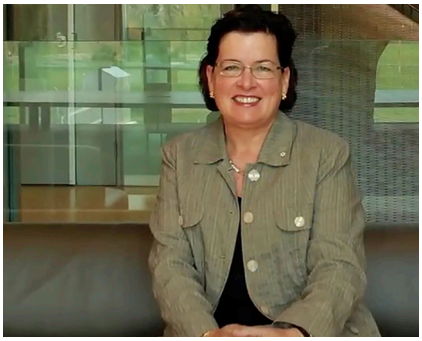
Cell Genomics

CellPress
OPEN ACCESS

Voices

Voices of GA4GH members: Collaborating in technology and policy development

The Global Alliance for Genomics and Health (GA4GH) is organized around eight Work Streams, where members develop technology standards and policy frameworks to enable the responsible sharing of human genomic and health-related data. These are implemented across 24 Driver Projects, reflecting real-world genomics initiatives. In this Voices, *Cell Genomics* asked GA4GH members to reflect on their engagement with GA4GH and how this has driven progress in open science and genomic medicine.



Bartha Maria Knoppers, PhD
McGill University Center for Genomics and Policy; Founding Work Stream co-lead, Regulatory and Ethics

A human rights framework

From the first planning meeting in snow-bound New York on January 28, 2013, to the visible excitement of the last in-person meeting in Boston, the vision for GA4GH and its realization is forever associated with shared commitment. For its socio-ethical and legal policy adherents, the building of the *Framework for the Responsible Sharing of Genomic and Health-Related Data* and its subsequent translation into 14 languages was a moonshot start. Policy wonks like to discuss. At one time, even the title of the Framework was an issue. Nevertheless, its content was achieved within 1 year, largely due to the work of Edward Dove (now co-chair of the Regulatory and Ethics Work Stream [REWS]). The leaders of GA4GH stood behind the *Framework*—not a mere decoration—by asking all applicants for membership to agree to it. International policy and tools were also built to translate the *Framework* into practical guidance. Sometimes "discussions" went on for a number of years as was the case for the "Return of Results" policy. The bridge from genomic research data to the other side—that is, to its possible use in the clinic—crosses a wide socio-cultural and ethical normative chasm. Now the REWS has completed a policy with which it should probably have started (*mea culpa*), one on public engagement and another is coming on equity, diversity, and inclusion. Irrespective, good faith and international participation and excitement in the "cause" remain high. Pandemic or not, a true global policy community exists. A face-to-face meeting in the snow in New York—better yet, Montreal—anyone?

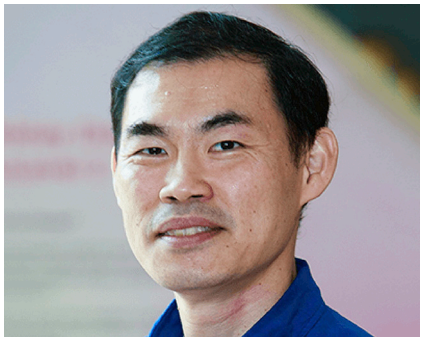
Danya Vears, PhD
Biomedical Ethics Research Group, Murdoch Children's Research Institute; Work Stream contributor, Regulatory and Ethics

Striving to make a difference through policy

I had just started my first postdoctoral research position at the Centre for Biomedical Ethics and Law, KU Leuven, Belgium, when I was invited to join the GA4GH REWS Pediatric Task Team in June 2015. I didn't know much about GA4GH at the time, but I was thrilled by the invitation to join an international network so soon after the completion of my PhD. What really excited me, however, was the opportunity to contribute to the development of policy that would have real world impact—something I feel very passionate about.

My experiences with GA4GH, to date, have certainly not disappointed. Within the Pediatric Task Team, under the leadership of Profs. Martina Cornel and Jan Friedman, I was involved in the development of "points to consider" documents relating to genomic newborn screening and pediatric data sharing. Within the Participant Values Task Team, I assisted Prof. Anna Middleton and her team in the development of the Your DNA, Your Say survey, which asked a range of stakeholders to share their views on whether they would donate their data and for what purposes. More recently, within the Return of Results Task Team, I led a systematic review on stakeholder perspectives on the return of results from genomic research. The results of this review helped inform the GA4GH 2021 Policy on Clinically Actionable Genomic Research Results, which was led by Dr. Anna Lewis and Profs. Robert Green and Bartha Knoppers. The findings are also informing the development of a framework to assist researchers in determining

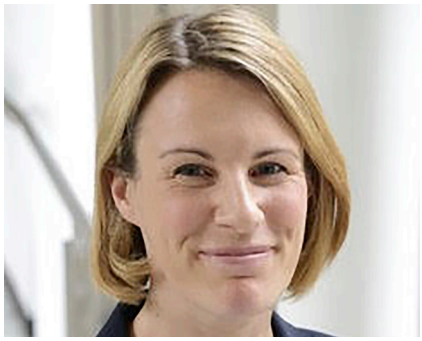
how to plan for the return of results, which will mean return practices are more consistent and ethically appropriate. I am incredibly grateful to have had the opportunity to work with so many wonderful colleagues over the past 6 years and look forward to continuing to contribute to the amazing work GA4GH is undertaking.



Calvin W.L. Ho, JSD, MSc, LLM, FRSPH
University of Hong Kong; Work Stream contributor, Regulatory and Ethics; subgroup co-lead, Accountability Policy

A call to engage: Consensus building from the bottom up

In 2020, the Hong Kong Genome Institute (HKGI) was established by the territory's government to implement the Hong Kong Genome Project (HKGP), an initiative to promote the development of genomic medicine and population health in Hong Kong. In setting up HGKI operational procedures, key issues included feedback of whole-genome sequencing results to participants, right of participants to withdraw, data security and privacy protection, data access mechanisms, and use of data. For this, we consulted the GA4GH Regulatory & Ethics Toolkit, which has been an important ready-to-use reference for internationally accepted standards on responsible sharing of genomic and health-related data. For instance, discussions and recommendations on setting up infrastructure to support the flow of data from clinical practice into research and establishing data access and accountability mechanisms that are appropriate to research settings are implicit in the considerations of the expert committee and likely to be evident in the policies and practices of the HKGI. Some of my colleagues and I have the double privilege of contributing to the work of GA4GH and the HKGI and have thereby been able to facilitate the cross-pollination of values, concepts, and practices. As an intrinsically open and collaborative enterprise, GA4GH provides the forum, mechanisms, and resources for all interested stakeholders to be involved and in a manner that is non-directive. The consolidation of international policies and standards of the GA4GH in this issue of *Cell Genomics* is a further step in this direction and is, in my view, exemplary of consensus building from the bottom up.

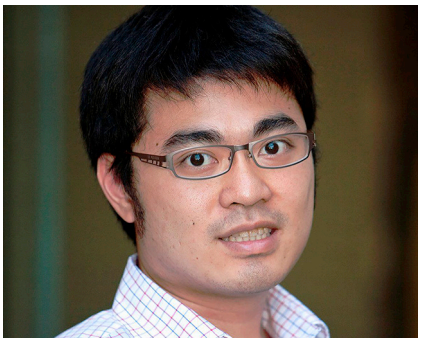


Anna Middleton, MSc, PhD
Wellcome Connecting Science, Wellcome Genome Campus, Cambridge; subgroup lead, Participant Values, Regulatory and Ethics Work Stream

Including public voices

Behind every genomic data point is a person who has thoughts and feelings about the way their data are used—this was the driving force behind suggesting that we set up the Your DNA, Your Say study. This project aimed to collate views from public audiences around the world so that the work of GA4GH could incorporate public attitudes, values, and beliefs into policy. After an invitation through the Regulatory and Ethics Work Stream , we started with a small group of social scientists and genetic counsellors all connected to GA4GH work, and the project has since expanded beyond and out of GA4GH and now has a collection of collaborators from 23 countries around the world. It has been so inspirational to lead a group of researchers, all very passionate about working directly with public audiences and all spearheading a sense of justice that an organization as large as GA4GH must have a public voice in its work. Together, we wrangled with really interesting questions, e.g., how to translate "DNA data" into languages that do not currently have these words, e.g., Twi and Ewe (two Ghanaian languages), and how to explain the complexities of genomic data sharing using metaphor and film. We now have 17 different languages of the survey and have collected 37,000 completed responses from representative public audiences across the globe and have shown that public audiences are very unfamiliar with the concepts of DNA, genetics, and genomics, and trust in the sharing of genomic data beyond our own doctor is very low. All of us within the Your DNA, Your Say project feel strongly that the aim of genomics is to serve society, and for this to be true, genuine partnership with society has to exist. This research shows that so much more needs to be done to bring public audiences on the journey to deliver the benefits of genomics, and as a social scientist and genetic counsellor myself, I'm always keen to connect with scientists and policy makers around the world who want to hear what public audiences think about the implications and application of genomic technology.

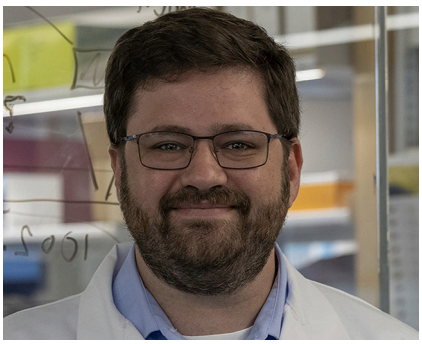
CellPress
OPEN ACCESS



Jacob Shujui Hsu, PhD
Graduate Institute of Medical Genomics and Proteomics, College of Medicine, National Taiwan University; Work Stream contributor, Equity, Diversity, and Inclusion

Finding supportive colleagues and increasing diversity

As a bioinformatician in human genome research, I first heard of GA4GH during a Curating the Clinical Genome workshop about exchanging genomic data for undiagnosed rare disease patients from different sources. I immediately realized how this platform would impact the world. GA4GH was born with a genuinely open science spirit; all events and materials are transparent. With rapid interdisciplinary collaboration as a community, GA4GH put efforts on aggregating all achievements together and providing community recommendations for many perspectives. The technical toolkits are open source, with collaborative efforts from international groups. It is beneficial to mitigate the inadequate academic resources for small academic groups, providing opportunities to follow common standards and to share local experiences. As a member from Taiwan, the Equity, Diversity, and Inclusion (EDI) working group provided sufficient opportunities for exchanging opinions at a monthly meeting with complete meeting minutes. They also set up an onboarding contact program to match newcomers with current members to minimize the onboarding barriers. Building infrastructure sometimes is not very credible in an academic environment, and so having the support from international colleagues was helpful. Currently, the publicly available genomic data from Asia are still underrepresented considering the population size worldwide. The transparency and standards developed by GA4GH are encouraging non-western societies to be more willing to share their data. In terms of data diversity and equity, this is a big step. I am so thankful to have met GA4GH in my early academic career and would love to grow together with GA4GH.



Alex Handler Wagner, PhD
Nationwide Children's Hospital and the Ohio State University; product co-lead, Variation Representation Work Stream; Driver Project Champion, Variant Interpretation for Cancer Consortium

Advancing genomic medicine through diverse perspectives

I first became involved in GA4GH through the Variant Interpretation for Cancer Consortium (VICC) Driver Project, as a postdoctoral researcher at Washington University in St. Louis. My research led me to participate in the GA4GH Genomic Knowledge Standards (GKS) Work Stream, where we are addressing challenges in genomic knowledge representation alongside other Driver Projects with shared purpose. Working within a GA4GH Work Stream is an engaging experience, where standards are developed through the shared challenges and diverse goals of the GA4GH Driver Projects.

My own experience with GA4GH highlights the commitment of the alliance to inclusion and engaging a diversity of perspectives. Despite my status as a newcomer and a relatively junior voice, I found that my contributions were welcomed and fairly considered during the development of the GKS standards; over time, I was even invited to co-lead the Variation Representation working group on the basis of these contributions. The GA4GH engagement with a diversity of perspectives helps us address global challenges in genomic and health data sharing and have made the GKS standards more robust to the myriad challenges in clinical genomic knowledge exchange. Today, as an independent investigator at the Institute for Genomic Medicine at Nationwide Children's Hospital, I am excited to see advancements in our ability to treat pediatric cancers driven by the Variation Representation Specification (VRS), reported in Wagner et al., in this issue (https://doi.org/10.1016/j.xgen.2021.100027), and emergent GKS specifications. I am hopeful for a future enabled by GA4GH, when our global collective knowledge of genomic variation and its role in clinical care are available to patients anywhere.

Mélanie Courtot, PhD
EMBL's European Bioinformatics Institute (EBI); product co-lead, Data Use Ontology

Building a global health knowledge ecosystem

Progress in multi-omics technologies has allowed us to generate big data; however, its usefulness in benefiting patients is severely limited by the inability to automatically understand and integrate it. To address this challenge, I build intelligent systems to leverage vast amounts of biomedical and biological data and answer research questions of importance to human health, biology, and society. Efforts such as GA4GH are critical not only in building specifications but also—and perhaps more importantly—in bringing communities together while training and mentoring junior colleagues.

GA4GH provides means to engage with a highly diverse group of people from all aspects of clinical and biomedical realms and a range of technologies and standards. My contributions to GA4GH have given me the opportunity to combine my love for knowledge representation with the expert input of diverse domain specialists. Leading the development of the Data Use Ontology has been a highlight: we built standards and tools to improve discovery and enable streamlined access to human controlled-access datasets, which have been adopted worldwide (see Lawson et al. in this issue; https://doi.org/10.1016/j.xgen.2021.100028). Seeing the direct impact of the products we develop is rewarding and motivating, and a strong testament to the GA4GH powerhouse: nowhere else could we as efficiently take new solutions from concept to impact via deployment at scale.

Braiding the strands of people, tools, and data together will realize the GA4GH vision toward "responsible sharing of genomics data to benefit human health" and contribute to building the global health knowledge ecosystem I dream of.

Melissa Haendel, PhD, FACMI
Center for Health AI, University of Colorado Anschutz Medical Campus; Work Stream co-lead, Clinical and Phenotypic Data Capture

Putting the patient back together again

As a basic science researcher by training, I have always sought to investigate the relationship between variation in the genome with phenotypic outcomes across developmental stages and environmental contexts to reveal mechanisms. This passion led me to create the Monarch Initiative, which aims to integrate genotype-phenotype data from across species to aid rare disease diagnosis. However, this has been near impossible in humans because of the siloization of genomic data from the clinical data that is necessary for interpretation of the genome. I began my GA4GH tenure at the 2nd Plenary meeting and became one of the few individuals involved in both the data and clinical working groups. I tried to translate clinical needs to the techy genomicists and possibilities for new kinds of answers to the clinical community. It was for this reason, I believe, that I was nominated to co-lead the Clinical and Phenotypic Work Stream. In this context, the GA4GH community has helped realize the dream of creating a "packet" of de-identified, lightweight, and computable set phenotypic features for use together with genomic data. The design of Phenopackets took 6 years! However, persistence and global partnerships have helped realize Phenopackets adoption in many different countries and for applications as diverse as rare disease diagnostics and biobanking. The value proposition for working together to "put the patient back together again" would not have been possible in any other community than GA4GH. I am honored to have served the GA4GH community for so many years and look forward to realizing the phenomics era together.

CellPress
OPEN ACCESS

Michael Baudis, MD, PhD
University of Zurich and Swiss Institute of Bioinformatics; Work Stream co-lead, Discovery; Driver Project Champion, SPHN & ELIXIR Beacon

Globally inclusive genomics

While the final approval of standards for the GA4GH ecosystem runs through a well-organized process, the way for individual projects to reach such a maturity is anything but. GA4GH works on the principle of voluntary contributions of time and expertise from scientists and practitioners in very diverse biomedical and computational fields, organized with the help of an essential, amazing core administrative team. Academic seniority plays a negligible role; continuing involvement, expertise in the topic at hand, and the ability to listen to—and accept—arguments made by your peers are most relevant.

As maintainer of a large open cancer genomics resource (http://progenetix.org), I was drawn to GA4GH because the current landscape of resources and research in this area is highly fragmented, does not reflect the whole range of relevant diseases, and lacks a balanced representation of populations from all over the world. The GA4GH principle of federated, globally inclusive approaches to genomic and health has proven successful, keeping me as an active contributor and leader in different parts of the organization—with a special interest in cross-Work Stream standards alignment—since its foundation. And I have learned a lot in the process.

Orion Buske, PhD
PhenoTips; product co-lead, Pedigree

Finding a home within a global community

My history with the GA4GH began in 2014 during my postgraduate work on patient similarity algorithms. Our database, PhenomeCentral, was one of several that had recently developed patient matchmaking systems to help researchers find additional families with the same rare condition, and the Matchmaker Exchange had just been founded to federate these efforts and exponentially increase the odds of finding matches. The systems needed a standardized application programming interfaces (API) to talk to each other, and we started working closely with GA4GH to ensure our API would be aligned with their early standards activities.

When I attended the 2nd GA4GH Plenary, I was immediately captivated by the openness and enthusiasm of the community "collaborating on interfaces and competing on implementations," as David Haussler would proclaim. I served as the technical liaison between the Matchmaker Exchange and the GA4GH until I completed my postgraduate studies. I then started a new role at PhenoTips, where our platform helps physicians collect and manage standardized phenotypic, genetic, and family history data. By this time, the GA4GH was like a family to me, so I set out to find a new home within the GA4GH that better aligned with my new role. The GA4GH's open culture made it easy to explore the various Work Streams and participate in as many calls as I could. After several months of exploration, the Clinical & Phenotypic Data Capture Work Stream quickly became my new home. When the need for a pedigree standard became clear, I jumped at the opportunity to co-lead this product with Grant Wood. This upcoming standard, like all the others at the GA4GH, owes its existence to the volunteer contributors from around the world and the support of its stellar secretariat. I am so grateful to be part of this wonderful community, and I look forward to seeing it and its impact continue to grow.

Andy Yates, MSc
EMBL's European Bioinformatics Institute (EBI); Work Stream co-lead, Genomic Knowledge Standards

Cross-pollination leads to better standards

GA4GH is a standards organization like no other and a truly collaborative organization. Ideas are allowed to cross pollinate between our focus areas and develop into fully fledged standards. One such idea was "service-info," a standard to enable the discovery and aggregation of services via computational methods. Service-info began life in the Cloud Work Stream and was later adopted by the refget specification to allow refget servers to describe their capabilities. Our secretariat team brought service-info to the attention of the Discovery Work Stream, which was developing a service registry standard and needed a mechanism for service implementations to describe themselves. A gap analysis performed at the time was unable to identify an existing standard that fulfilled this need. Service-info was a credible solution, and with additional input and development from GA4GH Driver Projects, we converted a custom solution into a standard. Any API that now implements the service-info standard can be aggregated and presented through service registry irrespective of them being a GA4GH standard. This underlines how GA4GH not only develops standards that serve to advance human health but can be used by any other domain.

Kathy Reinold, MSCS, MS Bioinformatics, PMP
Broad Institute of MIT and Harvard; Work Stream contributor, Data Use and Researcher Identities

Driving standards

What is most compelling about GA4GH is the validation of their standards through Driver Projects. Most often, standards don't lead, but, rather, they follow in the wake of innovation, which, through trial and error, establishes best practices. By convening collaborators and experts with a deep interest in the area of development and then putting the draft standards to use through Driver Projects, GA4GH is able to confirm that its standards are effective and relevant.

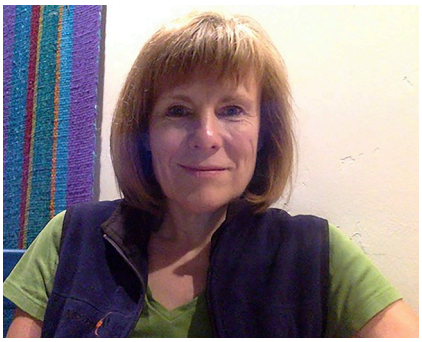
I first came across GA4GH when working in a large pharmaceutical company that was launching a major genomic center. To reuse data gathered across many clinical trials, the group sought a way to capture consent electronically to expedite data access requests. Fortunately, the Data Use Ontology (DUO) was well underway and served as a starting point for sharing data internally. Through the use of machine-readable requirements and the researcher's internal records, this group succeeded in automating parts of this process.

Jonathan Lawson
Broad Institute of MIT and Harvard; product co-lead, Data Use Ontology

A forum of experts to design and build the future

The exponentially increasing amount of human genomic data available for research and a similar increase in tools and researchers to analyze these data present us with two paths. Either we continue to allow once-analyzed data to rest in silos and be forgotten or we make these data accessible for compliant re-analysis with the latest breaking datasets and computational tools, driving innovative scientific discoveries that transform human health outcomes. Without an organization like GA4GH, our community is doomed to the former case, inefficiently leveraging data that could have driven scientific breakthroughs. The solution to avoid repeating this historical tragedy and maximizing the benefit of research requires global coordination in technology and policy. Software integrations and policy change are challenging enough in a single nation or institution, yet GA4GH remarkably created a forum in which community members nucleate conversations toward these solutions and then design and build them to unlock the vast potential for genomic data analysis at hand. Our work at the Broad Institute on the Data Use Oversight System (DUOS), reported in Cabili et al. in this issue (https://doi.org/10.1016/j.xgen.2021.100031), owes its success in part to GA4GH connections and ideas sparked in GA4GH meetings. This forum of passionate experts candidly discusses issues and collaboratively designs solutions with us, so we can make sure it meets the community's needs.

CellPress
OPEN ACCESS



Melissa Cline, MS, PhD
UC Santa Cruz Genomics Institute; Driver Project champion, BRCA Challenge

Driving toward federation

I originally joined GA4GH through my project. The BRCA Challenge Consortium was launched by GA4GH to improve the care of patients at risk of heritable cancers through international data sharing and collaboration. Our web portal, BRCA Exchange, was an early GA4GH Driver Project. Our codebase contains GA4GH Work Stream software whose functionality I personally helped motivate in some cases. When VRS was defined but not yet adopted, the Work Stream observed that Driver Projects, including mine, had a resourcing issue and provided a library, which I integrated promptly. We are now addressing the harder problem of aggregating evidence of variant pathogenicity. I sense great traction in federated analysis: analyzing sensitive patient-level data in its secure home environment, generating variant-level data that are less sensitive and can be shared more openly. While this offers new opportunities for sharing siloed data, it presents distinct challenges in diverse areas, ranging from cloud security to international regulatory compliance. As a leader of a small team, this can feel overwhelming. Yet as a member of this community, I benefit greatly from opportunities to exchange knowledge with experts across the realm, who are helping me define a path forward. This is why I remain involved in GA4GH.

Tiffany Boughtwood, MBA
Australian Genomics, Murdoch Children's Research Institute; Work Stream contributor, Data Use Ontology and Equity Diversity and Inclusion.

Joining a global community

Involvement in GA4GH has been an impactful, valuable process—both for me personally and for Australian Genomics. With the rapid advancements in health genomics internationally, the opportunity to come together as a global community to progress the development of common standards and approaches has enriched and accelerated our individual efforts. Further, participating in GA4GH Work Streams has provided the opportunity for nations like Australia to ensure that we're adopting international best practice and to further ensure that the standards developed are informed by the Australian experience and so will be fit for purpose for local health genomic implementation. From a personal and professional perspective, I have relished the opportunity to work with an extraordinary number of dedicated, talented people with expertise across the health genomic spectrum from all across the globe. The GA4GH should be celebrated for orchestrating and enabling this multidisciplinary, multinational collaboration, which progresses health genomics for the benefit of all people.

Alexander Kanitz, MSc, PhD
University of Basel & Swiss Institute of Bioinformatics; Driver Project champion, ELIXIR Cloud & AAI

Your contribution can make a difference

What impressed me most about GA4GH was how easy it was to get involved; how diverse the community is with respect to expertise, use cases, and domains; and how respectfully the contributors interact with one another. Among many other benefits, the Global Alliance has allowed us to propose projects as a Google Summer of Code organization, which led to us to being able to mentor a total of six students over the course of 3 years. These students have had a tremendous impact on our initiative growing as an open-source software community of its own and have helped us substantially to be able to set up a GA4GH Cloud-backed stack that enabled us to demonstrate, for the first time, how all four Cloud API standards could work together during the 8th GA4GH Plenary in 2020. The decentralized, international, and predominantly virtual mode of operation had also prepared us well for the COVID-19 pandemic, which, in turn, accentuates the critical importance of the Global Alliance's efforts to harmonize and facilitate data access and analysis across international boundaries.

Jordi Rambla, MSc
Centre for Genomic Regulation (CRG); Driver Project champion, ELIXIR Beacon

Volunteering as a driving force

The European Genome-phenome Archive (EGA) remit is to keep sensitive human biomolecular, phenotypic, and clinical data safe, secure, and available for the long term; fostering their reuse as much as possible while honoring the rules under which data have been consented. As our service is worldwide, having standards that help in expediting the submission, the access request, and approval process and an efficient consumption of the granted data are of great interest for the service. GA4GH is producing something for us in every aspect of such processes. Therefore, many EGA team members have been involved in different parts of the GA4GH product generation, and I, as EGA team lead, joined the effort in the initial GA4GH days. The initiative seemed promising given the broad distribution of people, organizations, and topics, and this caught my interest. I decided to join the Beacon project, which seemed a clever idea. Nowadays, I'm co-leading the version 2.0 specification. The volunteering willingness could be sensed everywhere, and that brings a positive spirit to the whole organization. It has ups and downs, as you struggle between work, volunteering tasks, and diverse time zones, but, at the end of the day, it proves worth the effort.