Meeting Report
Virtual Connect 2020

Global Alliance for Genomics & Health

This year, the 2020 GA4GH Connect meeting was held entirely virtually in response to global health concerns related to COVID-19. The meeting focused on two major areas: (1) collaboration across Work Streams and (2) development of the 2020 GA4GH Strategic Roadmap.

Table of Contents

1. Gap Analysis Summary
2. Working Sessions
    a. Assurance Levels
    b. Beacon V2
    c. Building Tooling
    d. DURI Connector
    e. Federated Analysis Systems Project (FASP)
    f. FHIR Implementation Guide
    g. Future of Variant Data Representation
    h. htsget: Interactions with DRS & other projects
    i. Key Management
    j. Machine Readable Consents
    k. Pedigree
    l. Phenopackets: Treatment Course + VRS Integration
    m. Sequence Annotation
3. Roadmap Sessions
    a. Clinical & Phenotypic Data Capture (Clin/Pheno) Roadmap
    b. Cloud Roadmap
    c. Data Security Roadmap
    d. Data Use & Researcher Identities (DURI) Roadmap
    e. Discovery Roadmap
    f. Genomic Knowledge Standards (GKS) Roadmap
    g. Large Scale Genomics (LSG) Roadmap
    h. Regulatory & Ethics (REWS) Roadmap
4. Report Backs

GAP ANALYSIS SUMMARY

Recording | Presentation | Gap Analysis Summary

Starting in 2019, Heidi Rehm and Andrew Morris conducted a Gap Analysis Survey to identify successes and gaps with GA4GH. In this opening session, Rehm and Morris shared key findings and themes from the Gap Analysis, and attendees voted on the top six areas that GA4GH should focus on addressing first. These are:

1. Collaborate with healthcare, EHRs, and HL7
2. Improve Interoperability and alignment between standards
3. Create tooling and testbeds to support implementation
4. Highlight end-to-end implementation demonstrations
5. Harmonize terminologies, data models, definitions of maturity, and versioning across the organization
6. Increase facilitation, training, and support for implementations

WORKING SESSIONS

Assurance Levels

Agenda & Presentations | Recording

The Data Use & Researcher Identities (DURI) and Data Security Work Streams met to raise awareness about assurance levels and shared a state-of-the-art solution for strong ID proofing used in ELIXIR. The teams aim to use the information discussed to improve the Data Security Infrastructure Policy and to consider assurance levels across Data Security and DURI products.

Beacon V2

Agenda & Presentations | Recording

The team presented the current status of Beacon v2 and explored how Beacon can work with other GA4GH standards developed by other Work Streams such as Clin/Pheno, Genomic Knowledge Standards (GKS), and Data Use & Researcher Identities (DURI). The team also discussed ways to better align products and initiatives within the Discovery Work Stream.

Building Tooling

Agenda & Presentations | Recording

The session focused on streamlining the process of standards development to make it easier for work streams to publish their standards. These include: 1) an automated documentation builder to convert OpenAPI specifications to HTML, 2) cross work-stream, consolidated approaches to compliance testing and reporting, and 3) human-friendly introduction to standards.

DURI Connector

Agenda & Presentations | Recording

The Data Use & Researcher Identities Work Stream met to discuss connecting the Data Use Ontology (DUO) and GA4GH Passports standards in data access request systems. The team collected user workflows to see how DUO and GA4GH Passports can fit into different processes to further streamline and automate data access processes.

Federated Analysis Systems Project (FASP)

Agenda & Presentations | Recording

The Discovery, Cloud, Data Use & Researcher Identities, and Data Security Work Streams met to further work on FASP—a project demonstrating the utility of GA4GH APIs working together in an ecosystem. The team saw an updated "golden demo" using the DNAstack/Google software, and implementers shared feedback on building systems that support these APIs. The team will continue to collect feedback and expand the current "golden demo" to include more data sources and complexity.

FHIR Implementation Guide

Agenda & Presentations | Recording

During the session, teams from Children's Hospital of Philadelphia (CHOP) and Johns Hopkins University shared the outcomes of test implementations of the Phenopackets-FHIR guide and a JSON-LD method for validating & roundtripping phenopacket to FHIR, respectively. Next, the team aims to refine Phenopackets-FHIR mappings.

Future of Variant Data Representation

Agenda & Presentations | Recording

The Variation Representation Specification (VRS) and VCF teams met to discuss similarities between their specifications, particularly around the construction and representation of CNVs and SVs. The teams will schedule regular touchpoints between the specifications on CNV/SV representation to harmonize where possible, and record similarities and dissimilarities in public-facing documentation.

htsget: Interactions with DRS & other projects

Agenda & Presentations | Recording

The session focused on extending features of htsget and discussing cross-Work Stream integration. These include: define service-info vocabulary for htsget endpoints, clarify Data Use & Researcher Identity (DURI)/Passport flows, extend interoperability testing to registered htsget servers, and explore opportunities for a FHIR Genomics htsget pilot.

Key Management

Agenda & Presentations | Recording

The Cloud, Large Scale Genomics (LSG), and Data Security Work Streams discussed the system architecture and rationale behind key management. The team arrived at two crucial points: 1) embed geographic distribution and national constraints in the system, and 2) make key management itself more manageable.

Machine Readable Consents

Agenda & Presentations | Recording

The group discussed the Machine-Readable Consent Drafting Guidance, which aims to eliminate ambiguity between data sharing consent language in consent forms and machine-readable data use terms (DUO) to enable more efficient and ethical data sharing. The community is encouraged to read the Machine-Readable Consent Drafting Guidance and comment on clarity, value, and feasibility.

Pedigree

Agenda & Presentations | Recording

During the session, the team shared a pilot case for managing integration of patient-reported and clinician-curated health data, cross-collaborated with the VCF team, and arrived at a consensus on data modeling priorities for the Pedigree standard. The team will continue work on a Minimum Core Data Set, to result in a new FHIR profile, and will aim to have a draft Pedigree standard ready by 8th Plenary.

Phenopackets: Treatment Course + VRS Integration

Agenda & Presentations | Recording

The Genomic Knowledge Standards (GKS) and Phenopackets teams met to discuss implementation of the Variation Representation Specification (VRS) alongside the Variant building block in Phenopackets. The team also identified a potential new cross-Work Stream effort on Treatment Course, with plans to continue discussion post-Connect.

Sequence Annotation

Agenda & Presentations | Recording

Large Scale Genomics (LSG) and Genomic Knowledge Standards (GKS) Work Streams discussed the current status of the Sequence Annotation standard. The group aims to harmonize existing approaches to feature-level annotation, and will work closely with the Variation Representation Specification (VRS) team to follow best data-modeling practices.

ROADMAP SESSIONS

With the launch of the first GA4GH Strategic Roadmap in 2018, GA4GH is now undergoing a Gap Analysis project to identify successes and gaps within GA4GH and to update the Strategic Roadmap. During Virtual Connect, each work stream met to discuss their Gap Analysis feedback and continue finalizing their 2020 roadmaps.
Gap Analysis Summary Document
Draft 2020 Strategic Roadmap

Clinical & Phenotypic Data Capture (Clin/Pheno) Roadmap

Recording | Draft Clin/Pheno Roadmap
Clin/Pheno identified next steps for advancing the Phenopackets standard and will aim to have a draft of the Pedigree standard ready by 8th Plenary. The team also reiterated the importance of connecting with labs, healthcare systems, and patients.

Cloud Roadmap

Draft Cloud Roadmap
For the Virtual Connect meeting, the Cloud Work Stream did not host a roadmap session—choosing instead to invest time in cross-Work Stream initiatives such as the Federated Analysis Systems Project (FASP).

Data Security Roadmap

Recording | Draft Data Security Roadmap
Data Security reviewed Gap Analysis feedback and agreed with most of the findings. They will aim to provide better implementation support and create guides for performing security audits. The team also received good cross-Work Stream feedback throughout the meeting, and aims to work with the REWS team on data sharing in the realm of rapidly evolving patient privacy standards and work with the Federation Analysis Systems Project (FASP) on security-by-design.

Data Use & Researcher Identities (DURI) Roadmap

Recording | Draft DURI Roadmap
DURI aims to continue advancing its key deliverables: the Data Use Ontology (DUO), GA4GH Passports, and computable Consent Forms (in partnership with REWS). The team is also developing a plan to connect these standards together to optimize the data access request workflow.

Discovery Roadmap

Recording | Draft Discovery Roadmap
Discovery discussed the advancement of key roadmap deliverables, including Networks (Service Registry Standard and development of a GA4GH Service Registry), Search

Implementation, Schemablocks, and Beacon v2, as well as ways to integrate across Discovery products. The team is also focused on integration across Work Streams, and is involved in Federated Analysis Systems Project (FASP) to demonstrate how Discovery, Cloud, and DURI Work Stream products work together.

Genomic Knowledge Standards (GKS) Roadmap

Recording | Draft GKS Roadmap

From the gap analysis, GKS will address content and technical gaps in its standards, as well as increase cross-Work Stream and pan-SDO communications. GKS noted no significant issues for proposed Variation Representation Specification (VRS) and Variation Annotation roadmaps, and has support for current activities and direction.

Large Scale Genomics (LSG) Roadmap

Recording | Draft LSG Roadmap

LSG focused on cross-Work Stream activities, including mapping data structures from the file formats group to GKS initiatives and exploring DURI-Cloud-LSG workflows. Moving forward, the team aims to conduct outreach to community tool developers for input and explore potential for clinical engagement.

Regulatory & Ethics (REWS) Roadmap

Recording | Draft REWS Roadmap

REWS has made progress on deliverables and has community support on key initiatives such as standard data licenses, machine readable consents, the Consent Task Force, the Your DNA, Your Say survey, patient engagement practices, and the GDPR and Data Sharing Forum. REWS also discussed developing a position statement on data sharing during the COVID-19 pandemic.

REPORT BACKS

Recording | Presentation

During the Report Backs, each session shared their key accomplishments from the meeting and next steps. You may use the following time stamps to locate a specific report back slide:

1. **01:22**: Regulatory & Ethics Update
2. **05:02**: Data Security Update
3. **06:13**: Large Scale Genomics Update
4. **09:23**: Genomic Knowledge Standards Update
5. **11:57**: Data Use & Researcher Identities Update
6. **14:56**: Discovery Update
7. **19:26**: Cloud Update

8. **21:28**: Clin/Pheno Update
9. **23:19**: Pedigree
10. **25:23**: Phenopackets: Treatment Course + VRS Integration
11. **26:48**: FHIR Implementation Guide
12. **29:02**: Key Management
13. **31:16**: Assurance Levels
14. **33:04**: Machine Readable Consents
15. **36:45**: Beacon V2
16. **39:34**: Future of Variant Data Representation
17. **42:04**: Sequence Annotation
18. **44:03**: htsget: Interactions with Other Projects
19. **49:38**: Building Tooling
20. **53:31**: Federated Analysis Systems Project (FASP)