Advanced Research Computing for Sociologists

María Sigríður Finnsdóttir, PhD
April 7th, 2026

Digital Research Alliance of Canada | Alliance de recherche numérique du Canada

Digital Research Alliance of Canada | Alliance de recherche numérique du Canada

High-performance computing.

High-performance computing (HPC) is also known as supercomputing. It can handle:

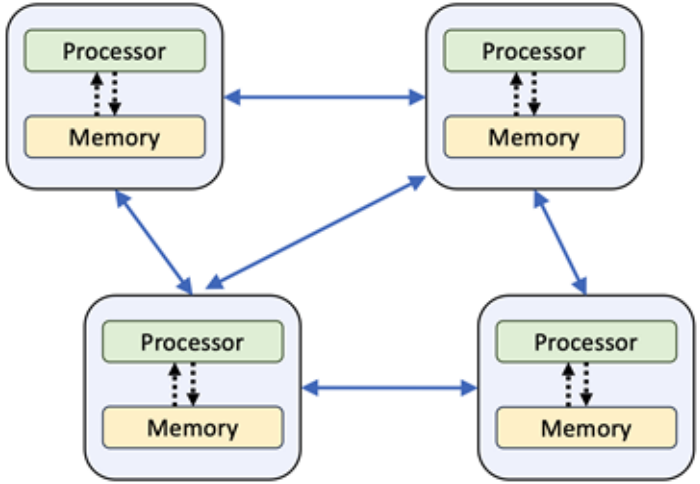
- Faster computation
- Larger problem
- More detailed simulations

HPC involves many CPUs together, which can mean:

- Many single CPU jobs running at the same time, or
- Many CPUs working in sync to solve the same problem

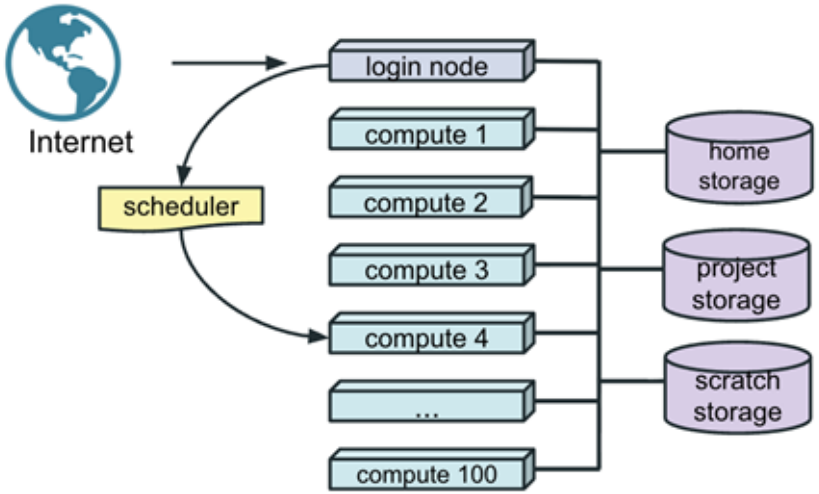
Either way, communication between cores is key!

Principle tool for HPC is the computing cluster.



Using an HPC cluster.

1. Log in from the internet
1. Upload data via the login node to shared storage
1. Set up and test your workflow on the login node
1. Submit a batch job to the scheduler
1. The scheduler chooses where to run the job
1. The system notifies the user when the job is complete



Dr. Ethan Fosse, Associate Professor of Sociology, University of Toronto

“We use the Digital Research Alliance of Canada's high-performance computing systems to examine how public discourse changes over time in large text data. Our current project analyzes newspaper coverage to identify how topics emerge, persist, split, and transform across periods, with the goal of understanding changing patterns of meaning. In doing so, we develop a novel unsupervised learning method for analyzing how meaning changes in text data, which contrasts with conventional methods, such as topic modeling, that assume meaning is time-invariant. This is sociologically important because it helps reveal how institutions classify the social world, how issue attention rises and falls, and how cultural and political boundaries are reproduced or reworked in public communication. High-performance computing makes our analysis possible by allowing us to run computationally intensive language models repeatedly across many temporal slices, compare alternative specifications, and work at a scale that matches the scope of the substantive questions. Our computational pipeline could be used to analyze a number of other important questions, such as those related to shifting narratives around inequality, polarization, migration, trust, health crises, or social movements.”

Digital Research Alliance of Canada | Alliance de recherche numérique du Canada

Cloud computing.

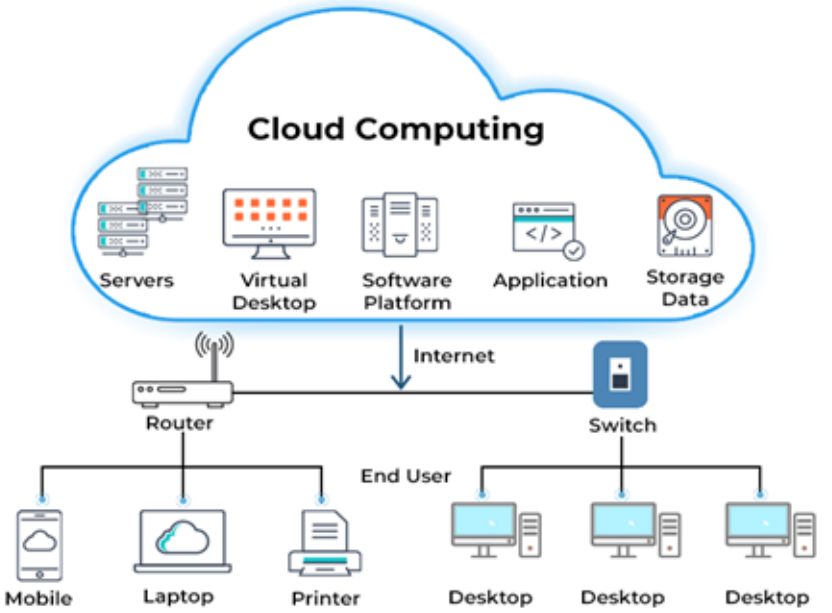
Cloud computing basically lets you use someone else's (more powerful) computer as your own.

Cloud computing can be used in a variety of ways, including:

- Web-hosting
- Bursts of on-demand computing

HPC is better for powerful computing, cloud for always-available computing.

Most common way to use the cloud is with a virtual machine.



Digital Research Alliance of Canada | Alliance de recherche numérique du Canada

Cloud computing: Virtual machines.

A virtual machine (VM) is a simulated computer, on which you can run programs on just like a real machine.

Cloud computing makes use of VMs to allow:

- Virtualized hardware configurations, which are different to the physical machine running the code
- Multiple instances on one machine
- Automated creation and deletion of virtual hardware

Dr. Michelle Maroto, Professor of Sociology, University of Alberta

“I’m currently working with confidential electronic health records data from PointClickCare, which is health tech company that provides software and support systems to skilled nursing facilities (SNFs) and long-term and post-acute care (LTPAC) settings across Canada and the United States. My work focuses on experiences of disability and health among people living in long-term care facilities and these data have a lot of fine grained measures for both. By working on the cloud, I am able to securely store this large dataset and quickly access the data when I need to. The Digital Alliance also provided helpful support for when PointClickCare was sharing the data with me. This allowed me to receive an access code and transfer the data over a secure server. Finally, my team also extends to researchers at the University of Toronto, through the Alliance and HPC cloud, we have been able to easily collaborate on research and data analysis. “

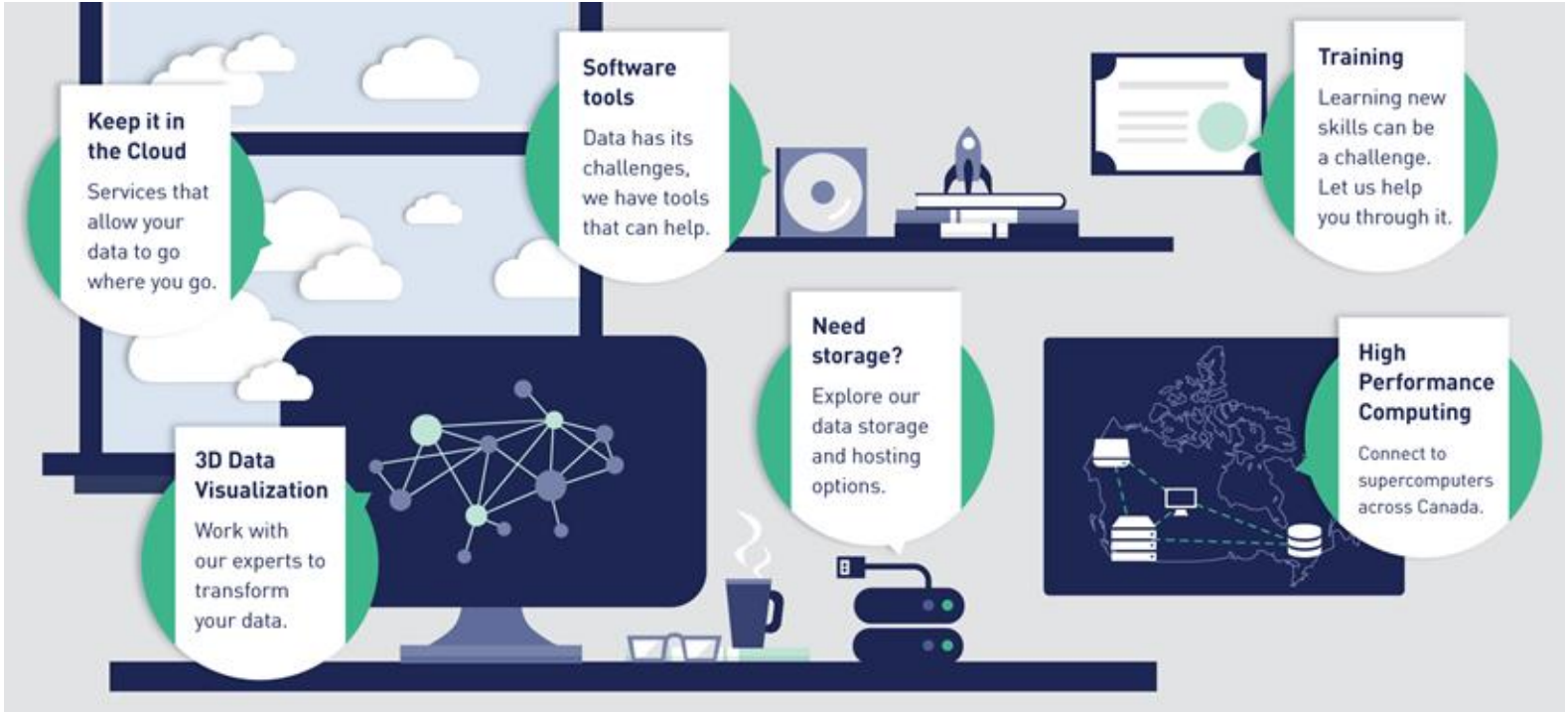
Digital Research Alliance of Canada | Alliance de recherche numérique du Canada

Digital Research Alliance of Canada.

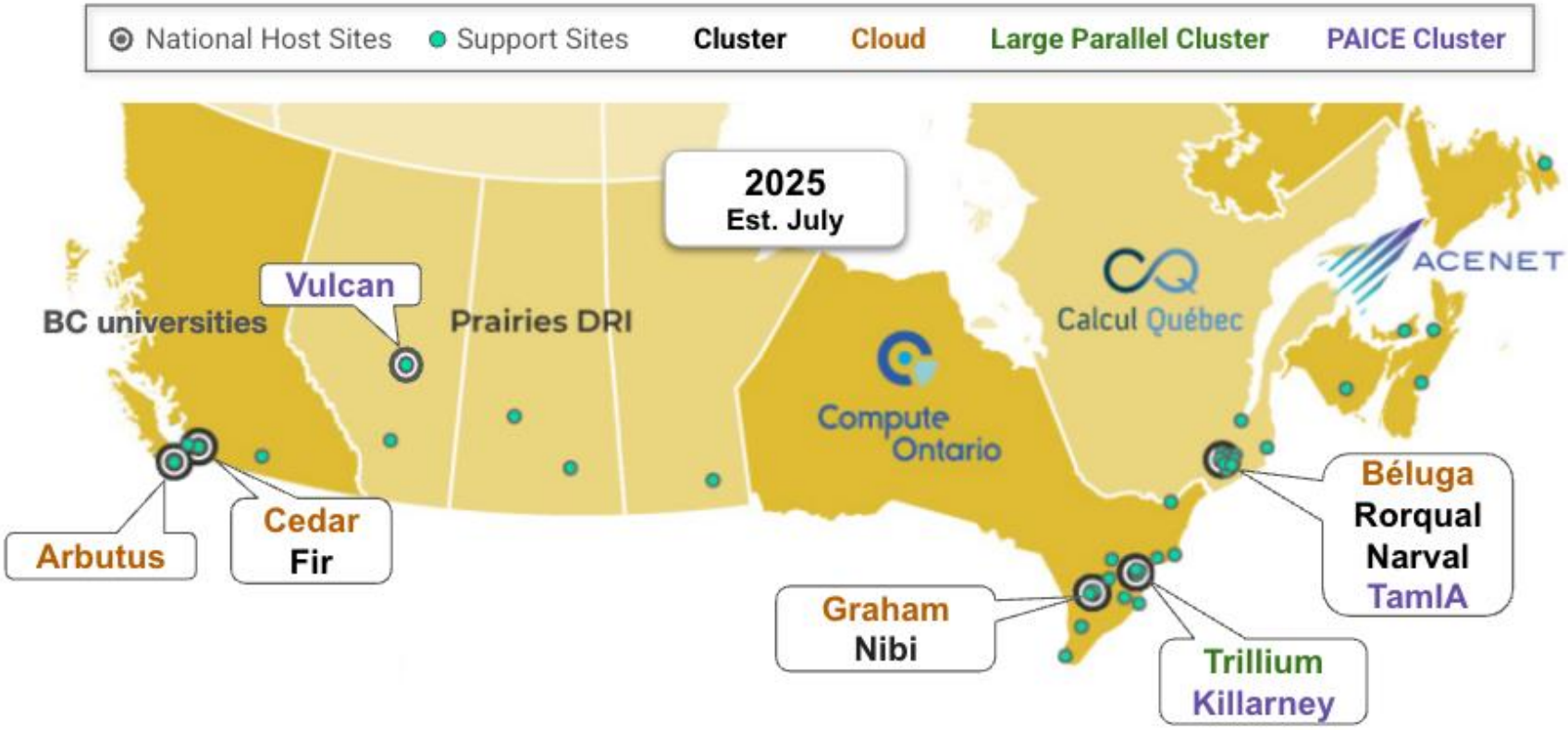
The Digital Research Alliance of Canada is a non-profit organization, funded by the federal government, that supports Canadian researchers in the areas of:

- Advanced research computing (ARC)
- Research data management (RDM)
- Research software (RS)

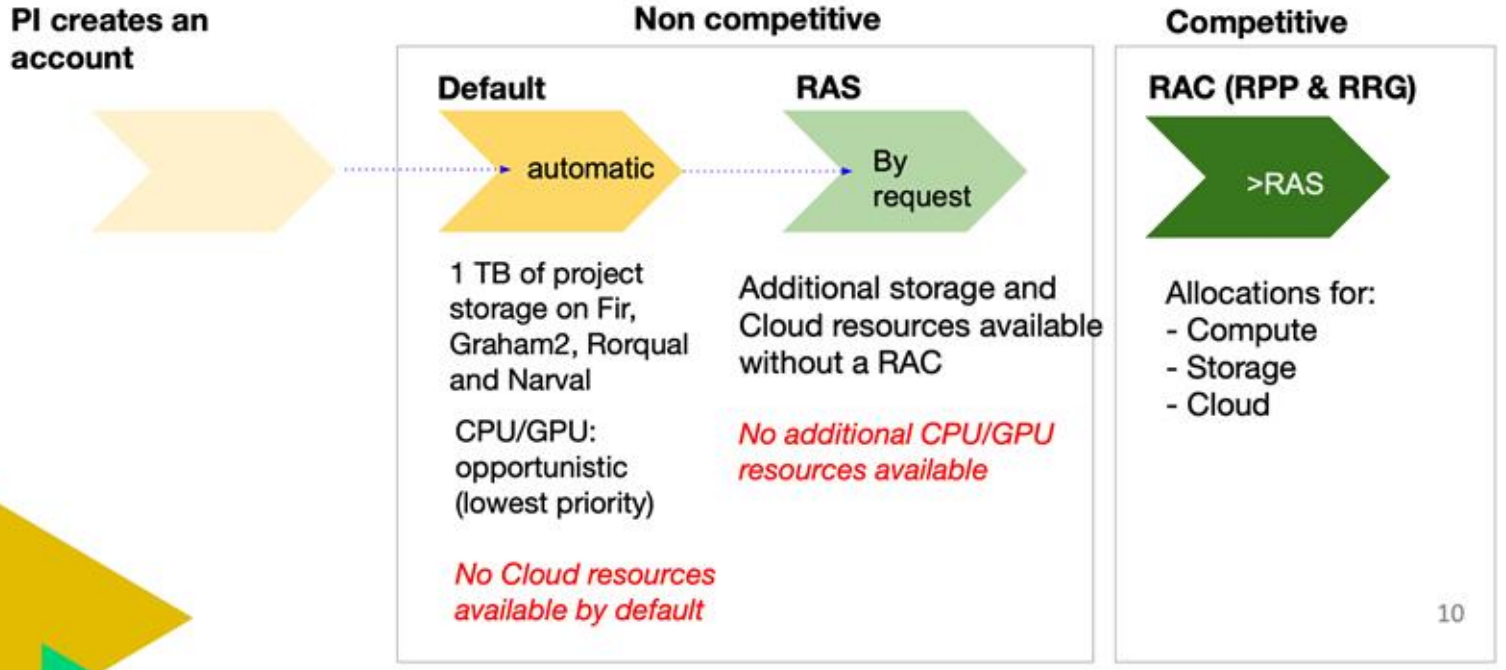
Digital Research Alliance of Canada | Alliance de recherche numérique du Canada



Digital Research Alliance of Canada | Alliance de recherche numérique du Canada



Digital Research Alliance of Canada | Alliance de recherche numérique du Canada


PI creates an account
Non competitive
Competitive
Default
RAS
RAC (RPP & RRG)
automatic
By request
>RAS
1 TB of project storage on Fir, Graham2, Rorqual and Narval
CPU/GPU: opportunistic (lowest priority)
No Cloud resources available by default
Additional storage and Cloud resources available without a RAC
No additional CPU/GPU resources available
Allocations for:
- Compute
- Storage
- Cloud
10

Storage.

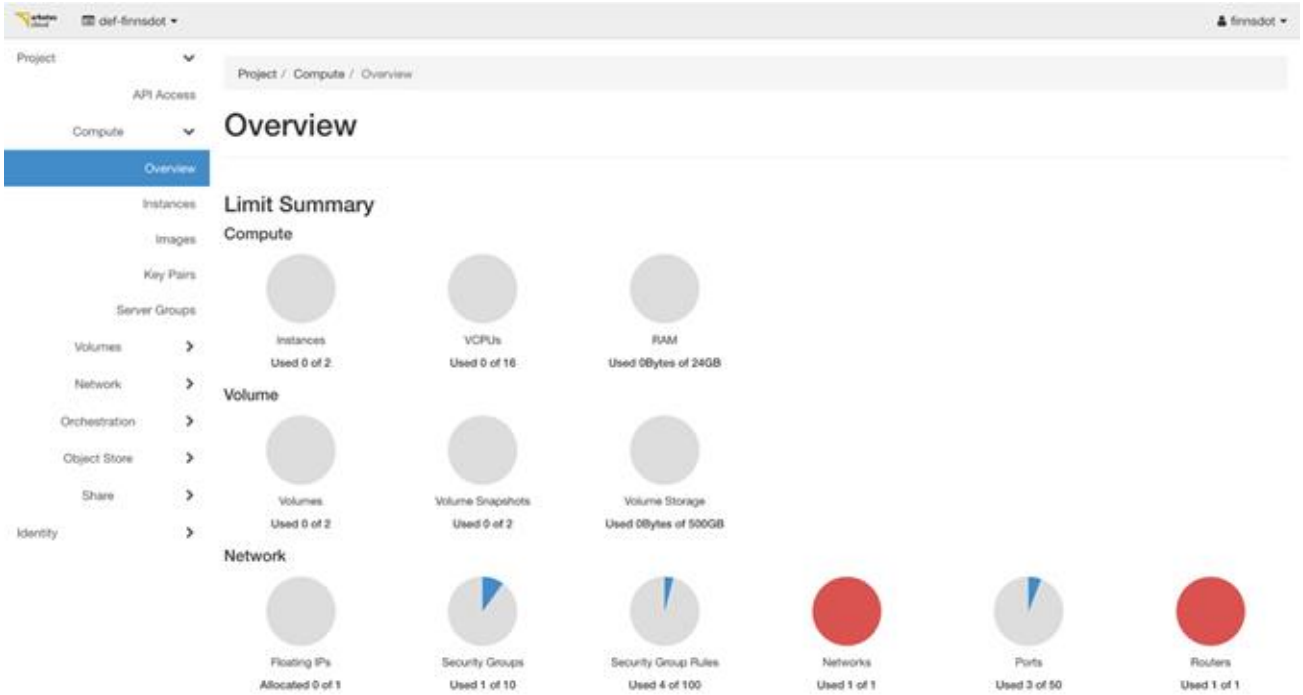
You can access 4 to 40 TB of project storage on the HPC clusters through the Alliance

Possible to access up to 100TB of nearline storage through RAS

All data stored with the Alliance is stored in Canada

Data can be securely moved on and off the clusters using Globus

Accessing the cloud.



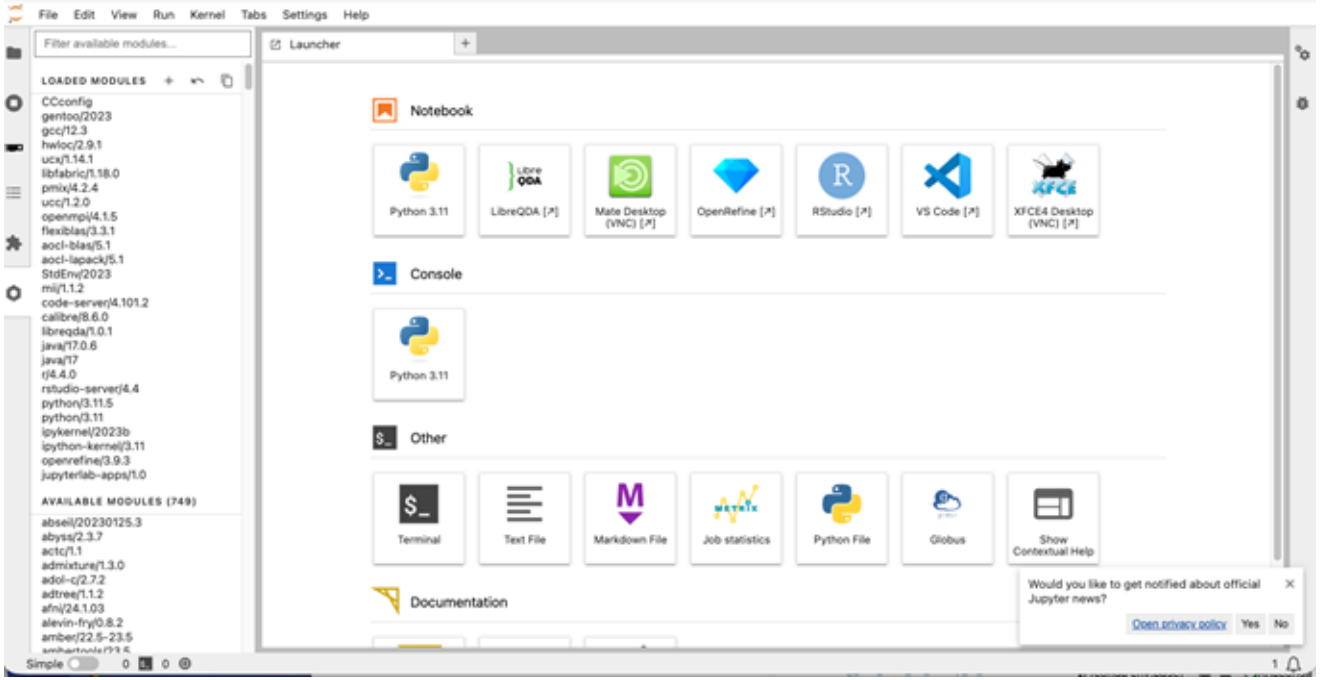
Digital Research Alliance of Canada | Alliance de recherche numérique du Canada

Accessing HPC.

There are two ways to access the Alliance's HPC resources:

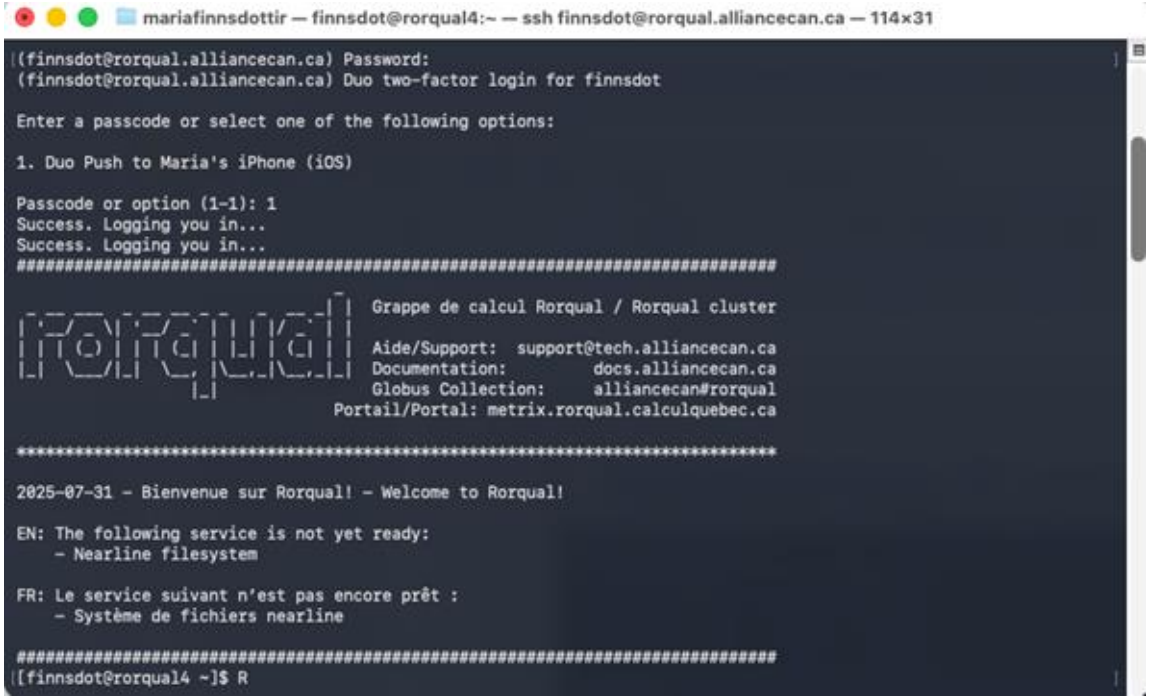
1. Through JupyterHub in your browser
2. Through the terminal/shell

HPC through JupyterHub.



Digital Research Alliance of Canada | Alliance de recherche numérique du Canada

HPC through the shell.



Further resources.

The Alliance's technical documentation

The Alliance includes a national HSS team
- You can reach us by emailing support@tech.aliancecan.ca with "HSS" in the subject line

HSS Spotlight Webinar Series on digital and computational research in the humanities and social sciences

WestDRI (SFU) video archive of training

Connect with other social science and humanities researchers and experts through the HSS Commons

Information on the Alliance's clusters
- Visit the page for the specific cluster to find its URL

Workshops and events are advertised on Explora

A workshop I made for using R on the HPC cluster

Thank you!

Questions?

maria.finnsdottir@westdri.ca