University of Maryland
Health Sciences and
Human Services Library

HS/HSL Updates - [HTTPS://WWW2.HSHSL.UMARYLAND.EDU/HSLUPDATES/]

DABS (Data and Bioinformation Stuff) Volume 1 Issue 8: Cloud Computing

Posted on February 26, 2021 - [https://www2.hshsl.umaryland.edu/hslupdates/?p=4710] by J-P Courneya - [https://www2.hshsl.umaryland.edu/hslupdates/?author=16]

The Center for Data and Bioinformation Services - [https://www2.hshsl.umaryland.edu/cdabs/] (CDABS) is the University of Maryland Health Sciences and Human Services Library hub for data and bioinformation learning, services, resources, and communication.

We are wrapping up another week (Feb 22 -26) of learning and growing at CDABS. Our adventures had us working on HPC (High Performance Computing) at IU (Indiana University) as part of their HPC Onboarding for Biologist workshop. The National Center for Genome Analysis Support (NCGAS) provides this HPC workshop to help new users learn about HPC resources available to them, other course offerings, and NCGAS services. This workshop and a video linked below had me thinking quite a bit about research computing particularly computing on the cloud. Folks let me say that computing on the cloud is becoming more pervasive in research computing. Knowing about this topic is worth your time since as researchers in the modern age we will be faced with having to use the cloud to do our computing more and more. Datasets are moving to the cloud. Software has already moved to the cloud. And some day our workstations may only be terminals to connect to our actual computers which exist on the cloud.

Check out these links for resources and learning about computing for genomics on the cloud.

1. This article, freely available in PMC (Pubmed Central), by Ben Langmead and Abhinav Nellore describes how cloud computing is used in genomics for research and large-scale collaborations, and argues that its elasticity, reproducibility and privacy features make it ideally suited for the large-scale reanalysis of publicly available archived data, including privacy-protected data. (10 minute read) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6452449/ - [https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6452449/]

2. Keynote from European Bioconductor Meeting 2020: Sehyun Oh – Bioinformatics On Cloud: How to leverage cloud-based resources for your bioinformatics works. (40 minute watch) https://youtu.be/bFvT4_fqpwE - [https://youtu.be/bFvT4_fqpwE]

3. The Seven Bridges Platform is a cloud-based environment for analyzing genomics data. Use the Platform to securely store, analyze, and share data amongst team members working both locally and globally. The Platform co-locates analysis workflows alongside genomic datasets to optimize processing. Read a

nd learn more at the SevenBridges knowledge center. (10 minute overview) https://docs.sevenbridges.com/ - [https://docs.sevenbridges.com/]

4. Terra is a cloud-native platform for biomedical researchers to access data, run analysis tools, and collaborate. The vision of Terra is to enable the next generation of collaborative biomedical research. There are several projects that exist independently on the platform – AnVIL, BioData Catalyst, and FireCloud, for example. Each project on Terra serves a unique research purpose, while still offering the benefits of the Terra platform to every user. (10 minute read) https://terra.bio/ - [https://terra.bio/]

5. Galaxy is an open source, web-based platform for data intensive biomedical research. The main Galaxy instance is an installation of the Galaxy software combined with many common tools and data; this site has been available since 2007 for anyone to analyze their data free of charge. The site provides substantial CPU and disk space, making it possible to analyze large datasets. You can even install your own Galaxy and choose from thousands of tools from the Tool Shed. (10 minute overview) https://galaxyproject.org/tutorials/g101/ - [https://galaxyproject.org/tutorials/g101/] (Galaxy Main) https://usegalaxy.org/ - [https://usegalaxy.org/]

Questions?

Contact: Amy Yarnell, Data Services Librarian and **Jean-Paul Courneya**, Bioinformationist — at data@hshsl.umaryland.edu - [mailto:data@hshsl.umaryland.edu] .

To read more of our content and stay informed please visit our communications page and fill out the form to subscribe. Subscribe here: https://www2.hshsl.umaryland.edu/cdabs/communications - [https://www2.hshsl.umaryland.edu/cdabs/communications.cfm]