NHGRI Genomic Data Science Analysis, Visualization, and Informatics Lab-space (AnVIL) Renewal Plans

Chris Wellington and Ken Wiley

National Advisory Council for Human Genome Research
May 16, 2022
What is AnVIL?

- Cloud-based
- Genomic
- Data sharing
- Analysis platform

AnVIL Program Renewal Timelines & Milestones

AnVIL funding: Broad Institute and Johns Hopkins University

- NHGRI Workshop: Future directions for the AnVIL
- Workshop report presented to Council
- GDSWG Meeting: Discussion of AnVIL renewal
- Concept presentation to Council

2018 | 2019 | 2020 | 2021 | 2022 | 2023
Two key components

Limited Competition RFA
Support and improve existing infrastructure
1. AnVIL of today
2. AnVIL of tomorrow
3. Why a limited competition RFA?

Open Competition RFA
Add clinical components
What is AnVIL today?
Data sharing

More than 4 petabytes, 600,000 subjects, 20 research networks

https://anvilproject.org/overview/data-consortia
Analysis platform

Terra: 284 public workspaces
Dockstore: 1,051 analysis workflows
Bioconductor: 2,083 software packages
Galaxy: 8,568 tools (incl. PharmCat)

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

HL7 FHIR

FedRAMP
FedRAMP certified

Implemented on Google Cloud Platform
Outreach and user engagement

- GSP/CCDG/CMG
- MAGIC Jamboree
- Genomic Data Science Community Network (GDSCN)
- Howard University Virtual Applied Data Science Training Institute (VADSTI)
- American Society of Human Genetics (ASHG) Workshops

---

Announcing the AnVIL Cloud Credits Program (AC2) Awardees

Posted: June 03, 2021

NHGRI’s Genomic Data Science Analysis, Visualization, and Informatics Lab-space (AnVIL) cloud genomics platform is pleased to announce the awardees of the pilot phase of the AnVIL Cloud Credits (AC2) Program.

**Awardees**

Seventeen proposals were received from 14 different institutions and of these, the AC2 Review Committee (AC2RC) has awarded 6 proposals with cloud credits. Those awardees include:
What will AnVIL be tomorrow?
What will AnVIL be tomorrow?

A multi-functional discovery platform for genomics
(from the AnVIL workshop report)

1. Increasing tool availability
2. Improving interoperability
3. Addressing barriers to cloud computing
4. Supporting the clinical research community
Increase tool availability

- Best practice workflows
- Simplify data harmonization and integration
- Support user-developed tools
- Cost estimators, “tools about tools”
- Improved searching
Improve interoperability

• NIH Cloud Platforms Interoperability Efforts (NCPI) https://anvilproject.org/ncpi

• Enable cross-platform:
  • Authentication and authorization
  • Data discovery
  • Dataset, workflow, and result exchange

NIH Researcher Auth Service
GA4GH Data Repository Service
Fast Healthcare Interoperability Resources
Address barriers to cloud computing

• Reducing cost barriers
• Simplifying data access
• Improving the user experience
• Supporting curriculum development
• Empowering limited-resource institutions
Why a limited competition RFA?
Why a Limited Competition RFA?

• Leverages NHGRI’s investments in AnVIL
• Minimizes user disruptions
• Allows new requirements
• Incorporates input from peer review
Mechanism of support

- Two Cooperative Agreement (U24) awards
- Five years, FY 2023 - FY 2027
- Limited to current AnVIL awardees

<table>
<thead>
<tr>
<th></th>
<th>FY23</th>
<th>FY24</th>
<th>FY25</th>
<th>FY26</th>
<th>FY27</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost</td>
<td>$6.5M</td>
<td>$6.5M</td>
<td>$6.5M</td>
<td>$6.5M</td>
<td>$6.5M</td>
</tr>
</tbody>
</table>