



Becoming a GA4GH Member

Are you Interested in joining the Global Alliance for Genomics and Health (GA4GH)? You've come to the right place! **In this document you will find information about all of the working groups within GA4GH.** Joining one of our working groups requires you to complete two easy steps:

Step 1

Select which group(s) you would like to join based on the information in this document.



Step 2

Email your group selection(s) to secretariat@ga4gh.org.

After you have completed these steps, you will be added to the relevant group mailing list(s). You will also be given the opportunity to participate in the [GA4GH onboarding program](#). This program pairs you up with an active GA4GH member for a one-on-one conversation that will help to orient you, provide you with helpful resources, and give you an understanding of how you can contribute.

Please note that by participating in GA4GH meetings, attendees agree to adhere to the [GA4GH Standards of Professional Conduct](#).

To ensure that you get the most out of this document, take a moment to review the following terms/concepts:

Terms/Concepts	Description
Work Streams (WS)	GA4GH WSs develop standards and tools that are designed to overcome technical and regulatory hurdles to international genomic data-sharing. Most WS meetings focus on high-level discussions, while in-depth standards development work is conducted within smaller subgroups.
Driver Project (DP)	GA4GH DPs are real-world genomic data initiatives that help guide development efforts and pilot tools. These global stakeholders advocate, mandate, implement, and use GA4GH frameworks and standards in their local contexts.
Work Stream Members	Any individual who signs up to a GA4GH WS Google Group. WS members are encouraged, but not required, to join in on GA4GH discussions both asynchronously via email or slack as well as during virtual face-to-face meetings.



Work Stream Contributors	Any WS member who has also made a meaningful contribution to a GA4GH standard, policy framework, or other deliverable. All GA4GH contributors are expected to provide meaningful, documentable input on at least one GA4GH standard, policy framework, publication, or another deliverable.
Work Stream Managers	Each WS has one manager. This individual is an employed member of the GA4GH secretariat. They assist with WS-related coordination and project management while ensuring that the activities of the WS align with the goals of GA4GH.
Work Stream Leads	Each WS has one to two leads. Leads are contributors within the community and are responsible for achieving the high-level goals and vision of their respective WS.

Now that you have an understanding of GA4GH terms/concepts, see below for descriptions and meetings schedules of all current working groups within GA4GH.



Active Work Stream and Subgroup Meetings

Last updated February 1st, 2022

Clinical and Phenotypic Data Capture Work Stream

- [Pedigree Subgroup](#)
- [Phenopackets Subgroup](#)
- [Cohort Subgroup](#)
 - [Cohort - Pheno](#)
 - [Cohort - MIACC](#)

Discovery Work Stream

- [Data Connect/Networks Subgroup](#)
- [Beacon API Subgroup](#)

Data Security Work Stream

- [Passports/ Authorization and Authentication Infrastructure \(AAI\) Subgroup](#)

Large Scale Genomics (LSG) Work Stream

- [Future of VCF Subgroup](#)
- [Htsget API Subgroup](#)
- [RNASeq Subgroup](#)
- [Refget API Subgroup](#)
- [File Formats Subgroup](#)

Genomics Knowledge Standards Work Stream

- [VCF/VRS Alignment Subgroup](#)
- [Sequence Annotation Subgroup](#)
- [Variation Representation Subgroup](#)
- [Variant Annotation Subgroup](#)

Data Use & Researcher Identities (DURI) Work Stream

- [Data Use Ontology \(DUO\) Subgroup](#)
- [Passports Subgroup](#)

Cloud Work Stream

- [Workflow Execution Service \(WES\) Subgroup](#)
- [Task Execution Schema \(TES\) Subgroup](#)
- [Data Repository Service \(DRS\) Subgroup](#)
- [Tool Registry Service \(TRS\) Subgroup](#)

Regulatory and Ethics Work Stream

- [Data Access Committee Review Standards \(DACReS\) Subgroup](#)
- [Bias in AI/ML Subgroup](#)
- [Genetic Discrimination Subgroup](#)
- [Consent Working Group](#)



Clinical and Phenotypic Data Capture Work Stream

Supports the clinical adoption of genomics by developing standard ontologies, best practices, and information models to represent and exchange clinical data and phenotype information for use in genomic medicine and research.

Work Stream Meeting (All members)	
At the main Clin/Pheno Work Stream meeting, you will be updated about the work being conducted in the subgroups listed below. The full WS meeting also poses an opportunity for discussion of new subgroups/working groups that are needed.	Meeting Schedule: Once every two months on Wednesdays at 20:00 UTC
Pedigree Subgroup	
The group aims to standardize the way systems represent family health history and pedigree information. This will allow patients and care providers to share this information more easily between healthcare systems and help software tools to use this information to improve genome analysis and diagnosis.	Meeting Schedule: Biweekly on Thursdays at 20:00 UTC
Phenopackets Subgroup	
The group focuses on the development of the Phenopacket standard. The Phenopacket Schema represents an open standard for sharing disease and phenotype information to improve our ability to understand, diagnose, and treat both rare and common diseases. A Phenopacket links detailed phenotype descriptions with disease, patient, and genetic information, enabling clinicians, biologists, and disease and drug researchers to build more complete models of disease.	No longer meeting for now as the v2 standard has been approved. Any topics/presentation related to Phenopackets will be discussed in the Full WS meetings.
Computable Cohort Representation Subgroup	
This subgroup aims to create a computable cohort representation standard to computationally describe a cohort and execute the cohort definition as a query across a Cohort Beacon network. This group has not yet formally proposed a deliverable, and it is still	Meeting Schedule: N/A



<p>exploring potential outputs.</p> <p>The group is further divided into the MIACC group and the Pheno group. Thus, this larger cohort group acts as a venue for those two groups to report back with updates. You will find details for the Pheno and MIACC group below.</p>	
Cohort - Pheno	
<p>This subgroup is working to capture the deep phenotypic characteristics as standard definition to execute against a baseline population to extract counts and statistics of the subjects within and across datasets.</p>	Meeting Schedule: Ad Hoc
Cohort - MIACC	
<p>The MIACC (Minimum Information About Computable Cohort) subgroup is working to define the sets of attributes required to describe a cohort consistently (eg. age, biological sex).</p>	Meeting Schedule: Ad Hoc

Discovery Work Stream

Creates a unified data discovery platform to accelerate genomics and clinical data access and utilization around the world.

Work Stream Meeting (All members)	
<p>At the main Discovery Work Stream meeting, you will be updated about the work being conducted in the subgroups listed below. The full WS meeting also poses an opportunity for discussion of new subgroups/working groups that are needed. Key strategic goals are discussed between group leads at this meeting.</p>	Meeting Schedule: Monthly on the second Wednesday at 15:30 UTC
Data Connect/Networks Subgroup	
<p>This subgroup focuses on the ongoing development and potential new implementations of the Data Connect</p>	Meeting Schedule: Biweekly on Wednesdays at 16:00 UTC



<p>API which was approved by the GA4GH Steering Committee in April of 2021. The Data Connect API is a standard for the discovery and search of biomedical data. It provides a mechanism for 1) describing data and its data models and 2) searching the data with the given data model.</p>	
Beacon API Subgroup	
<p>This meeting focuses on the development of the Beacon API. This API (sometimes extended with a user interface) allows for data discovery of genomic and phenoclinic data.</p>	Meeting Schedule: Biweekly on Tuesdays at 15:00 UTC

Data Security Work Stream

Creates technology standards and best practices for protecting data and services consistent with the GA4GH policy framework, including the development, customization, and adoption of standards for identity management, data security, privacy protection, and service assurance.

Work Stream Meeting (All members)	
<p>At the main Data Security Work Stream meeting, you will be updated about the work being conducted in the subgroups listed below. The full WS meeting also poses an opportunity for discussion of new subgroups/working groups that are needed.</p>	Meeting Schedule: Ad hoc
Passports/ Authorization and Authentication Infrastructure (AAI) Subgroup	
<p>The Passports/AAI Technical Working subgroup is a collaboration between the Data Security and DURi Work Streams. This group was formed with the aim of creating a small technical task force to discuss outstanding technical issues regarding v1.0 and migrating to a new version of the GA4GH Passports Standard named v1.2. AAI in security and DURi Passports are closely linked and therefore bring together those working in both areas.</p>	Meeting Schedule: for 2021 and 2022 this has been a weekly meeting between a small technical subgroup and then progress will be communicated out to the DURi Passports group and in a specially arranged workshop



Large Scale Genomics Work Stream

Creates standardized methods for accessing large-scale genomic data (reads, variants, and expression data) by file-based, API-based, cloud-based, and distributed access.

Work Stream Meeting (Leadership only)	
At the main LSG Work Stream meeting, leadership will meet to discuss the work being conducted in the subgroups listed below. The full WS meeting also poses an opportunity for discussion of new subgroups/working groups that are needed.	Once every 3 months
Future of VCF Subgroup	
This project focuses on improvements to the VCF format as well as looking into potential future formats to handle the increasing scale of large sequencing projects more efficiently.	Meeting Schedule: Every 4 weeks on Monday alternating between 14:00 UTC and 21:00 UTC
Htsget API Subgroup	
htsget is a protocol for secure, efficient and reliable access to sequencing read and variation data. This project describes ongoing improvements to the protocol and implementations to support 'two-dimensional' slicing of very large variant (VCF) datasets.	Meeting Schedule: Every 4 weeks on either Tuesday or Wednesday at 23:00 UTC and 17:00 UTC respectively
RNASeq Subgroup	
This subgroup focuses on the development of the RNAget API. This API describes a common set of endpoints for search and retrieval of processed RNA data. This currently includes feature level expression data from RNA-Seq type assays and signal data over a range of bases from ChIP-seq, methylation or similar epigenetic experiments.	Meeting Schedule: Biweekly on Wednesdays at 15:00 UTC
Refget API Subgroup	
This subgroup focuses on the development of the refget API. The refget API enables access to reference genomic	Meeting Schedule: Biweekly on Wednesdays at 15:00 UTC



sequences without ambiguity from different databases and servers using a checksum identifier based on the sequence content itself. Currently, refget API focuses on single reference sequences uniquely identified by their checksums.	
File Formats Subgroup	
This subgroup focuses on the SAM/BAM/CRAM file formats. It describes ongoing additions to SAM/BAM to improve efficiency and support new sequencing platforms. For CRAM the primary focus is on improved efficiency, random access for long reads and CRAM file sizes for short read data (with an expectation of 15-20% reduction). Additionally improved indexing, and especially index documentation.	Meeting Schedule: Every 4 weeks on Tuesdays alternating between 16:00 UTC and 21:00 UTC

Genomics Knowledge Standards Work Stream

Develops, adopts, and adapts standards-based components to enable the exchange of reference genomic information through common APIs, thereby enabling the downstream analysis of genomic data.

VCF/VRS Alignment Subgroup	
This meeting is to create alignment between the Variant Representation Schema and the Future of VCF teams from the GKS and LSG Work Streams respectively.	Meeting Schedule: Monthly on Wednesday at 20:000 UTC
Sequence Annotation Subgroup	
The Sequence Annotation subgroup is developing a common model describing annotated genomic features such as transcripts and genes for use by others within GA4GH and in the general community.	Meeting Schedule: First and third Thursday of every month at 15:00 UTC
Variation Representation (VR) Subgroup	
This subgroup discussed the VR specification - an extensible model for computational variation	Meeting Schedule: Biweekly on Mondays at 17:00 UTC



representation. The specification includes a data model, message schema, and methods for normalization of variants and computing identifiers.	
Variant Annotation Subgroup	
The VA Specification (VA Spec) will define extensible data models to support representation of diverse kinds of statements made about genetic variation, and the evidence and provenance supporting these statements. The specification will include information models, message exchange schema, and a formal framework for defining custom extensions to the core model. A more detailed description of these components can be found here .	Meeting Schedule: Every Wednesday at 16:00 UTC

Data Use & Researcher Identities Work Stream

Facilitates and enables the harmonization of researcher identities by defining who is a bona fide researcher and one or more identity providers that respects this definition and can provide a portable electronic identity. DURi will also generate a data use ontology that will be used to both state use restrictions as well as researchers' purposes.

Work Stream Meeting (All members)	
At the main DURi Work Stream meeting, you will be updated about the work being conducted in the subgroups listed below. The full WS meeting also poses an opportunity for discussion of new subgroups/working groups that are needed. We also discuss new strategic aims for DURi including Roadmaps and seek feedback from any implementers or those with use cases that could inform DURi work and standard development.	Meeting Schedule: Once every two months on Wednesdays at 20:00 UTC (3pm ET/8pm GMT/
Data Use Ontology (DUO) Subgroup	
This subgroup focuses on the (DUO) standard. DUO allows semantic tagging of datasets with restrictions and permissions about their usage, making them automatically discoverable based on the users' authorization level of users, or their intended data uses. This team focuses on developing a data use ontology to help determine if a	Meeting Schedule: Meetings temporarily paused until the new phase of DUO activities begin again.



researcher may be granted access to a particular data set.	
Passports Subgroup	
GA4GH Passport specification aims to support data access policies and procedures within current and evolving data access governance systems. This specification defines Passports and Passport Visas as a standardized method of communicating the data access authorizations that a user has based on either their role (e.g. researcher), affiliation, or access status.	Meeting Schedule: Every 6 weeks on Wednesdays at 15:00 UTC

Cloud Work Stream

Helps GA4GH Driver Projects take full advantage of modern cloud environments. Its initial focus is on ‘bringing the algorithms to the data’, by creating standards for defining, sharing, and executing portable workflows. Standards under discussion include workflow definition languages, tool encapsulation, cloud-based task and workflow execution, and cloud-agnostic abstraction of data access.

Work Stream Meeting (All members)	
Meeting Description: At the main Cloud Work Stream meeting, you will be updated about the work being conducted in the subgroups listed below. The full WS meeting also poses an opportunity for discussion of new subgroups/working groups that are needed.	Meeting Schedule: On the second and fourth Mondays of the month alternating between 18:00 UTC and 22:00 UTC respectively.
Workflow Execution Service (WES) Subgroup	
The WES API describes a standard programmatic way to run and manage workflows. Having this standard API supported by multiple execution engines will let people run the same workflow using various execution platforms running on various clouds/environments.	Part of full Cloud WS meeting
Task Execution Schema (TES) Subgroup	
The TES is a simple standardized interface for requesting	Part of full Cloud WS meeting



long running asynchronous compute jobs.	
Data Repository Service (DRS) Subgroup	
The DRS provides a mechanism to read (and in the future, write) data objects across object stores in a cloud-agnostic way.	Part of full Cloud WS meeting
Tool Registry Service (TRS) Subgroup	
The TRS API is one of a series of technical standards from the Cloud Work Stream that together allow genomics researchers to bring algorithms to datasets in disparate cloud environments, rather than moving data around.	Part of full Cloud WS meeting

Regulatory and Ethics Work Stream

Focuses on the ethical, legal and social implications of international data sharing. Building on a human rights framework, the Regulatory and Ethics Work Stream (REWS) aims to provide harmonized forward-looking consent and privacy policies, and anticipatory data governance models. If you have any questions about Regulatory and Ethics Work Stream, contact rews-coordinator@ga4gh.org.

Work Stream Meeting (All members)	
At the main REWS General meeting, you will be updated about the work being conducted in the subgroups listed below. Occasionally, members will participate in a voting process to approve policies and frameworks. The full WS meeting also poses an opportunity for discussion of new subgroups/working groups that are needed.	Meeting Schedule: Once every 1-2 months
Data Access Committee Review Standards (DACReS) Subgroup	
The DACReS work group leverages their expertise in bioethics, data privacy law, data science, and genomics to empirically investigate challenges and opportunities within Data Access Committees (via surveys and interviews) to develop an online, evidence-based Toolkit (“DACReS Toolkit”). We will build this resource through an	Meeting Schedule: Ad hoc



international, consensus process with DACReS members during meetings and stakeholder workshops.	
Genetic Discrimination (GD) Subgroup	
The GD group within REWS is a collaborative effort supporting the work of the Genetic Discrimination Observatory (GDO). The subgroup aims to create specific tools that can be used by the community to combat genetic discrimination and provides contributions to the ongoing work at the GDO.	Meeting Schedule: Ad hoc
Consent Task Force	
This group focuses on the collection, analysis, typology and harmonization of consent clauses from the different jurisdictions of the international genomic community. The goal is to provide sampled language to researchers and clinicians for clinical genetic testing, biobanking and population studies, genomic research, familial genomics and pediatrics.	Meeting Schedule: Ad hoc
General Data Protection Regulation (GDPR) Forum	
The forum on the GDPR (EU) 2016/679 was launched in August 2018 as part of the REWS. The purpose of the forum is to discuss and disseminate information regarding European data protection law – specifically the GDPR – and its impact on international genomic and health-related data sharing. Monthly one-page GDPR briefs provide concise and timely guidance on specific topics relevant to health researchers, institutions, ethics bodies, and data sharing initiatives internationally.	Meeting Schedule: Ad Hoc