Cloud Workstream: Overview

- The GA4GH Cloud Workstream helps the genomics and health communities take full advantage of modern cloud environments.

- Our initial focus is on “bringing the algorithms to the data”, by creating standards for defining, sharing, and executing portable workflows.
Defining, Sharing, & Executing Portable Workflows

Workflow Repository

Workflow Packaging
Cloud Workstream: Standards

- **defining portable workflows**: tool builders need to be able to package their tools for reuse.

- **sharing portable workflows**: tool builders need to be able to offer their tools for others to use, and tool consumers need to be able to discover the tools they need.

- **executing portable workflows**: once a tool consumer has selected a tool, they need to be able to run it in their preferred compute environment, pointing at input and output data in their preferred storage environment.
Cloud Workstream: Driver Projects

<table>
<thead>
<tr>
<th>pressing need</th>
<th>fast follow</th>
<th>maybe later</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genomics England</td>
<td>NCI DCF (GDC + Resources)</td>
<td>BRCA Exchange</td>
</tr>
<tr>
<td>HCA</td>
<td>NIH All of Us</td>
<td>VICC</td>
</tr>
<tr>
<td>EGA</td>
<td>EVA</td>
<td>ENA</td>
</tr>
<tr>
<td>AGHA (individual sites)</td>
<td>AGHA (overall)</td>
<td></td>
</tr>
<tr>
<td>ClinGen (var classification)</td>
<td></td>
<td>ClinGen (other)</td>
</tr>
<tr>
<td>CanDIG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICGC ARGO</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- transitioning from Containers and Workflow Task Team
  - evaluating where previous standards work (TRS, TES, WES, & DOS) fit Driver Project needs
  - will incorporate and prioritize relevant elements going forward
Cloud Workstream: Testbed

1. build a workflow portability testbed app
2. driver projects contribute one or two packaged workflows
   - workflows they intend to run for realz
   - together with test input data and an output verifier
3. driver projects run the testbed in their local environment
   - including contributing back portability patches if needed

Result: every test workflow runs in every environment
Value: demonstrate real-world usability / utility of new standards
Q&A