
Global Alliance
for Genomics & Health
Collaborate. Innovate. Accelerate.
GA4GH 2025
Annual Report
The rising tide of genomics
and health data

TABLE OF CONTENTS

Preface

Dear Colleagues,

When we launched GA4GH in 2013, the field was anticipating a flood of genomic data and a critical question: do we have the tools and infrastructure needed to make use of the world's growing amount of data and make precision medicine a reality?

Over the past decade, our community of GA4GH contributors, collaborators, and partners has worked to answer that question, building the technical standards and policy frameworks needed to responsibly store, analyse, and share genomic and related health data at scale. Last year, at the GA4GH 13th Plenary meeting in Uppsala, Sweden, we saw those tools being applied in practice: real organisations, real data, and real impact on patients' lives.

The data is here. The standards and frameworks are here. Now, we are working to extend their impact. That means continued engagement across sectors and around the world. GA4GH's mandate is global, and making our standards and tools work in local contexts is critical to ensuring that all people can benefit from advancements in genomic science.

I want to acknowledge that the past year has been especially challenging, with uncertainties in the global science, economic, and geopolitical landscapes. Even so, we have continued to prioritise convening, coordinating, and collaborating as a global genomics community. Thank you to everyone who has contributed to GA4GH. I look forward to continuing our work together to drive global impact in 2026 and beyond.

Best regards,

Peter Goodhand
CEO, GA4GH

Letter from the Chair

Dear GA4GH Community,

We are at a genuinely exciting moment in genomic medicine. As Peter mentioned, the data is here, the standards are maturing, and — perhaps most significantly — we are beginning to see what it looks like when the two come together in practice.

What stands out to me most from the past year is how much the landscape has shifted — not just in what we have built, but in what is now being asked of us.

I am sure that many of us would agree that the work we do is not easy, and last year brought further challenges: limited resources, shifting priorities, and a rapidly changing international landscape. This is precisely when standards matter most — when the ground is moving beneath our feet, shared infrastructure will keep us coordinated and connected.

Now more than ever, we have been prompted to adapt and evolve in real time. The rapid advancement of artificial intelligence in genomics is one example: GA4GH is responding to global innovation and technical developments within our new AI Work Stream set to launch at the April Connect 2026 meeting. We are also expanding how we engage and drive impact globally, through local networks and a growing National Initiatives Forum, because truly global standards must be relevant in all regional contexts.

The progress shared in these pages reflects real work by a remarkable community of contributors. I hope you find it to be a valuable reflection of what we have built together and a prompt for what comes next.

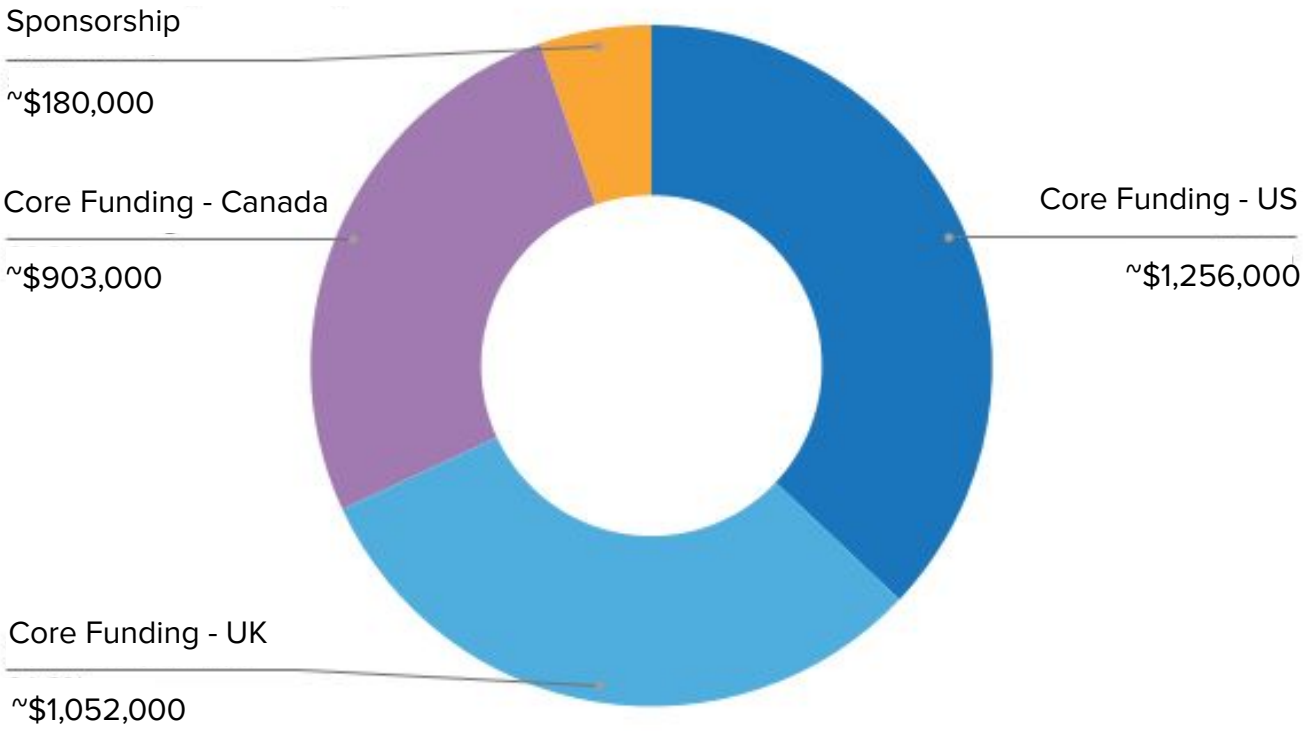
With gratitude for your continued dedication and support,

Heidi Rehm
Chair, GA4GH

Funding update

GA4GH's work is made possible through the sustained support of funders who share our commitment to responsible, global genomic data sharing. 2025 marked the close of our first formal five-year funding period — a milestone that reflects how far the organisation has matured. In GA4GH's early years, operations were sustained through a patchwork of ad hoc grants and Host Institution support. The structured multi-funder model developed over the past five years has allowed us to grow into a stable, globally active standards body with an annual operating budget of approximately $3.4 million USD.

In 2025, core funding was contributed roughly equally across three regions: Canada (32%), the United Kingdom (32%), and the United States (30%), with additional supporting funds making up the remainder. GA4GH's 23 staff members are distributed across five Host Institutions — the Broad Institute of MIT and Harvard, EMBL's European Bioinformatics Institute, the Ontario Institute for Cancer Research, the Wellcome Sanger Institute, and the Victor Phillip Dahdaleh Institute of Genomic Medicine at McGill University — with the majority of funding directed towards personnel, followed by meetings, travel, and software.



2025 funding sources in USD

We have received the commitment of the Canadian Institutes of Health Research (CIHR), whose new five-year grant of $1 million CAD/year — awarded in 2025 in support of GA4GH's global genomic data sharing work — provides a meaningful anchor as we navigate this transition. Canada's investment is a strong signal of confidence and leadership in what the GA4GH community has built and what we are capable of building next.

As we continue to sustain our progress and drive momentum, we recognise that we are undertaking the next funding period in a more challenging environment. Current grants are being extended into 2026 and 2027 to ensure continuity of operations while renewal discussions proceed, but the path forward for both US and UK funding remains uncertain. To secure a stable foundation for GA4GH's next chapter, we are actively engaging with existing and prospective funders to chart a path forward. We invite any interested parties to get in touch with us.

We thank all of our Core Funders for their continued partnership and commitment to making responsible global data sharing a reality.

Core Funders

CIHR IRSC
Canadian Institutes of Health Research / Instituts de recherche en santé du Canada

UKRI Medical Research Council

NIHR | National Institute for Health Research

NIH National Human Genome Research Institute

wellcome

Host Institutions

BROAD INSTITUTE

EMBL-EBI

OICR Ontario Institute for Cancer Research

McGill — VICTOR PHILLIP DAHDALEH INSTITUTE OF GENOMIC MEDICINE AT MCGILL UNIVERSITY

wellcome sanger institute

Supporting Funders

HDRUK Health Data Research UK

illumina

Welcoming new leadership

GA4GH product development is driven by leaders in the genomics and health community, whose expertise ensures products meet real-world needs. In 2025, the following individuals joined our leadership team.

GA4GH Vice-Chair: Nicky Mulder

Building on over a decade of involvement in GA4GH, Nicky Mulder, of the University of Cape Town, joined the GA4GH Executive Leadership Team as Vice-Chair of the Board of Directors to help guide the strategic direction of the organisation. Mulder brings a crucial perspective to GA4GH, working to integrate African genomic data into global research in a way that is equitable, non-exploitative, and supportive of the development of African research capabilities.

GA4GH Chief Product Officer: Sasha Siegel

Based at EMBL's European Bioinformatics Institute, Sasha joined as the GA4GH Chief Product Officer, driving product development by leading the Technical Team and collaborating with the community to ensure products are fit for purpose and best positioned for implementation and uptake. She has experience advancing operational thinking to translate scientific advancements into positive health impacts.

New Board Members

The GA4GH, Inc. Board of Directors helps drive the organisation's strategic mission. Four new Board Members were appointed in 2025, joining nine existing members.

David Glazer
Verily

Arcadi Navarro
University Pompeu Fabra; Centre de Regulació Genòmica

Patrick Tan
Precision Health Research, Singapore

Krystal Tsosie
Arizona State University

New Work Stream and Community of Interest Leads

Charlotte Barclay
Centre for Infectious Disease Genomics and One Health
Infectious Disease Community

Benjamin Berk
CareEvolution
Clinical & Phenotypic Data Capture Work Stream

David Bujold
Canadian Centre for Computational Genomics
Discovery Work Stream

Samantha Chill
Deloitte
Infectious Disease Community

Tom Conner
Broad Institute of MIT and Harvard
Data Security Work Stream

Miro Cupak
Discovery Work Stream

Toyofumi Fujiwara
Database Center for Life Science
Rare Disease Community

Ada Hamosh
Johns Hopkins University of Medicine
Rare Disease Community

Nishan Katuwal
Dhulikhel Hospital; Kathmandu University Hospital
Infectious Disease Community

Venkat Malladi
Quest Diagnostics
Cloud Work Stream

Brian O'Connor
Nimbus Informatics, LLC.
GA4GH Implementation Forum

Vasiliki Rahimzadeh
Baylor College of Medicine
Regulatory & Ethics Work Stream

Wafaa M. Rashed
Pan-African PGS Education and Research Initiative
GA4GH Implementation Forum

Kelly Shen
Simon Fraser University
Neuroscience Community

Andra Waagmeester
Johns Hopkins University School of Medicine
Rare Disease Community

Ma'n H. Zawati
McGill University Centre of Genomics and Policy
Regulatory & Ethics Work Stream

We thank former leads Tudor Groza, Randy McIntosh, and Dianne Nicol for their service to our community.

Global outreach and engagement

In 2025, GA4GH continued to develop a deliberately varied portfolio of outreach and engagement mechanisms — recognising that different contexts call for different levels of formality and structure, and that truly global standards have to work in local contexts. These range from structured membership forums and formal institutional partnerships, to conference satellite events and community-led local groups operating in local languages.

The National Initiatives Forum (NIF) remains the cornerstone of this approach, providing a structured mechanism for coordination and peer exchange among national and large-scale genomic initiatives. Through virtual and in-person meetings in 2025, NIF supported member initiatives in identifying shared challenges related to interoperability, governance, and implementation, and introduced national efforts to GA4GH standards, policies, Work Streams, and Communities of Interest.

NIF members

All of Us Research Program
AllForOne Canada's Precision Health Partnership
BIG MU
BGS
BIPMed
Canadian Genomic Data Commons
Serbia Centre for the Fourth Industrial Revolution
ClinGen Clinical Genome Resource
Danish National Genome Center
elixir
estonian genome center university of tartu
FINNGEN
gms
Genomics Australia
Genomics England
GHGA
HDRUK Health Data Research UK
Hong Kong Genome Institute 香港基因組中心
H3Africa
IndiGen Genomics for Public Health in India
Inserm
Institute of Molecular Biology
Instituto Nacional de Saúde Doutor Ricardo Jorge
International HundredK+ Cohorts Consortium
Matchmaker Exchange
Ministry of Health Manatū Hauora
NHS Genomic Medicine Service
Pan-Canadian Genome Library Bibliothèque génomique pancanadienne
PRECISE Precision Health Research Singapore
Qatar Genome
Nordic Alliance for Clinical Genomics
UCL Human Genomics

GA4GH staff and community members represented GA4GH at conferences and scientific meetings around the world throughout 2025, presenting on GA4GH standards and frameworks, participating in workshops and panels, and building relationships with researchers, clinicians, and policymakers working at the intersection of genomics and health. These outreach opportunities extend GA4GH's reach beyond its formal membership, introducing standards and tools to new communities and creating pathways for future engagement.

GA4GH strengthened partnerships with existing regional genomics networks to extend engagement beyond historically well-represented regions. For example, GA4GH hosted a satellite NIF meeting at the Asian Pacific Conference of Human Genomics in Indonesia, providing a forum for regional initiatives to align priorities and explore deeper involvement in GA4GH. Presenters from Singapore, Hong Kong, Taiwan, Thailand, Sri Lanka, Japan, and the Philippines offered regional perspectives on data sharing, capacity building, and standards adoption. GA4GH also had a presence at 22 global conferences.

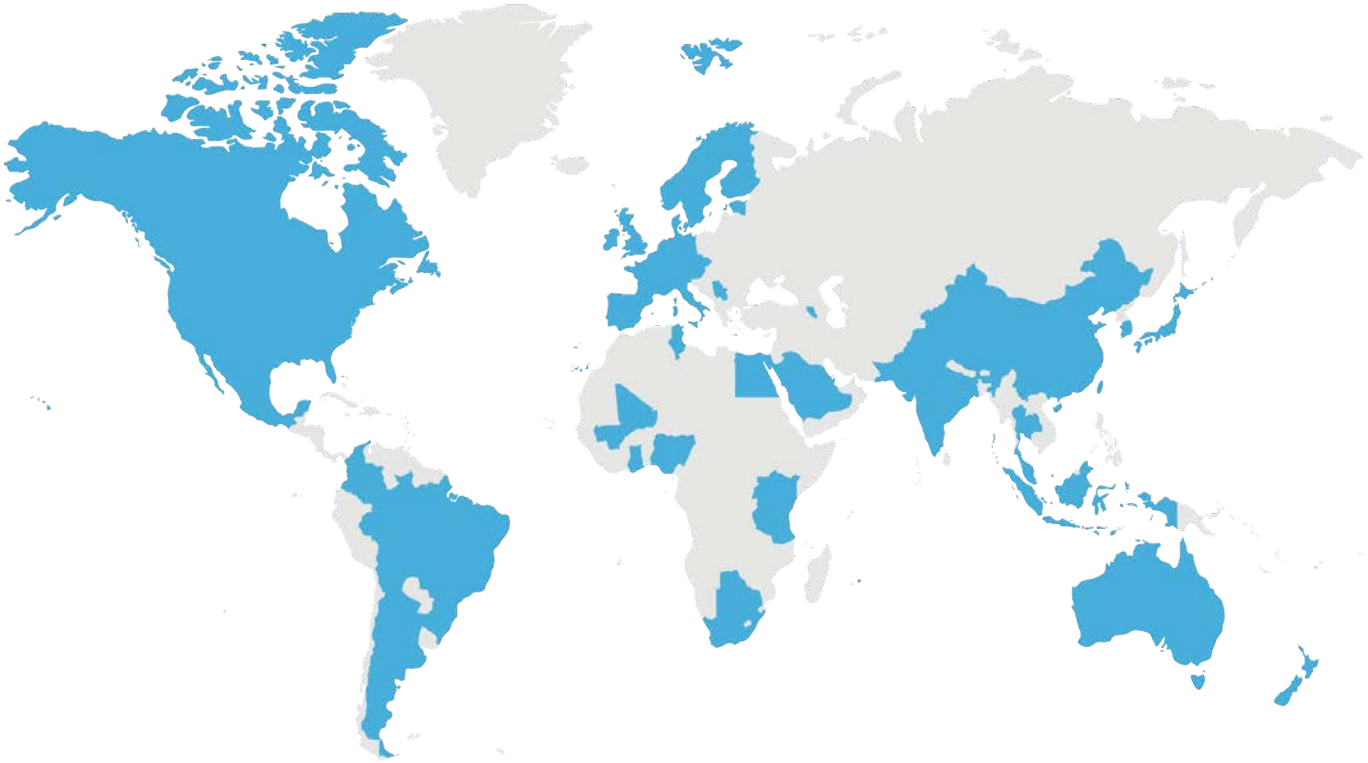
GA4GH also began exploring community-led local engagement models to support participation across different time zones, languages, and regulatory environments. One early example is a proposed Japanese rare disease satellite community, currently in development and led by Toyofumi Fujiwara (Database Center for Life Science, Research Organization of Information and Systems), which aims to mirror the activities of the global GA4GH Rare Disease Community while running in Japanese and adapting content to the Japanese research and clinical context.

Together, these efforts reflect GA4GH's commitment to equitable global engagement as an essential element of our mission.

2025 Snapshot

Our community

GA4GH contributors from across **51 countries** came together last year to participate in GA4GH activities, events, and meetings, fostering a global dialogue to advance standards, tools, policy frameworks, and collaborative efforts that support genomic and health data sharing.



GA4GH formed a Strategic Partnership with the Clinical Research Data Sharing Alliance (CRDSA) to align on genomic and clinical research data sharing standards.

CRDSA
Clinical Research Data Sharing Alliance

The GA4GH April Connect 2025 and 13th Plenary meetings welcomed

370+ in-person attendees

from **51** countries.

Our products and publications

6 newly approved products

Variation Representation Specification (VRS) v2.0 establishes standardised descriptions of variant evidence to enable reliable sharing of variant knowledge across diagnostic labs, research institutions, and knowledge-bases.

Refget Sequence Collections v1.0 standardises identifiers for collections of reference sequences to help researchers match datasets to reference sequences, verify data integrity, and reproduce analyses across institutions and platforms.

Categorical Variation Representation Specification (Cat-VRS) v1.0 enables search, analysis, and sharing of categories of genomic variants, connecting patient variants to existing knowledge and accelerating diagnosis and treatment decisions.

Variant Annotation Specification (VA-Spec) v1.0 provides a model to express knowledge associated with a variant, enabling researchers to quickly and reliably integrate variant knowledge across databases, sequencing labs, and health records.

Whole Genome Sequencing (WGS) Quality Control (QC) Standards v1.0 establish guidance for consistent, reliable, and comparable genomic data quality across institutions, which supports data sharing without data reprocessing or re-validating.

Experiments Metadata Checklist (Expmeta) v1.0 specifies the minimum information needed to describe how genomics experiments are conducted, allowing researchers to interpret, reproduce, and combine data across labs and studies.

2 new products in development

Human Exposome Data Standards aim to help researchers share and combine data on environmental exposures that affect human health.

Data Passports aim to to represent data access policies in a standardised and clear way using a visa-based governance framework.

15 academic papers

published by the GA4GH community, covering implementations of GA4GH products, efforts to advance data interoperability, thought leadership in the field of genomic data sharing, and more (see appendix).

9 guest blog posts

exploring public attitudes towards genomics, ethical considerations in genomic research, and challenges in piloting federated health data sharing (see appendix).

2025 timeline

18 February • Sasha Siegel joins GA4GH as Chief Product Officer

20 February • GA4GH publishes its first Annual Report

27 February • GA4GH Inc. welcomes four new board members

10 March • The Clinical Genomics Laboratory hosts its inaugural meeting

1 to 4 April • The April Connect 2025 meeting was held at the Broad Institute of MIT and Harvard

27 March • Variation Representation Specification v2.0 and refget Sequence Collections are approved as GA4GH products

15 May • Nicky Mulder joins the GA4GH Executive Leadership Team

12 June • Categorical Variation Representation Specification (Cat-VRS) v1.0 and Variant Annotation Specification (VA-Spec) v1.0 are approved as GA4GH products

24 June • GA4GH and CRDSA agree to a Strategic Partnership

6 to 10 October • 13th Plenary is co-hosted by SciLifeLab in Uppsala, Sweden

16 to 20 November • The Experiments Metadata Checklist (Expmeta) v1.0 and Whole Genome Sequencing (WGS) Quality Control (QC) standards v1.0 are approved as GA4GH products

GA4GH products in action

GA4GH standards and tools provide a blueprint for sharing genomic and health data in secure, responsible, and technology-forward ways. In 2025, there were powerful examples of GA4GH products in action — connecting genomic data to clinical care, protecting individual and community data rights, and enabling federated discovery across borders and health systems.

"Fundamentally, what standards do for any community is give them a place to start. It gives them common ground with which to communicate. If we can do that, then great science and impact will follow." — Sasha Siegel, GA4GH Chief Product Officer

NHS Genomic Medicine Service and Genomics England implement GA4GH products to integrate genomics into healthcare

In July 2025, the National Health Service (NHS) in England published a ten year health plan that aims to advance medical care with a goal of incorporating genomics into nearly fifty percent of all health care interventions.

Central to the plan is a “Unified Genomic Record” for every patient, linked to a “Single Patient Record” (SPR) — a comprehensive view of an individual’s health history that also gives patients control over who can access their data. Linking genomic data to this record holds real promise to advance predictive, preventive, and personalised health care.

Delivering this vision requires the NHS to scale its genomics capabilities so they are not only available in large hospitals, but reaching all the way down to the local GP surgery level. A country-wide data platform — built on GA4GH standards — provides the digital infrastructure to make this possible and continue to support the availability of consented data for research in the National Genomic Research Library.

The NHS Genomic Medicine Service, in partnership with Genomics England (a GA4GH Driver Project), has built a clinical genomics platform with GA4GH standards at its core. A federated database, implemented as a GA4GH Data Repository Service (DRS) server, is under development. This will act as a central data store from which any clinical lab in the network can request genomes relevant to their analysis.

The platform also incorporates htsget and GA4GH-maintained file formats (CRAM, BAM, VCF), allowing labs to retrieve only the portion of a genome relevant to a specific variant, rather than downloading an entire genome file.

Implemented in concert, GA4GH standards support faster, more efficient genomic analysis across the entire healthcare network. Looking ahead, the NHS is also beginning to implement Beacon v2.0 to query variant data, which will help facilitate rare disease diagnosis and identify patients for cancer clinical trials, directly advancing patient health outcomes.

Rakeiora Genomics Platform: protecting data sovereignty through GA4GH standards

The Rakeiora Genomics Platform, built in partnership with Māori communities in Aotearoa New Zealand, integrates genomic and clinical data with genealogical knowledge that has been passed down through generations. From the outset of the collaboration, there was a collective acknowledgement that sovereignty over the samples and data had to remain with the individuals and communities contributing data to the platform.

To honour that commitment, the technical team built safeguards that protect private and identifiable information into the platform, implementing GA4GH's htsget and Passports standards and aligning with the *GA4GH Framework for Responsible Sharing of Genomic and Health-related Data*.

Htsget ensures that even when a whole genome sequence is available, researchers can access only the portion relevant to their study, limiting exposure of personally identifiable genetic information without limiting scientific utility.

Access is further controlled through GA4GH Passports and the Authentication and Authorisation Infrastructure (AAI). Each researcher is assigned a securely-encoded "visa" specifying exactly which genomic datasets they are permitted to access. Participant consents and researcher approvals are managed through ELIXIR's AAI Resource Entitlement Management System (REMS), which is fully GA4GH-compliant.

The Rakeiora Genomics Platform demonstrates that globally harmonised standards can work alongside local cultural frameworks to ensure data security and sovereignty, while fostering more personalised, informed, and culturally-relevant care.

Read more on the GA4GH website:

GA4GH standards implementation in the Rakeiora Genomics Platform supports data security and sovereignty in Aotearoa New Zealand

By Jaclyn Estrin, GA4GH Senior Science Writer

29 January 2026

Denmark's Rigshospitalet: building a national genomic reference on GA4GH standards

Genomic Medicine (MDxCore) Rigshospitalet in Denmark is one of Europe's highest-volume clinical genomics operations — processing about 50,000 samples, 20,000 requisitions, and 10,000 clinical reports every year across cancer, rare disease, and infectious disease cases. Running clinical whole genome sequencing since 2016, the centre has built a full automated ISO-accredited sequencing and bioinformatics pipeline that takes a sample from sequencer to clinical report with minimal manual intervention — reducing complexity, errors, and time at every step.

To ensure that data generated through this pipeline can be shared, queried, and reused across institutions, the centre has adopted GA4GH standards throughout — most notably in DenGen, a national Danish genome frequency database built on a reference set of 2,211 unrelated genomes with more to come.

DenGen has adopted Beacon v2.0 for variant queries and aligned with the Variation Representation Specification for both single nucleotide variants and structural variants — giving Danish clinicians and researchers a standards-compliant national reference they can query across institutions.

The centre is also implementing Phenopackets for clinical metadata, which would allow data generated in the course of patient care to serve research and reuse purposes without additional effort.

At Rigshospitalet, data standards are treated not as a technical convenience but as safety infrastructure — as essential to clinical genomics as quality control, accreditation, and the ability to replace and improve individual components in a complex clinical operation. The ambition is for data generated in the course of patient care to serve multiple purposes simultaneously: the current patient, future patients, and the wider research community. Looking ahead, the centre is planning to expand its GA4GH adoption further — including the Data Use Ontology and Consent Clauses — and is contributing to broader efforts to align clinical bioinformatics practice across Europe through ISO and the European Health Data Space.

Canada takes a first step toward federated genomic data sharing

Canada has invested heavily in genomic research, but until recently, the resulting data assets have been largely siloed across projects, institutions, and provinces. Variations in data governance, jurisdictional regulations, and the absence of a national strategy have made large-scale data sharing challenging.

The Pan-Canadian Genome Library (PCGL) — a GA4GH Driver Project hosted at McGill University and funded by the Canadian Institutes of Health Research (CIHR), Genome Canada, and the Digital Research Alliance of Canada — is working to change that. PCGL is building a national infrastructure where genomic and health data can be ingested, processed through standard analytical pipelines, and made discoverable and accessible to authorised researchers. Its hybrid architecture encourages centralised data storage while also supporting federated nodes for datasets with specific sovereignty or jurisdictional requirements.

GA4GH standards are central to the design. The PCGL Research Portal is built on the Bento platform, a GA4GH standards-based open-source software for building genomic data sharing infrastructure. The portal implements Beacon v2.0 for federated data discovery, Phenopackets for its data model, and the Data Use Ontology (DUO) for data access governance. Internationally, PCGL datasets will also be discoverable through the Canadian Genome-Phenome Archive (CGA) — Canada's node in the Federated European Genome-phenome Archive (EGA) — linking Canadian data to the broader global genomics ecosystem.

Data is already flowing into PCGL, with the Quebec COVID-19 Biobank (BQC19) contributing the first dataset to enter the ingestion pipeline. The system is set to scale rapidly: Genome Canada has mandated that the 12 Canadian Precision Health Initiative (CPHI) sequencing projects it funds must deposit their data into PCGL. With governance shaped by an Indigenous Genetics Circle and patient partners, and alignment with GA4GH standards throughout its technical stack, PCGL is turning Canada's first step toward federated genomic data sharing into a working national infrastructure.

The International Fetal Genomics Consortium prototypes clinical data querying for fetal genomics

When a child is born, DNA abnormalities account for 20% of perinatal deaths or admissions to children's hospitals.

The International Fetal Genomics Consortium (IFGC), launched in 2015, aims to harness the progress in genomics to better care for families during and after pregnancy.

With a goal of accelerating diagnosis, improving care, and preventing stillbirths, members of the IFGC collaborated with an international network of hospitals and research institutes to create a data repository tailored to fetal genomics. They aggregated more than 3,000 prenatal cases, spanning fetal structural anomalies, stillbirths, and non-anomalous pregnancies. The majority of the cases are genome trios, composed of DNA sequenced from a fetus, the parents, and occasionally the placenta. The team has now piloted GA4GH Beacon v2.0 and Phenopackets to make the repository's data computable, queryable, and accessible for research and clinical use.

Leveraging a subset of 750 structural anomaly genome trios from the repository, the team built a prototype in the Broad Institute's Terra Platform, securely storing de-identified prenatal data in the Phenopackets format. The data was made queryable via Beacon v2.0, with data access governed by the GA4GH Data Use Ontology (DUO).

To standardise the data, phenotypes were extracted from ultrasound reports and converted into the Phenopackets schema. Each Phenopacket received a unique identifier, described features using Human Phenotype Ontology (HPO) terms, and included information on variants, providing the foundation for a searchable prototype.

Leveraging Beacon v2.0, the team then successfully implemented filtering capabilities for phenotypic features and variants, supporting queries by HPO terms. For instance, a researcher could search for male fetuses between 1 and 40 gestational weeks with a specific HPO term, and the prototype would return the number of records, the Phenopacket IDs, and the links to view the Phenopackets in Terra — offering a practical solution for controlled data access without exposing the Phenopackets directly.

This prototype demonstrates that GA4GH standards can support responsible, queryable data infrastructure even in a clinically sensitive domain. Next, the team is building a more comprehensive Beacon interface with a broader prenatal network, pointing towards a more connected future for prenatal genomics research.

Epic implements GKS tools to improve genomic variant interpretation in clinical workflows

Epic is a global healthcare software company that integrates patient data into a single health record that can follow patients wherever they receive care. Its users range from patients accessing their own health information to clinicians such as geneticists and oncologists, to health system researchers analysing health records to advance medical knowledge. Built with a commitment to interoperability, Epic's software facilitates health data sharing responsibly across healthcare organisations and thus plays a key role in industry efforts to unlock new insights from genomic data. The Epic team is currently implementing GA4GH's Genomic Knowledge Standards (GKS) to interoperate with knowledge-bases that are prevalent in clinical and laboratory workflows, matching variants to annotations.

In clinical genomics workflows, clearly defined variant representations and their associated knowledge enable genomic results to prompt personalised patient education and treatment plans. Epic has embedded standards throughout the journey from test order to clinical result. When a clinician and lab are housed within different organisations, Epic uses HL7 v2 to transmit the order and receive results automatically, extending the v2 standard to align with GA4GH data models, especially for handling complex variant and knowledge types.

Epic is developing a tertiary analysis workflow to match VCF data with current genomic knowledge-bases, using GKS standards.

In this workflow, assayed variants are represented as VRS alleles. Epic then runs a matching algorithm based on Cat-VRS to annotate each variant with knowledge across a wide range of knowledge sources represented with GKS, including ClinVar, ClinGen Gene-Disease Validity, gnomAD, Variant Effect Predictor (VEP), dbSNP, and the knowledge-base of organisations that are using Epic's software. Annotations are imported back into Epic using the Variant Annotation Specification (VA-Spec), in a standardised, shareable format that includes pathogenicity statements, gene-disease association, population frequency, molecular consequences, functional impact predictions, and variant level literature.

All of this gives pathologists and variant scientists a richer, more reliable foundation for genomic interpretation — reducing the time of case review, enabling knowledge exchange across software systems, and ensuring that patients benefit from reinterpretation as new knowledge about their variants emerges.

Singapore puts 110,000 genomes to work with standardised quality control

Singapore has sequenced more than 110,000 genomes through the first two phases of its National Precision Medicine (NPM) programme — one of the most successful national genomics efforts in the world. Making that data useful for research depends on being able to trust its quality, consistently, across every sample.

The Whole Genome Sequencing Quality Control (WGS-QC) standards, approved by the GA4GH Product Steering Committee in November 2025, provide exactly that: a defined set of quality control metrics ensuring clarity and consistency across tools, labs, and projects. Deployed by Precision Health Research, Singapore (PRECISE), WGS-QC is now applied across Phase I data (10,000 genomes) and is currently being rolled out across Phase II data, as the programme moves into its third phase: evaluating how genomics can inform national health strategies in practice. The PRECISE Genomic Innovation team is also working towards elevating these standards towards alignment with internationally recognised ISO-level quality frameworks.

The NPM programme's WGS data can be accessed via TRUST, Singapore's national health-related data governance framework and analytics platform. Authorised researchers can filter samples by various criteria, including QC metrics, to identify datasets best suited for their work and proceed directly to analysis without needing to redo quality checks, saving time, money, and resources. WGS-QC also establishes a common quality baseline for PRECISE's sequencing partners, streamlining quality assurance across the programme.

Looking ahead, the PRECISE Genomic Innovation Team aims to incorporate WGS-QC into its data processing pipelines to automate the delivery of standardised QC metrics, reducing manual effort and accelerating the pace at which Singapore's genomic data can be used for research and clinical care.

Updates, progress, and achievements from across the GA4GH community

Supporting implementation of GA4GH products

GA4GH is expanding activities to further support the implementation of its products. The GA4GH Implementation Forum introduced new ways to support adoption, and the GA4GH Tech Team is developing a platform to explore metrics on the impact of GA4GH products.

GA4GH Implementation Forum

The **GA4GH Implementation Forum (GIF)** aims to facilitate the adoption of GA4GH products and ensure interoperability through a collaborative approach. By considering diverse requirements and available resources, GIF aims to lower the barriers to implementing GA4GH standards in the spirit of building networks of knowledge instead of silos of expertise.

GIF "Ask Me Anything" (AMA) programmes

In 2025, GIF launched the **"Ask Me Anything" (AMA) programme** — a series of events that bring GA4GH product teams and current and future implementers together to discuss product implementation questions and challenges. The aim of the programme is to provide dedicated time and space to clarify technical understanding and facilitate implementation progress.

Three AMA programmes were held last year, featuring Beacon, Phenopackets, and Passports, Visas, and AAI.

Beacon
March 2025

Phenopackets
May 2025

Passports, Visas, and AAI
November 2025

GIF Projects

GA4GH Implementation Forum (GIF) Projects are community-led initiatives designed to address challenges in genomic data interoperability by implementing one or more GA4GH products. These projects focus on practical solutions to enhance data discoverability, sharing, accessibility, and overall integration across the genomics ecosystem. In 2025, GIF onboarded the following two new GIF Projects.

New Cloud-based BRCA Exchange variant analysis environment using GA4GH standards in Camber

The BRCA Exchange is a global resource for curated information on BRCA1 and BRCA2 gene variants associated with breast and ovarian cancer. To enable scalable, reproducible, and standards-based genomic analyses, this GIF Project aims to integrate BRCA Exchange variant data with GA4GH standards using Camber — a cloud-based scientific computing environment that allows researchers to query, analyse, and visualise data at scale. The project plans to implement GA4GH Genomic Knowledge Standards (GKS), the Tool Registry Service (TRS), and the Data Repository Service (DRS) to enable a platform for collaborative genomics research.

New Trusted Research Environment (TRE) Open Suite

The TRE Open Suite project aims to develop an open, extensible framework and software development kit for building and deploying the technical foundations for trusted research environments. A key innovation of the project is a universal matrix of security aspects that systematically addresses privacy, data ownership, software licensing, and artificial governance in federated environments. The framework also incorporates the concept of programmatic trust management, providing tools and interfaces that allow automated evaluation, verification, and enforcement of trust conditions across components and institutions. The project is currently focused on drafting the security framework and integrating GA4GH standards and trust-enhancing technologies.

Federated Variant-Level Matching (VLM) Project

Timely access to genotype and phenotype information from diverse sources would allow the clinical and research communities to achieve more accurate and faster diagnosis and intervention for patients. The Federated VLM GIF Project aims to support a large federated network of genomic databases with a focus on rare disease cohorts and clinical labs. This network allows researchers and clinicians to query variant-level genotypes and access individual-level phenotypic features upon obtaining a genotype match, utilising the Beacon v2.0 API. Currently, Federated VLM has connected four rare disease research platforms and one commercial software platform for clinical labs, with active expansion of the network underway. Best practices are in development to help guide the community in launching more nodes in the network.

Guest Blog Post Spotlight

Even with the progress we have been able to share, federated data sharing remains a challenging endeavor. In a guest blog post published on 18 September 2025, Lindsay Smith of the International Precision Child Health Partnership (IPCHiP) transparently conveyed the challenges IPCHiP has faced in piloting federated health data sharing.

Borderless data, boundless hurdles: navigating the challenges of international federated sharing

By Lindsay Smith

29 January 2026

In the world of international data sharing, there is a daunting chasm between promising innovations and their real-world application — a space where data sharing initiatives often falter, not for lack of vision, but due to the complex web of challenges that arise in implementation. Several challenges arose when piloting data federation within IPCHiP's flagship study, Gene-STEPS (Shortening Time of Evaluation in Paediatric Epilepsy Services).

On-premise vs cloud considerations

Institutions had different requirements for hosting data sharing software in a cloud vs an on-premise (on-prem) environment. It is possible to accommodate these differing requirements, but it would have been helpful to have the ability to easily compare notes. On-prem environments require greater isolation features than cloud, and while a cloud approach may be able to mitigate isolation concerns, it can be expensive to maintain even with a minimal set up.

Data security concerns

In order to run workflows in another institution's environment, we need open ports to the internet and significant infrastructure updates to support isolation or security management features. Open ports are not always feasible in an institution hosting health data, and these updates can be cost prohibitive.

If the cloud supporting data sharing software is set up in an environment isolated from the main data store, data needs to move to be accessible. Is this really a federative approach, if we have to move the data?

Querying connected datasets

User friendly interfaces and no-code solutions are needed to enable use by clinicians and researchers who are not trained in query methods. No-code solutions have limitations, and if a researcher wishes to query beyond simple filters, technical knowledge is required.

Data access

Once the datasets of interest are located, administration of data access requests was an unanticipated complication for some sites. For sites without a central data access committee or system, it was unclear how to triage requests appropriately.

Workflow languages

Preferences for particular workflow languages vary across bioinformaticians and sites.

Workflow cost management and computing resources

If a researcher triggers a workflow in another site's environment, it is unclear who pays for the workflow costs and how. To support cost management, there also needs to be agreement on what the workflow is intended to do at each site.

Results management

The ability to combine datasets and queries from across our sites is key to letting go of our "phone a friend" process. With four sites, and the potential for multiple variables of interest, it is unclear if data is stored or harmonised in a way that allows for complex queries.

By sharing experiences openly (both the good *and* the bad), engaging in collaborative pilot projects, and growing from the lessons learned so far, we are gradually building bridges over the chasm between innovation to application, one brick or beam at a time. I would call on other global data sharing collaborations to continue engaging in a collective dialogue about lessons learned, challenges faced, and opportunities to drive progress together. The potential rewards — better variant interpretation, improved analytical and diagnostic power, and the realisation of precision health — are well worth the journey.

Updates from the GA4GH Communities of Interest

GA4GH supports five Communities of Interest, offering forums for researchers, clinicians, and technical experts to define, discuss, and advance real-world use cases involving GA4GH standards and tools. By bringing the community together around specific domains, the Communities are uniquely positioned to bridge the divide between technical developers and clinical practitioners. In 2025, GA4GH onboarded new leadership to the Communities to steward and expand their work and activities.

Cancer Community

The Cancer Community aims to connect international cancer initiatives through interoperable standards and promote collaborative cancer research.

The Cancer Community Co-Leads contributed their insights to the article, "Cancer genomics and global collaboration," published by Open Access Government on 20 November 2025. The Co-Leads discuss the current state of cancer genomics and how the field may evolve in the future.

Clinical Genomics Laboratory Community

In partnership with ClinGen, the Clinical Genomics Laboratory Community (CGLC) seeks to facilitate knowledge exchange among global clinical laboratories to enhance the accuracy and efficiency of genetic testing and analysis for the betterment of patient care.

The CGLC held six meetings in 2025, discussing topics of relevance to the community. These include engaging the community in piloting the draft ACMG/AMP/CAP/ClinGen Sequence Variant Classification (SVC) v4.0 standards, showcasing somatic cancer variant classification and knowledge sharing across platforms, advancing the classification of variants in ACMG secondary findings genes, and the use of artificial intelligence (AI) and large language models (LLMs) in variant classification.

Infectious Disease Community

In partnership with the Public Health Alliance for Genomic Epidemiology (PHA4GE), the Infectious Disease (ID) Community aims to standardise the use of pathogen and host genomic data to better diagnose and treat infectious disease.

Charlotte Barclay (Centre for Infectious Disease Genomics and One Health), Samantha Chill (Deloitte), and Nishan Katuwal (Genome Sequencing Research Lab at Ghulikhel Hospital; Kathmandu University Hospital) joined as new Co-Leads of the Infectious Disease Community to renew ID Community activities and cultivate a forum for collaboration and knowledge exchange to advance global standards that address infectious diseases.

Neuroscience Community

In partnership with the Brain Research International Data Governance & Exchange (BRIDGE) project of the International Neuroscience Coordinating Facility (INCF), the Neuroscience Community bridges scientific discovery with the technical foundations needed for responsible and scalable data sharing in neuroscience research.

The Neuroscience Community held 16 meetings in 2025, focused on enabling more effective, large-scale use of neuroscience and biomedical data by addressing three interconnected challenges: how data is governed, how it is structured and standardised, and how people are trained to work with it. The discussions worked toward building a globally informed understanding of data governance practices and barriers, defining the technical foundations for interoperability through shared models and ontologies, and developing practical guidance and educational resources to support responsible use of emerging tools.

The Neuroscience Community drafted two white papers on neuroscience scientific education processes, titled: “Bridging the Neuro-AI Chasm: A Framework for Scalable, Contextually Adaptive Training Resources in Large-Scale Brain Data Science” and “The Neuroscience Education Readiness Checklist for Neuroscience Tools and Methods.” Both papers are pending publication.

Kelly Shen (Simon Fraser University) joined as a new Co-Lead of the Neuroscience Community.

Rare Disease Community

The Rare Disease Community explores real-world rare disease use cases through the implementation of interoperable standards, with an emphasis on clinical and patient perspectives.

The Rare Disease Community convened five times throughout 2025 to share tools and resources. Community members learned about Phenotips, a platform for structured phenotyping and clinical workflow integration; the Monarch Initiative, which discussed the importance of harmonising terminologies to connect fragmented data; and HiFi Solves, a global search engine that aims to enable federated, privacy-preserving data sharing.

The GA4GH Rare Disease Community also hosted a **Rare Disease Day webinar** on 27 February 2025 to showcase collaborative efforts to advance rare disease knowledge and enhance global strategies for patient engagement. Speakers included Daria Julkowska (Coordinator of the European Rare Diseases Research Alliance), Charlene Son Rigby (CEO of Global Genes), and Heidi Rehm (Chair of GA4GH).

Ada Hamosh (Johns Hopkins University) and Andra Waagmeester (Amsterdam UMC) joined as new Co-Leads of the Rare Disease Community. Toyofumi Fujiwara (Database Center For Life Science) was onboarded as the satellite community lead for Japan.

Updates from the GA4GH Work Streams

In 2025, Work Stream contributors continued to drive forward the development of standards, tools, and policy frameworks. Work Steam and Product Leads guided several new products through approval, pushed forward interoperability and implementation efforts, and launched new groups to address emerging challenges — from federated machine learning to environmental health data to the ethical dimensions of AI in genomics. In this section, please explore updates and highlights from each of the Work Streams.

Clinical & Phenotypic Data Capture Work Stream

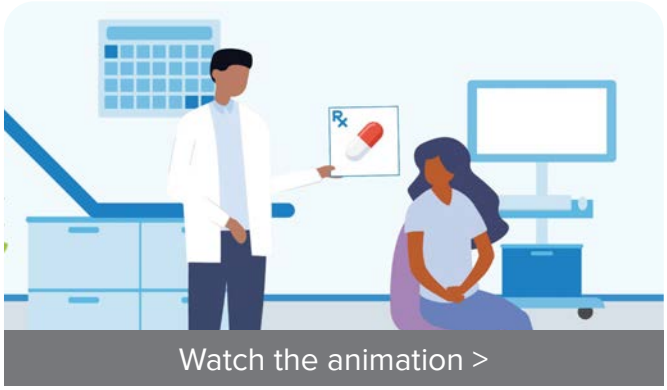
The Clinical & Phenotypic Data Capture (Clin/Pheno) Work Stream produces standards and information models to help clinicians and researchers describe and exchange clinical and phenotypic data.

Phenopackets provides a human and machine-readable way to structure clinical and phenotypic data about a patient or individual. In 2025, the product team focused on schema enhancements, improved documentation, and new tooling to support interoperability with other standards such as the Observational Medical Outcomes Partnership (OMOP) and Fast Healthcare Interoperability Resources (FHIR). The team also expanded product implementation and education, delivering a **webinar** for an educational series hosted by the Global Genomic Medicine Collaborative (GGMC).

Pedigree aims to allow computable exchange of family health history and representation of larger, more complex families. The product team advanced interoperability efforts to align this standard with Epic, FHIR, Phenotips, and Progeny. The team also continued Kinship Ontology (KIN) refinement.

A new study group, **Human Exposome Data Standards**, launched in 2025 to build standards for capturing and sharing data on environmental exposures that affect human health.

In collaboration with the GA4GH Communications team, the Work Stream developed an animation to depict how Clin/Pheno standards facilitate the representation and exchange of patient data to improve patient care.



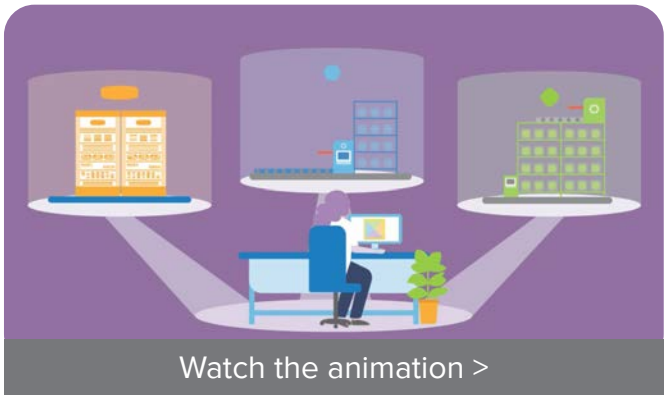
Cloud Work Stream

The Cloud Work Stream standardises how researchers bring algorithms to the data by defining, sharing, and executing portable workflows across any environment.

The Data Repository Service (DRS) provides a standardised set of data retrieval methods for analysis. In 2025, the Cloud Work Stream mapped 22 registered DRS servers across the GA4GH community, revealing an ecosystem with more than 72 petabytes of data. The product team also released DRS v1.5 in March 2025, adding fields for indicating cold storage, explicit cloud locations, object count and size metadata, and best practice guidance for using the GA4GH Data Connect standard with DRS.

The Cloud Work Stream also launched a new initiative to develop standards for **federated machine learning**, aiming to identify the technical, policy, and operational challenges of federated machine learning across distributed biomedical datasets; provide guidance and recommendations on these challenges; and support secure, standards-based orchestration of compute, data access, and model exchange while ensuring compliance and sustainability.

In collaboration with the GA4GH Communications team, the Work Stream developed an animation to depict how Cloud Work Stream products facilitate large-scale genomic data analysis by bringing “compute to the data.”



Data Security Work Stream

The Data Security Work Stream (DSWS) develops products and guidelines to safeguard privacy, data integrity, and service availability. DSWS also evaluates GA4GH products for risk or security implications.

In collaboration with the Data Use & Researcher Identity Work Stream, DSWS co-launched the Data Passports group, which aims to devise a computable, visa-based governance framework to represent data access policies.

To answer community questions, DSWS held an "Ask Me Anything" (AMA) event in collaboration with the GA4GH Implementation Forum (GIF). The AMA covered Passports and the Authorisation and Authentication Infrastructure (AAI), two products that facilitate secure data access.

DSWS participated in ISO-TC 215 meetings, providing updates on **Verifiable Credentials for genomics data** and **Passports** to Ad Hoc Group 2.

DSWS held collaborative discussions with several GA4GH groups, including the Beacon team on privacy considerations and the Open TRE Suite GIF Project on attested TLS.

Data Use & Researcher Identity Work Stream

The Data Use & Researcher Identity (DURI) Work Stream creates tools to communicate data access permissions.

DURI co-launched the Data Passports group with the Data Security Work Stream. This new initiative aims to devise a computable, visa-based governance framework to represent data access policies.

The Data Use Ontology (DUO) allows data stewards to tag datasets with permitted use terms that facilitate data discovery and access. The product is under new leadership, with discussions underway to explore updates and changes based on community needs.

Discovery Work Stream

The Discovery Work Stream helps researchers and clinicians explore, find, and search for genomic and health data.

The Experiments Metadata Checklist was approved by the GA4GH Product Steering Committee as an official GA4GH product. The standard specifies the minimum information needed to describe a genomic experiment — providing crucial context to help researchers understand the results.

Beacon enables researchers to discover new, relevant datasets without compromising dataset privacy or ownership. The team made progress in several areas.

- The product team released **Beacon v2.2.0**, incorporating enhancements and bug fixes. The team is also conducting ongoing privacy and security work.
- The product team launched a new **Metadata Scout initiative**, which aims to extend Beacon v2.0 with a lightweight, optional metadata model for discovering higher-level resources such as datasets, cohorts, biobanks, studies, and networks — in addition to individual-level genomic records.
- The **Beacon Implementers Forum** was launched to provide a community-driven space to share experiences, challenges, and innovations in Beacon deployment. The team held two forum meetings: the first took place in July 2025 and showcased **Serverless Beacon (sBeacon)**, developed by the Commonwealth Scientific and Industrial Research Organisation (CSIRO), and the second meeting in November 2025 highlighted **Molgenix EMX2's** work to integrate Beacon into its open-source data platform.

Genomics Knowledge Standards Work Stream

The Genomic Knowledge Standards (GKS) Work Stream aims to define a common language for computers to describe variants and their associated biomedical knowledge, making it easier to share and apply this information to improve patient health outcomes.

The GA4GH Product Steering Committee approved several GKS products in 2025:

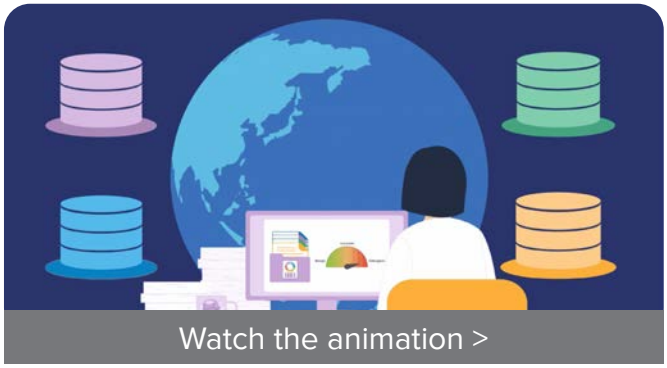
- The Variation Representation Specification (VRS) v2.0 expands upon the previous version by allowing for additional descriptive metadata, and advancing the representation of both "small" and "structural" variants.
- The Categorical Variation Representation Specification (Cat-VRS) v1.0 provides a data model and related tools to represent categorical variants in genomic knowledge-bases.
- The Variant Annotation Specification (VA-Spec) defines a modelling framework and machine-readable schema to represent statements of knowledge about variants.

GKS hosted several Open Houses, providing an entry point for the community to learn about the Work Stream, its products, and implementation examples.

- At the January 2025 Open House, the GKS team provided an overview of **VRS v2.0**. Melissa Cline (University of California, Santa Cruz) discussed the implementation of VRS in the **BRCA Exchange**, which supports VRS digests in its variant nomenclature system — enabling users to query the database using VRS digests and more efficiently compare information across repositories.
- The June 2025 session spotlighted **VA-Spec** implementations within the **Clinical Interpretation of Variants in Cancer (CIViC) knowledge-base** and **ClinVar**, a public archive hosted by the National Center for Biotechnology Information (NCBI).
- The November 2025 session spotlighted **Cat-VRS** and explored how **Epic**, a healthcare software company, has integrated GKS products into its internal data model and knowledge infrastructure.

The GKS Work Stream strengthened cross-standard interoperability and advanced real-world implementation across clinical and knowledge-base systems through implementer pilots, tooling, and community collaboration.

In collaboration with the GA4GH Communications team, the Work Stream developed an animation to convey the vision of the GKS Work Stream: to define a “common language” to share genetic variants and scale genomic medicine.



Large-Scale Genomics Work Stream

The Large-Scale Genomics (LSG) Work Stream develops and maintains products to describe, compress, store, encrypt, and transfer genomic data at scale.

The GA4GH Product Steering Committee approved two LSG products in 2025:

- Refget Sequence Collections establishes a unique identifier for an entire group of reference sequences, providing an unambiguous way to identify the collection.
- Whole Genome Sequencing (WGS) Quality Control (QC) describes a set of quality control metrics and their detailed definitions to facilitate the exchange of results across institutions.

Regulatory & Ethics Work Stream

The Regulatory & Ethics Work Stream (REWS) develops community-driven policies, tools, and standards that address regulatory and ethical considerations in genomic research and international data sharing.

In 2025, REWS convened eight active groups, exploring the ethical and legal angles of emerging technologies and data sharing approaches, methods to ensure the equitable application of genomics, and more. These groups are: Clinical Data Sharing; Data Visiting; Generative Artificial Intelligence (AI) for Genomic Data Sharing; Genetic Discrimination; Global ELSI: Building Global Capacity for Genomic Research; Genomic Sequencing in Newborn Screening; Pathogen Genomic Data Sharing; and Measuring the Uptake of Ethics and Data Governance Policies in Genomics.

REWS also publishes briefs and blog posts to inform policy, exploring regulations and their implications on data sharing and public attitudes towards genomic data sharing.

Health Data Sharing, Privacy, and Regulatory Forum

The Health Data Sharing, Privacy, and Regulatory Forum publishes briefs to explore and discuss established and emerging data-related regulations and their impact on data sharing.

Children's data protection and genomic research (part 1: general considerations)

By Michael J. S. Beauvais (University of Toronto)

23 January 2025

Children's data protection and genomic research (part 2: consent and lawful bases)

By Michael J. S. Beauvais (University of Toronto)

4 February 2025

[Will the UK participate in the European Health Data Space?]

By Miranda Mourby (University of Oxford)

17 June 2025

Public Attitudes for Genomics and Policy Forum

The [Public Attitudes for Genomic Policy forum] shares findings that relate to studies around public attitudes about how researchers collect, use, and share genomic and health data.

[Public perceptions of international genetic information sharing for biomedical research in China]

By Zhangyu Wang, Meng Wang, and Li Du (University of Macau)

16 January 2025

[The perceived risks of sharing genomic data with researchers]

By Richard Karlsson Linnér (Universiteit Leiden, Netherlands)

22 July 2025

[The role of the 'commercialisation effect' in shaping Australians' attitudes towards a hypothetical national genomic repository]

By Brad Elphinstone (Department of Psychological Sciences, Swinburne University of Technology)

30 October 2025

[Individual attitudes toward medical and genetic data sharing in Latin America and development of two new indexes]

By Gabriela Chavaría Soley, Ronald Alfaro Redondo, and Henriette Raventós (Universidad de Costa Rica)

20 November 2025

[Older research participants are motivated to receive genetic results for the benefit of younger relatives]

By Amanda Willis (Garvan Institute of Medical Research)

9 December 2025

Health Technologies Industries: Points to Consider Infographic

Explore key principles for responsible international genomic and health data sharing by health technology industries. This infographic is adapted from the 2024 resource "[International Genomic Data Sharing by Health Technologies Industries: Points to Consider]," developed by the Industry Core Group and the Centre of Genomics and Policy, part of the Victor Phillip Dahdaleh Institute of Genomic Medicine — a GA4GH Host Institution.

GA4GH events

April Connect 2025

The GA4GH April Connect 2025 meeting brought together members of the GA4GH community for four days of collaboration, technical advancement, and strategic discussion. 160 attendees convened at the Broad Institute of MIT and Harvard, with more individuals participating virtually. The meeting focused on driving actionable outcomes and advancing the GA4GH mission to ensure standards are effectively implemented, widely adopted, and make a meaningful impact in real-world settings.

160 participants

43 sessions

View meeting Report >

13th Plenary

The GA4GH 13th Plenary meeting, co-hosted by SciLifeLab, brought together members of the global genomics and health community to cultivate a dialogue around scaling genomic and health data sharing to enhance human health outcomes around the world. Nearly 300 attendees gathered in Uppsala, Sweden, and more than 300 participants registered to join virtually. The meeting centred on four core themes: the value of genomic medicine to advance health outcomes, the evolving role of artificial intelligence (AI) in genomics, the importance of collaboration, and the vital need to cultivate trust amongst all participants in the healthcare ecosystem.

The meeting was highlighted in *Technology Networks* and several blog posts from meeting attendees (Ewan Birney, Deborah Ekusai Sebatta, and Rukiya Mohamed Haji).

Learn more about the meeting in our press release and in a video produced by SciLifeLab.

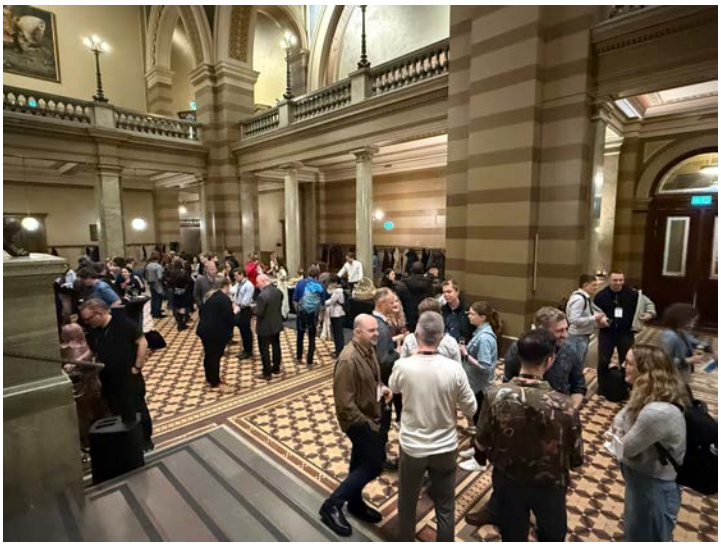
~300 participants

50 sessions

View Meeting Report >

Thank you for being a part of the GA4GH community

Our contributor community is the driving force behind the development and implementation of our products, standards, and policy frameworks. We want to express our gratitude to our contributors for their continued dedication, engagement, and partnership, as we work together to advance the responsible sharing of genomic and health data.



Are you interested in contributing or getting more involved with GA4GH? We invite you to **join our Work Streams and Communities of Interest**, or explore our **Open Calls page** for specific opportunities to provide your feedback and expertise.

Thank you to our funders

Thank you to our Host Institutions, Core Funders, Supporting Funders, and Assigned Expert Funders whose generosity has made our work, operations, and events possible. Your continued support sustains the infrastructure, community, and convening power that make GA4GH's work possible.

Core Funders

CIHR IRSC
Canadian Institutes of Health Research / Instituts de recherche en santé du Canada

UKRI Medical Research Council

NIH National Human Genome Research Institute

NIHR | National Institute for Health Research

Wellcome

Host Institutions

BROAD INSTITUTE

EMBL-EBI

McGill University
Victor Phillip Dahdaleh Institute of Genomic Medicine at McGill University

OICR Ontario Institute for Cancer Research

Wellcome Sanger Institute

Supporting Funders

HDRUK Health Data Research UK

illumina

Assigned Expert Funders and Employers

BROAD INSTITUTE

Epic

Genomics England

Matchmaker Exchange

NIH National Human Genome Research Institute

Nationwide Children's

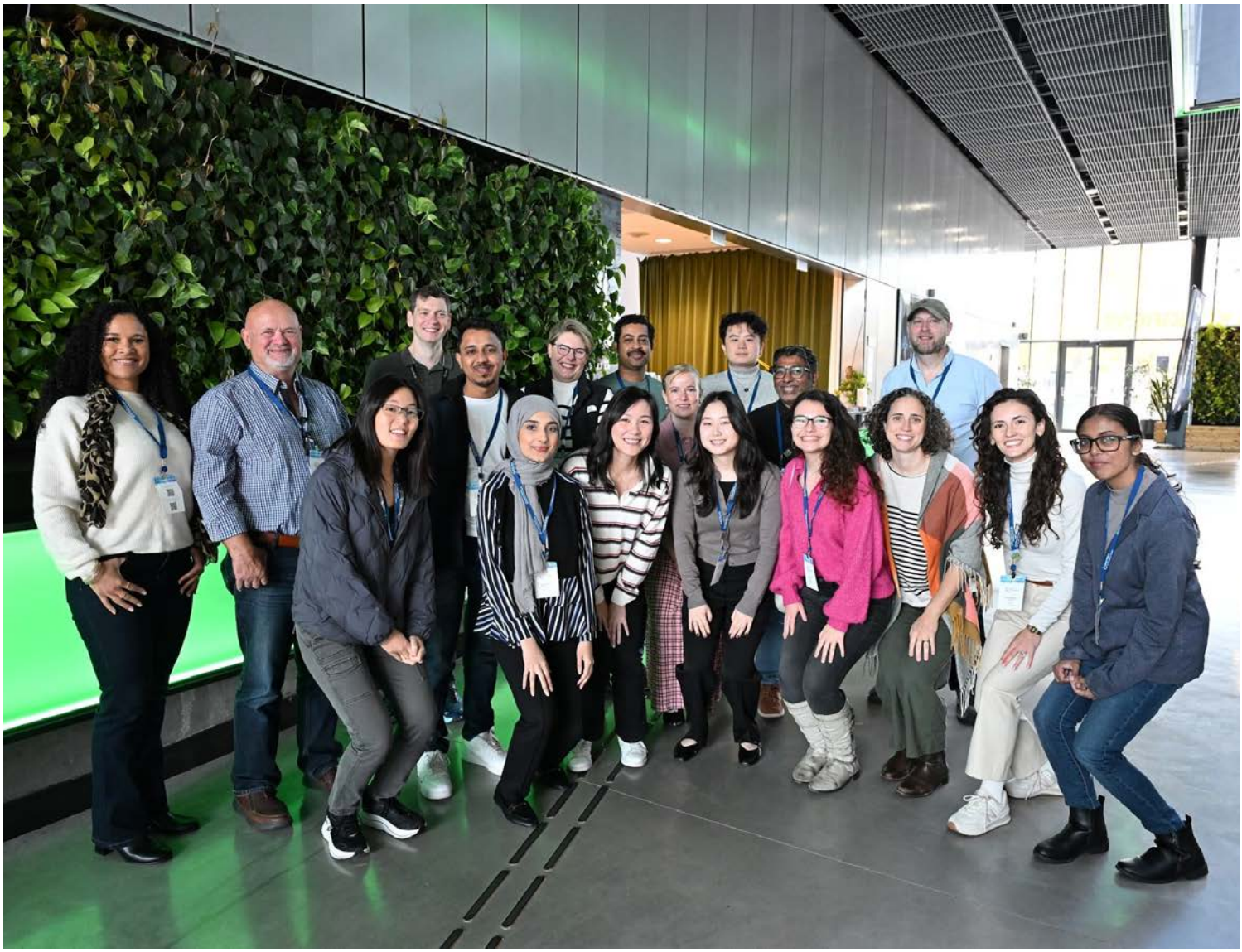
University of Cape Town

The University of Melbourne

VICC

Thank you to GA4GH staff

We are fortunate to have a dedicated team that drives the daily activities and sustains the forward momentum of the Work Streams, Communities of Interest, and broader organisational operations. So much behind-the-scenes work is thoughtfully and strategically managed by GA4GH staff, and we thank each of you for continuing to coordinate, contribute, and advance the work of our organisation.



Appendix

Academic papers

Bridges, Yasemin, et al. “Towards a standard benchmark for phenotype-driven variant and gene prioritisation algorithms: Pheval - phenotypic inference evaluation framework.” *BMC Bioinformatics*, vol. 26, no. 1, 22 Mar. 2025, https://doi.org/10.1186/s12859-025-06105-4.

Chen, Weijie, et al. “Multimodal Data Curation via interoperability: Use cases with the medical imaging and Data Resource Center.” *Scientific Data*, vol. 12, no. 1, 1 Aug. 2025, https://doi.org/10.1038/s41597-025-05678-2.

Chimirri, Leonardo, et al. “Consistent performance of large language models in rare disease diagnosis across ten languages and 4917 cases.” *eBioMedicine*, vol. 121, Nov. 2025, p. 105957, https://doi.org/10.1016/j.ebiom.2025.105957.

Connolly, John J., et al. “The International Health Cohorts Consortium (IHCC) advances population health research and Genomic Discovery.” *Communications Medicine*, vol. 5, no. 1, 21 Aug. 2025, https://doi.org/10.1038/s43856-025-01026-y.

D’Altri, Teresa, et al. “The Federated European Genome–Phenome Archive as a global network for Sharing Human Genomics Data.” *Nature Genetics*, vol. 57, no. 3, Mar. 2025, pp. 481–485, https://doi.org/10.1038/s41588-025-02101-9.

Graefe, Adam S., et al. “An ontology-based rare disease common data model harmonising international registries, FHIR, and phenopackets.” *Scientific Data*, vol. 12, no. 1, 8 Feb. 2025, https://doi.org/10.1038/s41597-025-04558-z.

Graefe, Adam S. L., et al. “RareLink: Scalable Redcap-Based Framework for Rare Disease Interoperability Linking International Registries to FHIR and Phenopackets.” *Nature News*, Nature Publishing Group, 18 Nov. 2025, www.nature.com/articles/s41525-025-00534-z.

Howley, Caitlin, et al. “The expanding Global Genomics Landscape: Converging Priorities from national genomics programs.” *The American Journal of Human Genetics*, vol. 112, no. 4, Apr. 2025, pp. 751–763, https://doi.org/10.1016/j.ajhg.2025.02.008.

Joly, Yann, et al. “A call for a global cyberbiosecurity framework in Genomics.” *Nature Genetics*, vol. 58, no. 1, 10 Dec. 2025, pp. 9–12, https://doi.org/10.1038/s41588-025-02447-0.

Lawson, Jonathan, and Thomas Conner. *GA4GH Data Passports: Proposing a New Technical Specification for Data Policy Representation*, 2025, https://doi.org/10.2139/ssrn.5372874.

Nahal-Bose, Hardeep K., et al. “The ICGC Argo Data Dictionary for Standardizing Global Cancer Clinical Data.” *Scientific Data*, vol. 12, no. 1, 20 Nov. 2025, https://doi.org/10.1038/s41597-025-06068-4.

Rekerle, Lauren, et al. “GA4GH phenopacket-driven characterization of genotype-phenotype correlations in mendelian disorders.” *The American Journal of Human Genetics*, vol. 113, no. 1, Dec. 2025, pp. 57–70, https://doi.org/10.1016/j.ajhg.2025.12.001.

Rye, Claire E., et al. “Rakeiora Genomics Platform: A Pathfinder for Genomic Medicine Research in Aotearoa New Zealand.” *Journal of the Royal Society of New Zealand*, vol. 55, no. 6, 24 Mar. 2025, pp. 2481–2505, https://doi.org/10.1080/03036758.2025.2469626.

Thaldar, Donrich, et al. “Communicating clearly about data sharing in genomics.” *Human Genomics*, vol. 19, no. 1, 14 July 2025, https://doi.org/10.1186/s40246-025-00784-z.

Zhao, Hangjia, and Michael Baudis. “*pgxrpi*: An R/bioconductor package for user-friendly access to the Beacon V2 API.” *Bioinformatics Advances*, vol. 5, no. 1, 26 Dec. 2024, https://doi.org/10.1093/bioadv/vbaf172.

Guest blog posts

Beauvais, Michael J. S. “Policy Brief: children’s data protection and genomic research (part 1: general considerations).” *The Global Alliance for Genomics and Health*, 23 January 2025.

Beauvais, Michael J. S. “Policy Brief: children’s data protection and genomic research (part 2: consent and lawful bases).” *The Global Alliance for Genomics and Health*, 4 February 2025.

Elphinstone, Brad. “Public Attitudes for Genomic Policy Brief: The Role of the ‘commercialisation effect’ in shaping Australians’ attitudes towards a hypothetical national genomic repository.” *The Global Alliance for Genomics and Health*, 30 October 2025.

Karlsson, Richard. “Public Attitudes for Genomic Policy Brief: The perceived risks of sharing genomic data with researchers.” *The Global Alliance for Genomics and Health*, 22 July 2025.

Mourby, Miranda. “Policy Brief: will the UK participate in the European Health Data Space?” *The Global Alliance for Genomics and Health*, 17 June 2025.

Smith, Lindsay. “Borderless data, boundless hurdles: navigating the challenges of international federated sharing.” *The Global Alliance for Genomics and Health*, 18 September 2025.

Soley, Gabriela Chavaría, Redondo, Ronald Alfaro, and Raventós, Henriette. "Individual attitudes toward medical and genetic data sharing in Latin America and development of two new indexes." *The Global Alliance for Genomics and Health*, 20 November. 2025.

Wang, Zhangyu, Wang, Meng, and Du, Li. "Public Attitudes for Genomic Policy Brief: public perceptions of international genetic information sharing for biomedical research in China." *The Global Alliance for Genomics and Health*, 16 January. 2025.

Willis, Amanda. "Older research participants are motivated to receive genetic results for the benefit of younger relatives." *The Global Alliance for Genomics and Health*, 9 December 2025.

Webinars hosted by GA4GH

In 2025, the GA4GH community led and participated in workshops and webinars to drive advancement of GA4GH products, as well as become more embedded in the GA4GH community. Alongside the Work Stream Open House meetings and GIF "Ask Me Anything" sessions highlighted earlier in the Annual Report, GA4GH also hosted:

Fireside Chat with Serena Scollen and Tim Hubbard**, 11 July 2025**
GA4GH Fireside Chats bring together experts in the genomics and health space. This conversation featured Tim Hubbard, Director of ELIXIR, and Serena Scollen, Head of Human Genomics and Translational Data at ELIXIR.

GA4GH 101 Webinar: Building a Global Framework for Precision Medicine**, 25 September 2025**
Angela Page, GA4GH Chief Strategy and Engagement Officer, and Sasha Siegel, GA4GH Chief Product Officer, provided an overview of GA4GH, our products and tools, and opportunities to get involved in the community.

Global Alliance
for Genomics & Health