

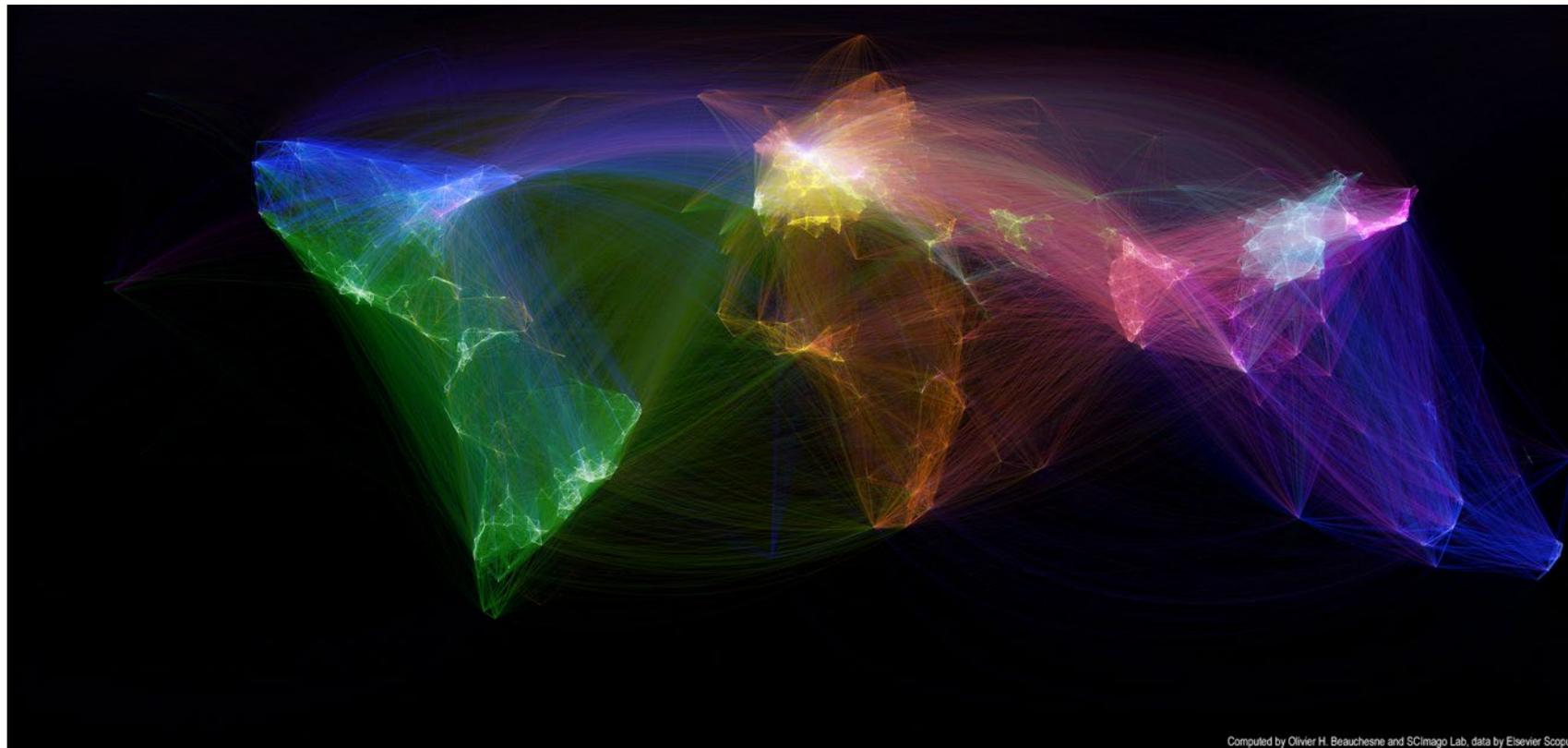


**Global Alliance**  
for Genomics & Health  
Collaborate. Innovate. Accelerate.

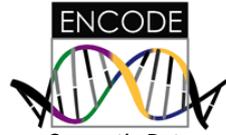
# The World of Human Genomics is Changing

Peter Goodhand  
CEO

# Map of Scientific Collaborations from 2008-2014



# International Collaboration and Data Sharing over 30 years



Consortia Data Release, Data Use, and Publication Policies (2009)



National Institutes of Health  
NIH Genomic Data Sharing Policy (2014)



Human Genome Project (1990)



Bermuda Principles (1996)



1000 Genomes Project (2008)



NHGRI

Policy for Release and Database Deposition of Sequence Data (2000)



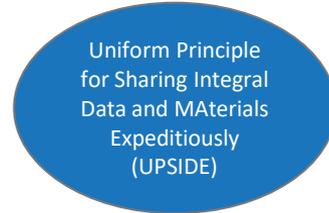
International Cancer Genome Consortium (2008)



National Institutes of Health  
Policy for Sharing of Data Obtained in NIH Supported or Conducted Genome-Wide Association Studies (2007)



Personal Genome Project (2005)



(2004)



International HapMap Project (2002)

# Challenges in 2013

---



Global Alliance  
for Genomics & Health

- Data typically in silos: by type, by disease, by country, by institution
- Analysis methods are non-standardized, few at scale
- Approaches to regulation, consent and data sharing limit interoperability
- Research and healthcare not aligned

**Don't act:** an overwhelming mass of fragmented data, such as electronic medical records in many countries

**Collective Action:** achieve the interoperability of the www or global telecommunications / smartphones

# Key Milestones



Global Alliance  
for Genomics & Health

2013

Planning

- First meeting in New York where idea for GA4GH is conceived
- [White Paper](#) is published describing the need for genomic data sharing

2014 - 2016

Building  
Momentum

- GA4GH formal launch
- Explore opportunities in genomic data sharing with 4 Working Groups
- Publish the *Framework for Responsible Sharing of Genomic and Health - Related Data*
- Develop three demonstration projects

2016 - 2018

Strategic  
Shift

- Ewan Birney joins as the 3rd Chair of GA4GH
- Launch the next strategic phase, “**GA4GH Connect**”
- Release the first four standards under the new model: htsget, refget, WES, and Beacon API



The Global Alliance for Genomics and Health aims to accelerate progress in genomic science and human health by developing standards and policies for responsible and secure genomic and health-related data sharing.

## GA4GH aims to...



Enable **international** data sharing



Promote sharing across the **translational continuum**



Encourage technology-enabled **federated approaches**



Promote **interoperability**



## GA4GH achieves this by...

- **Creating and maintaining** standards and protocols for data exchange
- **Convening** stakeholders
- **Catalyzing** sharing of data
- **Acting** as a clearinghouse
- **Fostering** innovation
- **Committing** to responsible data sharing

# The GA4GH Ecosystem



Global Alliance  
for Genomics & Health



## Global Alliance members include:

- Universities and research institutes (22%)
- Academic medical centers and health systems (10%)
- Disease advocacy organizations and patient groups (4%)
- Consortia and professional societies (13%)
- Funders and agencies (5%)
- Life science and information technology companies (46%)

# How GA4GH Works



Global Alliance  
for Genomics & Health

		Real-World Driver Projects									
Technical Work Streams	Discovery	✓		✓		✓		✓			
	Large-Scale Genomics		✓		✓		✓		✓		
	Data Use & Researcher IDs	✓		✓		✓	✓				✓
	Cloud		✓	✓						✓	
	Genomic Knowledge Standards		✓				✓	✓	✓		
	Clinical & Phenotypic Data Capture	✓			✓	✓	✓				✓
Foundational Work Streams	Regulatory & Ethics	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Data Security	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Partner Engagement

# Work Streams vs. Driver Projects

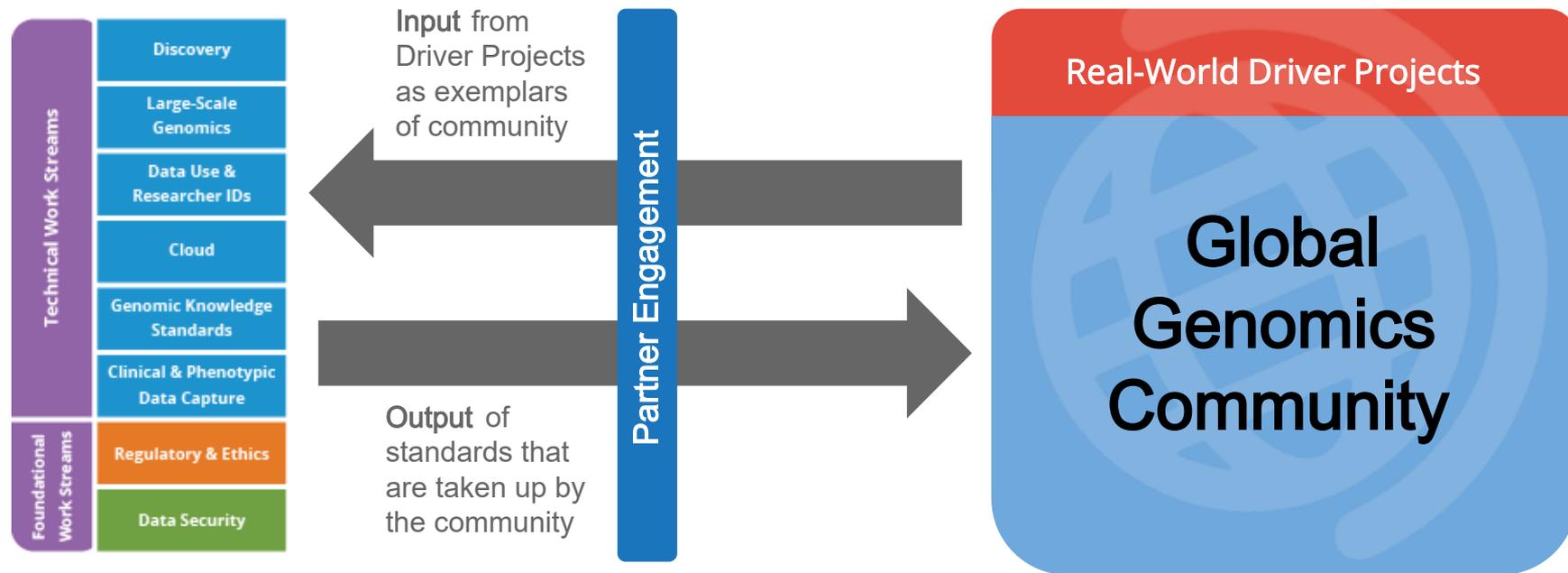


Work Streams	Driver Projects
<ul style="list-style-type: none"><li>• Internal to GA4GH</li><li>• Deliver standards and policy frameworks based on the Strategic Roadmap</li><li>• Run by 2 volunteer leads within the community</li><li>• Contributors come from a variety of projects and organizations</li></ul>	<ul style="list-style-type: none"><li>• External to GA4GH</li><li>• Provide input towards the Strategic Roadmap and standards development</li><li>• Contribute FTE resources to Work Streams for standards development</li><li>• Pilot implementations for new standards</li></ul>

# GA4GH and the Community



Global Alliance  
for Genomics & Health





## Findable

- Beacon API
- Data Use Ontology
- refget API
- Search API
- Service Registry Prototype
- Tool Registry Service (TRS)

## Accessible

- Authentication and Authorization Infrastructure
- Data Repository Service
- Data Use Ontology
- Researcher ID & Bona Fide Status

## Interoperable

- Phenotype Representation
- Phenopackets/FIHR
- Pedigree Representation
- Genetic variant file formats
- Read file formats
- RNASeq Expression Matrix
- RNASeq API
- Crypt4GH
- Variant Annotation
- Variant Representation
- Task Execution Service
- Testbed interoperability demonstration
- Tool Registry Service
- Workflow Execution Service

## Reusable

- htsget Streaming API
- refget API
- Variant Annotation
- Workflow Execution Service
- Testbed interoperability demonstration

# GA4GH 2019 Driver Projects



Global Alliance  
for Genomics & Health



All of Us Research Program  
United States



Human Cell Atlas  
International



EU CAN  
International



Australian Genomics  
Australia



ICGC-ARGO  
International



Autism Sharing Initiative  
International



BRCA Challenge  
International



Matchmaker Exchange  
International



EpiShare  
International



CanDIG  
Canada



Monarch Initiative  
International



GEM Japan  
Japan



ClinGen  
United States



National Cancer Institute (NCI)  
United States



Swiss Personalized  
Health Network  
Switzerland



ELIXIR Beacon  
Europe



TOPMed  
United States



European Joint Program  
For Rare Diseases  
Europe



ENA/EVA/EGA  
Europe



VICC  
International



H3Africa  
Pan-Africa



Genomics England  
United Kingdom

# Framework for Responsible Sharing of Genomic and Health -Related Data



Global Alliance  
for Genomics & Health

## Aims

- Foster responsible data sharing
- Protect and promote the welfare, rights, and interests of groups and individuals who donate their data
- Provide benchmarks for accountability
- Establish a framework for greater international data sharing, cooperation, collaboration, and good governance
- Serve as a dynamic instrument

## Foundational Principles

- Respect Individuals, Families and Communities
- Advance Research and Scientific Knowledge
- Promote Health, Wellbeing and the Fair Distribution of Benefits
- Foster Trust, Integrity and Reciprocity

The screenshot shows the title page and introductory sections of the document. At the top center is a purple circular icon with a white scale of justice. Below it, the title "Framework for Responsible Sharing of Genomic and Health-Related Data" is displayed in bold black text. Underneath the title, it says "Last Updated: 9 DECEMBER 2014". On the left side, there is a "Table of Contents" section with a list of items: Preamble, Purpose and Interpretation, Application, Foundational Principles, Core Elements for Responsible Data, Sharing, Implementation, Mechanisms and Amendments, Acknowledgements, and Appendix 1. Below the table of contents is a "CONTRIBUTORS" section. Further down is a "Downloads" section with a blue button that says "DOWNLOAD PDF". Below that is a "Download in other languages" section with a list of languages: Arabic / Chinese / French / German / Greek / Hindi / Italian / Japanese. On the right side, there is a "Preamble" section with the heading "Download in other languages" and two paragraphs of text. The first paragraph discusses the importance of data sharing for biomedical research and the role of the framework. The second paragraph describes the framework's development under the auspices of the Global Alliance for Genomics and Health and its mission to accelerate progress in human health.

## Examples of other genomic data sharing initiatives

- International 100K+ Cohort Consortium
- International Common Disease Alliance
- UK – Biobank
- European Million Genomes
- CINECA – European and Canadian Cohorts
- US - Million Vets Program
- W.E.F. - “Breaking Down Barriers” - health data
- Sweden – Clinical > research
- Denmark – Clinical > research

# Genomics data and health, in Canada...



**Global Alliance**  
for Genomics & Health  
Collaborate. Innovate. Accelerate.



Matchmaker  
Exchange



**CARE**  
for**RARE**



*All for One*



- Centralized Public (open) Genomic Knowledge bases
- Data Safe Havens
- Data Commons – trusted, controlled single repository of multiple data sets and compute resources
- Hub and Spoke federation – common data elements, structures, access and use rules
- Linkage of distributed and disparate data sets

# 150+ Genomic Data Initiatives Globally



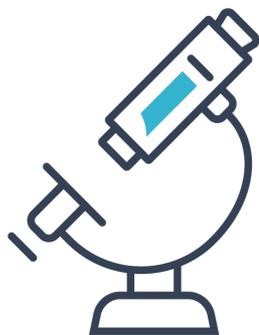
Global Alliance  
for Genomics & Health



Clinical/Genomic  
Medicine

40

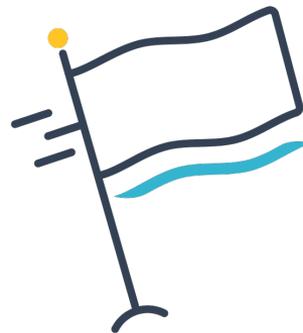
Initiatives



Research

70

Initiatives



National

21

From 15 Countries



Cohorts

64

Globally

# Research to Medical

## Research

- English as language
- Lightweight legal
- Identical/similar systems
- Open data
- Publications
- Grant-funding



## Practicing Medicine

- National language
- Heavy legal framework
- Very different systems
- Closed data
- Not published
- Contract-funding



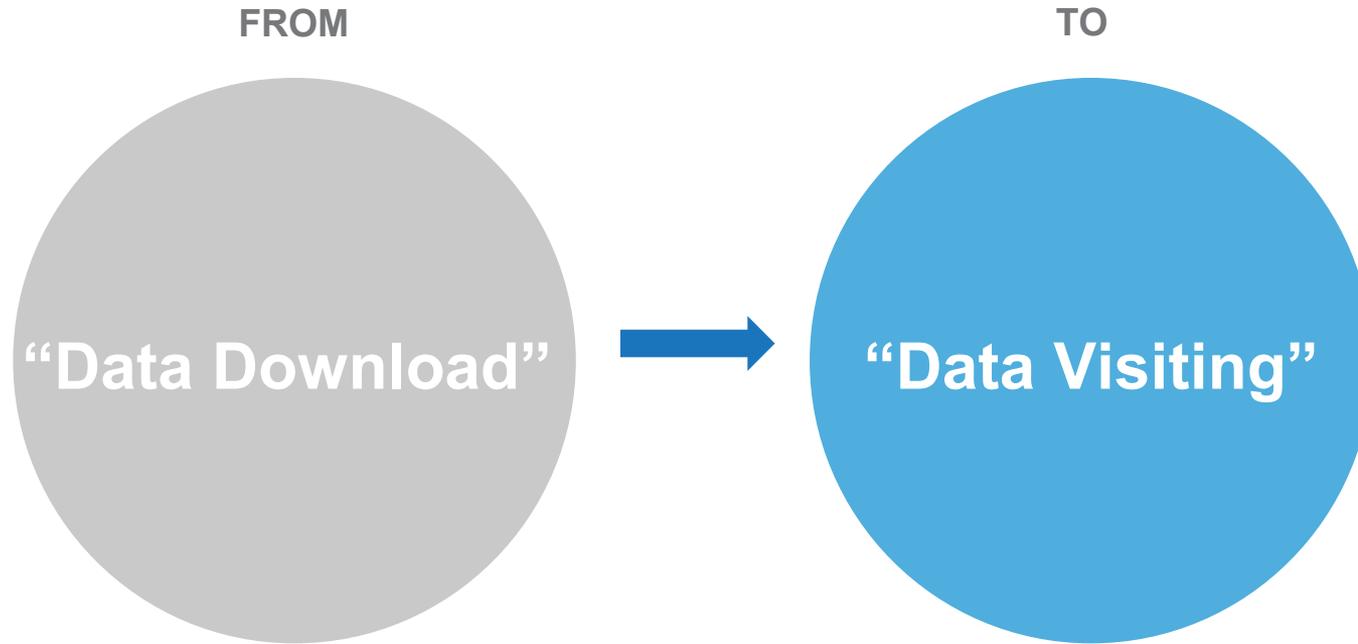


*“A grouping of autonomous organisations and datasets with a centralised control”*

Federation allows us to....

1. **Move analysis to data** , not aggregate data close to each researcher
2. **Have broad, reciprocal data access methods** which respect national processes and patient consent
3. Transfer methods and skills into the **healthcare sector**
4. **Leverage healthcare data** to make more discoveries on humans

# A New Data Sharing Paradigm



# Alignment of Standards Across the Continuum

