HARVARD
T.H. CHAN
SCHOOL OF PUBLIC HEALTH

Onboarding Information for New Faculty & Academic Appointees on:

- **Research Computing**
- **Data Security**
- **Data Management**

Welcome to the Harvard Chan School! Research computing is a critical resource for investigators across all departments. Below is some basic information on research computing and related resources available at School, University, and beyond. Much of the information contained in this document can also be found on the School's Research Computing web page: https://www.hsph.harvard.edu/research-computing/.

For more information or help connecting with contacts for resources, please contact: Michele Sinunu, Director of Research Platforms in the Office of Research Strategy and Development (ORSD).

Research Computing Clusters

Harvard Faculty of Arts and Sciences Research Computing (FASRC) – the School's preferred provider for Research Computing is FASRC. Standard services at FASRC are FREE for eligible Harvard Chan School PIs, including the Cannon high-performance computing cluster. New users may find it helpful to view videos on getting started on the cluster.

If a PI's storage needs exceed 20 TB, the request must be approved by Deane Eastwood, HSPH IT Chief Information Officer. Storage requests can be made directly to FASRC who will contact Deane for approval.

FASRC is suitable for data that fall in Harvard Security Levels DSL1 or DSL2 (see **Data Security**, below). Data that fall in DSL3 should be stored and analyzed in the FAS Secure Environment (FASSE). For data that require more stringent data security measures, please contact HSPH IT (Deane Eastwood or Matt Ronn) to discuss options.

HMS O2 – The other main computing cluster at Harvard is HMS O2. If collaborators are at HMS or data are hosted at HMS, some faculty choose to use the HMS O2 resources, which may incur a cost to PIs.

Other Research Computing Related Resources at Harvard

The **Harvard Cloud Program** is managed by Harvard University Information Technology (HUIT), which provides access to University-negotiated services. Harvard-associated cloud accounts can be created via this link. The University has negotiated agreements with Amazon Web Services (AWS), Google Cloud Platform (GCP), and Microsoft Azure. Negotiated terms vary by provider, but all have immediate contractual protections for your work (i.e. terms negotiated by Harvard's Vendor Management Office). If you access cloud computing via a Harvard account, you are strongly encouraged to go through Harvard even when there are no discounts available for the contractual protections. Please note that standard cloud environments are not compliant with many federal data security policies such as FISMA and FedRAMP. Ensuring your preferred cloud solution is set up to be compliant with University and sponsor data security requirements can be complex. Please contact Andy Ross, the School's senior information security officer, when you are setting up a cloud environment.

Harvard Institute for Quantitative Social Services (IQSS) is a social science research center that provides a research computing environment, data science services, and trainings. IQSS also hosts the Dataverse, an open source research data repository.

The **Harvard Center for Geographic Analysis** supports research related to geospatial technologies and methods through consultation, data, and software.

The Harvard Chan based **Quantitative Biomedical Research Center** (QBRC) core facility provides large scale biomedical data mining, management, pattern recognition, and software engineering support.

The **Harvard Chan Bioinformatics Core** provides training on basic data skills and analysis of high-throughput sequencing data; and next-generation sequencing experimental design and analysis support.

Data Security

It is critical that Harvard Chan School researchers familiarize themselves with University and sponsor policies and requirements related to data security. Basic information and links are provided on the School's research data security information page. Andy Ross, the Harvard Chan School's senior information security officer, is available to answer questions and assist in navigating requirements.

Note that researchers who work with data that are sensitive or require a Data Use Agreement (DUA) must submit data management plans via the new Data Safety Application for review prior to accessing the data, and must complete an annual online Data Security Training Course.

In some instances, PIs or members of their research teams purchase their own equipment if University solutions do not meet their needs. Faculty purchased servers, laptops, and other equipment present a considerable risk to both the investigator and University systems. Please work with HSPH IT Director of Infrastructure Services Matt Ronn **prior** to purchasing to ensure PI-purchased equipment is physically secure, has a patch management plan, is encrypted, and is protected from cyberattacks to ensure the integrity of research data and the security of the Harvard network. All server equipment at Harvard must meet minimum security criteria before being allowed on the network. Hardware should be centrally purchased from Harvard vendors with appropriate warranties and service plans, and should be securely housed in the Harvard Chan School data center.

Data Management

The Longwood Medical Area Research Data Management webpage has a number of resources for research teams including:

- Harvard Medical School Biomedical DMP Template on the DMPTool, a free web-based tool to assist the creation of funder or university specific plans
- Biomedical data lifecycle
- Trainings and events

The University has developed a research support webpage focused on resources related to research administration and compliance, research computing, and research data and scholarship. Note that some of the resources and tools listed are only available to community members of specific Harvard schools.