

# CCAST support for AI-related research and teaching at NDSU

**NDSU**

CENTER FOR  
COMPUTATIONALLY ASSISTED  
SCIENCE AND TECHNOLOGY

Information Technology Division, North Dakota State University

[ndsu.edu/ccast](https://ndsu.edu/ccast)

August 23, 2024

# What is CCAST?

Center for Computationally Assisted Science and Technology

- Pronounced “*C-cast*”.
- A department within the Division of IT.
- Founded in 2003 as the Center for High-Performance Computing (CHPC).
- Changed to “CCAST” in 2008.
- Datacenter and main offices in Research 2, on-campus offices in Quentin Burdick Building.
- 7 full-time staff.



# What does CCAST do?

- High-Performance Computing (HPC)
- Data Storage and Transfer
- Scientific Software
- Training
- Internships
- Education
- Consulting
- Proposal Writing
- Research



# High-Performance Computing

- CCAST maintains the **largest academic supercomputing facility** in North Dakota:
  - 12,000+ CPU cores
  - 50+ TB aggregate RAM
  - 90 Nvidia GPUs
  - 2+ PB fast data storage
  - High-speed (40-100 Gb) network
- Almost all equipment purchased with external money (NSF MRI, ARPA, other grants).
- Most resources are **free** for all faculty, staff, and students in ND.



# Data storage and transfer

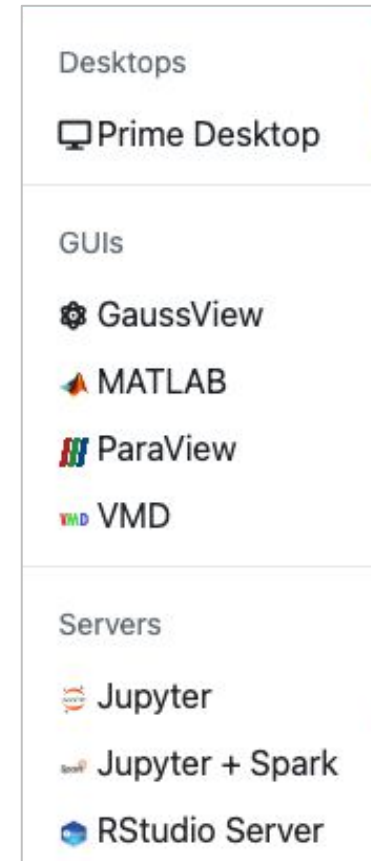
- CCAST and IT provide access to **10 PB** of storage capacity for research data:
  - **2 PB** high-performance storage
  - **6 PB** tape backup/archive storage (can grow up to 19 PB)
  - **1 PB** research data archive (Windows fileshare, ability to grow to multiple PB)
  - **500 TB** long-term storage (for EPSCoR ND-ACES project)
- CCAST also hosts NDSU's Globus service for large data transfers.
  - For example: In 2022, CCAST migrated 50 TB of research data from Google Drive to one of the above locations.





# Scientific software

- CCAST maintains several collections of scientific software, including:
  - Hundreds of scientific software modules on our HPC systems.
  - About a dozen (and growing) graphical/interactive applications via the OnDemand web platform.
- Newly hired **Scientific Software Specialist** (Nov. 2023).



# Training

- **CCAST Advanced Research Computing Training Program** offered each semester, open to all faculty, staff, and students in ND.
- Since 2019, **more than 600 faculty, staff, and students** have participated in the program.
- **Topics covered:**
  - Introduction to HPC
  - Linux for HPC
  - Parallel computing
  - Machine learning
  - Bioinformatics
  - Python and R
  - Materials modeling
  - Quantum computing
  - and more...



# Internships

- CCAST provides paid and unpaid internships.
- Interns gain skills in:
  - Linux system administration
  - Scientific computing
  - Software deployment
  - Electronics/IoT
  - and more...
- In the past 5 years, CCAST has provided internships to **40+ students.**





# Consulting, education, proposals

- CCAST provides a variety of other services:
  - Consulting with individual research groups on advanced computing needs.
  - Letters of support and facilities descriptions for funding proposals.
  - Dedicated FTE time (usually grant-based) for development of custom solutions.
  - Focused training for individual groups/departments.
  - Guest lectures and hands-on exercises for courses with a computational focus.

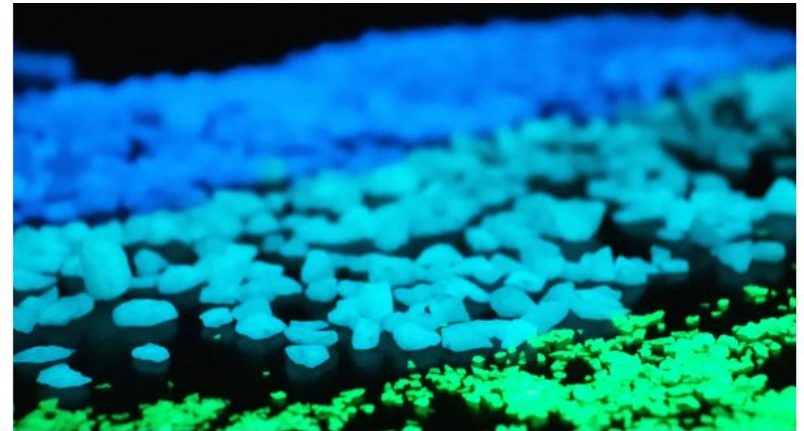
# Research

- CCAST staff have research experience in multiple domains:
  - Bioinformatics
  - Chemistry
  - Computer Science
  - Materials Science
  - Microbiology
  - Physics
  - Statistics
- **40+ years** combined experience in high-performance and scientific computing.

NEWS NORTH DAKOTA

## 30 years after discovery, NDSU researcher finds how the material in products keeps glowing

Khang Hoang has found what makes a pigment in luminous paint continue to glow, long after the lights are out.



Strontium aluminate stones from CORE Glow of British Columbia, Canada give off blue, green and aqua glows. Courtesy: CORE Glow

<https://www.inforum.com/news/north-dakota/30-years-after-discovery-ndsu-researcher-finds-how-the-material-in-products-keeps-glowing>

# CCAST support for AI at NDSU

	<b>Available now</b>	<b>In progress</b>	<b>In the future</b>
<b>Datacenter infrastructure</b>			
<b>Software and tools</b>			
<b>Training and support</b>			

# Datacenter infrastructure for AI

## Available now

- Datacenter power and cooling for up to 1 MW IT load.
- 90 Nvidia GPUs – A100, A40, A10, L4, P100, RTX.
- 2+ PB HPC data storage for large datasets.
- High-speed 100 Gb networking.

# Datacenter infrastructure for AI

## Available now

- Datacenter power and cooling for up to 1 MW IT load.
- 90 Nvidia GPUs – A100, A40, A10, L4, P100, RTX.
- 2+ PB HPC data storage for large datasets.
- High-speed 100 Gb networking.

## In progress

- 2<sup>nd</sup> row datacenter expansion to add more rack space.
- Procurement of new and larger GPUs, e.g., Nvidia H100/200.
- Evaluation of NVLink for fast GPU-GPU communications.



# Datacenter infrastructure for AI

## Available now

- Datacenter power and cooling for up to 1 MW IT load.
- 90 Nvidia GPUs – A100, A40, A10, L4, P100, RTX.
- 2+ PB HPC data storage for large datasets.
- High-speed 100 Gb networking.

## In progress

- 2<sup>nd</sup> row datacenter expansion to add more rack space.
- Procurement of new and larger GPUs, e.g., Nvidia H100/200.
- Evaluation of NVLink for fast GPU-GPU communications.

## In the future

- NSF MRI proposal for new AI-focused cluster.
- Exploration of new GPU architectures, e.g., Nvidia Grace-Hopper, AMD Mi-X GPUs, etc.
- All flash data storage systems, e.g. VAST, Weka.

# Software and tools for AI

## Available now

- Popular AI/ML frameworks, e.g. TensorFlow, PyTorch.
- Tutorial for training neural networks on CCAST.
- GPU-enabled Jupyter Notebooks for interactive AI.

# Software and tools for AI

## Available now

- Popular AI/ML frameworks, e.g. TensorFlow, PyTorch.
- Tutorial for training neural networks on CCAST.
- GPU-enabled Jupyter Notebooks for interactive AI.

## In progress

- Developing more tutorials and guides for popular tools.
- Deploying more “batteries included” workflows.
- Expanding menu of preconfigured Jupyter kernels.

# Software and tools for AI

## Available now

- Popular AI/ML frameworks, e.g. TensorFlow, PyTorch.
- Tutorial for training neural networks on CCAST.
- GPU-enabled Jupyter Notebooks for interactive AI.

## In progress

- Developing more tutorials and guides for popular tools.
- Deploying more “batteries included” workflows.
- Expanding menu of preconfigured Jupyter kernels.

## In the future

- Distributed DL workflows for running large models across multiple nodes.
- Custom AI environments for the classroom.
- Platforms for real-time AI inference using pretrained models.

# Training and support for AI

## Available now

- Regular training program in advanced research computing each semester.
- Staff expertise in AI/ML – Scientific Software Specialist hired Nov. 2023.
- Individual support for users struggling with AI workflows.



# Training and support for AI

## Available now

- Regular training program in advanced research computing each semester.
- Staff expertise in AI/ML – Scientific Software Specialist hired Nov. 2023.
- Individual support for users struggling with AI workflows.

## In progress

- Dedicated workshops in applied AI/ML for researchers in all fields.
- Adding a second Research Facilitator in Fall 2024.
- Hiring more student interns in Spring 2025.

# Training and support for AI

## Available now

- Regular training program in advanced research computing each semester.
- Staff expertise in AI/ML – Scientific Software Specialist hired Nov. 2023.
- Individual support for users struggling with AI workflows.

## In progress

- Dedicated workshops in applied AI/ML for researchers in all fields.
- Adding a second Research Facilitator in Fall 2024.
- Hiring more student interns in Spring 2025.

## In the future

- Evaluating needs for staff AI and data science experts.
- Potential for regular AI training program separate from standard HPC training.

# Tips for faculty and staff

## **1. Discuss AI needs with CCAST.**

If you plan to make use of CCAST HPC resources for AI research and education, set up some time to talk with us about your needs. In particular, if you plan to use CCAST in the classroom, let us know of your needs a semester in advance so we can make sure you have the resources you need prior to the start of your course.

# Tips for faculty and staff

## **2. Participate in CCAST HPC training programs.**

HPC skills are essential for conducting AI-related research at scale. The best way to develop these skills is to participate in our training programs, offered every Fall, Spring, and Summer semester. CCAST training programs are open to all faculty, staff, and students at NDSU, and their collaborators at other institutions.

# Tips for faculty and staff

## 3. Contact CCAST Support for help.

Adapting existing AI workflows to run on HPC clusters can be difficult and is a frequent source of frustration for AI researchers. CCAST staff have substantial experience deploying AI/ML software for HPC, and we are here to help. Review our Knowledge Base articles and tutorials on AI/ML and email us at [ndsu.ccast.support@ndsu.edu](mailto:ndsu.ccast.support@ndsu.edu) if you're having trouble getting AI codes to run properly.



# NDSU

CENTER FOR  
COMPUTATIONALLY ASSISTED  
SCIENCE AND TECHNOLOGY

Learn more at [ndsu.edu/ccast](https://ndsu.edu/ccast)

or

Contact us at [ndsu.ccast.support@ndsu.edu](mailto:ndsu.ccast.support@ndsu.edu)