

# GA4GH 7th Plenary Meeting Report

---

## TABLE OF CONTENTS

1. [INTRODUCTION](#)
2. [OCTOBER 21](#)
  - a. [Plenary Meeting Welcome](#)
  - b. [Opening Remarks](#)
3. [OCTOBER 22](#)
  - a. [Breakout Sessions](#)
    - i. [Introduction to GA4GH](#)
    - ii. [Collaborative Genomics Standards Development](#)
    - iii. [Data Access and Usage](#)
    - iv. [GenoPri](#)
    - v. [National Initiatives](#)
    - vi. [ELIXIR::GA4GH Strategic Partnership: How We're Working Together](#)
    - vii. [Patient Perspectives](#)
    - viii. [The Future of Technology in Genomics](#)
  - b. [Keynote: TransFAT: Translating Fairness, Accountability, and Transparency into Data Science Practice \(Julia Stoyanovich\)](#)
  - c. [GA4GH Updates](#)
    - i. [Overview and 2020 Strategic Roadmap](#)
    - ii. [Genomic Data APIs](#)
    - iii. [htsget API V1](#)
    - iv. [refget API V1](#)
    - v. [RNAget](#)
    - vi. [Joint Q & A](#)
    - vii. [File Formats](#)
    - viii. [Crypt4GH](#)
    - ix. [CRAM](#)
    - x. [Future of VCF](#)
    - xi. [Joint Q & A](#)
    - xii. [Cloud APIs](#)
    - xiii. [Tool Registry Service API](#)
    - xiv. [Workflow Execution Service API](#)
    - xv. [Data Repository Service API / Workflow Execution Service API](#)
    - xvi. [Joint Q & A](#)
    - xvii. [Data Use & Research Identities \(DURI\)](#)
    - xviii. [DURI Overview](#)
    - xix. [Data Use Ontology V1](#)
    - xx. [GA4GH Passports V1](#)
    - xxi. [Authentication & Authorization Infrastructure V1](#)
    - xxii. [Joint Q & A](#)



- xxiii. [Clinical and Genomic Data Sharing](#)
- xxiv. [Variation Representation V1](#)
- xxv. [Phenopackets V1](#)
- xxvi. [Beacon API V1](#)
- xxvii. [Joint Q & A](#)
- xxviii. [Ethics, Privacy, and Security](#)
- xxix. [Data Security Infrastructure Policy V4](#)
- xxx. [Regulatory & Ethics Policy Framework](#)
- xxxi. [GA4GH Copyright & IP Policy](#)
- xxxii. [Joint Q & A](#)

#### **4. OCTOBER 23**

- a. [Yakup's Journey to Hope](#)
- b. [Keynote: Advancing Medicine through Data Sharing and Collaboration on a Global Scale \(Heidi Rehm\)](#)
- c. [Panel: Collaborating Across Traditional Boundaries](#)
  - i. [Overview](#)
  - ii. [Partnerships from a WSI / Wellcome Genome Campus Perspective](#)
  - iii. [Open Targets Initiative and Building Public-Private Partnerships](#)
  - iv. [Infectious Diseases and Data Sharing During Outbreaks](#)
  - v. [Joint Q & A](#)
- d. [Panel: Diversity in Genomics](#)
  - i. [Overview](#)
  - ii. [Novel Methods to Engage Inaccessible and Disadvantaged Communities](#)
  - iii. [Clinical Use of Current Polygenic Risk Scores and Need for Diversity in the Genetic Pool](#)
  - iv. [Ethics of AI](#)
  - v. [Joint Q & A](#)
- e. [Panel: Future of Genomic and Health Data](#)
  - i. [Overview](#)
  - ii. [Protecting Privacy and Security of Genomic Data](#)
  - iii. [Building Standard Interpretation Data Alliance Using Blockchain and Modern Cryptology](#)
  - iv. [Advancing Health Data Interoperability](#)
  - v. [Joint Q & A](#)
- f. [Keynote: The Power of Data Sharing in Genomics and Beyond \(Eric Lander\)](#)

#### **5. ACKNOWLEDGEMENTS**

- a. [7th Plenary Programme Committee](#)
- b. [Funding Partner Recognition](#)
  - i. [Host Institutions](#)
  - ii. [Core Funders](#)
  - iii. [Gold-level Funding Partners](#)
  - iv. [Silver-level Funding Partners](#)
  - v. [Bronze-level Funding Partners](#)

## INTRODUCTION

The GA4GH 7th Plenary Meeting in Boston, USA brought together more than 450 individuals representing more than 250 organizations across 27 countries for two days of keynotes, talks, updates, and workshops focused on advancing development work for the immediate data sharing needs of the genomics and health community. The live stream had over 140 unique views across all live streamed plenary sessions. The GA4GH 7th Plenary Meeting was covered by [Bio-IT World](#) and by [GenomeWeb](#).

## OCTOBER 21

The GA4GH 7th Plenary Meeting opened with brief remarks from current GA4GH Chair Ewan Birney and founding GA4GH Chair David Altshuler, now Executive Vice President of Global Research and Chief Scientific Officer at Vertex Pharmaceuticals.

### Plenary Meeting Welcome

*Ewan Birney (EMBL-EBI)*

[Presentation video](#)

Birney identified four goals for the 7th Plenary meeting: to share updates on GA4GH technical standards and policy frameworks, to support the adoption and use of GA4GH deliverables, to celebrate initiatives external to GA4GH, and to coalesce the GA4GH community and the larger genomics and health communities. He also unveiled the GA4GH standards approved in 2019.

[Download Slides](#)

### Opening Remarks

*David Altshuler (Vertex Pharmaceuticals)*

[Presentation video](#)

Altshuler, the GA4GH Founding Chair, discussed the history of GA4GH and its progress as an organization, as well as his vision for the future of GA4GH. Altshuler began by describing the day—January 28, 2013—that 50 people from eight countries met for the first time in New York, USA to discuss the upcoming mass of genomic and health data that would be collected across many stakeholders. Looking forward, Altshuler hopes that GA4GH and the greater genomics and health data communities will approach genomics and health data sharing in the clinical context, address the diversity problem in genomics today, and build public trust in data sharing.

## OCTOBER 22

Day 2 of 7th Plenary began with a wide variety of breakout sessions and workshops. Julia Stoyanovich of New York University gave the day's keynote presentation and the audience learned about several new and upcoming GA4GH standards through a "firehose" session in the afternoon.

### Breakout Sessions

#### Introduction to GA4GH

*Peter Goodhand (OICR, GA4GH), Adrian Thorogood (McGill University, GA4GH), Melissa Konopko (Wellcome Sanger Institute, GA4GH), Rishi Nag (EMBL-EBI, GA4GH), and Lindsay Smith (OICR, GA4GH)*

[Session video](#)

GA4GH Staff provided an overview of the organization, addressing the GA4GH mission, ecosystem, milestones, and key concepts. Panelists explained how the GA4GH Connect Strategic Roadmap is guided by the needs of real-world initiatives in genomics and health ("Driver Projects") and achieved through the contributions of Foundational and Technical Work Streams. Each of the GA4GH Work Stream Managers detailed the roles, responsibilities, and accomplishments of their workstreams. The session concluded with audience questions answered by GA4GH Staff.

[Download Slides](#)

#### Collaborative Genomics Standards Development

*Robert Freimuth (Mayo Clinic), Adam Berger (NIH), Bron Kisler (NIH), Kim Pruitt (NIH), Soichi Ogishima (Tohoku University)*

Speakers discussed the benefits of creating and using global standards in genomics and healthcare, as well as some key challenges involved in genomics and health standards development and adoption, such as alignment with national and international standards organizations, standardizing both tangible (e.g., biosamples) and intangible (e.g., genomic data) objects, and ensuring data is FAIR (findable, accessible, interoperable, reusable). Speakers provided examples of standards development and adoption within the United States National Institutes of Health (NIH) and the GENome Medical alliance Japan (GEM Japan), and cross-standards development organization (xSDO) collaboration between GA4GH, ISO, and HL7.

[Download Slides](#)

## Data Access and Usage

*Sarion Bowers (Wellcome Sanger Institute), Tommi Nyrönen (CSC-IT Center for Science), Melanie Courtot (EMBL-EBI), Jonathan Dursi (Hospital for Sick Children), Maarten Kremers (SURFnet)*

Speakers discussed data access and usage concepts in the context of standards and infrastructures for data access and restriction. The panel provided an introduction to data access and usage concepts, including federated identity management, federated authentication, and interederation. They explained the vision behind the GA4GH Data Use Ontology (DUO) standard, which can determine data access restrictions for researcher queries. Panelists also outlined infrastructures, such as CanDIG, that provide secure, distributed authentication and authorization to query and access data across geographic and governmental boundaries.

[Download Slides](#)

## GenoPri

*Program Co-chairs: Xiaoqian Jiang (UTHealth), Madeleine Murtagh (University of Newcastle)*

Speakers during this two-day, externally-organized workshop discussed topics including: the roles of data users and data providers in data privacy and security; global privacy regulations for genomic and health data; the push for an international data protection regulatory framework; automated data access and researcher identification; models for preserving privacy of genomic and health data from de-identification to data access queries; tools for secure and private sharing, processing, and analyzing of sensitive data; dynamic control of data privacy rights; data encryption methods; and a reflection on the history of genomic data privacy.

Speakers and session abstracts available at [genopri.org](http://genopri.org).

## National Initiatives

*Kathryn North, Zornitza Stark, Clara Gaff, and Oliver Hoffman (Australian Genomics), Per Sikora (Genomic Medicine Sweden), Hidewaki Nakagawa (GEM Japan), Mavis Machirori and Madeleine Murtagh (University of Newcastle), Heidi Rehm (ClinGen, Matchmaker Exchange), Ellen McDonagh and Augusto Rendon (Genomics England (GEL)), Sergi Beltran (ELIXIR, EJP-RD), David Glazer (NIH All of Us Research Program), Sabine Oesterle (Swiss Personalized Health Network (SPHN))*

In its fifth annual educational meeting, the GA4GH National Initiatives group brainstormed common needs for standards across their organizations, continued to identify areas necessitating collaboration and sharing of resources and expertise, discussed the development of a National Initiatives Toolkit to share available resources, such as minimal clinical datasets and consent

forms, and discussed plans to formalize the National Initiatives consortium within GA4GH. During the session, the group also presented accomplishments to date.

[Download Slides](#)

### **ELIXIR::GA4GH Strategic Partnership: How We're Working Together**

*Mikael Linden (CSC - IT Center for Science), Juha Törnroos (ELIXIR Beacon), Gary Saunders (ELIXIR Europe), Melanie Courtot (EMBL-EBI), Alexander Kanitz (University of Basel), Dylan Spalding (EMBL-EBI), Jordi Rambla (FUNDACIO CENTRE DE REGULACIO GENOMICA), Susheel Varma (EMBL-EBI)*

GA4GH has collaborated with ELIXIR, the European research infrastructure for life sciences, since 2014. This collaboration has developed into a Strategic Partnership, established in 2019 to help coordinate standards development and European engagement in GA4GH. The session focused on demonstrating implementations of GA4GH standards across the ELIXIR network and at national nodes, discussing the ELIXIR collaboration with other GA4GH Driver Projects to advance standards development and highlighting the facilitation of interoperable standards from various GA4GH Work Streams.

[Download Slides](#)

### **Patient Perspectives**

*Ada Hamosh (Johns Hopkins University), Tiffany Boughtwood (Australian Genomics), Illona Schelle (Inspire2Live), Anna Middleton (Wellcome Genome Campus), Richard Milne (Wellcome Genome Campus), Sonia Vallabh (Broad Institute), Erik Minikel (Broad Institute)*

Panelists discussed the importance of including patient perspectives in health research. The Your DNA, Your Say survey results (from 37,000 participants across 22 countries) were introduced, elucidating the roles that patient trust and understanding play in health data sharing and research. For example, two panelists' experiences with rare genetic prion disease caused them to change careers, now working as research scientists aiming to understand, prevent, and treat their family's disease.

[Download Slides](#)

### **The Future of Technology in Genomics**

*Andy Yates (EMBL-EBI), Moran Cabili (Foundation Medicine), Sumit Jamuar (Global Gene Corp), Patrick Kemmermen (Princess Maxima Center), Marco Roos (Leiden University Medical Center), Barend Mons (Leiden University Medical Center)*

### [Session video](#)

Panelists discussed the role of modern technology in genomics and medicine at various institutions around the globe. Speakers noted the increasing presence of technology in genomics data from collection and storage, to sharing and accessing data in the future and where they expect future technological advancements in the coming years, including machine learning and artificial intelligence (AI), user interfaces (UI) for genomic and health data providers, and analytics platforms for researchers.

### [Download Slides](#)

## **Keynote: TransFAT: Translating Fairness, Accountability, and Transparency into Data Science Practice**

*Julia Stoyanovich (New York University)*

### [Keynote video](#)

Stoyanovich discussed ways in which data science can reveal bias in many facets of life—from the criminal justice system, to employment opportunities, to same-day Amazon delivery availability. However, as Stoyanovich notes, the data itself cannot always reveal where the bias originates or why it is present. Stoyanovich encouraged attendees to think about how applications of data science can affect other people or groups, to be transparent with the algorithms used, and to provide a summary of the means for collecting and processing data or to release the data itself when possible.

### [Download Slides](#)

## **GA4GH Updates**

### **Overview and 2020 Strategic Roadmap**

*Ewan Birney (EMBL-EBI, GA4GH), Heidi Rehm (Massachusetts General Hospital, Broad Institute, GA4GH)*

### [Session video](#)

Speakers introduced the GA4GH Product Approval Process, explaining the path of a deliverable from prototype, to proposal, submission, and eventually approval as a GA4GH Standard. GA4GH Driver Projects guide product development by the GA4GH Work Stream Contributors through dynamic engagement with partners to encourage contribution and eventual adoption. As a co-lead of the GA4GH Gap Analysis effort, with Andrew Morris, Rehm outlined the planned updates to the 2020 Strategic Roadmap which will integrate findings from consultations with

Driver Projects and gaps recognized by contributors. The overview concluded with an animation depicting a brief description of GA4GH as the global genomics standards organization.

[Download Slides](#)

## **Genomic Data APIs**

htsget API V1

*Mike Lin (mlin.net LLC)*

[Presentation video](#)

The htsget API is a GA4GH Approved Standard web protocol for serving high-throughput sequencing datasets, such as BAM, CRAM, and VCF. The speaker explained htsget as analogous to a bridge, connecting file servers and web APIs. The protocol begins with an initial request (GET), followed by a response ticket (JSON), fetching data (GET), and concatenating these to a final result. Htsget enables retrieval of sensitive data and can authenticate and authorize users.

[Download Slides](#)

refget API V1

*Andy Yates (EMBL-EBI)*

[Presentation video](#)

The refget API is a GA4GH Approved Standard for accessing reference genomic sequences without ambiguity from various databases and servers. The API can be used in aligning sequencing reads, calling variations, annotating genes, and interpreting variations for reliable, reproducible genomic analysis. Refget is compatible with CRAM, the genomics compression standard, and has a forthcoming serverless AWS implementation without a database.

[Download Slides](#)

RNAget

*Sean Upchurch (California Institute of Technology)*

[Presentation video](#)

The RNAget API is a forthcoming product for retrieving RNA-derived data at scale. The API creates a matrix from slices of samples or features in an RNA data matrix, and works with both expression and continuous RNA data. RNAget harmonizes existing APIs, such as ENCODE and



the Expression Atlas, to increase interoperability and scalability, reduce deployment efforts, and anticipate emerging user needs.

[Download Slides](#)

Joint Q & A

*Mike Lin (mlin.net LLC), Andy Yates (EMBL-EBI), Sean Upchurch (California Institute of Technology)*

[Q&A video](#)

**File Formats**

Crypt4GH

*Alexander Senf (EMBL-EBI)*

[Presentation video](#)

Crypt4GH is a GA4GH Approved Standard file format that eliminates the need to decrypt data for use, and allows data to remain encrypted throughout transport and storage. Crypt4GH uses personalized encryption with an encrypted header for sharing, which mirrors envelope encryption and saves storage space. The file format also uses symmetric encryption for the data itself, in a way that leaves encrypted data byte-position compatible with original data and indexes.

[Download Slides](#)

CRAM

*Rob Davies (Wellcome Sanger Institute)*

[Presentation video](#)

CRAM is a GA4GH Approved Standard file format for compressed data storage and retrieval. CRAM is an open source file format which is free to use. It is compatible with SAM and BAM workflows, but is faster to write and 40-70% smaller than BAM and saves on storage costs compared to other file formats. The CRAM format is used by many genomic data programs around the world, and has benefits for both large and small scale data storage.

[Download Slides](#)

## Future of VCF

*Yossi Farjoun (Broad Institute)*

### [Presentation video](#)

There is an increasingly large amount of genomic data, which means variation callsets are growing rapidly in size, and are mostly stored in incompressible arrays. Different ideas for approaching these problems exist, but necessitate a data-driven decision. Variant Call Format (VCF) is a text file format for representing genetic variation data composed of alleles or genotypes.

### [Download Slides](#)

## Joint Q & A

*Alexander Senf (EMBL-EBI), Rob Davies (Wellcome Sanger Institute), Yossi Farjoun (Broad Institute)*

### [Q&A video](#)

## Cloud APIs

### Cloud Work Stream Overview

*David Glazer (Verily)*

### [Presentation video](#)

The GA4GH Cloud Work Stream is creating a suite of interoperable APIs to “bring the algorithms to the data” with the goal of helping the genomics and health communities take advantage of modern cloud environments. The Work Stream creates standards for defining, sharing, and executing portable workflows.

### Tool Registry Service API

*Denis Yuen (Ontario Institute for Cancer Research)*

### [Presentation video](#)

The Tool Registry Service (TRS) API is a GA4GH Approved Standard for listing and describing available tools for exchanging, indexing, and searching both standalone containerized tools and workflows with multiple tools, described with workflow languages (e.g., CWL, WDL, Nextflow). The API currently has implementations through Dockstore, Biocontainers, and Agora.

### [Download Slides](#)

Workflow Execution Service API

*Brian O'Connor (University of California Santa Cruz)*

### [Presentation video](#)

The Workflow Execution Service (WES) API is a GA4GH Approved Standard for sending workflows to compute environments, bringing algorithms to data. The developers described plans to continue encouraging implementations of WES, generate a WES registry, and enhance the API to improve WES and (forthcoming) DRS integration

### [Download Slides](#)

Data Repository Service API / Workflow Execution Service API

*Brian O'Connor (University of California Santa Cruz)*

### [Presentation video](#)

The Data Repository Service (DRS) API is a forthcoming standard for looking up data objects and fetching their content via multiple protocols. DRS currently has live implementations on servers such as Terra and Seven Bridges, as well as the Swagger/Open API Client, and planned implementations on many other servers and clients. Future plans for DRS include generating a DRS registry, developing a comprehensive testbed, and promoting implementation and use throughout the community.

### [Download Slides](#)

Joint Q & A

*David Glazer (Verily), Brian O'Connor (UCSC), Denis Yuen (OICR)*

### [Q&A video](#)

## **Data Use & Research Identities (DURI)**

DURI Overview

*Tommi Nyrönen (ELIXIR)*

### [Presentation video](#)

The GA4GH DURl Work Stream is composed of two parts: Data Use and Researcher Identities. The Data Use Ontology, a GA4GH Approved Standard, satisfies the data use portion of the workstream, while the proposed GA4GH Passports standard satisfies the researcher identities portion. The speaker also shared an animation about federated human data access in Europe.

[Download Slides](#)

Data Use Ontology V1

*Melanie Courtot (EMBL-EBI), Jonathan Lawson (Broad Institute)*

[Presentation video](#)

The GA4GH Approved Standard Data Use Ontology (DUO) allows data depositors to unambiguously tag their data use with limitations, allows researchers to make discovery and access requests, and allows data access committees to triage data access requests faster. DUO is currently implemented by eight institutions and counting, and future updates to DUO involve providing DUO-compliant consent clauses for Institutional Review Boards and other collaborations with the GA4GH Regulatory and Ethics Work Stream.

[Download Slides](#)

GA4GH Passports V1

*Craig Voisin (Google Inc.), Stephanie Dyke (McGill University)*

[Presentation video](#)

The forthcoming [GA4GH Passports standard](#) enables a standardized approach to authenticate and authorize researchers for access to data using GA4GH Passports and Visas. The specification provides guidance on encoding [Registered Access](#) to data and was developed in collaboration with the forthcoming Authentication and Authorization Infrastructure standard.

[Download Slides](#)

Authentication & Authorization Infrastructure V1

*David Bernick (Broad Institute)*

[Presentation video](#)

The forthcoming GA4GH Authentication & Authorization Infrastructure (AAI) standard works hand-in-hand with GA4GH Passports to allow organizations to pass on proof of access permissions from other organizations in a secure way. Tokens and identity/authorization are

packaged to provide for downstream trust between institutions without a centralized authority, enhancing cross-institutional sharing without compromising security. The AAI has been implemented with ELIXIR AAI and Beacon, and Google Identity Concentrator (IC) and Data Access Manager (DAM).

### [Download Slides](#)

Joint Q & A

*Craig Voisin (Google Inc.), Stephanie Dyke (McGill University), Melanie Courtot (EMBL-EBI), Jonathan Lawson (Broad Institute), David Bernick (Broad Institute)*

### [Q&A video](#)

## **Clinical and Genomic Data Sharing**

Variation Representation V1

*Reece Hart (Reece Hart Consulting)*

### [Presentation video](#)

The GA4GH Approved Standard Variation Representation (VR) aims to standardize the definition, sharing, and identification of biological variation within and between systems. The standard has been implemented and adopted by the BRCA Exchange, ClinGen Allele Registry, and VICC Meta-knowledgebase.

### [Download Slides](#)

Phenopackets V1

*Melissa Haendel (Monarch Initiative, Oregon State University, Oregon Health & Science University)*

### [Presentation video](#)

The GA4GH Approved Standard Phenopackets is a way to share case-level phenotypic information data spanning clinical and research systems. The standard has been implemented in EUCANCan, NIH UDP, EU SOLVE-RD, PhenoTips, EBI Biosamples, Cafe Variome, AMED BioBank Network, and is useful in national genomic medicine initiatives and other genomic medicine organizations.

### [Download Slides](#)

Beacon API V1

*Jordi Rambla (CNAG-CRG, EGA, ELIXIR)*

[Presentation video](#)

The GA4GH Approved Standard Beacon API can be implemented as a web-accessible service that users may query for information about a specific allele. ELIXIR Beacon and ELIXIR AAI implementations allow data owners to light Beacons at different tiers of data access: public, registered, or controlled. Working with GA4GH Driver Projects, Beacon API developers hope future versions will provide greater information about variants and allow Beacon users to access case data.

[Download Slides](#)

Joint Q & A

*Reece Hart (Reece Hart Consulting), Melissa Haendel (Monarch Initiative, Oregon State University, Oregon Health & Science University), Jordi Rambla (CNAG-CRG, EGA, ELIXIR)*

[Q&A video](#)

## **Ethics, Privacy, and Security**

Data Security Infrastructure Policy V4

*Jean-Pierre Hubaux (EPFL)*

[Presentation video](#)

The recently approved Data Security Infrastructure Policy (DSIP) defines guidelines, best practices, and standards for building and operating the data security infrastructure recommended for stakeholders in the GA4GH community. The DSIP builds on the GA4GH *Framework for Responsible Sharing of Genomic and Health-Related Data* principles and enforces the GA4GH Privacy and Security Policy. As a living document, the DISP will be revised and updated over time in response to changes in the GA4GH Privacy and Security Policy, and as technology and biomedical science continue to advance.

[Download Slides](#)

Regulatory & Ethics Policy Framework

*Bartha Knoppers (McGill University), Madeleine Murtagh (Newcastle University), Adrian Thorogood (GA4GH)*

### [Presentation video](#)

The speakers shared updates on the recent review and re-approval of the GA4GH *Framework for Responsible Sharing of Genomics and Health Related Data*, as well as updates to the GA4GH Consent Policy and the GA4GH Data Privacy and Security Policy. The Framework is available in 13 translations and is founded on the universal human right to share in the benefits of scientific advancement. The Consent Policy is founded on consent being an open, on-going process and aims to promote autonomous decisions while emphasizing international data sharing. The revised policy places greater emphasis on transparency with data subjects. The Data Privacy and Security Policy provides guidance to protect and promote the security, integrity, and availability of data and services, and the privacy of individuals, families, and communities whose data are processed.

### [Download Slides](#)

GA4GH Copyright & IP Policy

*Ray Krasinski (Philips), Jorge Contreras (University of Utah)*

### [Presentation video](#)

The forthcoming GA4GH Copyright & IP Policy is intended to encourage open and collaborative participation in standards-development, under clear expectations, while ensuring that GA4GH has clear rights to adapt and distribute contributions under copyright law. The Policy will also ensure GA4GH Work Products are accessible to the world under clear licenses. If approved, the Policy would immediately bind individual contributors as a condition of participation in standards-development, and will be a condition of becoming a GA4GH Member Organization.

### [Download Slides](#)

Joint Q & A

*Jean-Pierre Hubaux (EPFL), Bartha Knoppers (McGill University), Madeleine Murtagh (Newcastle University), Adrian Thorogood (GA4GH), Ray Krasinski (Philips), Jorge Contreras (University of Utah)*

### [Q&A video](#)

**OCTOBER 23**

## **Yakup's Journey to Hope**

*Sergi Beltran (CNAG)*

[Presentation video](#)

The European Joint Programme on Rare Diseases (EJP RD) is a GA4GH Driver Project that uses genomic and health data sharing to advance researchers' and clinicians' understanding and treatment of rare disease. The EJP RD uses many GA4GH standards to accomplish their goal, as shown in a video in which doctors and researchers from many countries and institutions collaborate and share data to diagnose and treat a child named Yakup with a rare neurogenetic disease.

## **Keynote: Advancing Medicine through Data Sharing and Collaboration on a Global Scale**

*Heidi Rehm (Massachusetts General Hospital, Broad Institute, Harvard Medical School)*

[Keynote video](#)

Rehm discussed the ways in which healthcare and medicine are advanced when researchers and clinicians collaborate and share their data on a global scale. Providing examples of projects aiming to advance medicine through genomics and health data sharing such as GnomAD, ClinGen Resource, and Matchmaker Exchange, Rehm revealed how a lack of data sharing blocks the understanding, diagnosis, and treatment of pathogenic variants, and can misrepresent benign variants as potentially harmful. Rehm mentioned the successes enabled by greater diversity and more widespread sharing of data, while also reminding the audience of the challenges in sharing genomic and health data today, such as underrepresentation of populations, public unwillingness to share data, and inefficiency of current computational infrastructure for large scale variant calling.

## **Panel: Collaborating Across Traditional Boundaries**

### **Overview**

*Kathryn North (Australian Genomics), Julia Wilson (Wellcome Sanger Institute), Ewan Birney on behalf of Ian Dunham (EMBL-EBI), Bronwyn MacInnis (Broad Institute)*

[Panel video](#)

[Download Slides](#)



Panelists discussed the implications of institutional boundaries, for example at the Wellcome Genome Campus, and how institutions can collaborate with other organizations to approach problems. Approaches to supercede boundaries across organizations, fields, geography, culture, and species can prevent problems (e.g., prioritizing targets prevents drug failure in the clinical trial phases) or resolve problems (e.g., applications of pathogen genomics in infectious disease can inform public health outbreak responses).

### **Partnerships from a WSI / Wellcome Genome Campus Perspective**

*Julia Wilson (Wellcome Sanger Institute)*

[Presentation video](#)

[Download Slides](#)

### **Open Targets Initiative and Building Public-Private Partnerships**

*Ewan Birney on behalf of Ian Dunham (EMBL-EBI)*

[Presentation video](#)

[Download Slides](#)

### **Infectious Diseases and Data Sharing During Outbreaks**

*Bronwyn MacInnis (Broad Institute)*

[Presentation video](#)

[Download Slides](#)

### **Joint Q & A**

[Q&A video](#)

### **Panel: Diversity in Genomics**

#### **Overview**

*Madeleine Murtagh (University of Newcastle), Calvin Ho (National University of Singapore), Alicia Martin (Broad Institute), Mavis Machirori (University of Newcastle)*

[Panel video](#)

Panelists discussed the repercussions of limited diversity within genomics and health research and clinical care settings. With poor diversity in genomics and health, deep medicine and AI are likely to further racial and ethnic gaps and inequities in healthcare, and the use of polygenic risk scores and GWAS are increasingly inaccurate trait predictors with low diversity. Panelists concluded that the lack of diverse representation in research, of both participants and staff, will perpetuate poor diagnosis of diverse populations in clinical settings.

### **Novel Methods to Engage Inaccessible and Disadvantaged Communities**

*Calvin Ho (National University of Singapore)*

[Presentation video](#)

[Download Slides](#)

### **Clinical Use of Current Polygenic Risk Scores and Need for Diversity in the Genetic Pool**

*Alicia Martin (Broad Institute)*

[Presentation video](#)

[Download Slides](#)

### **Ethics of AI**

*Mavis Machirori (University of Newcastle)*

[Presentation video](#)

[Download Slides](#)

### **Joint Q & A**

[Q&A video](#)

### **Panel: Future of Genomic and Health Data**

#### **Overview**

*Laszlo Radvanyi (OICR), Meng Yang (BGI Big Data Center), Jean-Pierre Hubaux (EPFL), Teresa Caban Zayas (U.S. Department of Health and Human Services)*

[Panel video](#)

[Download Slides](#)

Panelists discussed challenges of working with big data, with topics covering a broad range including: privacy, security, complexity, storage, public availability, access mediation, ethics, regulation, and standardization. The panelists also posed potential technical, legal, and organizational solutions to these challenges, and shared their visions for the future norms as big data becomes an increasing presence. From electronic health systems, to blockchain based data sharing, to cross-border data alliances, these solutions were proposed from multi-national perspectives including Switzerland, China, and the United States.

### **Protecting Privacy and Security of Genomic Data**

*Jean-Pierre Hubaux (EPFL)*

[Presentation video](#)

[Download Slides](#)

### **Building Standard Interpretation Data Alliance Using Blockchain and Modern Cryptology**

*Meng Yang (BGI Big Data Center)*

[Download Slides](#)

### **Advancing Health Data Interoperability**

*Teresa Caban Zayas (U.S. Department of Health and Human Services)*

[Presentation video](#)

[Download Slides](#)

### **Joint Q & A**

[Q&A video](#)

### **Keynote: The Power of Data Sharing in Genomics and Beyond**

*Eric Lander (Broad Institute)*

[Keynote video](#)

Lander provided an overview history of important events in data sharing, beginning with the Human Genome Project, to the discovery of human disease genetics, and now to the mapping of cancer genetics. Lander explained how progress in genomics and health was facilitated through

human cohorts, low-cost and efficient sequencing technologies, advances in data science, and biobank meta analysis—but ultimately would not have been possible without data sharing.

[Download Slides](#)

## ACKNOWLEDGEMENTS

### 7th Plenary Programme Committee

Thank you to the Programme Committee who made this meeting possible:

- Robert Friemuth, Mayo Clinic, USA (Committee Chair)
- Tommi Nyrönen, CSC – IT Center for Science, Finland
- Heidi Sofia, National Human Genome Research Institute, USA
- Stacey Donnelly, Broad Institute of MIT and Harvard, USA
- Hidewaki Nakagawa, Riken Institute, Japan
- David Siedzik, Broad Institute of MIT and Harvard, USA
- Nicola Mulder, University of Cape Town, South Africa
- Madeleine Murtagh, Newcastle University, United Kingdom
- Sarion Bowers, Wellcome Sanger Institute, United Kingdom
- Jeff Gentry, Foundation Medicine, USA
- Peter Goodhand, Ontario Institute for Cancer Research, Canada

### Funding Partner Recognition

Thank you to the GA4GH 2019 Funding Partners whose support made GA4GH 7th Plenary possible.

### Host Institutions



### Core Funders

Funding for this meeting was made possible in part by a supplement to U41HG006834 from the National Institutes for Health. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention by trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

**Gold-level Funding Partners**



**Silver-level Funding Partners**



**Bronze-level Funding Partners**

