

# The Cloud Work Stream Roadmap 2026 and Beyond

**Work Stream / Subgroup: Cloud**

[SELECT A GA4GH GROUP](#)

**Date (Day/Month):** Friday, 10 October at 11:00 - 12:30 CEST

**Time (CEST):** 11:00 - 12:30 CEST

**Session chair(s):**

- Name **In-person** : [Brian D OConnor](#)
- Name **Virtual** : [Venkat Malladi](#)

**Link to slides:** [P GA4GH Connect slide - The Cloud Work Stream Roadmap 2026 and ...](#)

---

## Session description:

This session will focus on initiatives that the Cloud Work Stream will push forward over the next year. We will start the session by a check-in on our progress so far in 2025. And then open the discussion to explore how we evolve our current initiatives (data and federated compute) in 2026 and beyond. We will also examine specific API development goals and relationships with other Work Streams and initiatives, such as the GA4GH Implementation Forum (GIF).

## Meeting aims:

- Review how we did on our 2025 roadmap
- Evolve our roadmap and initiatives in 2026
- Set specific API development goals in 2026
- Discuss our relationship with other GA4GH efforts

## Resources and links:

- [W 2025 Cloud WS Roadmap.docx](#)
- [GA4GH Cloud Work Stream: Minutes and Actions \(Nov 2020-present\)](#)

## Agenda:

	Agenda item	Presented by	Item type	Time
1.0	Introduction	Brian & Venkat		5 min
1.1	How did we do on our 2025 roadmap initiatives? - High-level, accessible, interop vision			
2.0	Evolving our roadmap & initiatives for 2026	Brian & Venkat		5 min
3.0	API Goals in 2026 - Key few features each will work on with 2026 - <i>We can pause for brief <b>Questions</b> after each API</i>	Group		50 min
3.1	DRS	Brian / Michael		10 min
3.2	TRS	Denis		10 min
3.3	TES	Venkat / Kyle		10 min
3.4	WES	Patrick / Christopher		10 min
3.5	New APIs - MeshCards - PFB	API Champions Michael		10 min
4.0	<b>Discussion</b> - Next Steps for 2026 & Collaboration with other GA4GH groups - What <a href="#">initiatives</a> do we carry forward? - What new API goals do we set for 2026? - How does our work relate to trust and passports? AI efforts? - Who will <a href="#">drive 2026 initiatives forward</a> ? - What connection points do we need with others?	Brian & Venkat		30 min

	<ul style="list-style-type: none"> <li>- Collaboration on our APIs with Drivers and other Work Streams?</li> <li>- GIF and the Registry, how do we spotlight our work?</li> <li>- How does our work fit in with GA4GH AI/ML initiatives?</li> <li>- <i>And more!</i></li> </ul>			
--	---	--	--	--

**Zoom recording :**

<https://us02web.zoom.us/rec/share/h2rmbbb2ZmliAqizxxSaS3EQ6aCnBCbuH9prEtBlkSdXc2JLU0Us08uKsDfQilRd.VTDepkAXr3tG-ueY>

**Passcode: 2T\*+E66X**

**Summary:**

The Cloud Workstream has evolved from developing individual cloud abstraction APIs (TRS, WES, TES, DRS) into advancing an integrated ecosystem supporting FAIR data discovery and democratized compute. In 2025, key achievements include the expansion of the DRS network to 23 servers hosting 72 PB of data, the launch of the Federated Compute Initiative, and updates to DRS 1.5, TRS, and WES to enhance interoperability and cloud integration. Emerging efforts such as Mesh cards and PFB aim to strengthen data and metadata federation across systems. Strategic themes emphasize federated compute, standardized access control, fair-use principles, and AI-based approaches for data sharing and reproducibility. Next steps focus on publishing DRS and WES papers, developing PFB specifications, hosting open house and hackathon events, defining the Federated Compute and Mesh initiatives, and fostering cross-stream collaboration on access control, ontology, and cost-sharing models.

## **High level key takeaways (5 overarching points from the key takeaways)**

### **· Transition from API Development to an Integrated Cloud Ecosystem**

The Cloud Workstream has evolved from its early focus on developing core APIs—TRS, WES, TES, and DRS—into a cohesive framework that integrates these standards into a functional ecosystem supporting seamless data and workflow interoperability.

### **· Strategic Priorities for 2025 and Beyond**

The Workstream is advancing two major goals: promoting FAIR data discovery and access through DRS, and democratizing compute via GA4GH APIs. Initiatives such as the Federated Data Initiative, which now connects 23 servers and 72 PB of data, exemplify this forward-looking focus on scalable and federated infrastructures.

### **· Ongoing Development and API Interoperability**

Continued enhancements across APIs strengthen the ecosystem—DRS 1.5 introduces richer metadata and self-reporting features, TRS expands support for multiple workflow languages and AI integration, WES emphasizes interoperability and validation tools, and TES aligns with broader compute management objectives.

### **· Emerging Standards and Expanded Interoperability**

New tools such as Mesh cards and the Portable Format for Biomedical Data (PFB) are enabling more flexible and interoperable data exchanges across federated systems. Plans to formalize PFB specifications and encourage wider adoption signal growing maturity and standardization within the Cloud Workstream.

### **· Collaboration, Community Engagement, and Cross-Stream Integration**

The Workstream emphasizes inclusive collaboration through open houses, hackathons, and joint sessions, while exploring cross-cutting themes such as federated compute, standardized access control, fair use principles, and AI-based approaches to data sharing. Partnerships with new contributors, including those working on semantic web and AI technologies, further strengthen integration across the GA4GH ecosystem.

## Key takeaways:

### 1. Cloud Workstream Overview and Evolution

- The Cloud Workstream began eight years ago to develop cloud abstraction APIs — TRS, WES, TES, and DRS — which became GA4GH standards around four to five years ago.
- The focus has now shifted from developing individual APIs to using them together as an integrated ecosystem.

### 2. 2025 Achievements and Strategic Focus

- Two central initiatives for 2025:
  - FAIR data discovery and access through DRS.
  - Democratizing compute via GA4GH APIs.
- The DRS map now includes 23 servers and 72 PB of data.
- The Federated Compute Initiative launched with a bi-monthly working group and a dedicated Slack channel.

### 3. API-Specific Updates

- DRS 1.5 released, adding cloud metadata, storage status, and self-reporting.
  - 2026 goals: DRS publication, registry collaboration, and geolocation support.
- TRS received maintenance updates and seeks new collaborations, including AI integration and support for multiple workflow languages.
- WES is focusing on interoperability with DRS and TRS, validation tools, and a planned community paper (potentially through a write-a-thon).
- TES aligns with broader compute resource management goals.

### 4. Emerging Standards and Interoperability

- Mesh cards describe federated systems of repositories; an implementation exists in Gen3.
- PFB supports tabular data interoperability and is used in Gen3 and other systems.
- Plans include creating formal PFB specifications and expanding community engagement.

### 5. Engagement and Collaboration

- Open houses and GIF spotlights proposed to welcome newcomers and celebrate progress.

- Hackathons and collaborative sessions emphasized as key to innovation.
- Champions are needed to lead specific initiatives and maintain momentum.

## 6. Strategic and Cross-Stream Themes

- Growing focus on federated data and compute ecosystems and a potential mesh initiative.
- Proposal to adapt self-assessment tools for cloud repositories.
- Exploration of data stewardship and cost-sharing, including compensation models for data producers.
- Recognition of the need for standardized access control and data-use ontologies across APIs.
- Interest in developing fair use principles and AI-based scoring to improve data sharing and reproducibility.

## 7. Collaboration with New Contributors

- Engagement with Valmar and others exploring semantic web and AI technologies that complement existing GA4GH infrastructure.

### **Next steps:**

#### **1. Publications and Specifications**

- Draft and submit DRS and WES papers.
- Develop and circulate PFB specifications for community review.

#### **2. Community Engagement**

- Launch open house sessions and AMA (Ask-Me-Anything) events.
- Organize WES write-a-thon and hackathons focused on federated compute and interoperability.

#### **3. Initiative Development**

- Define the structure and goals for Federated Compute and Mesh Ecosystem Initiatives.
- Identify champions to lead and coordinate efforts.

#### **4. Cross-Stream Collaboration**

- Collaborate with AI and Data Use & Governance workstreams on access control, ontology, and fair-use scoring.
- Coordinate with registry and metadata teams for DRS geolocation and integration.

#### **5. Technical Enhancements**

- Extend TRS to support AI models and multiple workflow languages.
- Build validation services for WES and improve interoperability testing.

- Develop frameworks for managing compute and storage usage across APIs.
- 6. Long-Term Exploration**
  - Evaluate cost-sharing and incentivized data contribution models.
  - Refine self-assessment tools for cloud repositories.
  - Explore multi-use ontologies for fine-grained data access control.

## **Meeting minutes:**

### **History and Overview of the Cloud Work Stream**

- Brian provides a brief history of the Cloud Work Stream, starting eight years ago with the need for cloud abstraction APIs.
- Key APIs mentioned include the Tool Registry Service (TRS), Workflow Execution Service (WES), Task Execution Service (TES), and Data Repository Service (DRS).
- These APIs became GA4GH standards about four or five years ago.
- Brian emphasizes the importance of using these APIs together rather than individually.

### **2025 Initiatives and Achievements**

- Brian outlines the two main initiatives for 2025: sharing and finding data with FAIR standards using DRS, and democratizing access to compute using GA4GH APIs.
- Venkat joined as co-lead in 2025, shifting the focus from individual API evolution to the integrated use of APIs.
- The DRS map was created, showing 23 DRS servers worldwide with a total data footprint of 72 petabytes.
- The federated compute initiative was launched, with a new working group meeting bi-monthly and a Slack channel for engagement.

### **Individual API Updates and Goals for 2026**

- Brian transitions to individual API updates, starting with DRS.
- DRS 1.5 was released, adding features such as cloud metadata, storage status, and self-reporting of object counts.
- The 2026 goals include drafting a DRS publication, collaborating with the registry team, and incorporating geolocation.
- Brian highlights the need for champions to drive new features and improvements.

## **TRS and TES API Updates**

- Dennis provides an update on TRS, noting a maintenance release and dependency updates.
- TRS has a few minor open issues and is exploring new initiatives and collaborations.
- Venkat discusses the need for TRS to support different workflow languages and AI integration.
- The discussion includes the potential for TRS to describe LLM models and integrate with AI systems.

## **WES API and Future Plans**

- Valentin outlines the goals for WES in 2026, including increasing interoperability with DRS and TRS.
- The need for a validation service and validation clients for WES is highlighted.
- Martin suggests writing a WES paper to present the API to the community and increase participation.
- The idea of a write-a-thon to collaborate on the WES paper is proposed.

## **New Cloud Standards and Mesh Cards**

- Michael Lukowski provides updates on new cloud standards, including Mesh Cards and PFB.
- Mesh Cards are used to describe entire federated systems of data repositories, with a working implementation in Gen3.
- PFB is an interoperability standard for tabular data, used in Gen3 and other systems.
- The need to create specifications for PFB and engage the community is discussed.

## **Open Houses and Collaboration Opportunities**

- Brian proposes open houses to make the Cloud Work Stream more accessible to newcomers.
- The concept of GIF spotlights is introduced to celebrate small victories and share achievements with the community.
- The importance of hackathons and collaborative work sessions is emphasized.
- Brian highlights the need for champions to drive initiatives and collaborate with other work streams.

## **Discussion on Initiatives and Collaboration**

- The group discusses the structure of initiatives, with some suggesting a focus on federated data and compute.
- The idea of creating a Mesh Initiative to support ecosystems is proposed.
- The need for open houses and “ask-me-anything” sessions to engage the community is reiterated.
- The importance of cross-stream collaboration and addressing governance challenges is emphasized.

## **Discussion on Adapting Self-Assessment Tools for Cloud Repositories**

- Vaso introduces the idea of adapting self-assessment tools to better serve cloud repositories and questions the feasibility of outcomes and criteria.
- Brian expresses enthusiasm for the idea and suggests it could facilitate cross-pollination across work streams.
- Brian invites Graham and Venkat to share their thoughts, indicating a collaborative approach to the discussion.

## **Exploring Data Stewardship and Cost-Sharing in Cloud Environments**

- Vaso introduces the concept of building an ecosystem where data producers are compensated for their data's use, similar to YouTube content creators.
- Another speaker expands on the idea, emphasizing the importance of controlling access to expensive storage and compute resources and proposing standardized constraints.
- The need for a consistent theme across all APIs to manage compute and storage usage effectively is highlighted.

## **Standardizing Access and Control in Cloud APIs**

- Brian agrees with the need for a richer way of describing data use and access credentials, suggesting that a multi-use ontology could be beneficial.
- He emphasizes exploring these concepts further within Cloud and other work streams.
- Brian notes that the discussion has significantly influenced their thinking about initiatives for 2026.

## **Discussion on Complementing Existing Ecosystems**

- Wollmar discusses involvement in projects exploring technical choices, including GA4GH and semantic web technologies.
- Valmar expresses interest in complementing the existing ecosystem with additional features and technologies.
- Brian appreciates Valmar's input and notes that it aligns with the AI Work Stream's focus on identifying missing features in cloud APIs.

### **Proposing Fair Use Principles in Data Sharing and Reproducibility**

- Venkat introduces the concept of fair use principles in data and code sharing, proposing the use of AI to assess the fairness of data-sharing practices.
- He suggests applying this principle across GA4GH and other platforms to improve visibility and reproducibility in analyses.
- Brian supports the idea and encourages further exploration of fair use principles in cloud environments.