

# HMS Research Computing Overview

**PROVIDING COMPUTATIONAL TOOLS AND  
EXPERTISE ACROSS THE RESEARCH PIPELINE**

Research Computing  
Harvard Medical School IT

4/13/2022

# RC Mission Statement

---



Collaboratively working with researchers to identify, design, and deliver a wide range of technological solutions at the ever-changing forefront of biomedical research

# What RC Does for You

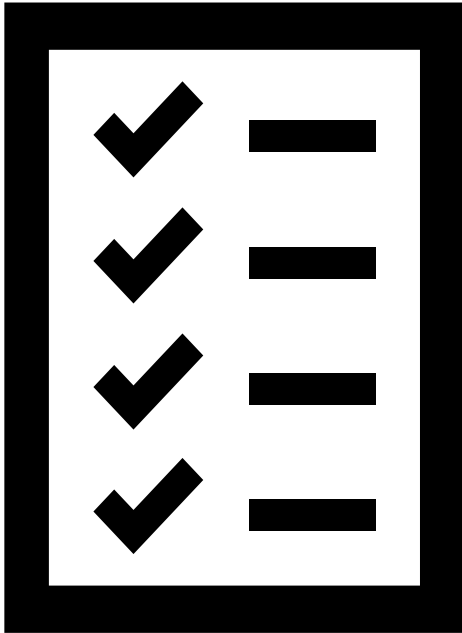
---

- **Operate** – Provide reliable IT services for research
- **Consult** – Help researchers use IT
- **Train** – Teach researchers to use IT
- **Support** – Assist researchers when there is a problem
- **Innovate** – Prototype IT to advance the research mission



# In Other Words...

---



- Help researchers evaluate and choose effective solutions
- Help researchers access and use those solutions efficiently
- Build and operate solutions when needed
- Empower researchers to do good science

# Who We Are

---



# Agenda

---

- Consulting and Training
- High-Performance Computing
- Research Data Management
- Storage Solutions
- Research Imaging Solutions
- Research Applications and Software
- Research Computing Core



# Consulting and Training

---

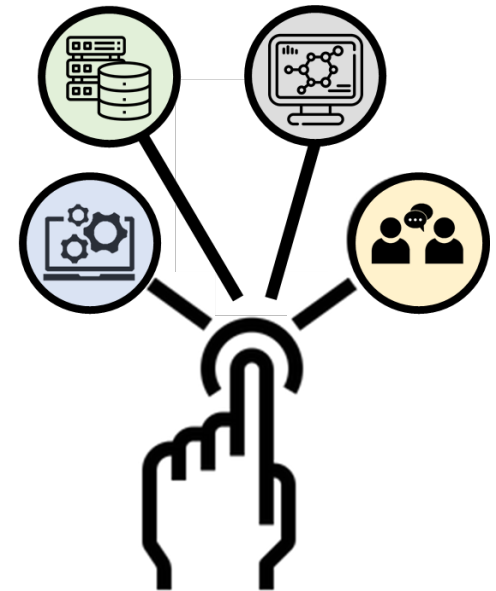
- 2-hour workshops, 30-minute info sessions
- 1-1 meetings to:
  - Plan a project, estimate resource needs
  - Design a pipeline
    - microscope -> analysis -> archive
  - Troubleshoot a job on O2
  - Get help using imaging or data transfer software



# HPC and Bioinformatics

---

- Access to the O2 cluster – compute, GPUs, high-memory workstations, storage, interactive sessions
- Advice on how to use resources efficiently
- "My jobs are taking too long to start running!"
- Software, databases
- Expertise in building workflows, scripting, and bioinformatics





# Research Data Management

---

- Collaborate with researchers to better organize, manage, and store research data throughout the various stages of the data lifecycle
- Develop automated methods for migrating data between storage platforms
  - Advance and refine processes to move infrequently accessed data to long term storage
- Create and maintain data management tools and resources to prepare data for sharing and reuse

# Research Data Management

## Resources

- Research Data Management Onboarding Checklist
- Research Data Management Offboarding Checklist
- Harvard Biomedical Research Data Management Website



<https://datamanagement.hms.harvard.edu>

# Storage Solutions

- HMS offers several storage options that allow users to store data in different places, each with distinct behaviors, performance, and means of access
- **Active**
  - Compute (O2)
  - Collaborations
- **Standby**
- **Cold** (future offering)

	Scratch	Active Compute	Collaborations	Standby Standard
<b>Purpose/Used For</b>	<ul style="list-style-type: none"> <li>Transient files used during a single job on the HMS High-Performance Compute (HPC) Cluster.</li> <li>Temporary (days to weeks) high performance storage for data that can be easily regenerated.</li> </ul>	<ul style="list-style-type: none"> <li>Active research data that is frequently accessed, modified, or computed against.</li> <li>Run many analyses simultaneously on a High-Performance Compute (HPC) Cluster.</li> </ul>	<ul style="list-style-type: none"> <li>Active research data that is frequently accessed, modified, or computed against.</li> <li>Share documents and files with colleagues, both within and outside of your department.</li> </ul>	<ul style="list-style-type: none"> <li>Inrequently accessed data, directly available for reference, retrieval, or analysis.</li> <li>Can operate as an intermediary location, to organize and prepare research data for long-term retention, as required.</li> </ul>
<b>Filesystem Path(s)</b>	/r/scratch3	r/Atlas1   r/Atlas2   r/groups	/v/scratch/files	/t/standby/files   r/standby
<b>Write Speed</b>	High	High	Medium	Medium
<b>Read Speed</b>	High	High	Medium	Medium
<b>Access From</b>	<ul style="list-style-type: none"> <li>O2 compute cluster</li> <li>O2 transfer cluster</li> </ul>	<ul style="list-style-type: none"> <li>O2 compute cluster</li> <li>O2 transfer cluster</li> </ul>	<ul style="list-style-type: none"> <li>Windows</li> <li>Mac OS</li> <li>Linux</li> <li>O2 transfer cluster</li> <li>Online Storage Tool</li> </ul>	<ul style="list-style-type: none"> <li>Windows</li> <li>Mac OS</li> <li>Linux</li> <li>O2 transfer cluster</li> </ul>
<b>Policy &amp; Eligibility</b>	<ul style="list-style-type: none"> <li>Allocation amount dependent on lab needs and available resources.</li> <li>10TB per user</li> </ul>	<ul style="list-style-type: none"> <li>Currently accepting storage requests from groups with a primary appointment with an HMS Quad-based pre-clinical department.</li> <li>Allocation amount dependent on lab needs and available resources.</li> <li>If not eligible, please contact HMS Research Computing to discuss further options.</li> </ul>	<ul style="list-style-type: none"> <li>Need to have at least two co-owners who can add, edit, and remove files as well as grant additional user access.</li> <li>If you want access to an existing collaboration, you need to have the manager or owner of the collaboration grant you access.</li> </ul>	<ul style="list-style-type: none"> <li>Allocation amount dependent on lab needs and available resources.</li> </ul>
<b>Protection &amp; Retention</b>	LOW No snapshots or backups. Data will be deleted 30 days after last access.	HIGH Snapshot Recovery (90 days) Disaster Recovery (Off-site)	HIGH Snapshot Recovery (90 days) Disaster Recovery (Off-site)	HIGH Snapshot Recovery (90 days) Disaster Recovery (Off-site)
<b>Cost to User</b>	No Cost	No Cost	No Cost	No Cost
<b>Cost to HMS</b>				
<b>Harvard Data Security Level</b>	Up to Level 3	Up to Level 3	Up to Level 3	Up to Level 3
<b>Request Storage</b>	Additional information available on the Research Computing Confluence Webpages	Submit a Storage Request	Submit a Storage Request	Submit a Storage Request
<b>Contact Us</b>	RC Consultants • rchelp@hms.harvard.edu	HMS Research Data Management Team • rdm@hms.harvard.edu	HMS Research Data Management Team • rdm@hms.harvard.edu	HMS Research Data Management Team • rdm@hms.harvard.edu

Disclaimer:

- Storage offerings may change based on product offerings and researcher feedback; we will continue to update the community on changes and improvements.
- HMS will continue to explore market offerings and leverage existing partnerships to develop the future Cold Storage option. Additional information will be made available as the storage offering evolves. We recommend that labs continue to identify data to move to Cold Storage, ensuring easier transition once the storage offering is finalized, data identified for Cold Storage will be moved to Standby in the mean time.
- On these storage offerings not meet your storage needs? We're always interested in receiving feedback, please reach out to Research Data Managers at rdm@hms.harvard.edu
- Last updated: 2020-09-05



**HARVARD**  
MEDICAL SCHOOL

DEPARTMENT OF  
Information Technology

# Research Imaging Solutions

---

## Overview

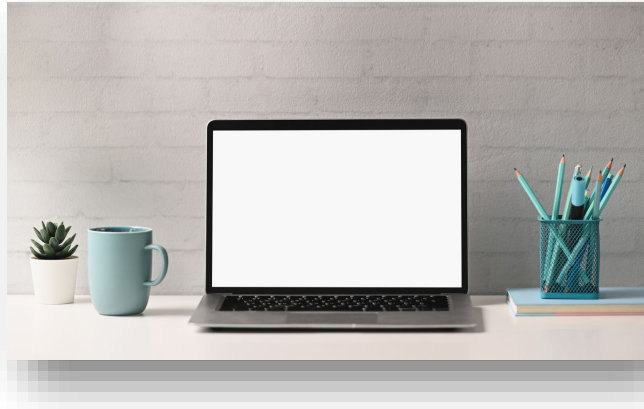
- Works to identify ways to streamline image preparation so researchers can maximize their time at the bench, writing grants and developing discoveries.

## We provide:

- Guiding researchers to effectively communicate their science
- Research Imaging Training and Seminars
- 24/7 Access to Imaging Resources and Guidelines

# Research Applications and Software

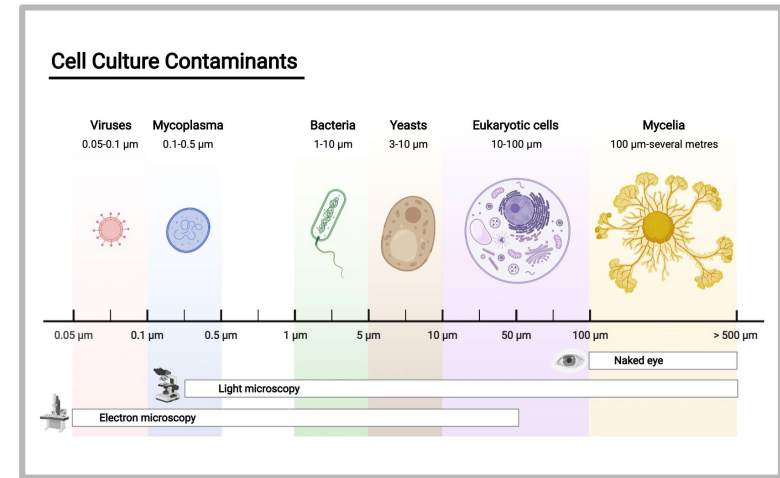
---



- Access to scientific desktop software
- Solutions, applications, and software packages
- Focused on enabling and enhancing HMS researchers' ability to perform their research

# Research Application Platforms

- BioRender (\$)
  - A web platform for creating and sharing illustrative scientific figures for publications and communications using a scientifically accurate image library
- Data Transfers via Globus
- eLabNext Platform
  - Electronic Lab Notebook



# Research Applications Platform(2)

- OMERO

- Visualization platform for the management of microscope images and metadata

- REDCap

- A web-based Electronic Data Capture (EDC) application designed for building and managing online surveys and data capture for research and clinical studies

- Research Data Visualization Platform – Pilot

- Shiny applications hosting via RStudio Connect system. Supporting R and Python code and libraries within Shiny applications

# Research Software

---

## Scientific Desktop Software

- ChemOffice
- DNASTar Lasergene
- FlowJo (\$)
- Geneious
- GraphPad Prism (\$)
- MATLAB (also on O2)
- SnapGene (\$)
- JMP Genomics
- JMP Pro
- LabVIEW





# Research Computing Core

---

## Overview

- A set of billable services provided by Research Computing and HMS IT
- Promotes deeper collaboration across the greater Harvard biomedical research ecosystem
- Establishes transparent and sustainable IT services for our research community

# Research Computing Core

---

## Billable Services

- O2 Cluster and Storage
  - Researchers whose PIs **do not** have a primary or secondary faculty appointment in an [HMS Basic and Social Science Department](#) are billable.
- Software like BioRender, FlowJo, GraphPad Prism, SnapGene
  - Only offered to Harvard faculty, staff, and students
  - Visit the RC website to learn about eligibility

# The Most Important Slide!

---

Contact RC early and often:

- When new members join the lab
- Before starting a project / buying hardware
- When you're wondering how to do something
- When something doesn't work
- A 5-minute email may save time, frustration, \$
- If you don't ask, we can't help



Thank you. Questions?

[rchelp@hms.harvard.edu](mailto:rchelp@hms.harvard.edu)

<https://it.hms.harvard.edu/rc>