times) - In landscape analysis/defining scope stage, nothing has been formally proposed to GA4GH Steering Committee yet - Landscape analysis spreadsheet - Use case collection form - Scope document - Meeting minutes	Lindsay	2 min
3.	GA4GH Schema Library: Requirements Collection - Feedback Document	Susan Fairley, Jeremy Adams, Francesca Frexia, ALL	45 min
4.	A.O.B.	All	10 min

Minutes:

SF: How we might re-use or align schemas in GA4GH. At minimum make it easier to find where these things are. There have been previous attempts at this. This eventually was passed along to TASC - our technical alignment subcommittee. The discussion at TASC came to a few conclusions: some sort of mechanism for document schemas across GA4GH would be helpful, and we likely want this to be something on the lighter end. Promote and support model reuse. What are the requirements and what are the potential problems in using it. Potential for better documentation and reuse in GA4GH.

This discussion has stemmed from Schemablocks, a potential solution that was developed previously.
GW: Are we already seeing some problems with multiple work streams doing different schemas?

SF: There are areas where we will have the same or similar entities popping up in one standard or another. In the future, if these schemas could more easily map to one another it would enable better interoperability. Some of our standards are building out and will touch upon work seen in other work streams.
JA: Central common platform or place where WSs can define or publish schemas that can easily be picked up by other WSs. One example could be pagination with APIs.
GW: Other SDOs have all experienced and had to deal with this. HL7 v3 created a reference information model, and all other v3 messages would have to use those data elements.
SF: May be the case where it will take some time to build out. Would look to encourage people to reuse, but where they have to build something else they could do that. Ask if there is a need for a new model and if we could reuse.
AH: How granular, when talking about schema, do you mean?
SF: See schemablocks as a jumping off point for this discussion.

OB: This sounds great, some of that process has been going on in the phenopacket ecosystem. Making sure we're aligned and only adding to the existing schemas that were created. Additional level of how we are communicating this, how to document this. Benefits right now of the schemas expanding, whenever there is an expansion, that triggers pairwise or three way conversations across Ws, this will help
RC: There's a space of are you defining a suite of APIs, vs how are WS defining or handling this information. What's the best practice across GA4GH. If you want to talk about a pedigree, here is the appropriate piece of information you ought to use. Talking about different types of tools - DRS is a service, vs pedigree or phenopackets which is more of a model
SF: When we start to talk about pedigree or phenopackets, I'm sure there are different things going on in Cloud. The topics of files and how we talk about and reference them will come up in a number of different places. Comes back to what we want the scope to be.
RC: Hitting the point where we're crossing WSs, requires a lot more interaction. Have had difficulty finding information I needed previously. I'm trying to do x, what standards will be useful
GW: Start with a front-end website, once we have tht library of components, people can put together a list of components for use cases

JJ: APIs are one thing, schemas are lower than that. The schemas themselves are hard to get them to work together. Eg. with phenopackets, with one version we have it develop before VRS, for version 2, we deprecated that and addin VRS. Then we had a problem with technology mismatch - Phenopackets is protobuf, VRS is JSON schema. The VRS team wrote a converter so we could import it into the schema and reference it. Can do it, but involves a lot of cooperation. Curated set of standards that have been approved?

Mismatch between JSON and protobuf, they don't convert that easily
JJ: Would love to talk to TASC about how we did it with VRS and Phenopackets. Could use something like LinkML, way of defining a schema salad. Defines it in a metalanguage that can convert to other formats.

FJ: This problem has arised again and again. One project worth looking into is SchemaBlocks. Having a fundamental reference implementation would be helpful as well. ISO and W3C have a specification language they use. Open-stand.org. Would shy away from anything that gets into too specific
JJ: The standards we publish have reference documentation that is agnostic, and I think it should remain that way. A more agreed upon way of specifying would help with implementation

Document is open for comment - hoping to move forward by the end of August/early September.

GW: Come back to Clin/Pheno later and see where we're at.

SF: Look at scheduling some time at the October connect meeting, which could facilitate discussion across the WSs. Once we get through the requirements gathering, we need to collate and share that back.Please let us know if you'd like to be a part of this.

2021-07-27: Computable Cohort Representation

Minutes: The minutes for Computable Cohort Representation can be found here.

2021-07-22: Pedigree Subgroup

Chair: Orion Buske
Attendees "Name (Project Affiliation)": Elizabeth Janes (GA4GH), Katherine Johnston (H3Africa), Bing Feng (UUtah), Alejandro Metke (AGHA/CSIRO), Kathryn Van Diemen (TrakGene), Jules Jacobsen (Monarch/QMUL), Meen Chul Kim (CHOP)

	Actions Arising	Assigned To	Deadline
1.	Set up a task board for implementation demo	Lindsay Smith	
2.	Confirm with Allison about testing with KF data	Meen Chul Kim	
3.	Reach out to GEL (Antonio) about their pedigree standardization	Lindsay / Orion	
4.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH 9th Plenary registration - Please note that a GA4GH Schema Library: Requirements Gathering discussion will take place at the full WS meeting on July 28	Grant Wood, Orion Buske, Lindsay Smith	5 min
2.	Reviewing action items: - Add issues from the redcap/open pedigree implementation to Github (AM/DC) - Add mod to the pedigree tools repo to enable pulling from the relationship terminology directly (MCK)	All	5 min

	- Talk to Bob about best method of engaging with the HL7 community (GW)		
3.	Pedigree Implementation Demo - Draft Diagram - Key Next Steps - GA4GH Staff Support	All	30 min
4.	A.O.B.	All	10 min

Minutes:

Draft Diagram & Demo Planning
OB: The conversion tooling carries the format in some way. Nail down those key points as a proof of concept for plenary. It helps clarify what the minimum points are that we can demonstrate. Bing - what capacity do you have to implement CPF within the conversion tooling was well, which might be easier than adding to the tool.
BF: Can convert between ped and bodicea. Thinking of bodicea to input, and then write a function to output into cpf format, that can be used in the cosegregation tool.
OB: From my perspective, anything that shows enhanced interoperability through systems that would have an otherwise harder part of talking. Showing a way to bring other tools into the system is valuable.
JJ: It's hard to make new data sources and import them, so I'd agree with that.
OB: Are you saying it would be easier if we built the ability to export into bodicea format, that is preferable, and you would tool to convert, vs either adding a direct ability to support cpf into the conversion tooling, or that you a converter from phenopacket or FHIR with pedigree into cpf.
BF: Depending on the time. My tool can directly read bodicea. In the long run, would like it to read GA4GH Pedigree format.
OB: Is there any mods you do to the bodicea format?
BF: Bodicea uses traditional proband definitions, so it's not ideal but can take it for now.

OB: What are your thoughts from Phenopackets? Right now there is a branch, do we host in some other place with a description of how it is used?
JJ: It's mostly ontology terms right? Could use native phenopacket ontology terms, could create a pedigree format definition. To put into the schema itself would need to make a branch and merge it in.
OB: Want to make it easy for people using phenopackets to also receive pedigree information.
JJ: In the model we have a family which includes a collection of family members which can be phenopackets. Right now we have a phenopackets pedigree which is basically ped in JSON format. Could make GA4GH pedigree an option. Would need to get people to the correct format (v2.1) of phenopackets.
OB: Any preferences for the different options?
JJ: The format itself has been settled down quite well.
OB: The main pieces are in place yes.
JJ: Absolute easiest thing to do would be to create a standalone schema
OB: It's a wrapper around stuff in phenopackets
JJ: Composing the pedigree ontology from bits of phenopackets
OB: To repeat back - instead of changing current phenopacket family, we would define a new top level message and then have it import and pull in existing phenopacket message definitions as needed to fill that out.
JJ: Right. If you look at phenopacket base package, you might need to define something like a relationship.
OB: Recommend that it would be a basic phenopacket family.pedigree
JJ: It would be an either or option.

JJ: Have done an example for how you would write the MME model with phenopackets, could do a similar thing with Pedigree - https://github.com/phenopackets/phenopacket-schema/blob/v2/src/test/proto/mme.proto
OB: Importing base and referencing those same ones.

MCK: Ran a test on the mod. Created a ticket regarding pulling the ontology files. Will add that feature next week to the repo. It's running well!

AM: I saw you commented on one of the issues for the family relations ontology. I will be tagging some folks.
OB: Any news on naming from the Monarch folks?
AM: I think that was from the original plan of incorporating the ontology, but we should probably rename to something that doesn't clash.
KJ: KIN? For Kinship.

OB: From the exomiser side, we would either use phenopacket family, or fall down to ped
JJ: Exomiser is coming out in hopefully in a couple of weeks, and that's going to have initial phenopackets support for v1 phenopacket. It also supports ped natively. In terms of a conversion, unless there is a java tooling. Most of the tooling is in java?
AM: Yes
OB: And Meen has python covered.
JJ: The converter is written in node and the tests happen in Java. It probably wouldn't be so hard to write something to get into a ped file. At the moment, we can take pedigree from ped or family.
OB: Is this realistic? Or aspirational and we should be dialled down to say look we're running exomiser on data coming through this ecosystem
JJ: I would want to say exomiser for that. Would be handy to have some java tooling available for that.
OB: If we have more GA4GH resources, it might help flushing out the converter in a few different languages.

OB: One of the questions around feedback, what is the scope we would like to have for that tooling? Do we want to have n by n in terms of conversions that are possible? Is it around being able to convert in or convert out. Sense of their relative priorities? Right now there is input of a variety of format and output into ped or GA4GH pedigree format. Bing mentioned export into Bodicea would be a high priority.
BF: Ped also works for proof of concept.
JJ: For exomiser I just need the import.
OB: It almost worth adding the export of bodicea if there is richer data there. Show use of data beyond ped to show better analysis, representation, etc. Beyond ped, those other cancer occurences, those are used by your tool?
BF: Yes. Bodicea is better than ped.
OB: Ok, let's do that.
JJ: If in an ideal world the bodicea isn't necessary anymore, would that enable better analysis in the future?
BF: Bodicea is not as good as GA4GH, and it's a moving target. I want it to be able to read the GA4GH format.
OB: In order to meet that gap, Bing you would be importing in phenopackets or FHIR context.

JJ: Ask genomics england about how they represent pedigree? Ask Antonio about standardizing. They use a slightly different representation. They do have twinning.
OB: Is there any data exomiser or congenica would like to be getting that they aren't already?
JJ: cip is it's own data structure. Not much more than direct family relationships.

MCK: Let me talk to Allison about KF. We have KF portal, and the release protocols are open data. Any pedigree files I can contribute I can double check.
OB: Is there any practical use in that project for the first version of this format?

MCK: Definitely the protocols are interesting. Is there any specific kind of requirements you are thinking of?
OB: Export/Ingest or internal use for which you think adopting that FHIR IG would be useful or testable within the KF world
MCK: Yes, we already have ingested family relationships using a pretty similar approach. Only difference is that we use observation resource type. Internally still having a discussion about observation vs family member history.
OB: Is there a way we can present to the world a unified face or have some conclusion for observation vs family history by then.
MCK: On top of Allison's confirmation, I believe I can do both. Reingest with the pedigree IG model. Could demo fmh version.
OB: Request between now and plenary would be is there a way KF can show the use or retinterpretation with the new IG.
MCK: Does the size of data matter? We have a variety of studies with a variety of relationships? Difference in orders of magnitude. Some studies are only trios, some are multi.
OB: Use the multi-generational obnoxious ones that wouldn't be straightforward with ped.

2021-07-13: Computable Cohort Representation

Minutes: The minutes for Computable Cohort Representation can be found here.

2021-07-08: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": David Conlan (CSIRO), Elizabeth Janes (GA4GH), Bing Feng (UUtah), Meen Chul Kim (CHOP), Alejandro Metke (CSIRO/AGHA), Sean Garin (NHGRI), Kathryn Van Diemen (TrakGene), Melissa Cook, Mamana Mbiyavanga (H3Africa)

	Actions Arising	Assigned To	Deadline
1.	Add issues from the redcap/open pedigree implementation to Github	Alejandro Metke and Dave Conlan	
2.	Add mod to the pedigree tools repo to enable pulling from the relationship terminology directly	Meen Chul Kim	
3.	Talk to Bob about best method of engaging with the HL7 community	Grant Wood	
4.	Vision for interoperability pilot - put together a diagram	Lindsay Smith, All	

Global Alliance
for Genomics & Health

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH 9th Plenary registration - Adding to the Github readme	Grant Wood, Orion Buske, Lindsay Smith	5 min
2.	Redcap plugin implementation feedback	Dave Conlan, Alejandro Metke	15 min
3.	Preparing implementation demos for Plenary (Oct. 12 - 14) - What groups do we need to reach out to? - Meeting currently scheduled for Oct. 12 @ 2:30pm PT / 5:30pm ET / 10:30pm GMT / 8:30am AEDT (1.5 hrs)	All	20 min
4.	Pedigree tooling repo updates	Orion Buske	10 min
5.	A.O.B.	All	10 min

Minutes:

Redcap Plugin Update
AM: Found a few issues doing this implementation. Will add these formally as issues at some point.
DC: https://dconlan.github.io/open-pedigree/index.html
The basis of the redcap plugin uses Orion open pedigree, forms it in a new window. Open Pedigree. Have already done a version of an export to FHIR. The link above is a fork of the project. Where it's different is in the export section is a button for GA4GH FHIR. Will save to your harddrive and open with an editor. It does have my code for hooking up to Ontosever, will do searches for ontologies. Been trying to do full circles, so if I export can I import. Ran into a few issues: one of them is the comments field. Put the comments in the family member history records, but now it represents a relationship so the comment doesn't really fit anymore. Another problem is all the issues around dying early (eg. decreased, stillborn, miscarriage). No deceased string in the record. Didn't know what to do with not in contact.
'n' represents multiple children, not sure how to represent that in the FHIR spec. Can mark a relationship with consanguinity, but there is a slight difference from what the spec does from what the editor does. The spec says only two generations, so will only mark if it's a great grandparent that is shared.
AM: The choice we made around the choice of resource to represent the pedigree itself. At some point we were thinking of representing a system where something would evolve. Composition is immutable which might not be ideal. Talked about our own resource for pedigree, but that is the most difficult one. Can add these to the Github issues for discussion.
MCK: Exporting to GA4GH FHIR resources - assuming we have multi generation family, then when computing the relationship between patients, how is it computed? Is it starting from proband or one step proximity?
DC: Spec says you can create a record for a relationship, but I haven't been linking all the relationships. I've been linking parents and partners, for tree building and consanguinity. With a big tree, it can be hard to go up and down and say what the relationship is.
MCK: Now, it generating parent child and spouse relationships. Second question, based on the GA4GH FHIR spec and the composition resource, we can only set one proband. Given a multi generational family with multiple probands, how would it be handled?
DC: It would only do the one proband.

GW: Talked about decreased not having a string. I know in the minimal dataset table we talk about decreased and there's a FHIR family member history deceased resource. Most people would use that as a boolean element.
DC: That's what I used in the old system. But with the way the model worked, a family member history record doesn't represent an individual as such, it represents a relationship.
AM: The nodes are patients, and patient doesn't have a deceased string. Easy way is to extend patient and a comment.
GW: Saw in the box there was a box for genes, what goes in that box?
DC: Pulled from SNOMED
AM: Could hook up to HUGO notation.
DC: The phenotype uses HPO, but can configure and target any terminologies you want.

DC: Was given a patient question form to fill out, plan is to have each field associated with a FHIR path to where its associated with the pedigree. Use the FHIR path to build the pedigree diagram.
MCK: The redcap plugin, is that the one adding a form, or is exporting data.
AM: Pretty much both.
DC: Bring back an svg drawn by the editor. That's an extra inside the FHIR record, there's another section for pedigree diagram and it contain the svg as a document reference.
AM: Looking into how to bind the contents of the editor more closely, so you could send out a file that a patient fills out that could pre-populate the diagram.
GW: Are there users looking at this?
AM: Team at royal brisbane, mostly of genetic counsellors. Doing cardiac-related genomics work.

Pedigree Tools Repo
OB: New repo for pedigree tools. This is a starting point for setting up tooling related to converting between pedigree formats. Current functionality is you can take data from ped linkage, gedcom and boadicea and coert into GA4GH pedigree format. Serialized in JSON format. Can also serialize into protobuf. From there, now have a file you can store, share or import in the tool to convert in some other format.
Meen has some initial code in Python that he might be willing to contribute. Have multiple different programming languages or environments and allow people to have the tools they need.
MCK: I'll make a mod so you can pull from the relationship terminology directly.
OB: Relatively straightforward to use the protobuf implementation he laid out for phenopackets. In theory, as this gets used, all the various tools and integrations will evolve quite easily.
This is a public repo, everyone should be able to see and access it. Anyone can see the code and fork it and make a local version, can also contribute to suggest changes.

OB: When is the right time to engage outside members?
GW: HL7 is waiting for me to come back and formally share what GA4GH has done.
AM: Would be good to have a first version. Depends if you want to tie in with a formal HL7 process. Would be valuable to get some feedback now. Logistically not sure how to align GA4GH and HL7 processes.
GW: Is there a balance where we can separate feedback on the process? Get feedback but not the formal process until we're ready.
AM: I would organize to present this work at an HL7 meeting and point interested people with the Github.
GW: Will talk to Bob about this.

Implementation Demo Planning
AM: Present something similar to what Jules and I did with Phenopackets. Pick a use case. Show a few systems sharing data rather than individual efforts.
GW: Bing is this something we can work together to accomplish.
BF: Wondering what the requirement is? If grant is settled, I can write some code but depends on if it is funded.

OB: There are a few different groups and way to frame data collection coming into pedigrees. Redcap/open pedigree is one resource, Phenotips can offer that as another one, not sure if Trakgene could implement something by October. Systems that could get information from users. Then we have some tooling that allows for converting between different formats or getting it into this new standard. Ideally, that would be those visualization applications, also have analytic applications representable - like Bing's tool. In theory, if we wrote a script and add support for CPF into this tooling, we can bring that into the ecosystem and show that by having a common point we can make interop easier. Could discuss with Jules and get exomiser to consume that.
We have a few flavours of the conceptual model, and going between them is going to be important. Curious is phenopacket on FHIR and that work can somehow be leveraged or abused to help with this process.
AM: Don't think so. Hard coded. Could think of something more sophisticated but potentially a lot of work.
OB: At worst, it's adding import and export into this command line tool.
AM: We're in a better position because we started with a conceptual model. At a high level shouldn't be mismatch. Other challenge is maintaining the mapping. If we make changes would need to automate them to the implementation.
Interesting thing we could do, we have a chatbox used in genetic counselling. It asks questions about family history, if we could turn that into a representation of a pedigree. Could turn that into systems that consume.
GW: Really excited about the vision. Would like us to diagram that vision and start working towards it. On Oct.12, even share where we're at, even if we're not fully there yet.

2021-06-29: Computable Cohort Representation

Minutes: The minutes for Computable Cohort Representation can be found here.

2021-06-15: Computable Cohort Representation

Minutes: The minutes for Computable Cohort Representation can be found here.

2021-06-10: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Hannah Calkins, Alejandro Metke (CSIRO/AGHA), Bing Feng (UUtah), Mamana Mbiyavanga (H3Africa), Meen Chul Kim (CHOP), Melissa Cook

	Actions Arising	Assigned To	Deadline
1.	Share initial feedback on the Redcap plugin implementation	Alejandro	Next meeting
2.			
3.			

4.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH BED v1.0 is now open for comment - Cohort Representation Subgroup schedule starting next week	Grant Wood, Orion Buske, Lindsay Smith	5 min
2.	Milestone 3: Implementations (Continued) - Bing's use case - What should a pilot include? What do implementers need? How can we best support them?	Orion Buske, Bing Feng, All	35 min
3.	Relationship terminology updates	Alejandro Metke	5 min
4.	Pedigree tooling repo updates	Orion Buske	5 min
5.	A.O.B.	All	10 min

Minutes:

Implementations/Pilots
We need min. 2 implementations for the Pedigree standard. What is required to do something like this? How can we support implementers?
BF: Trying to map the GA4GH minimum set with the data requirements for cosegregation. Most of the requirements have been fulfilled with the minimum set. Don't know how much we need to do.
GW: Draft document that would show the data model in FHIR, and that could be something Bing could follow based on the data he's using?
AM: I can share the IG that has documentation and examples on how to use it. The Redcap plugin could be one of the two server implementations. Should focus more on testing the spec rather than worrying too much about the GA4GH requirements.
OB: There is an implementation within Phenopackets. If there was a consumer implementation with Exomiser that could be another, maybe with PhenoTips as well.
LS: Is there going to be a standalone JSON version?
OB: Yes, planning on setting that up. Direct replacement for ped or other custom JSON formats. Getting the tooling in place and getting the PhenoTips bits out of it. It should out the box support importing different formats, including our format, and could support exporting into GA4GH pedigree.
GW: Bing would not have to write any code or create a database schema, but instead we could do this as a mapping that shows our spec on one side and how it maps to Bing's.
OB: There are two ways of looking at this - is there an implementation from which Bing benefits from the technical implementation (either FHIR, Phenopackets or direct), or recommendations about the data and fields.
GW: Proposing the latter as a first step, what do people think?
AM: Would be a shame not to implement it eventually.
OB: Good effort, but not sure we could call it an implementation.
BF: We have a team to do transference of data and they can also do the coding. I don't have the funding to ask them to do it for me, I do have a student that could learn but it would take some time.

BF: I have minimum and more extended requirements. For sex - there are different definitions in the minimum set, think it should be defined as phenotypic sex.
OB: In genetic counselling, the gender is extremely important, then the phenotypic sex at birth, in addition to having karyotypic sex which is seen as something different.
GW: In our data set we have biological sex and the gender. Are we talking about adding a third element for phenotypic sex, which can be different?
BF: My understanding is that biological sex is broad.
AM: In ontologies, biological sex would be the parent concept, with genetic and phenotypic sex being the “children”. From FHIR modelling perspective, it’s relatively easy to add.
GW: Do we need to add a third or redefine it note it includes phenotypic sex?
OB: We want to narrow the definition. From biological to phenotypic or phenotypic at birth.
BF: I would vote for phenotypic, because biological is too broad
OB: I don’t see it being feasible to choose one based on the broader needs of driver projects. Is there a common case? The more explicit representations possible but more involved in terms of the data model. It could make sense to make it more general. In terms of scope, most of the ones we;ve talked about have related to genomic analysis.
BF: Administrative sex is good for 99% of cases
GW: In the document we have an appendix, I’d like to add a paragraph about this into 8.4

BF: Can this format support multiple probands in one pedigree?
OB: The way we framed it, right now the answer is no. There was the proposal about distinguishing the concept of a family and the concept of a pedigree. One being a collection of individuals and relationships, and the other being the view of in the context of a particular health care process. Could have one family, but multiple pedigrees that reference that family but have different probands.
BF: The proband is variant based, added stress on the user, would like the proband to auto select the proband from the consultand. Consultand doesn’t change. The program can use this consultand information to find the nearest relatives
OB: My understanding that there is the desire to represent multiple individuals that were the first to test positive.
GW: I added in the notes in the document that a proband type may be needed.
BF: FPTP will make my program easier
GW: There are some problems we won’t find out until we do the testing.

Multiple birth status is on the individual rather than on the pedigree as a whole. One outstanding question there is that there isn’t a good FHIR resource to handle that.

OB: I didn’t see anything laid out in the document that it is able to be more than one thing
GW: Feel free to add any text that might make that more clear
BF: Is that using a list of proband IDs?
GW: Or proband ID types
OB: Is the expected value a type or an identifier for an individual?
GW: Combo of both, you’d start with a type and the identifier. Added that an extension to include a proband type is needed. Repeating field.
BF: Suggest to not make the relationships too complicated. Have not seen a situation where the pedigree structure is all correct. When I have a structure and do an experiment to infer relationships, I always find errors (sees full sibs when they’re really half-sibs).
OB: Avoiding recording more obscure relationships because the confidence is low? Or keep the structure itself simple?
BF: Keep the structure simple.

OB: Right now, there is a duplication of data. Have to change data in two different places. If father ID and mother ID is laid out, having a relationship to proband laid out is redundant in some ways
GW: We had relationship codes because of legacy tools that were using those as opposed to IDs
OB: If we require IDs, does it make sense to document the relationship to proband
GW: No… but not sure if people have legacy data that they want to put in there
MCK: If you require mother and father, computationally speaking we can infer the complex relationships. Won't need extra column for relationship. But there is always some legacy, we may need an additional layer of software.

OB: It sounds like you were starting to work on redcap, would you be able to provide some feedback?
AM: Yes, will have to talk to Dave
OB: Even if there is input on practically plugging into a tool, or changes we could do
MCK: The Vanderbuilt team just released v1.1 of the redcap FHIR plugin
AM: I'll have to have a look, it's much more integrated because they have the ability to change things right into redcap

2021-05-27: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Meen Chul Kim (CHOP), Mamana Mbiyavanga (H3Africa), Bing Feng (UUtah), David Hansen (AGHA/CSIRO), Melissa Cook, Alejandro Metke (CSIRO/AGHA)

	Actions Arising	Assigned To	Deadline
1.	Set up a readme in the Pedigree model repo	Lindsay, Orion	Done - readme set up but not populated with info
2.	Continue discussions on Bing's implementation next meeting		June 10
3.	Set up a pedigree tooling repo	Orion	
4.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping	Grant Wood, Orion Buske, Lindsay Smith	5 min

2.	Relationship Terminology - Java project update - Cont. discussion on plans for OBO Foundry & Java Project	Alejandro Metke, Melissa Haendel, Orion Buske	15 min
3.	Milestone 3: Implementations - What's required? - Bing's slides	Orion Buske, Bing Feng	15 min
4.	Pedigree Tooling Repo	Orion Buske	15 min
5.	A.O.B.	All	10 min

Minutes:

Relationship Terminology

AM: This a basic Java project. In the resources are the main ontology itself. Authored using OWL function. First attempt at creating the ontology based on the spreadsheet, with ideas from the family history knowledgebase. Easier to look at this in protoJ. If you navigate to the test folders you get an abstract test with utility methods. The first example is a validation method, test case with the Simpsons family. Set up using axioms and add them to the ontology and run the reasoner and check that it is not consistent. The other test is inferring relations that are not in the graph, using the logical encoded in the ontology. Also working on extending the app called FHIR OWL that auto transforms OWL ontologies into FHIR. Includes the data objects hierarchy.
OB: Really exciting. How hard will it be to switch over the acronym?
AM: Should be straightforward.
[Shared screen on Shrimp ontoserver] The big hierarchy here is around the relationships. Takes into account the hierarchies that would have been ignored in the previous version. Idea is to keep it updated automatically.
OB: Keep it as a free for all for now, then move into following proper ontology rules.
AM: Yes, will be in flux for a little while.
OB: Perfect, we can just make PRs against it.

OBO Foundry - need to determine when would be the best time to do things. OBO has a single ontology where they keep all of the relations and that makes sense because most of the ontologies are class based. If you're going to start a new ontology you're going to use most if not all of the existing relations because those are typically similar in this domain.
OB: Does it make sense for one concept to be represented twice, one as a class and one as a relationship? There is the concept that is parent that could be defined, vs the relationship which is parent of.
AM: SNOMED takes the approach of the separate hierarchy called the metadata hierarchy of classes where they represent relations as concepts.
OB: https://bioportal.bioontology.org/ontologies/FHHO/
AM: Will check this out. From a purely functional point of view it won't make much of a difference.
OB: Might be a question for Melissa.
AM: Whole point is when to do it. Doing it upfront has disadvantages, harder to make changes. Would be great to share in Bioportal but should also think about sharing it with SNOMED. Would want this to be used in the clinical space as well.

Pedigree Tooling Repo

OB: Haven't managed to put things into a repo yet. Important to keep the messaging around Pedigree up to date. Have these pieces that are coming together so well, having this terminology will allow us to take family

health histories straight from an EHR, potentially multiple of them, and interpret, analyze and convert into coherent pedigrees for analysis. It is the bridge that is the difficult bit to do without a terminology and the tooling around it. Given the stakeholders we have and the data being generated, we have richer health data and that interpretation is necessary - that analysis and merging of reports. Communicating and better articulating that could be improved.
GW: GA4GH standards are purpose-built, but I've always hoped our work will be able to be applied to a bunch of different use cases and a number of stakeholders.
OB: Needing to be clear on who those stakeholders are and what those needs are. Make sure it's clear where those lines are drawn and who we are solving problems for. Articulated once as part of the product approval process, and are re-articulating it to ourselves as we develop and at plenaries. Where do we keep up to date, a summary of our self-image. Where to provide the overview that we keep up to date.
LS: Create a readme in the Pedigree Github Repo.

OB: Does anyone have experience with OWL parsing in Javascript?
MCK: I have experience in parsing with Python, but there wouldn't be a lot of difference.

Milestone 3: Implementations
OB: What timelines and deliverables are reasonable for interoperability.
BF: User of the pedigree format, do analysis for risk prediction, cosegregation, pedigree drawing. I want to convert our current format to GA4GH format and map our requirements. We only care about genetic relationships, not adoption. Age is complicated for cancer. We define proband as the first person tested positive. Proband may not be the consult depending on the gene of interest. In the future we want to include more information like environmental risk factors, polygenic risk scores, type of genetic testing.
OB: Do you end up with synch issues because the age of last follow up changes, but ages relative to each other might no longer be correct at a point in time? If the age field records the age at last follow up, and you see one of the people later, so their age goes up but everyone elses stays the same.
BF: Relative ages are not correct, yes.
Mapped to the minimum dataset, you have everything I need but a few things need more clarification. For disease, there is age of onset, but for cancer it's often age of diagnosis. Age of person, not sure whether age of ascertainment or age of last followup.
OB: The current status, after the initial alignment, we went and simplified the pedigree model quite a bit. There will be requirements that won't be supported in the initial model because they're more on the patient essence side. We removed observations from the current version of the conceptual model. Phenopacket does have the birth year and age of diagnosis, as does FHIR.
GW: Age of last follow up is an extension. We need to think in terms of extending the minimum core dataset. Don't want to people to think that's everything. Last follow up age is a nice example to add to the document.
BF: How to code for multiple twins? The genotype, need it as negative, het or homozygous. I don't know how it is coded in the coded format but can adapt my program to it.
GW: In one of the appendix sections we talk about things we'd like to add, and genotypic data is part of that. What I think we should do is look at the HL7 model and borrow from that.
BF: There is ancestry in the minimal set, but population is not quite the same thing as ancestry. Been using the cancer incidence in five continents (CI5). Better to separate the countries because of screening guidelines. Gnomad also has population codes.
GW: For multiple twins, it's connected to one individual person. Is it one individual part of a multiple birth.
BF: Using risk scores could minimize the data formatting requirements. Environmental risk scores could also be quite doable.
OB: I wonder whether there is a conversation to be had around risk scores being integrating into phenopackets. Probably could be covered in FHIR as well. Some of these things are straightforward to represent. In terms of the conceptual model, most parent protocols may already be able to represent much of this.

BF: Can do an extension in the future.
GW: Maybe in the next call, talk about how we take what Bing is doing and what we need to do with this pilot. Enough to show to GA4GH that we've matched the requirements of showing a pilot.
OB: Are there any upstreams systems you expect data to be coming from?
BF: Genetic counselling clinics in the cancer institute. BODICEA.
Will link the presentation in the notes.

2021-05-26: Full Clin/Pheno WS

Chair: Melissa Haendel, David Hansen
Attendees "Name (Project Affiliation)": Francis Jeanson, Ian Green, Gideon Giacomelli, Ashley Hobb, Susheel Varma (HDR UK), Hannah Calkins, Moni Munoz Torres (Monarch Initiative), David Hansen, Nicky Mulder, Stephanie Li, Peter Robinson, Orion Buske, Steve Laurie (CNAG-CRG), Ayhan Serkan ŞIK, Grant Wood, Francesca Frexia, Nebiyou Petros, Alejandro Metke, Melanie Courtot, Jules Jacobsen, Melissa Cook, Alex Felmeister, Robert Carroll, Mamana Mbiyavanga, Bing Feng, Mogomotsi Matshaba

	Actions Arising	Assigned To	Deadline
1.	Interested in joining the Cohort Representation Subgroup? Please let lindsay.smith@ga4gh.org know!	All	
2.	Please add comments to the Phenopackets Github issue on Treatment	All	ASAP (Phenopackets is nearing the end of the PRC process)
3.	Reach out to the EDI team (stephanie.li@ga4gh.org, ashley@dnastack.com, neerjah.skantharajah@ga4gh.org) if interested in volunteering for the Onboarding Pilot	All	

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping	David, Melissa, Lindsay, Moni	5 min
2.	EDI: Onboarding Pilot	Ashley Hobb	15 min
3.	Subgroup Updates: Phenopackets v2.0 - Latest Documentation - Product Review in Process - Open for Comment closes MAY 27	Peter, Jules	5 min

4.	Subgroup Updates: Pedigree Standard - Minimal Dataset Document - Draft model - FHIR IG - Relationship Terminology	Orion, Grant	5 min
5.	Subgroup Updates: Cohort Representation - Use Case/Requirements Collection - Please let us know if you'd like to join the mailing list for this group (cohort-representation@ga4gh.org) - Meeting Schedule TBD	Melissa, Lindsay, Melanie, Susheel	5 min
6.	Open Discussion	All	20 min
7.	A.O.B.	All	5 min

Minutes: (under construction)

EDI Onboarding Pilot

AH: Pilot project as part of the EDI group within GA4GH. The idea is that if we have diversity in our teams and structures that will be included in our standards. Ensure operations in GA4GH and the ways we work together consider EDI. Three main areas of focus: onboarding, participation levels, and equitable standards review. In the onboarding focus, we have a work stream guides pilot. Often there is a lot of overwhelming background and content. This can be intimidating because a lot of the GA4GH community has been working together for years. Bringing in WS guides to help people feel oriented. Think of WS meetings as dinner parties. You are not expected to give the whole background of GA4GH, provide documents, etc. -> this is the role of the WSM or GA4GH staff. Recruiting volunteers now! (MMT and MC raised hand).

Q/ Normally when we talk about diversity, historically discriminated minorities often don't come to groups like this. Is there something we can do to make GA4GH more attractive to people from disadvantaged populations?
A/ Really good point. Might be more focused related to outreach rather than onboarding. We decided to do onboarding first, because if we do the outreach to get the interest and don't have the right soft landing for interested contributors to feel welcome than that could be a wasted effort. In less resourced populations or countries there might be other barriers besides awareness and interest. Great question that we haven't tackled yet.

Phenopackets

JJ: Haven't done much recently, currently working through PRC comments. Need to respond to the spreadsheet in the next week or so.
Had a good meeting with the VICC group, one issue raised that we might need to make some changes about, regarding the chemo and hormone therapy treatment elements. Depending on what the actual substance is you're using, the actual treatment itself depends on what it's being used for. You could have a steroid that could be used as a chemo or hormone therapy or just a treatment depending on what the intention was. Modelling was based on ICGC-ARGO but may have to backtrack and make the treatments more generic. Will open an issue on the Github tracker for people to comment on - see here. Should be sorted before we finalize v2.0.

Pedigree

GW: Our efforts are focused on the second half of our activities creating the data model (first was the minimal dataset).
OB: Next steps are around distributional questions with the minimal dataset document and the pedigree relationship terminology - looking at best practices for dissemination. Will be starting a repo for pedigree related tooling that we'll talk about on the call tomorrow.
GW: Approaching the point where we need to demonstrate some pilots, get some volunteers from the GA4GH community.

Cohort Representation
LS: Welcome Susheel and Melanie, Cohort Subgroup Leads. Meeting timing still TBD.
SV: Having been talking about this project as part of HDR UK and the Beacon project. Not 100% sure where it actually fits yet. Landscape analysis for where this might fit in collaboration with Beacon and Phenopackets. Not intending to re-do anything that has already been done.
JJ: In Phenopackets, we have a very bare bones Cohort representation. Unclear how Phenopackets, Pedigree and Cohort will all fit together. Would be good to have some ideas on where and how this will all work. Had feedback from the PRC about the lack of modelling with the Cohort bit of phenopackets. If we don't have it in there, people won't have the ability to group Phenopackets by some arbitrary characterization.
SV: Collecting use cases right now, will connect back with Phenopackets and Cohort. See where we can extend with existing resources or whether we need to create something new. My use case is to extract COVID positive patients from multiple datasets. Discover and then extract. There will be other use cases.
JJ: Thinking more about extracting cohort information from some thing?
SV: First discovery - firing off a query into a network (ie. Beacon). Coming back with a set a patients and we can then extract a patient as a CSV file.
JJ: Having a cohort as a message where can go "this is the cohort that responds to your query" is still a valid thing to have.

GW: You're supporting ped in phenopackets. Could phenopackets v2.0 could include a link to another type of pedigree, which is the model we're working on. Don't have to recreate the whole schema, but have the ability to link out.
JJ: Yes, always been the intention. Implement the pedigree spec. Add to family message and have as a stand-alone package.
PR: A lot of these standards are being purpose built. Currently our pedigree representation in phenopackets is just a ped file. I'm not too familiar with the use cases the Pedigree standard will get - trying to figure out how to use it.
LS: GA4GH is working on making the material across our work streams more navigable. Clin/Pheno has a landing page here, but it's relatively bare bones until we can get a well-developed page up on ga4gh.org.

PR: Conceptual alignment of standards and implementation. GA4GH groups are using different flavours of things that are not 100% compatible with each other. What we did is implement a version of VR in Phenopackets protobuf. Is that the type of thing we want to do, or have one software implementation to do everything? GA4GH hasn't committed to one, but we as a group should develop a strategy for that.

SV: In Cloud, there was a strong opinion to have loose coupling, layering standards together.
PR: Con - Have to have more complicated software to use GA4GH

GA4GH might want to look at FHIR as a strategy to solve that problem, could turn into resources and profiles.
- However, this may not work well for the research community
- Research community cannot keep itself separate from the clinical community, need to find ways to work well together and share the same data and not have have data silos, bridge those gaps

- From someone sitting on both the clinical and research side, GA4GH needs to look at interoperability between standards and data. Approach of layering could be the best approach, like Smart on FHIR we could have GA4GH on FHIR, have a suite of standards sitting on FHIR for clinical or sitting on top of OMOP for research applications.
- FHIR is a framework you can use to build your own solutions. Can use it to build research solutions.
- Need to be mindful of community doing genomics that do not live in the EHR world, FHIR has a large barrier to entry. GA4GH cannot impose it on researchers. Need to have IGs that allow you to get knowledge out of FHIR.
- Any schema we build, no one is using, the tooling support to do the IG is needed.
- We as a community can work towards putting things together but is our responsibility to train and enable the people who are experts on the FHIR side. Partnership.

EDI Comments
DH: This onboarding is really just the first step in what I think GA4GH is doing in the EDI space. I thought the presentation at GA4GH Connect was fabulous. Would be great for everyone to drop Lindsay a note about getting involved. Need to be a welcoming organization. Build credibility to get people wanting to join.

[From Chat]
GG: CDISC Examples
- https://dataportal.answerals.org/search
- https://www.gaaindata.org/data3/explore/login.jsp;jsessionid=06D9AE18AE34A192E89A9F8A6B7D6B51

MMT: I think that, in the context of genomics research in GA4GH, there is no point in designing or building a schema only with FHIR in mind. And I like FHIR! I just don't think that it currently is capable of doing everything we need.
FHIR cannot do:
- Deep phenotyping (soon FHIR-Phenopackets!)
- Clinical research variables/CDEs (needed by all the commonses)
- Pedigree representation
- Genomic and other omics data
- Geospatial/social determinants
- Survey instruments
- etc.

FHIR needs content models to put in the containers. There are many clinically relevant data types that folks are working on aligning or implementing within FHIR (the things listed above)
We have other standards for most of these things, but they are used outside the EHR and/or in combination with EHR data. The community is working on this, but it is not yet ready.

2021-05-13: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Alejandro Metke (AGHA/CSIRO), Meen Chul Kim (CHOP), Zornitza Stark (AGHA), Moni Munoz-Torres (Monarch), Hannah Calkins, Soichi Ogishima (GEM-J), Jules Jacobsen (Monarch/QMUL), David Hansen (AGHA/CSIRO)

Global Alliance
for Genomics & Health

	Actions Arising	Assigned To	Deadline
1.	Create a new Github organization for Pedigree	Lindsay, Moni	DONE - see https://github.com/GA4GH-Pedigree-Standard
2.	Explore readthedocs as a publication option for the Minimal Dataset Doc	Grant	
3.	Set up Java project for the pedigree ontology	Alejandro	
4.	Add Soichi's comments about the ontology to the ontology repo	Orion, Soichi	
5.	Start a new tooling repo	Orion	

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping	Grant, Orion, Lindsay	5 min
2.	Getting the terminology into a proper valueset/ontology in the new repo	Orion, Moni	15 min
3.	Planning implementations for Milestone 3 based on the FHIR IG and Phenopacket branch	Orion, All	25 min
4.	Publishing the FHH Minimal Dataset	Grant	10 min
5.	A.O.B.		5 min

Minutes:

GW: The minimal dataset document is pretty much ready to go for a version 1.0. Last item is these two areas where we can add information about the model or about the implementation guide.
OB: In terms of it being living vs being citable by others, where would it live? Do an archive.org type thing?
JJ: Have had good success with phenopacket and vrs via readthedocs and having versioning. Can also download them via pdf. Can also see the change edits through commits.
GW: Since the audience for the document will go beyond the GA4GH world, is Github still a good spot?
MCK: This version 2.0 was written with link ml?
JJ: No, this is protobuf.
OB: Is this more like documentation or more like a publication? Living makes it more like a document.
GW: Great plan.

OB: Spend some time on getting the terminology into a proper valueset. Once it's hosted there it will make it easier to pull comments. This morning, we received some strong feedback from Soichi on specifics of the structure and ways that could be improved.
MMT: I can look into whether has done anything.

AM: I have a current version of an ontology, not complete but based on your spreadsheet. I haven't gone through all of the items, maybe a quarter through. I can quickly show the progress.
It's an OWL ontology that I'm editing by hand given the size it not too big. Only a few classes. The mitochondrial donor - it was listed as a biological parent, but for the department of health in the UK, it's not considered a parent (legal or biological), so maybe classify under relative.
OB: There are some different approaches we might want to take with this. For which purposes do we want biological attributes that have an implication on the health of the fetus or child.
AM: When you look at the consequences, need to decide what you're modelling. In terms of how to put it into a repo, propose creating a Java project where you would import the OWL api libraries and dependencies for a reasoner, then add the OWL ontology as one of the resources. Display unit test cases that you can run in a normal CI environment. Every time a new test case is developed you can run all of the tests and or if you make changes you can make sure you're not breaking anything.
JJ: The guys from Monarch suggested the ontology development kit (https://github.com/INCATools/ontology-development-kit), specifically for setting up something in Github. I don't know the advantages of using this over your own thing, but they are happy to walk us through it.
AM: I took a look at it, the toolkit is a wrapper for other languages. It has lots of tooling that allows you transform spreadsheets into content. My concern is that it's geared to ontologies that are heavy on classes, might be a lot of infrastructure for an ontology that will mostly be updated manually. Tests need to interact directly with the reasoner, whereas the toolkit uses SPARQL.
OB: That would be amazing Alejandro if you're up for setting that up. Soichi - I would love to add your comments as an issue in the Github.

OB: Pick up the conversation about milestone 3, the planning for initial implementations of the pedigree model, and using that as proof of concept. I think there were a number of people who joined the call last week who were interested in using it in some shape or form. In particular, systems for collecting and recording pedigree information that could use this standard as a pedigree export.
A number of different tools are on the upstream side, and then tools that help us complete different workflows. Exomiser seems like a good opportunity to test out consumption of the pedigree model.
JJ: It'll be after Phenopackets approval and once release 13 is out. It would be nice to have a separate module like a converter that could do this. That would make it really simple to put into Exomiser. Could make a test build to convert into ped format.
Something that could convert from this intro phenopackets could also work great.
OB: Seems like it would represent a flow on the phenopacket side. If we wanted to represent a flow on the FHIR side that demonstrates an implementation, what might that look like? One that comes to mind is looking at the Redcap plugin as a loop. It uses FHIR but isn't being served, doesn't really matter what that format is.
AM: We could use it, if we start using the pedigree plugin for redcap, we have a repo that collects the data. Part of the work is tooling to translate from redcap into FHIR. Valid as a way of showing flow but is really only passing through.
OB: When it gets converted into the clinical FHIR database, which profile or implementation does that go through.
AM: It's probably already similar to what we have.

SO: How about plink? If we can adapt to plink it will spread rapidly if it becomes available.
OB: Would you suggest we try to get it natively supported or use a tool?
SO: Set up a converter. Or, if we can include the developer of plink. Many researchers use plink.
OB: Plink uses a ped style format for importing. It sounds like we're coming up with the utility for a basic tool script that can accept data in a FHIR or phenopacket format and output it. We do have some code from the open pedigree. If people are ok with a node-based, javascript-based tool we could use that.
JJ: I think any sort of tool that will convert from one format to another will be really useful.
OB: I can start us off with a bit of a project. Where on Github?

MMT: Let's make a new organization for the Pedigree work.

OB: One of the action items is to create a repo for tooling around pedigree conversion. This will solidify some of the use cases and tooling. I was going to propose populating that repo with a basic node project with some code around file format conversions and things like that. Broadly, if anyone has some energy to help with that coding, especially on the FHIR side.
MCK: I'm in.
OB: Could be any language.
MCK: Python and Javascript. Please commit your base code so I can build off of it.
JJ: I can volunteer some kind of pedigree to phenopackets conversion but I'm not a Javacript person.
SO: I am also interested in helping.

2021-05-13: Computable Cohort Representation

Minutes: The minutes for the kick-off meeting of Computable Cohort Representation can be found here.

2021-04-29: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Katherine Johnston (H3Africa), Alejandro Metke (CSIRO/AGHA), Hannah Calkins (CHOP), Melissa Cook (NCI), Bingjian Feng (UUtah), Steve Laurie (Solve-RD), Yoshida T (NCC), Kathryn Van Diemen (TrakGene), David Hansen (CSIRO/AGHA)

	Actions Arising	Assigned To	Deadline
1.	Add the family history terminology to the new repo	LS/OB	
2.	Interested and willing to test out the pedigree standard? Please let us know!	ALL	
3.	Aim for Plenary 2021 for the interoperability pilot		

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - Review of action items from last meeting - Start a Google Spreadsheet to collect use cases for the family relationship terminology (OB/LS) → see https://github.com/ga4gh-cp/family_histor	Grant, Orion, Lindsay	5 min

	y_terminology/discussions - Prepare and share a few ontology prefix ideas to vote on, explore setting up a formal ontology (MH)		
2.	Pedigree FHIR IG - How can we help?	Alejandro, Meen, All	5 min
3.	Milestone 3: Interoperability Pilot - Which implementations will pilot it? - What is the data flow we can demonstrate?	All	45 min
4.	A.O.B.	All	5 min

Minutes:

OB: Discussions has the ability to vote ideas up, useful feature for aggregating or bringing up the most important user stories. Create a category of user stories to start with, have a template. We'll configure this up.

FHIR IG Updates
AM: Had resolved most issues, except the one to add more examples. If anyone has any additional examples please submit. We also set up continuous integration so everything should be automated.
OB: When we move the terminology out of the spreadsheet, the process of updating that FHIR IG, that should be straightforward?
AM: Yes, the current codesystem was created externally to the IG. There are ways of defining simple codesystems inside the IG if you have a list of codes. In this case it's like a reference to an external codesystem.
OB: So then I guess what we would be able to do is use the computational use cases that people provide in this new Github discussion space to look at what computational capabilities we want to then do on that FHIR model.
AM: Yes. I think the IG is independent to that implementation. The IG is just checking that the FHIR format that comes in complies with the restrictions that are put in the IG. We'd need a demo to run the computation, depending on what they are.

GW: At some point, I think we will want to engage the HL7 working group and have them review. What do we want to do when we engage with this group? Just show what we've done and present as a model, or do we want it to be submitted through HL7 to become an approved FHIR profile.
AM: Would be great to submit, but also depends on timing of when to do it. If you could do it in two steps. If we engage too early we might be subject to much more process, I think if we could do an initial release that is GA4GH only that matches the protobuf or JSON format. Then, we do the more complicated step of HL7 approval.
GW: Agree. Don't want HL7 to slow down or disrupt anything we're doing.

Third Milestone
OB: Taking stock of where we're at in the overall implementation plan. For milestone 3 we're talking about an interoperability pilot. When we were laying out the plan in the original proposal, now would be a good time to share this with everyone again. Is this still accurate, reasonable?
First, was an implementation of a test suite, make sure that a certain serialization of data conforms to this specification. Have a base level of verification that is provided out of the box, I think we're well on our way

here. Could have a tool that verifies conformance, and the internal consistency of the data that is represented in it. Are there any goals that come to mind for what that implementation might look like? Or what the conformance would include?
GW: A conformance or testing tool is critical to the future adoption of what we're doing. Ensures people are generating compliant messages. Wonder if this is a tool that already exists we could apply, or is this a tool we'd have to develop ourselves?
OB: There are use cases that are like a library that will be used computationally for either sending or receiving data to verify messages. There's also a lighter touch, where as part of the standard where we have a service where you drop a file and it runs a check. Could imagine using existing libraries and tools for existing implementation guides.
GW: Doing this is a web service is the best way to go.
AM: There is also support in FHIR. You might also want to, as part of the use cases, check nonsensical relationships that you can't check with syntactic checks that IGs do.
OB: I think Phenopackets has a similar mechanisms. FHIR goes deeper into verifying the actual terms, and then there may be some shared features there.

OB: The main piece of this milestone is around needing to demonstrate implementations of the spec that ideally pass compliance tests across this chain of care and across stakeholder groups. Wanted to get feedback from volunteers, or making proposals. Who are we going to work with on this. Originally, we looked at EHR, patient-facing vendors, clinical-facing vendors, clinical decision support systems and research systems (ie. Kids First). There are various ways we can imagine flows of information between some or all of these groups. It's a question of which ones we want to expand to, clinical system to clinical system is probably a less common use case.
AM: Happy to adapt the plug-in to become a reference, but would be good to show interop with other systems. Most of the data we deal with was drawn on paper before. I know TrakGene is used in some places, but not interoperating with them.
KJ: I need to check with Guida on how he generates his data. I know he is probably our best bet for someone who is collecting pedigree data. Also has an interest in collecting the rest of his data in redcap. There are one or two other groups that collect pedigree data but it might be paper diagrams.

HC: I am at CHOP, not on KF team but I work with them on a number of things. Use to be an EPIC admin.
OB: We can wait until Meen is able to respond with how KF might respond.
KVD: We have a potentially few use cases in place. TrakGene sites that are trying to share research data for projects.
BF: A lot of the data we have now is minimum.
OB: In an ideal world, I see there being a few opportunities for research to clinical system transfer, maybe some clinical decision support tools. Would it be worth reaching out CancerIQ?
GW: I think it would be to reach out to them again.
OB: Figuring out the best way in PhenoTips to demonstrate some of this. Thinking about the GA4GH Connect efforts where different groups demonstrated different parts of the process. Jules feedback on phenopackest would be really useful, use in exomiser would be a really great bridge.
GW: Have some sort of a sandbox we can all play in. Our family health is the tool at intermountain, internal tool not available to anyone outside of Intermountain.
OB: In terms of sandbox, can you elaborate on how that would work?
GW: Place where we can take the app and add the capability to the app to generate pedigrees and send off t someone doing the same thing on their end.
OB: Is this technical infrastructure? Private network on AWS? Is this just a document where we make URLs publicly available?
GW: All of those are possible.

OB: I was curious whether there was engagement on your side in participating in this milestone for implementing draft spec and showing interop with another system?
SL: Haven't had a chance to look through the documentation. If it was something clear and simple and wanted to use a test case from the research environment. Might be best to work through Jules. We can generate phenopackets and send them out but not sharing them right now because no one has requested them. Would be an easy thing for us to do.
OB: We have a pretty healthy number of people willing to do some draft implementations. Important to show it's less about the nodes and more about the edges. The way in which we think about demonstrating this. There is almost a milestone 3.3 where there is a free and open source implementation that could be used to collect data in this format in the first place. There are entry arrows.

What sort of timelines do we feel is reasonable?
Aim for 2021 Plenary to complete the interoperability pilot.

2021-04-22: Phenopackets Subgroup

Minutes: The Phenopackets minutes and agenda can be found in this document.

2021-04-15: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Meen Chul Kim(CHOP), Alejandro Metke (CSIRO/AGHA), Kathryn Van Diemen (TrakGene), Melissa Cook (NCI), Jules Jacobsen (QMUL/Monarch), Moni Munoz-Torres (Monarch), Melissa Haendel (Monarch), Bing Feng (UUtah), Hannah Calkins

	Actions Arising	Assigned To	Deadline
1.	Update meeting time to 1pm PT/4pm ET/9pm BST/10pm CEST/6am (Fri) AEST, etc.	Lindsay	Done
2.	Start a Google Spreadsheet to start collecting use cases for the family relationship ontology	Orion / Lindsay	
3.	Reach out to Matt Brush about setting up the ontology	Melissa	
4.	Prepare and share a few prefix ideas for the subgroup vote on	Melissa	
5.	Change affected to optional (remove 'not recommended')	Orion	
6.	Comment on the FamilyMemberHistory vs	ALL	ASAP

Global Alliance
for Genomics & Health

	Observation Github Issue		

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - New member introductions - Determine a new meeting time?	Grant, Orion, Lindsay	5 min
2.	Pedigree Connectathon Debrief	Orion	15 min
3.	Updates on FHIR IG alignment - Draft FHIR IG - Issue #2	Alejandro, Meen	10 min
4.	Management of the relationship terminology - Next Steps	Orion	25 min
5.	A.O.B.	All	5 min

Minutes:

[Lindsay to update the Subgroup calendar to start the meeting 1 hour later]

Connectathon Debrief
OB: Great opportunity to connect with more people and give some exposure to the work we're doing. Received some additional example Pedigrees from Guida. Valuable to go through the process of presenting and get additional feedback. Broadly, I think the questions that people had were mostly addressed. Part of the next steps fit well into the action items for today.

Terminology
OB: How to use and build upon a existing terminology, how to version and maintain.
MH: I think this a really important and great work. The content is richer than just pedigree, it's family relationships. More broadly useful than just pedigree, can get into public health and epidemiology contexts. There isn't a good ontology for family relationships for computational use. There is some overlap with NCIt, can coordinate with them. If we want to create into a full ontology, we can come up with a integration strategy with credit to everyone here. Can also be standalone as well.
The relationships between humans would sit beside things for birds and ants. Could use those properties in the creation of a familial ontology itself. Could create a new Github repo that we populate with the right bits. We can manage content in the Google Sheets since it's not a large file. Could also do Github tickets. Not too complex in the classes and relationships.
OB: Alejandro, you expressed interest in the bridging and reasoning. Does that align with what you were imagining?
AM: Yes, definitely. I think it's important to have it standalone and integrating. Also important to work in parallel what use cases we want out of the computational properties that defines the best modelling. I have some idea but would be good to formalize. For example, would like to infer additional relationships based on a limited set that you capture. Validating that a set of relationships makes sense and don't contradict each other.

OB: That sounds great. Is having those relationships being defined more semantically something that could bridge between FHH FHIR world where it's arbitrary between pairs of individuals. Multiple siloed networks of relationships.
AM: You're right, that might be a good way to bridge and get into the format we're proposing. You also mentioned that there are different styles of kinship, which could be another use case.
OB: Would it make sense to start a Google Spreadsheet and invite people to fill it out?
AM: That should work.
OB: With modelling, do you see all of these different relationships as classes or are they on the property side.
MH: I think most are relationships, very property rich in that sense. I like to put it in an ontology file format so I can run the reasoner and see how things go. Doing this doesn't preclude any final decisions, just getting it in a computable structure we can poke at. Start with Alejandro's idea of collecting use cases. That kind of very specific query, could also be in addition to more general use cases. Can build those simultaneously to build out the file. The queries are pretty easy.
AM: Totally agree, take use cases, and turn into test cases with the ontologies.
OB: What is the technical requirement that we should get to that would make it easy to keep phenopackets and FHIR up to date with the terminology.
MH: They would just call upon a released version. We can make different flavours for different uses, but would both refer to the same released version.
AM: We have software that takes an OWL file and produces a FHIR codesystem. Would need some work to ensure the FHIR codes are classes.
MH: That's a really cool enhancement, we have the same problems. FHIR thinks of things as codesets but they're not the same thing semantically.
AM: SNOMED just replicates as concepts, but I don't like that approach.
OB: In terms of next steps, who are the right people to connect with.
MH: Task Matt Brush - both GA4GH and ontology engineering. Help get it set up and then we can contribute.
JJ: Nico (?), one of the authors of the paper you linked to.
MH: Not an ODK problem, only 30 or so rows. Modelling to think about. I think Matt is good if we can convince him to do it.
OB: Anything else you recommend to do right now?
JJ: For directionality, have a suffix. So instead of parental, say parent of. Treat like properties. Need a different prefix. The naming of it.
MH: We need a name and prefix everyone is happy with. Shouldn't use REL, it's not used elsewhere but it's more family relationships. Naming is less critical than identifier prefix.

FHIR IG
AM: Not a lot of discrepancies between Meen and I's implementations. Decision to base relationship on observation or family member history. Need more people to chime in. Grant - was going to suggest you comment given your experience in the space. Both work, it's just trying to figure out which make more sense. Both have pros and cons and represent the same thing at the end.
MCK: Alejandro agree to agree.
GW: If I had to put on my HL7 hat, family member history has been part of the normative FHIR standard for several years. Was part of FHIR 2.0. Would choose family member history method.
AM: I think we only have the one outstanding issue.
OB: Meen, you have taken more an observation based approach, just Grant's response clarify things?
MCK: THe use of family member history makes sense in the context of clinical data modelling. It seems like the standpoint I am taking is to view the relationship more like a generic family relationship if that makes sense.
GW: From your point of view, if we did a model using observation, what benefits would it have over family member history.
MCK: Family member history has a property called patient whereas by default it misses a subject attribute. Without using an extension, this profile cannot associate patient with other people.

AM: The original design of family member history did include an extension for that purpose.
MCK: It seems like the profile misses the linkages between people. For rendering family history, my standpoint is a relationship between people.. Family member history is valid under a specific condition
OB: To summarize, family member history conceptually is the data you receive from a patient about their FHH, and you want to decouple that from the derived concept of a relationship between two people.
GW: What if we think about doing the family member history IG first? And then, after doing that, could go back and look at what type of model you're thinking about.
MCK: It's just a matter of the FHIR profile used. Have no objection.
GW: Would like to pursue your ideas, but suggest we coalesce around the family member history one first.

OB: In terms of next steps, does that mean there are any adjustments to the model or just the documentation around it between now and the next meeting?
AM: Already in the IG, will clarify in the issue.
MCK: Add more complex family pedigrees. I suggested to cover all of the examples Lindsay provided.
AM: Those will be additions to repo. I haven't pushed the changes to the repo but have organized the FSH files.

OB: Jules posted an issue around the affected status and the way we're describing the clinical condition that the pedigree is represented around.
JJ: In individual and pedigree, if we had a proband who was the subject of an investigation and they were marked as affected with a condition that you don't know. In order to investigate A and B condition, the pedigree itself is the relationship between these people… adding bits in the wrong place. Extra bit being the proband. What we need is a context.
OB: Interesting conceptual question, is a pedigree something that is separated from the context in which the pedigree is collected. Or, is that an intrinsic part of the pedigree. It informs what data is collected early on.
JJ: Date I think does belong with the pedigree. Things like the consultand. We didn't reach a conclusion about proband/consultand.
OB: Proband is the index case for a particular pedigree - this has a clear use for sections of genomic medicine. On the cancer side the concept of proband is not directly useful, and want a separate proband for the variant. The second model makes sense for a richer condition. Having a proband without a condition is a thing that happens if you don't specify the condition. In general, pairing a condition or concept with a proband.
OB: Deferred consultand, move to deeper provenance. And should couple proband to the reason code.
JJ: Proband is property on individual.
OB: Could be that, or the pedigree has probands for different conditions.
JJ: Relationships between individuals don't change much (biologically).
OB: Is the pedigree the collection of those biological relationships.
JJ: The context is why you collecting, who you're looking at and why.
OB: I always think of pedigree as having a proband on it.
JJ: The original point was about affected vs recommended.
OB: Will change the affected status to optional. Bing, would like to loop you in on this to connect with the FPTP.
BF: I think this format is for the minimal use case. So for my analysis would like to extend this to include the type of cancer and the variants detected. Then use my program to assign proband itself, won't rely on this.

GW: I think we keep in our mind how we use proband will be different depending on context. Proband is whoever we do the risk assessment for. Can re-orientate that pedigree to look at someone else.

Global Alliance for Genomics & Health

2021-04-08: Phenopackets Subgroup

Minutes: The Phenopackets minutes and agenda can be found in this document.

2021-04-01: Pedigree Connectathon

Chair: Orion Buske, Grant Wood
Attendees "Name (Project Affiliation)": Javier Diez (Fujitsu), Toyofumi Fujiwara (DBCLS), Charles Keenan (UToronto), Meen Chul Kim (CHOP), Katherine Johnston (H3Africa), Jules Jacobsen (QMUL/Monarch), Melissa Cook (NCI), Hannah Calkins (CHOP), Bing Feng (UUtah), Alejandro Metke (CSIRO/AGHA), Mamana Mbiyavanga (H3Africa), Asif Chinwalla (CHOP), Nishita Thota (GEL), Guida Landoure(H3Africa)

	Actions Arising	Assigned To	Deadline
1.	Starting merging the two FHIR IGs together - create issues on Github	Meen, Alejandro	
2.	Reach out to Grant, Orion or Lindsay if you have example data you are working with that you are comfortable sharing or example use cases for that data you are willing to share	ALL	
3.	Brainstorm names for the relationship ontology	ALL	

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - Welcome to all first time attendees! - The GA4GH Pedigree Subgroup meets bi-weekly on Thursdays @ 12pm PT/ 3pm ET/ 8pm BST/ 9pm CET/ 6am (Fri) AEDT, etc.	Lindsay Smith	3 min
2.	Brief overview of the GA4GH Pedigree Standard - Minimal Dataset Document - Draft Conceptual Model - Anonymized Pedigrees (Thank you Zornitza and Guida!)	Orion Buske	5 min
3.	Interoperability with HL7 FHIR - Open Discussion and Q&A	All	25 min

	- Draft FHIR IG		
4.	Interoperability with Phenopackets/Protobuf - Open Discussion and Q&A - Draft Phenopackets Implementation	All	25 min
5.	A.O.B.	All	2 min

Minutes:

Pedigree Overview
OB: Set of individuals who are part of that family with demographic information, and a set of relationships that pair those individuals together based on a taxonomy. Beyond that there is a bit of structuring. This is the starting point, and we intend to build on this for additional use cases for GA4GH

FHIR HL7
AM: I haven't made the changes with respect to last changes in the model. You'll find the mapping decisions, and the draft model. Can use Github to add issues. I developed examples using the FHIR shorthand syntax to create instances. Would be great if people want to implement the pedigree examples using the IG.
MCK: For individual I used patient, adding three extensions. Also used pedigree using component. Using observations to model relationships is the biggest difference between mine and Alejandro's proposal. Tried to model using directive relationships. Alejandro used family member history profile to model family relationships, whereas I used observation. I recognized the infection status as an attribute on an individual.
OB: In terms of anyone who is joining the call, are there people who are interested in the FHIR implementations that tried it out?

OB: It seems like one of the main challenges we're facing is the codifying of relationships and pedigree data that might otherwise be under described. Is anyone able to represent a tool where you would want to receive pedigree data in a structured format.
BF: I'm writing a co-segregation tool that would like to use this format.
OB: Is there any assistance or guidance we can provide?
BF: Run my tool in C++. Would like a tool that can assist with C++ programming.
JJ: I'd be interested in pedigree for Exomiser. And for phenopackets we do have c++ bindings. This doesn't have the pedigree stuff yet, this is just the v1 phenopackets schema. If you want to check out the stuff we're talking about.

HC: We are starting to be in an ingestion role. Have some thoughts about proband and familial relationships, which can sometimes be conflated.
OB: Useful conversation to have right now.
HC: One of the things we are needing to do is we're intaking data form different projects and need to pool together and look at re-use across. At times we want to use the pedigree info in a broader sense. Proband is really talking about a individuals relationship to a set of files, and not necessarily describing the specifics. Proband and father - don't know if proband is male or female. Helpful if we could explicitly state the proband.
OB: Can cause problems with merging pedigrees.
HB: Useful to understand that's the proband.
OB: In the conceptual model, the proband is a pointer to a particular person within a pedigree, and is independent of the inner family relationships.
MCK: We partially encounter these issues. We often get information from a trio, and more than 95% of cases the child is the proband.

OB: I think we're able to support that use case right now.
AC: Integration into a larger set of pedigree. We get many individual projects that are individual cohorts, and often want to roll up into master cohort. It might be good to discuss version updates, track provenance.
OB: There is a piece of that challenge we're trying to tackle right now, the representation of that data in whatever collection system you're using and backend system. The overall tool ecosystem and database system for merging that information. The model that we've been building would help with aspects of that, in particular the de-coupling of individuals from relationships.
BF: Question about merging. What happens to the proband when you merge.
OB: Would depend on use case - more a business model question rather than a data model question.
JJ: Concept of a proband is completely dependent on what you want to do with the pedigree.
BF: I need to adjust for ascertainment. If there are two pedigrees, I need to control for both. Would need to know both probands and have multiple probands in the pedigree. Don't have to be affected.
GW: In my clinical world, proband is a temporary of the person in the pedigree who is the focus of the risk assessment.
JJ: Exactly how we use it in Exomiser as well.
OB: How do you distinguish that from consultand.
GW: Consultand in my experience we don't use that much. Might be some overlap between those two terms.
OB: What we see clinically is that one patient is the proband, wouldn't move the proband.
JJ: Proband is generally the first person we've seen with the condition.
OB: For Bing's use case, the fundamental difference is that it's per varint or per condition. The change that needs to happen is as those conditions are coded in directly, have a boolean property for this person is the first person tested.
KJ: Proband is kind of label that you apply, and comes from whoever is collecting the pedigree, and is exchangeable depending on who is using the pedigree and for what purpose. Consultand is more permanent than proband, and should be moveable. How important is that proband ID to the pedigree data. Seems like it's more about the people looking at the pedigree than the data itself.
OB: Two cases: case of downstream analysis and in the clinical genetics setting, the proband is if the family is represented by one person in that family, the proband.
JJ: Are we trying to put too many bits of information in. The pedigree itself doesn't change, it's the label that changes. Depends on the context in which the pedigree is being used.
OB: Sounds like to the extent there might be an abstraction between the family and a pedigree, it's an use case specific bundle of people and relationships.

<u>Phenopackets/Protobuf</u>
JJ: In phenopackets, pedigree is phenopackets, we take the proband and consultand out of that. The general layout: we have a phenopacket that is an individual. We've also got a family, equivalent to interpretation.request ID that we have in Genomics England. The proband is the one you're interested in. The pedigree is a ped file essentially with a little bit more structure. Can build a map of individuals via their maternal and paternal lines. Instead of having a direct type of relationship in the individual itself.
OB: If someone were wanting to get started with this, they would go to the branch, build it, and then import in whatever their preferred language is and use that protobuf interface.
JJ: Precisely.
OB: Did anyone start experimenting with this, or look at it and have any questions?
JJ: I had one look at the latest pedigrees where there were a lot of consanguineous relationships. There's no flag for that in the pedigree itself.
OB: For those use cases, the consanguinity doesn't need to be explicit because it's logically implied. Did incorporate a change for a subtype of partner for consangeous partner so you can explicitly label it. If anyone has example data they are working with that they are comfortable sharing or example use cases for that data that they are willing to share.

Global Alliance
for Genomics & Health

2021-03-25: Phenopackets Subgroup

Minutes: The Phenopackets minutes and agenda can be found in this document.

2021-03-24: Clin/Pheno Full

Chair: Melissa Haendel, David Hansen
Attendees "Name (Driver Project Affiliation)": Grant Wood, Moni Munoz-Torres (Monarch Initiative), Algy Taylor, Ian Green, Arthur Hermann (Kaiser Permanente), Heikki Lehväslaiho (ELIXIR-FI), Peter Robinson (Monarch Initiative), Ayhan Serkan ŞIK, Zornitza Stark, Jules Jacobsen (Monarch Initiative), Núria Queralt Rosinach (EJP RD), Kathryn Van Diemen, Francesca Frexia (CRS4/BBMRI/EJP-RD), Bron Kisler (ISO Genomics), Katherine Johnston

	Actions Arising	Assigned To	Deadline
1.	Reminder to add any comments to the Phenopackets v2.0 schema NOW - the deadline for GA4GH SC (April 24) is approaching!	ALL	April 9
2.			
3.			

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - New Clin/Pheno Full Schedule: Wednesdays, every other month @ 12pm PST/3pm EST/8pm GMT/9pm CET/5am (next day) AEDT, etc. (19:00 UTC) - *effective post-daylight savings changes* - Genomics in Health Implementation Forum (GHIF) - upcoming webinars/workshops, incl. Pedigree Connectathon (see item below) - Interest in Cohort Representation? Please let Melissa or Moni know so we can coordinate a	Melissa, David, Moni, Lindsay	5 min

	smaller planning group.
	- Any new members? Welcome! 🎉
	  - **Arthur Hermann with Kaiser Permanente.** Arthur sits on the HL7/FHIR Genomics Working Group and is very interested in the work which this workstream has been doing.
	  - **Francis Jeanson** is aiming to join as much as possible the ClinPheno workstream to contribute where possible. Francis is also a long time member of the Data User and Research Identity (DURI) workstream and looks forward to sharing insights across! Francis is an informatics and analytics professional working with the **Ontario Brain Institute**, based in Toronto Canada.		

2.	Pedigree Standard Updates - Minimal Dataset Document - Draft Conceptual Model - Draft FHIR IG - Draft Phenopackets Implementation - Connectathon Registration - April 1st	Lindsay, Orion, Grant	5 min
3.	Phenopackets Updates - Documentation - Outstanding Github Issues (see 'GA4GH Discussion' label) - V2 is aiming to submit to GA4GH SC by April 24th - if you have any comments for the linked Github issues above, please do so now!	Jules, Peter	5 min
4.	Using Phenopackets in Gen-o: Assisting clinical multidisciplinary teams in assessing the results of a genomic interpretation - Slides	Algy Taylor (Manchester University NHS Foundation Trust)	30 min
5.	A.O.B.	ALL	5 min

Minutes:

Pedigree Updates

OB: Conceptual model, working on implementation in FHIR and Phenopackerc/protobuf. Part of the approach is to make sure it works in the context to other standards that contribute rich individual level information. Connectathon on April 1st - work on implementing in practice. Invite anyone interested in implementing the standard (either protobuf, FHIR or other) to join.

Phenopackets
JJ: Working on adding a quantitative traits elements, issues #261. Added a draft pedigree model in, as well as VRS. Now is the time to add any comments if you have any, as we're aiming to wrap this up by end of April for GA4GH SC.

Using Phenopackets in Gen-o
AT: National health service in the UK has been pushing the idea of mainstreaming genomic medicine, so ordering a genomic test is the same as ordering a blood test. The clinical process is a little different than what you might see in the US. The lab will come back with the results, usually a brief discussion between the lab and clinical team if it's an open-shut case, if not, it becomes a multidisciplinary team discussion. Once you get to some diagnosis, they are returned to secondary care with a clinical interpretation of those results and a care plan. The process is quite work heavy, and will have problems if there is an increase in patients. Realization that a lot of our data was not stored in a rigorous way. Previous, MDTs was exclusively presented on PowerPoint. Came at it from two angles - what did we need to capture from the clinician and what do they need to make a decision, and a focus on whole exome seq. We begun Gen-o development in 2017 and its routine use in 2018. Working with IBM on a redesign, with data largely stored using phenopackets.
The phenopacket model fit in perfectly with DOB, age, sex -> all map perfectly. We have added a few fields like name and address, NHS number. The only bit missing is taxonomy but we deal with humans!
We store inheritance pattern, some details of affected family members, along with the ability to upload pedigree drawings. The family record again maps really well. Additional family members we do intend to add in the near future. The one real diverge has been pedigrees - clinicians draw out the pedigree by hand as they're speaking to them, very little appetite to electronically encode it after it's been drawn.
For phenotypes, we've split the collection into two sections: observed by clinicians and through investigations (ie. images). Encoded these with HPO terms. Although as a database schema it doesn't really match the standard, it can easily be converted back into it. We've not yet gone into the phenotypic modifiers with severity and onset. Only other things we slightly diverged from was the negated boolean option, we have a enum type present that allows present, absent or unknown but clinician feels like it's important information.
Can enter the variant and it provides an pathogenicity score. In the database schema, we can can follow the interpretation model to the letter. We've gone straight to using HGVS, and also we include the supporting evidence that's gone into classifying the variants. Lastly, in the interpretation there is a resolution status, with this piece of software dealing with workflow. Because the report is sent to secondary care, it's useful for the geneticist to give a plain text summary of what the phenotypes mean, what's the interpretation, so there is a spot for those summaries. Cohort, family, gene, disease are on the roadmap to be implemented by the summer. Out of scope is biosample and htsfile.

MH: Really fantastic, impressed with the utility of what you have done. At a high-level, what the big win and what the challenge? It sounds like taking micro schema bits was straightforward, but putting them all together was harder.
AT: Big win was having a data schema and there were some things from prototype versions we didn't have a solution for - most things were solved by the schema. There was some mental gymnastics with trying to fit some things in. Generally only positive things for the schema, been a really big help.
HL: This software is all internal to your institute? The only thing that comes out of it is the summary, which is free text, or are you passing structured information that would be recorded in some other hospital system.
AT: Another thing good about using phenopackets is that it seemed like there was some kind of mapping

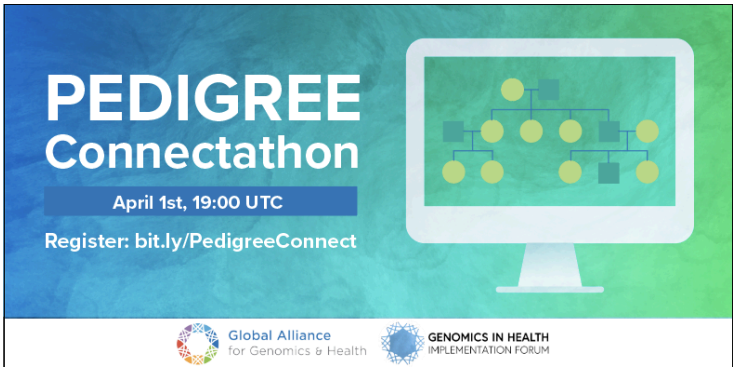
between phenopackets and HL7. Not an impossible task to integrate with EPIC, which should come to our hospital at some time. Building all this without any infrastructure behind us. Unfortunately will be on paper for some time. There are varying degrees of how digitally mature they are, have to aim at the lowest barrier.
HL: Is software proprietary?
AT: At present, it's been developed proprietary in that we've been designing it for our workflow, not looked into commercializing.
HL: The flow and the logic is really informative, and it would be great if you wrote a paper about it.
AT: Have heard talk about a paper being written. Expect it will happen at some point, everyone would benefit if data is being sent in a way that everyone can understand it.
MH: Great idea, we've been talking about deployment, people need to go through their experiences to harden the standard. Thinking about designing an API for phenotype matching, match across systems and give better structures for exposing that information. How to help implementers deploy something like that across systems once phenopackets has been implemented.

2021-03-18: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Driver Project Affiliation)": Meen Chul Kim (CHOP/KF), Kathryn Van Diemen (TrakGene), Zornitza Stark (AGHA), Alejandro Metke (CSIRO/AGHA), David Hansen (CSIRO/AGHA), Bing Feng (UUtah), Jules Jacobsen (QMUL/Monarch), Guida Landoure (H3Africa)

	Actions Arising	Assigned To	Deadline
1.	Create Pedigree branch in Phenopackets Github	Jules Jacobsen	
2.	Brainstorm names for the relationship ontology	ALL	
3.	Create issue on vital status vs life status	Orion Buske	
4.	Share conceptual model updates with Meen, Alejandro and Jules prior to the Connectathon	Orion Buske	
5.	FHIR Zulip Chat - start conversion on whether we're using FHIR appropriately with regards to patient and family member history	Meen Chul Kim	
5.	Share a de-identified pedigree with the group - Use case involving multiple wives	Guida Landoure	

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - Connect Meeting Report	Orion, Grant, Lindsay	5 min

	- Pedigree Connectathon - April 1st - Registration (for people who are not already part of the Subgroup) 		
2.	Updates to the Minimal Dataset Document	Grant Wood	15 min
3.	Cleaning up rough edges of the conceptual model	Orion Buske	15 min
4.	GA4GH Pedigree Ontology	Orion Buske	15 min
5.	Any updates to FHIR IG comparison efforts? (Time permitting, could save this for the Connectathon in two weeks)	Alejandro, Meen	5 min
6.	A.O.B.		5 min

Minutes:

Minimal Dataset Document

GW: A few comments came in, thank you for that. At the bottom there's a bit of new text added about how to read the tables. One things I did highlight is that there are some required and optional elements, and how required is what is needed to make the structure of the family tree. Added an appendix section at the end as well. Went through the tables and all of the text - highlighted in red means there is still a gap there. I have one section that is not complete - add some information for the reader on the conceptual model. Capturing the data elements talks about the tools and formats needed to collect FHH information. Appendix B includes the other elements to consider. Appendix C has supplemental information including information on Phenopackets and efforts in other groups.

Cleaning up the Conceptual Model

OB: Merge the pull request in so there is something marked. A few questions I wanted to run by people. In terms of splitting out the ontology/taxonomy into a separate repo, I think it would be helpful to align on the name of this. I wanted to open the floor but the first idea that came to mind was Human Relationship Ontology. Is that too narrow?
ZS: Biological relationships?
OB: There are non-biological relationships in there.
JJ: What's the format?
OB: Question of whether it should go in as OBO or OWL. Any preference of Has X vs Of X?
JJ: Can go either direction. How are we going to manage bidirectional relationships in that case? Two edges to each node in both direction?

OB: You could. I'll choose one. Last two points, one was around changing life status to vital status to align with KidsFirst.
ZS: Haven't come across this term yet
JJ: Also in ARGO
OB: In terms of clinical genetics, you come across life status more? I would vote to leave it then.
MCK: We picked vital status from SNOMED.
OB: In absence of a strong opinion to change, just leave as it.
JJ: Might be worth taking a moment to align. We based ours on ICGC-ARGO.
DH: I couldn't find vital status in snomed.
MCK: Sorry, from NCIt, has both vital status and life.
OB: I will create an issue to open up this for discussion. Last conversion was about identifiers. MB was advocating for CURIEs. Worried that without providing implementation guidance it's not clear on what the prefix should be.
JJ: Gene names you can enforce, but for an individual person there's no such thing.
OB: Looks like alternative IDs are a list of CURIEs
JJ: The ID is an arbitrary identifier
OB: One arbitrary, and lots of CURIEs
JJ: That doesn't make a lot of sense, I don't think that's feasible. I agree that CURIE should be supported, but in terms of types it's just a string.
OB: Can update to support that. Those were the only to-dos in terms of cleaning up. I will create an open issues for some of these. Will send a note to Meen, Alejandro and Jules about the specific changes. Hopefully update within the next week or so, so we can have some stability for the Connectathon. Mostly just wording based changes.
AM: Yes, should be fine.
JJ: Which bit specifically?
OB: A few points that came out as rough, like collected to date. Being clear about identifier vs ID.

<u>FHIR IG Updates</u>
MCK: The IG development has been in my backlog. In a sprint right now that ends tomorrow. Plan is to get back to it next week.
OB: Is the outcome that AM and MCK will have an aligned IG?
MCK: We should regard AM's IG as the base. Should open up some discussions about several issues Alejandro pointed out, then merge.
OB: Pull requests coming.

OB: There was one conversation last time about patient on family member history, are we using FHIR in a reasonable way? Has anyone done that, and if not, Meen can you do that?
MCK: Sure

JJ: I showed the implementation. For Zornitza's examples, would be helpful to have a volunteer. I'll make a branch that can be sent around.
OB: One other modelling related topic I'd like to discuss. From an analysis perspective, anticipating the new hire in GA4GH and building validation tooling. It seems like a common workflow will be extracting parent-child relationships from a pedigree that has been transferred. Common case we want to drive towards is where pedigrees are completely and robustly structured as possible. The one challenge is that the information is all pooled together in one place, and distinguishing one parent vs non parent relationship, using taxonomy to interpret. Does this make the barrier from an implementation perspective too high on the receiving side, when that could be simplified by making the common case explicit?
JJ: Easing use will ease adoption. Depends on how complete you want the pedigree to be. SHould relationships say mother and father as opposed to parent?

OB: One is that the relationship itself has a high level code on it, that is independent of the detailed one, allow user to quickly filter without having to interpret the ontology. Traverse the bounds of the connected graph - trade off.
JJ: These things won't be enormous in terms of nodes and edges. If you want to make it explicit in saying mother of, father of.
OB: At Connect, in Search, we came up with these queries that people are likely to ask, like find me this persons parent or child. Is there a value in this property, does there exist a relationship where these are the two entities involved
ZS: THere is value in this, it's a common question I ask - is this mother and father, and are they affected. Then scramble around to find that information because we only have IDs.
OB: Do we want to make that 80% easier
JJ: So we have a child note, one says mother, one says father, and it has an ID
OB: Reason for not doing with just an ID is the problem with adding metadata to that information. Extensibility is the separate relationship concept. If you tie to raw ID property, can't add information to that edge.
JJ: Depends on what the most important use case is.
MCK: In KidsFirst, we implemented the section option. It covers most of the cases, more than 90% of case. But, we are moving towards the first option now. We decided we needed to allow query accessibility more
OB: Your representation of those links, those edges, do you have simple mother father or something richer
MCK: We have grandfather, great grandmother, etc. We're planning to add all the possible edges.
GL: Are you thinking about families we see in our culture here, like multiple wives.
OB: The current model makes that easy to represent. If you have a few moments, would you be able to share a deidentified pedigree with the group?
GL: No problem

2021-03-11: Phenopackets Subgroup

Minutes: The Phenopackets minutes and agenda can be found in this document.

2021-03-01: Virtual GA4GH Connect 2021

The annual GA4GH Connect meeting took place on March 1-4, 2021. Please find the meeting report, including meeting minutes and slides, here.

2021-02-25: Phenopackets Subgroup

Minutes: The Phenopackets minutes and agenda can be found in this document.

2021-02-18: Pedigree Connectathon

Chair: Orion Buske, Grant Wood

Global Alliance
for Genomics & Health

Attendees "Name (Driver Project Affiliation)": Alejandro Metke (CSIRO), Moni Munoz-Torres (Monarch), David Conlan (CSIRO), Asiyah Lin (NIH), Meen Chul Kim (CHOP), Eric Torstenson (VUMC), John Grimes (CSIRO), Kathryn Van Diemen (TrakGene), Brian Walsh, Jules Jacobsen (QMUL/Monarch), Alex Felmeister (Illumina), Dietmar Fernandez (CRG)

	Actions Arising	Assigned To	Deadline
1.	Clean up rough edges of conceptual model	Orion	Next Meeting (March. 18)
2.	Compare generated FHIR IGs	Alejandro and Meen	Next Meeting (March. 18)
3.	Review GA4GH Pedigree Ontology and note any preferences for capitalization, IDs, no white space, etc. - Create repo in Github to track this more formally		

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH Connect registration (March 1-4th) - GHIF (March 9 & 10th)	Lindsay, Orion, Grant	2 min
2.	Brief summary of Pedigree Standard status	Orion	5 min
3.	Question/issues in building pedigrees using protobuf or FHIR - Example pedigrees - FHIR IG / documentation - Draft conceptual model	ALL	50 min
5.	A.O.B. - 2nd Pedigree Connectathon scheduled for April 1st, 2021		3 min

Minutes:

FHIR IG
AM: Interested in getting feedback from people who might have looked at the draft FHIR IG. Requested Zornitza and team to build pedigrees that were more complex. Dave has already reported a bug in the examples, Bart is female.
OB: It says the following instances are used to implement these examples and it has these links to particular patients. How do those work in the context of FHIR and the individuals? Unsure about the mapping between deceased patient in FHIR vs life status of individual and how they relate.
AM: One of the interesting bits. Just notice that the documentation might be out of date. Two options for life status - one is to add an extension and add a code system. Main issue is that unborn was not in the common ontologies. In the actual implementation ended up using FHIR boolean representation. Quite common to use in

FHIR. Get into philosophical Qs about whether you can have a unborn patient that is deceased, and life start. The deceased attribute in FHIR is common, and it didn't make sense to create a full extension so that drove the decision here.
DC: The open pedigree editor has a field called gestation age, when I wrote my version, if there was a gestation age given, I put it in as a string.
AM: We might want to have an extension for gestation age. That wasn't in the current conceptual model.
OB: Age was enough of a separate thing. Robustly and independently captured by FHIR and phenopackets. How to handle for unborn individuals?

MCK: I tried to model the first pedigree figure in FHIR. 4 generation families with different conditions. I got halfway through in terms of implementing individuals and relationships. Still need to work on pedigree model. I picked the profile called composition for the base. I was trying to reflect this proposed model in FHIR as exactly as possible. I picked patient and observation in order to model relationship. Used FSH/Sushi. I implemented the relationships partially from the proband - parents, parent sibling, grandparents.
OB: What prompted you to do everything relative to proband rather than discrete direct relationships that compose pedigree?
MCK: We can add relationships between every node, I considered the composition profile, and that profile has an attribution called subject. Not sure whether it is a proper approach but I considered that the proband.
AM: Main issue is that doesn't reflect conceptual model in that it wasn't mean to be proband centric.
OB: Without reworking all of the relationships, you could rework and subset the relationships around another individual in the family. If all relationships are relative to the same individual, those must be increasingly ambiguous the broader that gap. It means that you can't just subset those relationship entities in order to get a reduced or zeroed in representation of the family.
MCK: I'm not yet super familiar with the composition resource. As a starting point, started with the proband to reduce complexity. Haven't reviewed AM's IG yet.
OB: Would be helpful for the next step to align the results from your evaluation with the IG that AM put together. Are there any other things you found in doing the modelling work that were difficult or confusing?
MCK: Relationship was pretty straightforward.
AM: Is your version of the IG publicly available? Would like to see the decision you made and how they compare to what I made? Would like to see relationships as well.
MCK: This individual is the proband. I tried to first use the default attribute if relevant, and tried to add extensions to cover the rest of the attributes proposed. Added affected status as boolean, and used HL7's gender identity. We usually received de identified PHI, usually no way for us to put PHI data in FHIR.
AM: The affected status is an observation in my IG, with goal of minimizing the number of extensions used.
MCK: I used observation based on FHIR IG that says observation profile can annotate any demographic characteristics. Code is fixed. When relation is used, it should be represented from relative to individual. Subject is the proband and focus could be parent. Only allow codeable concept using Orion's valueset.
AM: I didn't use observation, I used family member history resource. The same FHIR spec recommends using observations if there isn't a more specific resource already available. Used one of the existing extensions, familymemberhistory patient record. Already have the link, the nodes and most of the work in that profile was getting rid of ambiguous information. Forbid filling out duplicate information. Can use as a relationship object.
MCK: I consider using family member history as well. I tried to pick FHIR profile that supports every kind of core attribute. I didn't end up choosing family member history, without using this extension, there's no way to point to the proband. It points to the consultant?
AM: It points to subject of the relationship.
OB: I thought that the way that resource usually defined, the patient property always refers to the proband. Family member history always report to the same proband.
AM: Good point. It's in the alternate name of the patient attribute. I also hesitating in using this, but after discussion the conclusion was that it was a poor documenting decision for the original use of this resource.

Something we could clarify in the IG. It's clearly poorly named, the alternative name is proband. But if you look at the main description it's just saying this family member history is the patient.
OB: Is there a risk of running into a problem where existing family members historys where the patient does refer to the proband.
AM: I think it's unavoidable. If someone is using the original model, they'll have most of the fields that are forbidden in our use, so will have to transform. Agree with Moni that we raise in the Zulip chat.
MMT: As we clarify the concept, and give semantics it will help present the model we're trying to follow.
AM: Having the conceptual model and referencing that will helpful, it gives the semantics.
OB: Sounds like MCK and AM should consult and confer. Is it realistic to come up with a unified FHIR IG? [yes!]
AM: Open some issues in the FHIR IG github to track.
MCK: I will complete the pedigree part of the implementation. I will put the IG on my personal Github.
AM: Using the pedigrees for both implementations.

Phenopackets
OB: Directionality of the labels. As defined, the term parent on it's own doesn't quite fit well with RDF or graph representation. By switching to a graph-based model, getting direction for these relationships is important for the properties of the resource to the labels.
JJ: I've done a bit of modelling for putting this into a phenopacket. I took the conceptual model and I put all these properties into some classes. [shared screen]. You could define all the bits but there was no way of defining where the relationships should live. It doesn't live in the pedigree or the relationship itself.
OB: Good point. You have these many to one relationships between the relationship class and pedigree and the individual class and the pedigree. Two ways this can be represented. One is to basically put a pedigree ID within the individual and the relationship, and now you have three separate tables which would be a flat structure. Might be preferable for search. Other way is a list of individuals and list of relationships where those properties aren't explicitly laid out.
JJ: I just created another object to hold a list of relationships.
AM: In FHIR IG, everything ends up in the pedigree that is a composition. Which effectively would be adding those relationships as a containment in the pedigree.
JJ: Makes sense to have it as something in here. Still need the individuals. Makes sense to have everything bundled in pedigree itself.
OB: Exactly, set.individuals and the list and set.relationships and the list. Then everything is self contained. The intent is that it's the way it would work from an object or resource based model. Case where it might be better for some modelling to be flat but referencing each other. I am anticipating someone asking for a CSV.
JJ: Formatting issue.
OB: Right, not a conceptual model issue.

OB: Is it regular way to have plural on repeated fields?
JJ: Reads as "get.individuals: at the end.
I didn't know what to put as sex.
OB: I created an issue for the discussion, what is the preferred valueset for use to choose these core terms from. It could be the FHIR administrative valueset.
JJ: No preference. Anything is better than boolean male, female.
MCK: How much granularity do we want to bring. If we use karyotypic sets we can annotate more clinical set.
JJ: PP has phenotypic and karyotyic sex. Phenotypic corresponds to FHIR administrative gender.
OB: Try to align as much as possible for the standard expectations within the clinical risk assessment use cases. As far as a I know, in genetics the general practice is sex at birth. So calling it sex is reasonble, using administratie gender is reasonable but some semantic friction. Could see the value of allowing richer representation.

JJ: I would go against that. One is ped, because it has a limited valueset, so have to map something down to that valueset. Also, there is a clinical thing with sex at birth. The sex here is meant to be the phenotypic sex. I think we should keep the sex part simple and correspond to what software uses.
OB: Might make the initial learning curve with the standard a little bit harder than it needs to be to start with. Look to v2 with an event model with time based observational information.
JJ: This collectedAt seemed odd, day collected or collection date might be easier, spells out type more easily.

JJ: The little trio worked really well. Reason collected was slightly bizarre, had to say what reason collected really meant.
OB: Recommendation for the same of that property? Our goal is 80% of use cases, and many ped use cases have a particular condition being investigated that is a boolean indicator. Handle that cleanly and then focus on the future.
OB: Let's do some brainstorming for collectedat and reasoncollected, remaining.
MCK: Condition seems to suffice to me.
JJ: The proband is the person who has the condition?
OB: Proband is usually the first person tested positive, or first person seen by or referred to a particular clinic. While consultand is the person being currently tested if different from the proband. If that's not clear from the definitions then we definitely need to be more explicit.
AM: Proband is always static, the consultand can change.
OB: In practice, we run into groups that say there are families with 4 probands. As long as we are explicit in what we mean.

OB: In terms of feedback, in order to make this a proper valuset, preferences for IDs, capitalization, no white space, etc. If someone could create an issue with recommendations for it to look more proper that would be helpful. We could switch to OWL file.
AM: I did a FHIR codesystem that you can browse in Shrimp. I haven't updated.
JJ: OWL or OBO in Github.
OB: Separate repo.
AM: It would concepts, not relationships. If you want reasoning, it's a different type of ontology.

GOALS: Clean up rough edges of conceptual model, and getting those three pedigree examples implemented in Phenopackets and FHIR.

2021-02-11: Phenopackets Subgroup

Minutes: The Phenopackets minutes and agenda can be found in this document.

2021-02-04: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Driver Project Affiliation)": Moni Munoz-Torres (Monarch), Bing Feng (UUtah), Alejandro Metke (AGHA/CSIRO), Meen Chul Kim (CHOP), Alex Felmeiste (Illumina), Kathryn Van Diemen (TrakGene), Dietmar Fernandez (EGA), Zornitza Starke (AGHA)

Global Alliance
for Genomics & Health

	Actions Arising	Assigned To	Deadline
1.	Create a Github issue for sex and gender	Orion	
2.	Set up a space in the Github repo for a FHIR IG	Lindsay	
3.	Continue to review the minimal dataset document. **Please add your approval to the comment to the top of the document**.	ALL	
4.	Connectathon scheduled for **Feb. 18th** (next meeting). Calling all technical folks!	ALL	

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - GA4GH Connect registration (March 1-4th) - GHIF (March 9 & 10th)	Lindsay, Orion, Grant	5 min
2.	Minimal Dataset Table - Resolving most recent comments	Grant	20 min
3.	Pedigree Connectathon - Feb. 18th, during a scheduled Subgroup meeting - Larger GHIF connectathon April 1st	Orion	10 min
4.	Issue #7 - Representation of Life Status	Group Discussion	10 min
5.	Issue #8 - Identifiers - Pending feedback from TASC team	Group Discussion	10 min
6.	A.O.B.		5 min

Minutes:

GW: Hoping people will still go to the document and review. Make your vote on the comment at the top of the document. Some comments about creating a proband type. We have an ID, that can be optional. To further qualify a proband beyond an ID number. If you got to the research subject resource in FHIR, status has different codes for a research subject. This was described some time ago as an option, if people think it's still a good thing to keep we will, or if that's too confusing then we can remove it.
If the proband ID is not required in a particular use case - an use case example may be that complete pedigree is collected, but when shared is proband is not selected if a receiver wants to looks at different probands.
Other comments around individual ID. There were ideas for an individual ID type - missing relationships, hidden for privacy reasons, etc. Not replacing the ability to ID a person within a pedigree, but there may be other identifiers that may be helpful and we could distinguish those with a type.
MCK: Mostly asked for clarification, and your answers were very helpful.
GW: Don't know if individual type is all that valuable, over time through piloting and experience we can discover if they have value.

MMT: Makes sense to me, I think it will be useful.
OB: Is some of the confusion about whether the type of individual vs the type of identifier?
GW: Should be identifier.

GW: MCK found a FHIR page that talks about other options for patient gender and sex. This is one of the problems we constantly have with standards. There are different options to implement that standard. Because this document is not an implementation guide, we don't seek to tell people how to implement.

Question about consanguinity - why make a redundant field. I guess it could be inferred depending on how big the pedigree goes. Some pedigree don't even include first cousins.
OB: We run into that alot. Simpler pedigrees like trios but want to pass consanguinity information with it.
ZS: Agree

GW: Question about recording medical interventions. Further in the document we have other data elements to consider. There is a FHIR resource.
For REA, as a starting point, suggested use of particular ontologies. Most people collecting FHH today are using their own ontologies, so they're not looking for someone to tell them what to use. Trying to let people know that this is an implementation question, multiple ontologies already out there.
MCK: Totally agree.

GW: Lastly, multiple birth FHIR resource. I understand the resource starts with a boolean, and then has an integer after that define the order of the birth. That use case was odd to me. My experience is boolean yes or no, but then was it twin or triplet. Maybe need a new FHIR resource.

OB: Idea was to block off some time to bring together technical people focused on FHIR and phenopackets and connect up to the first version of the standard. Proposal is to hold this on our regularly scheduled call in two weeks (Feb. 18). - (MMT sent out invitation to Jules & Peter).
Touching on the GHIF meeting, we'll have the opportunity to share what we're working on. Invite these large national projects to have technical representatives join the connectathon. Hold that in early April. Have the mini connectathon internally, and be in a good place to connect with external groups later. Would align with our regular call slots.

OB: Will make an issue for representation of gender and sex. Want to take a moment to get into this. Within the standard we have two fields, one for sex and one for gender. The intention is to allow for difference in storage of those two data elements. In particular to make sure the standard at a fundamental level can represent either of those concepts or atleast an administrative gender and the ability to a degree that allows us to say this individual is male to female transgender. We can describe that level of specificity so we can take advantage of/promote/drive the adoption of new representations of those individuals in pedigree visualization. At a fundamental level, what we really have is an administrative gender, if you're going to have one field to represent sex or gender it's your catch all term. If the individual is drawn as a square or circle, that is what it will be based on. If someone self-identifies as a gender that doesn't match that, and the other is if there was some genetic testing, or we know because they were the father, there is some bioglogical sex component that doesn't match administrative gender.
MCK: For clarification, the attribute sex is for birth sex? Whereas gender is reported.
OB: The intent is that the field called sex is the biologial sex if it is in contrast to the person's gender either self-report or administration.The gender field is what dictates the outside of the shape of the node, administrative gender. Propose we update the definitions.
MCK: The conceptual difference is pretty clear.

KVD: I think it's clear the way it is, given that biological sex is a required field. I've got an optional qualifier if a genetic counsellor records something different from that. We don't have codes for transgender yet. Have a flag to check gender. Having a specific field is perfect.
OB: In general, the information people would make a pedigree on is gender and not sex.
KVD: Can vary. We may build the initial pedigree on sex and go back and change those. Various acronyms. Ongoing debate on whether pedigree should reflect biology or what's going on socially.
OB: So from a GC perspective, in general it's actually sex and not gender that's basis for pedigree. But that sex is not tested biological sex but rather the administrative gender of that person
KVD: Yup that's fair. Often depends on where people are in the pedigree. Short answer is it's tricky.

OB: The way it's currently written is aligned with KVD. Can we make this clearer in general? Or anticipate this potential confusion around a field being called sex.
MCK: Maybe just provide some details in the documentation.
OB: Yes, add some examples. Sounds like it might be useful to highlight one field for the primary use of drawing a pedigree.
KVD: I would do one required and one optional.
OB: Right now both are optional. Could say that sex is required. Is anyone opposed?
[No]

OB: We'll want to figure out what the right terminology is for that. I'll put up an issue and put out a call for terminologies.
We had an issue discussing life status. We're posing the quesiton about whether this is status that should be associated with an individual vs an observation that's made about an individual at a certain point in time, such as that based on age. My sense is that the broader goal is to have life status information observation based or event based. For first version makes sense to defer events, but having some life status information to specify if an individual is alive or deceased. Either we leave life status as it is, as an attribute on an individual recognizing that it may have to be updated over time, or we remove it and make sure it is part of the next iteration where we add some mechanism for described health events and life status observations.
MCK: I am ok with keeping both. Even in FHIR, when you look at patient profile, I recognize life status is not just an object.
OB: Propose we keep as an attribute on the individual, allows more value for common use cases, with the recognition that we can provide more fine grained life status information down the road.

OB: Last topic is more on identifiers. MB is strongly encouraging us to use CURIEs. Answer is that it gets more complicated. That model works well for concepts, but the placeholder for identifiers which is a property on the individual is more like a list of IDs. I don't know what the prefix of the CURIE would correspond to here, where we're referencing individuals in a health or research setting.
AM: I think this might be implementation specific. We'd be ok to leave it as an ID, and leave implementation to the implementers of the spec. Maybe just say ID.
MMT: We should think about the architecture now. Wait and see what the TASC team recommends.
OB: May still be useful to distinguish an internal ID where these is no expectation it extends beyond this data model.
AM: Agree, at this level the distinction is relevant.
OB: I can make that more explicit in the documentation.

AM: Spot in the Github repo for the implementation guide?
LS: I'll send you an email and we can sort this out.

Global Alliance
for Genomics & Health

2021-01-28: Phenopackets Subgroup

Minutes: The Phenopackets minutes and agenda can be found in this document.

2021-01-21: Pedigree Subgroup

Chair: Orion Buske, Grant Wood
Attendees "Name (Driver Project Affiliation)": Grant Wood**,** Moni Munoz-Torres (Monarch), Melissa Cook (NCI), Bing Feng (UUtah), Alejandro Metke (AGHA/CSIRO), David Hansen (AGHA/CSIRO), Dietmar Fernandez (EGA-CRG), Amanda Spurdle (AGHA/BRCA), Alex Felmeiste (Illumina), Zornitza Stark (AGHA), Meen Chul Kim (CHOP)

	Actions Arising	Assigned To	Deadline
1.	Remove the repeated element for relative relationship	Grant Wood	
2.	Add more descriptive text to adoptive status to clarify that it is 'adopted in', parents not biological	Grant Wood	
3.	Review the minimal dataset and let Grant know if you approve/don't approve (please provide rationale for no approval)	ALL	Feb. 4th
4.	Interested in drafting out the FHIR IG? Please let Orion or Alejandro know	ALL	
5.	Share GA4GH Connect details when available (re: Pedigree connectathon)	Lindsay	
6.	Create Github issue for life status and event code	Orion Buske	

	Agenda Item	Speaker	Time
1.	Welcome, Housekeeping - New minutes document! - Google Summer of Code (GSoC)	Lindsay, Orion, Grant	5 min
2.	Minimal Dataset Table	Grant	20 min

3.	Draft Technical Model	Orion	20 min
4.	Next Steps: Review and Voting	All	10 min
5.	A.O.B.	All	5 min

Minutes:

Minimal Dataset
GW: Want to get to a point where we decide that this document is ready for a formal first draft status. Ready to share with other WSs and outside groups to get feedback. Some steps to go through before we say it's ready for public viewing. Document is not meant to be a standard, it is a recommendation or something that people can use as an informative document when looking at what to do with FHH initiatives. Document contains a minimum core dataset that we recommend FHH tools should collect. Idea is that at some point, organizations might have the use case to share this information with some other computing system, some other organization. A lot of research organizations used the PED format, and many have realized that format is limited in terms of the data elements it can collect. If we add to those data elements, we should agree on what the definitions of those data elements are. Added in some ideas about long-term goals and the methods we took to develop this document. Living document, so we will add to it over time.
Decided to divide up the table that contains the data elements. One is information on the pedigree, and one is on the individual. I have gone through this table and did what I felt was my last edit. I still have left information about some gaps that exist. Column for ped and column for mapping to FHIR as it exists today. HL7 has their working group meeting next week and I will be presenting this to them, and will share where there are some gaps with what we have and what they have. Hopefully will encourage that group to consider some ideas.
OB: In terms of multiple birth, just about distinguishing between types of births. What is the order vs how many individuals are there in that multiple birth.
GW: With multiple birth status question, want to know if you are a twin or triplet or more, but don't ask the order. Just my experience but don't want to limit others.
ZS: I generally agree. Only two: focus right after birth when there aren't names. Other is during a pregnancy, the twin often get labelled with ultrasound. Can be important if they're sampled. Shortly after birth it becomes less important. Overall we shouldn't worry about it
GW: We can always change this. Orion you also had a comment about adoptive status?
OB: For boolean, make sure we're explicit with adoption whether that means adopted in vs adopted out. Make sure it's clear who is being listed as the parent and whether that's birth vs adopted.
AS: Have relative relationships and relative first name and last name. In our setting we could present it as relative relationship but wonder if instead it could be last name or ID. Sometimes when they come to us they're stripped of names. Would not be able to use unless there was a separate row.
GW: The data elements people can pick and choose from.
AS: There's mother and father ID but no ID for relative other than name.
GW: Go under individual ID.
AS: So for the related individual, the relative relationship there is relative to who? There's a repeated field for family members, but then that family member is ID'd by name.
GW: Relative to whoever the proband is. Want the information to be able to be reordered for a another proband in the family. Individual is each ID in the pedigree, not the proband.
AS: Why are there then repeated elements for relative relationships. I thought individual ID was the proband, then repeating things for each relative.
OB: Individual could have multiple relationships to the proband that you would want to describe with different codes.

GW: I think we originally had this before we split up pedigree and individual. And now that's it' split, we don't need it. The relationship is needed but the repeated field part is not.
AS: Exactly, only need one.
GW: Good catch. I will remove that later and make sure there isn't any other confusing parts.
AS: Do you need the relationship? Don't need if the have a mother ID or father ID.
GW: It's optional.
AM: Mother and father relationships could be eliminated.
AS: We want the IDs.
AM: In the new model it's in the relationships. Meant to represent the node and the relationship tells how they are connected.
AS: For each person you repeat who mother and father are so you can cross-check.
AM: With graph model just need the node and ID and the relationship would tell you how it's connected to other nodes.
OB: The intent is probably to be agnostic to particular way things are recorded. This is independent from the practical considerations to the model where it is linked to ID.
GW: Where we want to get to a standard way of implementing that will be in Orion's model, as far as how the data is stored and shared. I think this question as we get to the place where people are getting to the implementations they might discover some things that we can come back and change.
Back to adoptive status, we want a boolean that says the person is adopted or not. If it's true, it means the person is adopted. True means in.
OB: So mother and father ID corresponds to the roles of mother and father rather than biological.
GW: I'll add some clarity here.

GW: We want to acknowledge that there are different ways and tools for how data can be captured. And so we want people to see this as the only way to capture data. Multiple use cases. Have a bit of information on REA, with all the conversations we had there, there are so many uses and methods that we didn't want to tell people a particular method but provide information about the different methods. Suggest we give a code to these different methods, so that when we share data people can codify what they're using. Section at the bottom for other data elements to consider like sharing status, not part of core set. Data can be post-coordinated, use case specific data elements that can be added to over time. A pedigree could have had some analysis done, and you may want to include the result of that analysis - pulled that out form HL7 as something to include down the road.
Sense a lot of people haven't had a full review, review over the next and send an email to me about approval.

Draft Technical Model
OB: This is now linked into the minutes. Moved over into Github and to the benefits of pull requests. This is to bring everyone up to speed with where the proposed basic initial conceptual model is. Started building out a more comprehensive pedigree model. We will get there, but moving in stages is the best way forward. Moved back to smaller chunks and building out from there. We have the pedigree, individual, and relationship. Have a good foundation with the event model built with Bing's model. Makes sense to build that in the next iteration, but explicitly removed it from this version to start at a solid core model that we can then build these pieces on to, such as reported information which would have provenance information associated with it. Additional information that's attached to an individual vs right now where the individual has the essential core information which is assumed to be true in the context of a clinical pedigree. We'll get beyond that into additional reported information by family members, Don't think event is necessary for this first version especially given in the context of proof of concept in FHIR or in phenopacket, those other bits of information like DOB and current age are able to be captured by those native formats anyway. Added value is adding that richness of who's reporting and how that relates but that's a longer discussion.

Starting with the point that says we're defining what a pedigree is, what a relationship is and what an individual is, and making the big step beyond ped that this is a graphical model, that these relationships are entities in and of themselves, which we can talk about, add more info to and build on top of.
Most things are optional with the exception of an ID for an individual and the individual relative. The most useful place to collaborate on the ontology is a dedicated spreadsheet. 26 relationship terms that spread biological and legal relationships, including mapping to HL7 role codes and NCIt. Open to comments or concerns about this being an improvement over what is there, while making it clear that we would build the event model. Consistent with the gedcomx format as well.
Does anyone feel it is insufficient to be a starting point, like a v1? This is the thing we go and proof out in practical implementations. This is then the core starting point that is simple enough to be understandable and communicable.
[No concerns raised]
OB: Propose we merge in the pull request, have that as a starting point and then proceed with the implementations we had proposed previously. AM and MCK, what are your thoughts on next steps on prototyping this or taking it for a test drive.
[Orion's Pull Request to Add draft pedigree model: https://github.com/ga4gh-cp/pedigree/pull/6]
AM: For the FHIR side we'd have to start with a draft implementation guide and I'm happy to help with that. Would then need to test it in some sort of connectathon. Happy with the model and should be straightforward to have a draft.
MCK: Happy to help with AM.

GW: Going back to the connectathon idea, I know there's the HL7 one. Is that what you were thinking of?
AM: Something for us to decide. Could leverage HL7 meetings, but also could organize a smaller GA4GH thing.
[Lindsay to keep everyone up to date about GA4GH Connect]

AM: Couple of important decisions to be made with FHIR resources. Once agreement is reached, implementation shouldn't take too long. Might run into something unexpected, but shouldn't take too long.
OB: One question about the concepts used by this initial version - life/death status - should it be an enumerable type or use an existing terminology? Question to be answered through the modelling process. Encourage everyone to review the model and add issues.
MMT: Is there a ticket for the issue about status?
OB: No, but will create one and will also will make one for event code.
Orion's Pull Request to Add draft pedigree model: https://github.com/ga4gh-cp/pedigree/pull/6