



NSF LEADERSHIP-CLASS
COMPUTING FACILITY

The Texas Advanced Computing Center (TACC), the NSF Leadership Class Computing Facility (LCCF), and how we support Astronomy

Dr. Jennifer Schopf

Large Scale Science Liaison, LCCF

Texas Advanced Computing Center, UT-Austin

May, 2026

TACC
TEXAS ADVANCED COMPUTING CENTER

 **TEXAS**
The University of Texas at Austin



What is TACC?

- Texas Advanced Computing Center, part of the University of Texas at Austin
- The largest NSF-funded national computing center for open science (plus NIST, NASA, NIH, DARPA, DOD, etc.)
- *Research* Computing
 - High Performance/ Supercomputing
 - AI
 - Analytics/Big Data
- 200+ Dedicated staff
- Altogether, ~20k servers, >1M CPU cores, 1k GPUs
- About seven billion core hours over several million jobs per year – for 3,000 projects and ~40,000 users per year.

*While we are a national provider, we have ***by far*** the most computing resources of any University in the United States— in speed, diversity, support, real science capability, etc.*



Who uses TACC?

- Around the country, users doing unclassified research at more than 400 institutions use TACC on over 3,000 projects for year
 - Including a number of startups; large industry use us more for tech pathfinding
- Since it's inception in 2001, TACC has had well over 120k users – 100k+ of which are students
 - >30k use the resources in any given year
- TACC users have **4 Nobel prizes**, many Gordon Bell prizes, and countless “first of its kind” computational achievements
- On our 4 largest systems, our many users submit 477 jobs an hour
 - That's a new submission every 7.5 seconds, 24x365
 - A “job” can be 1-2,048 servers, for up to 48 hours of runtime

How We Support Users

- Provide researchers with:
 - Computing, data, AI, software capabilities to support their research
 - The expert help to be able to use it
 - In the ways you want to consume it
 - Help with grants/strategy
- We also help institutions:
 - Recruit top faculty and top students
 - Win large grants, leveraging our capabilities



The NSF Leadership Class Computing Facility

- US NSF funded Major Research Equipment Facility Construction (MREFC)
- A more capable follow-on to the current (aging) NSF computing systems
- A more holistic and collaborative view of how we support “leadership applications”
- A long-term investment to match the science missions we support



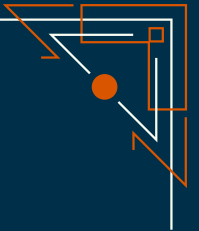
Primary LCCF Constituencies

- The current NSF large-scale simulation users (US NSF ACCESS)
- The NSF Large Facilities Community (MREFC's – all fields)
- NSF Mid-Scale infrastructures current and future (SAGE, FABRIC...)
- Emerging new users in new fields of science, particularly in AI
 - AI Institutes will be a key clientele, as will NSF NAIRR researchers
- The CI community beyond the users:
 - Research CI staff at universities, other NSF-funded centers
 - CS researchers
 - Colleagues at DOE, NASA, DOD and other labs.

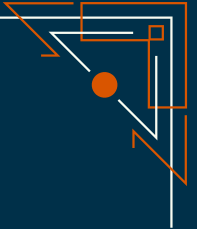
Horizon (The First LCCF System)

- Accelerated: 2,000 GPU nodes with a Grace socket and two Blackwell GPUs, 800Gb Infiniband (non-blocking).
- CPU: 4,750 Vera-Vera nodes, with two 88-core Vera sockets, 400Gb Infiniband.
- All nodes built in the new Dell Integrated Rack infrastructure
- Shared filesystem, Infiniband connected, roughly 400 usable PB, 100% Solid State. (8/16TB/s W/R)
- Primary compute capability will be :
 - ~300 PF double precision (10x Frontera) - Goal 15-20x for most apps
 - More with emulation??
 - ~10 EF "TF32" precision
 - ~20 EF bfloat16 precision
 - ~80EF INT4/TF4, if you are into that kind of thing

Horizon Racks



- As we are using the NVIDIA Blackwell “NVL-4” shelf, we aren’t limited to 72 GPUs per cabinet.
 - Weigh ~6,000lbs (~2,700 kg) with liquid and rear doors
 - ~2 Honda Civics, depending on your trim package
- 28 GPU racks will have 144 GPUs, and come in at about 215KW per rack
 - About half up and running (mid May 2026)
 - Initial public use expected ~July 2026
- Horizon CPU (Vera) racks will come a little later, but in roughly the same form factor – 72 dual-socket nodes per cabinet, >100KW per rack
 - Expect 66 CPU Racks
- Plus 34 storage, switching, management racks, 128 cabinets total
- Estimated ~13MW



Active Storage

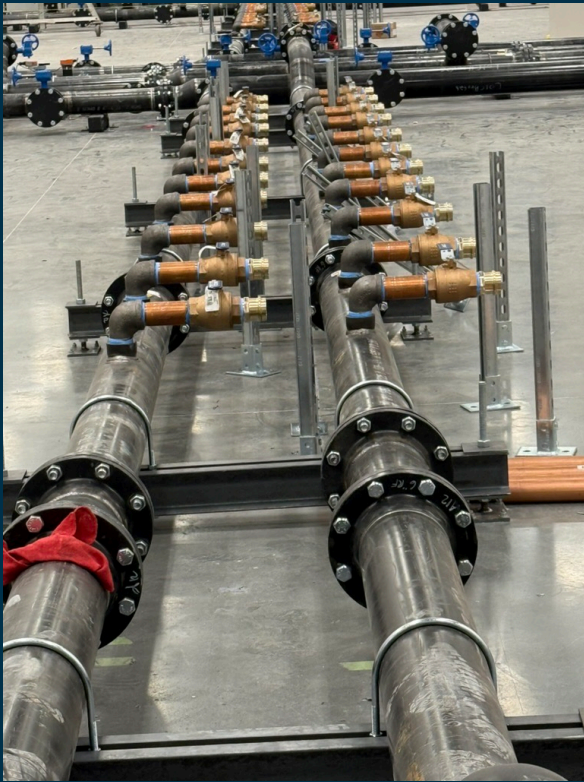
- Horizon storage- 400PB *all solid state*
 - From a bandwidth perspective, we will have 25x the total of all three scratch filesystems provided on Frontera (TACC's last large offering)
 - But for the small file access patterns we see most of the time, we think we will see even more to most apps
 - We expect the interruption time to do file system issues (even the short non-fatal ones) should decrease significantly
- The primary Horizon storage is VAST
 - We also have deployments of VDURA and WEKA for some other parts of TACC in the existing datacenter

Archive System Update as well

- >1Exabyte Tape - Ranch
 - 2018 - 100PB
 - 2012 - 10PB
 - 2006 -1PB
- People ask why we don't worry about deleting old data
 - Data from last ~3 years: 90% of all data
 - All data generated by humanity since the dawn of time until ~3 years ago: 10% of all data.



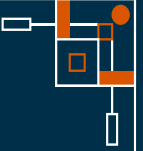
Our New Facility: Built to support 250KW/cabinet



September 2025
Underfloor cooling install



October 2025
Raised floor install

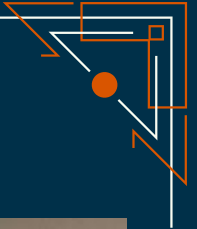


We are currently deploying

- As of January 1. . .



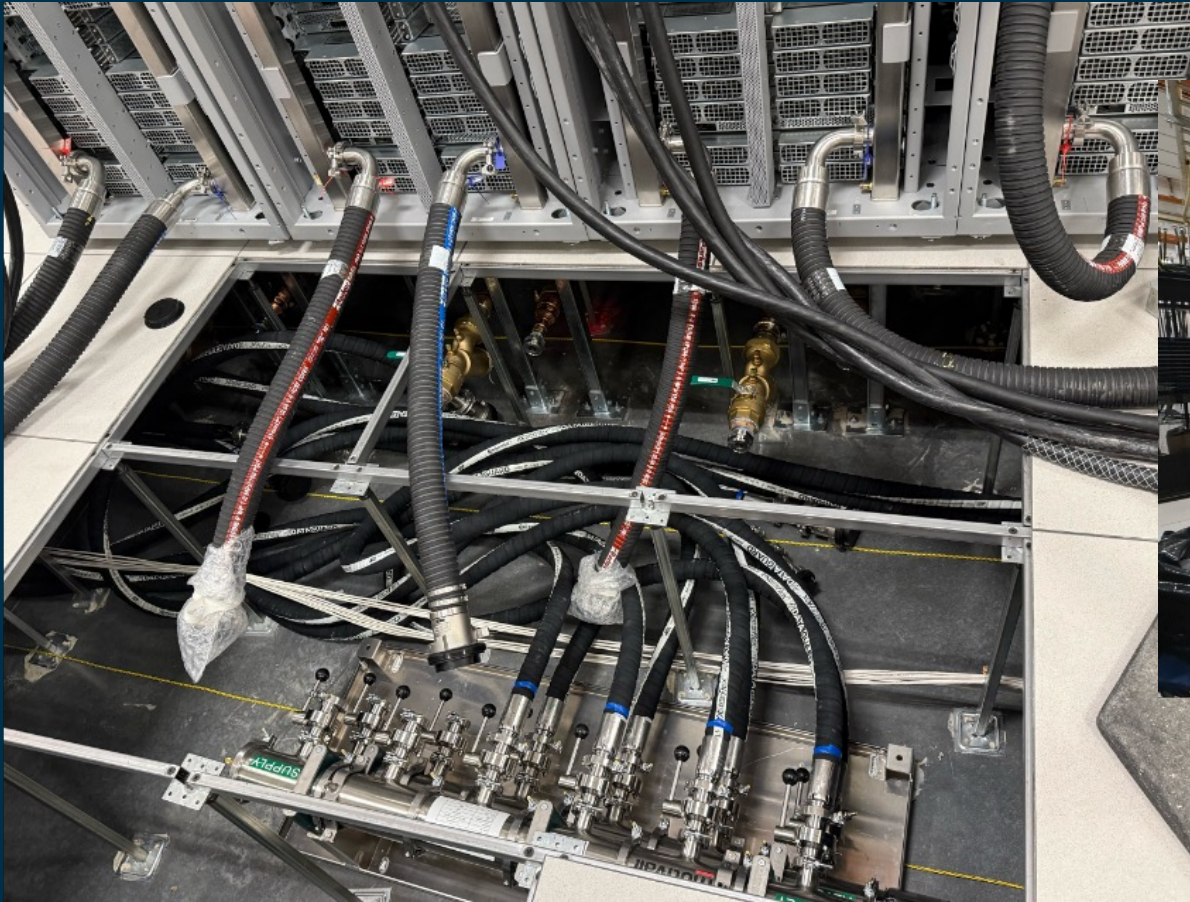
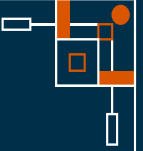
March 2026

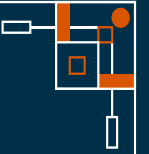


March 2026

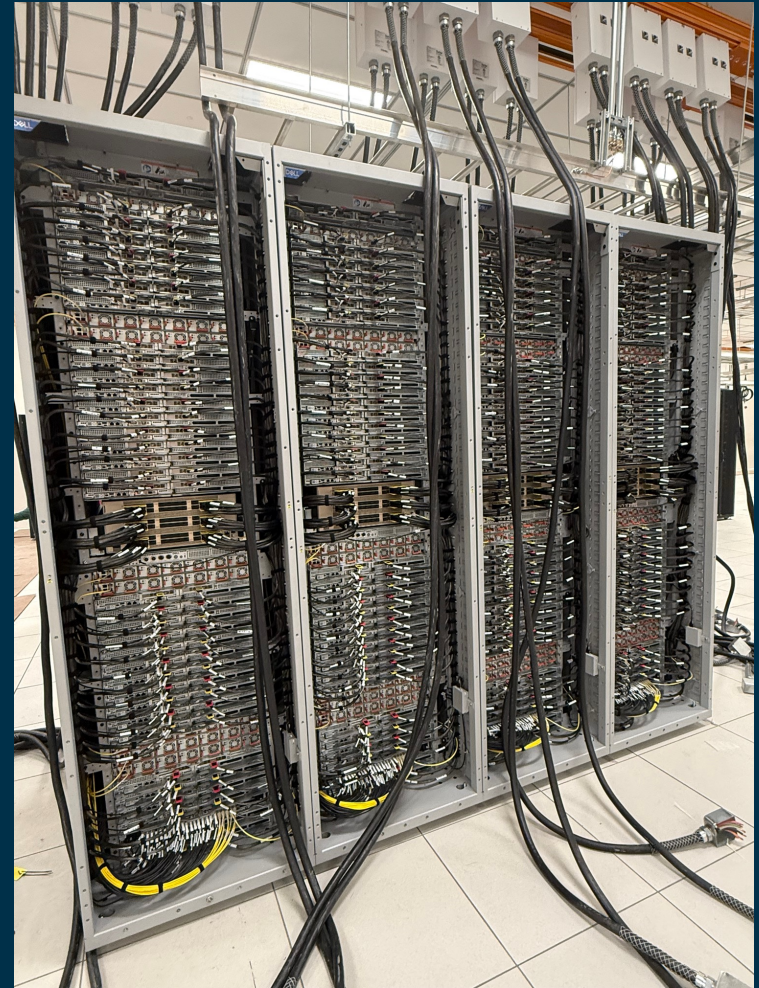


22 Racks on the Floor 4/1/26

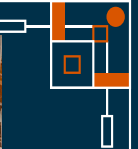


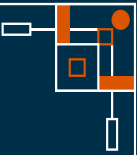


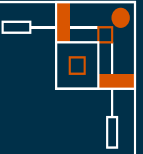
Early May 2024



Next: John Desantis

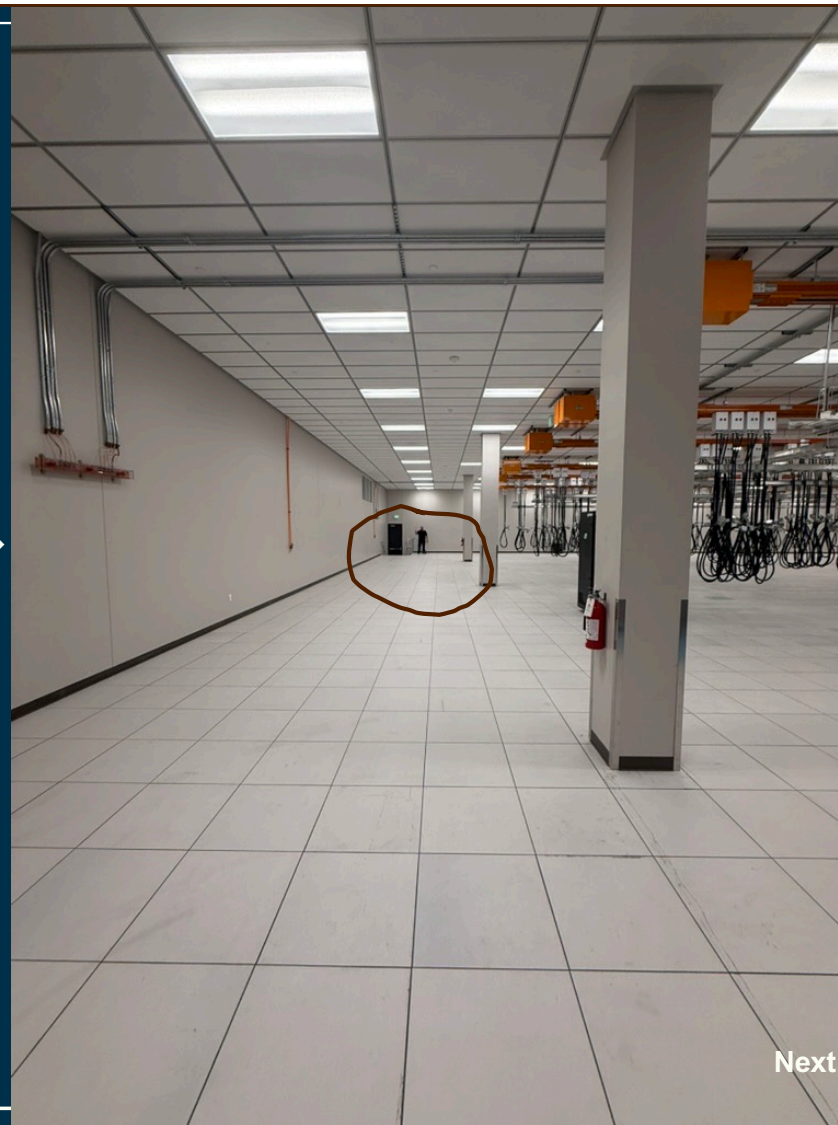






- Still lots of space to fill

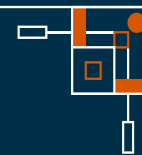
- For scale, that's Nathaniel ----->



Next: Jonelle

Justin

31



Different Acceptance Criteria- Characteristic Science Applications (CSA)

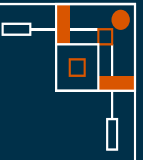
- CSA: Cutting-edge software projects built to tackle critical challenges that will define the facility's success
 - Purpose-built for the challenges scientists face and for the way scientists work
- Instead of designing supercomputers around abstract benchmarks (eg LINPACK), CSAs are built around real science applications
- Design not to achieve a particular performance benchmark, but to accomplish new science at unprecedented scale. Performance expectations for Horizon
 - >10X Peak of Frontera
 - >>12x "Science" by our CSA metrics.
- ~140 proposals, 30 examined in first round, downselecting to ~10

CSA Projects

General Area of Science	Application	Project Name
Astronomy and Astrophysics	Athena++	Astrophysical Fluid Dynamics at Exascale
	ChaNGa	Evolution of baryons and galaxies across the age of the Universe
	IceTray	Multi-Messenger Astrophysics with ICECUBE
	Enzo-E	Accelerating cosmological simulations of the first galaxies through deep learning
Biophysics and Biology	NAMD	Molecular Mechanisms of Viral Infection
	ensemble	Viruses in Respiratory Aerosols
	WESTPA/AMBER	Phase Separation of Disordered Proteins
Computational Fluid Dynamics	ensemble	Multiphysics simulation of a full hypersonic vehicle
	PSDNS	Large-scale DNS
Geodynamics and Earth Systems	AWP-ODC	Seismic simulation for hazard management
	CESM2	Coupled Earth-Atmosphere models
	CM1	Supercell thunderstorms and tornado prediction
	ISSM	Data-driven and physics modeling of ice dynamics
	SeisSol	Off-fault inelastic processes and fluid effects in earthquake simulation
	rhea	Bridging short and long time scales in global plate tectonics
Materials Engineering	EPW	Quantum materials engineering at the exascale
	MuST	Electron localization in materials
	PARSEC	Quantum calculations for the optical and dielectric properties of aqueous liquids
Other Applications	Grover	Detecting misinformation and social biases
	MILC	Lattice QCD for flavor physics

How does this change the machine?

- We're designing LCCF and Horizon for Science
 - CSA problems: only about 40% likely to use GPU acceleration by 2026
 - Different tuning point than DOE systems, for example
- Our mix looks like some other "national" systems
 - LUMI, KAUST, LRZ, etc.
- We chose to put 60% of our dollars into CPU – which translates to only about 20% of the performance

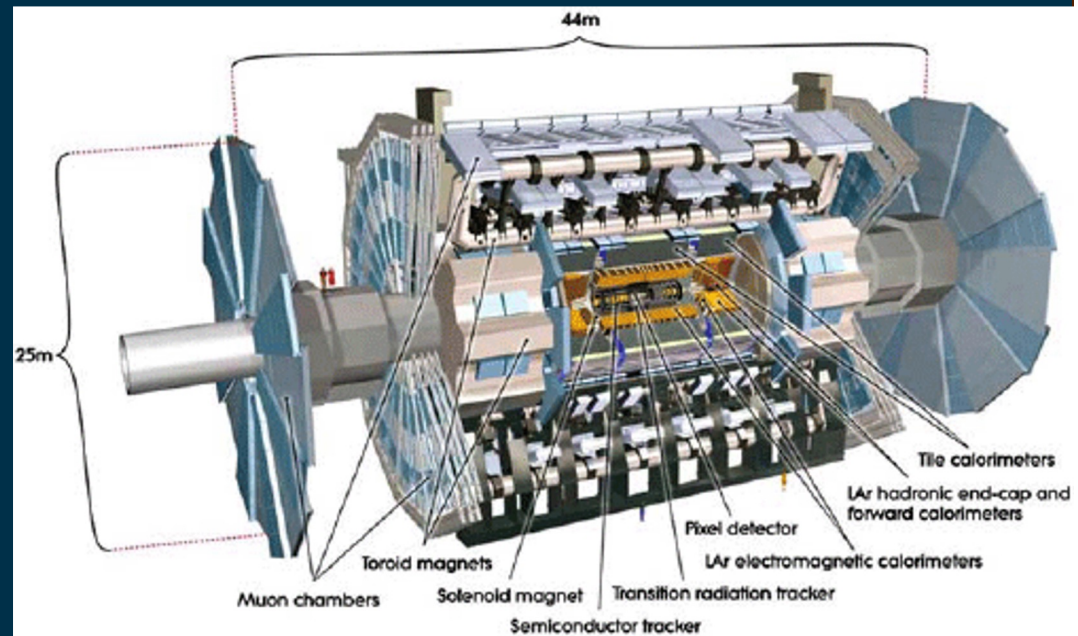


LCCF Goal is focused on real support for real science

-LCCF Science Liaison is now reaching out to all MREFCs, NSF MidScales, and other large science projects to learn what kind of support LCCF can offer to research teams

Large Hadron Collider ATLAS/CMS Instruments

- Over 6M SUs granted on Frontera for research across these projects (Plus on many other machines through XSEDE/ACCESS, dating back well over a decade).
- 500TBs of data flow for ATLAS in last year
- 3.34PB of data flow for CMS in last year



NCAR (NSF Geo)

- Live tape backup of full storage
 - 15PB of data transferred over 3 months from NCAR Wyoming Data Center to TACC
- Ongoing work to have TACC resources CSEM certified
 - Researchers can burst to TACC when Derecho isn't large enough for runs
- Possible shared work with tuning software for new architectures
- <https://ncar.ucar.edu/>



ASIDE

- We do work to improve large end-to-end data transfers
 - Orthogonal to other services
 - TACC doesn't need to be an end point
- NCAR transfers went from 1st PB 2 weeks, last PB 4 days
 - Bottle neck was tape cache at TACC, not the network or disk
- We have a public PS and Globus DTN testers as well

Laser Interferometer Gravitational Wave Observatory (LIGO)

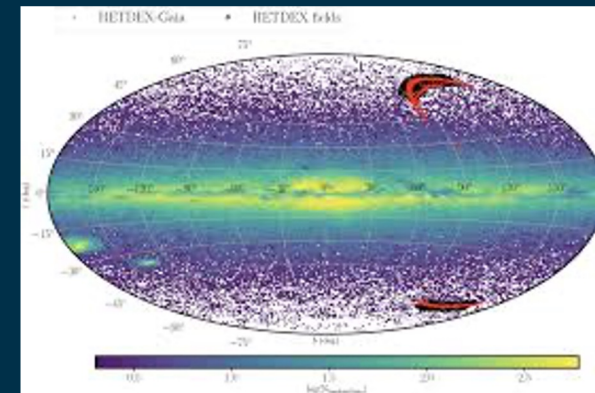


- 2.8M SUs on Frontera recently
- Many allocations on many machines through the years



Hobby-Eberly Telescope Dark Energy Experiment (HetDex)

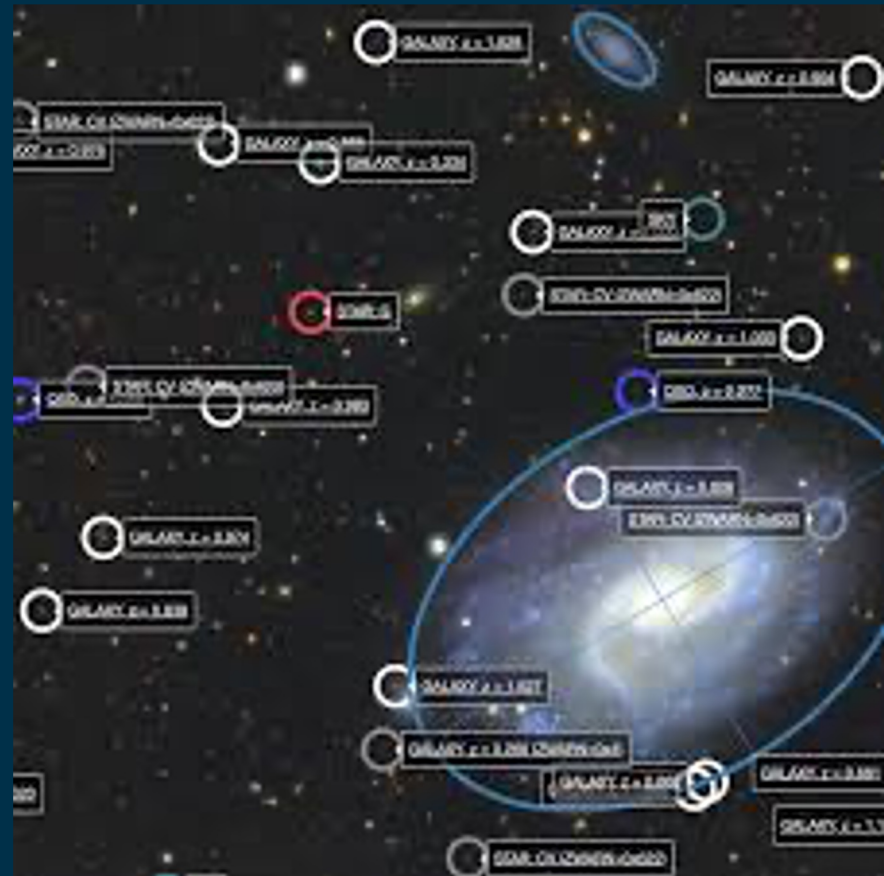
- An experiment to produce the largest map of the cosmos and uncover the nature of dark energy
- over 500TB now, special instrument, all the data and processing here
- <https://hetdex.org/>
- *HETDEX has received financial support from the State of Texas, the United States Air Force, the National Science Foundation and the generous contributions of many private foundations and individuals.*



Cosmic AI – contact Niall Gaffney

- Joint with NOIRlab and NRAO
- Transformative AI to increase the accessibility of astronomy data and knowledge for researchers, students, and the public
- Building next-generation AI tools to accelerate discoveries related to the most basic human question: Where do we come from?

<https://www.cosmicai.org/>



Next Gen Very Large Array Telescope (ngVLA)

- Consulting contract to assist in design of next generation data center and processing pipeline
- Joint work with software tuning, benchmarking, hardware design
- Joint work on data transfers (expected data rate >40PB/month)
- <https://ngvla.nrao.edu/>



Vera Rubin Observatory

- Spoke with Leanne Guy and her team just last week
- TACC as possible DAC for VRO data
- Possible support to tune codes to run on LCCF machines, work on tuning workflow



LCCF Focus on Science Support

- Yes, we'll have the biggest NSF funded machine in the US
 - Goal is to support research of all types
- Focus on science – and how we can help all of you achieve your science goals
- Contact: Jennifer Schopf

jms@tacc.utexas.edu



NSF LEADERSHIP-CLASS
COMPUTING FACILITY