Democratization and Cloud-Based Genomic Data Sharing: *Views of Platform Developers and Early Adopters*

Sarah C. Nelson[1], Jacklyn M. Dahlquist[2], & Stephanie M. Fullerton[2]

1–Biostatistics, University of Washington; 2 –Bioethics and Humanities, University of Washington School of Medicine

Background

- Genomic research rapidly shifting to cloud-based data sharing and analysis
- Numerous NIH-funded platforms are being built (e.g. AnVIL, BioData Catalyst, All of Us Research Hub, Kids First)
- Unclear tradeoffs for users and data stewards

Methods

- Qualitative interviews with 43 key informants including platform developers, funders, and classes of early adopters
- Interview guide covered data governance, platform attributes, and changes in work

Results

- Developers across platforms emphasize intent to promote a more diverse range of users and research uses
- Early adopters more equivocal, citing challenges including costs, training, and availability of software and data

Future directions

- Exploring distributed responsibility and data stewardship in the cloud
- Understanding experiences of underrepresented scientists to identify strategies to achieve data science equity

Developers aspire for cloud platforms to **democratize genomic science**, but early adopters recognize **numerous challenges**

Scan to learn more in our April 2023 publication, "Cloud-based biomedical data storage and analysis for genomic research: Landscape analysis of data governance in emerging NIH-supported platforms."

Key Informant Demographics

|	Developers* (n=21)	Users** (n=22)
Age (mean, SD)	46 (9.6)	45 (11.4)
Female (n, %)	10 (48%)	13 (59%)
Race and ethnicity (n, %)†		
White	18 (86%)	17 (77%)
Asian/Asian American	3 (14%)	4 (18%)
Hispanic/Latino or mixed race	3 (14%)	1 (5%)
Black/African American	1 (5%)	0

*Representing 5 cloud and 1 non-cloud sharing platform
**Majority (50%) analysts, also including PI data contributors, IRB officials, and funders
†Participants able to identify as multiple; categories with 0 not shown; "mixed race" category by participant request

Example quotes

"One thing that cloud platforms have the capacity to do is level the playing field. One can be in a less well-funded institute and still have access to all the computational facilities that anyone would have at the most advanced institutions." – Developer (ID 14)

[Accessing shared data through cloud platforms is] "still the opposite of democratizing…I have to know somebody to get access to the data. That's not democracy. That's a monarchy." – User (ID 99)
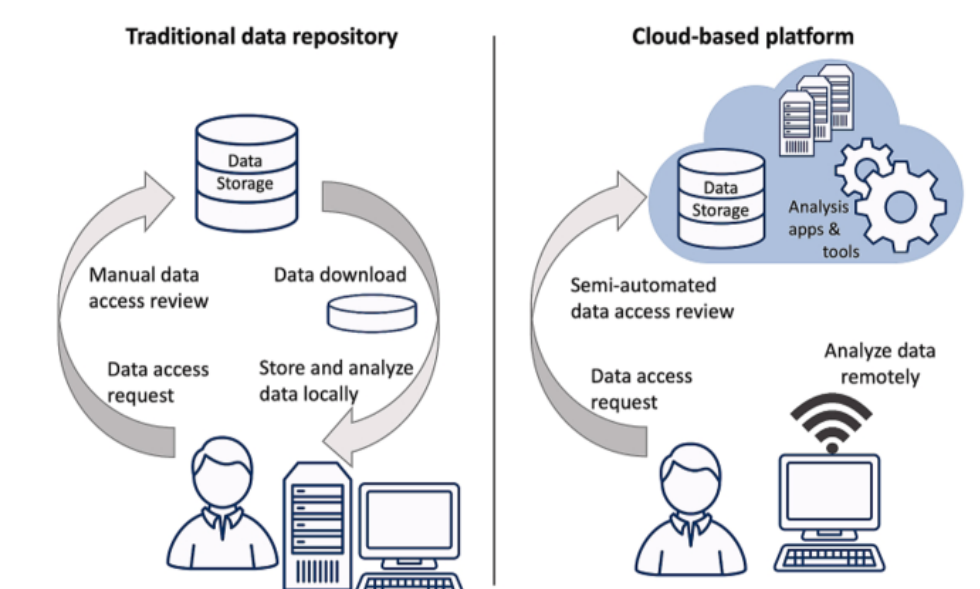
Prior vs. cloud-based data sharing