HPE DSI Townhall Meeting

October 4, 2021







Research





Research Computing Data Core

Carya



10K CPU cores & 327K GPU cores

64 Nvidia Volta V100 GPUs (32Gb/e)

Interconnect, Mellanox HDR Infiniband, 100Gb/s Line Rate



1,560 TB of shared hard-disk based storage

122 TB of shared flash storage





5,704 CPU cores

8 nodes with 2 NVIDIA P100 GPUs each

4 nodes with 8 NVIDIA V100 GPUs each



725TB NFS storage

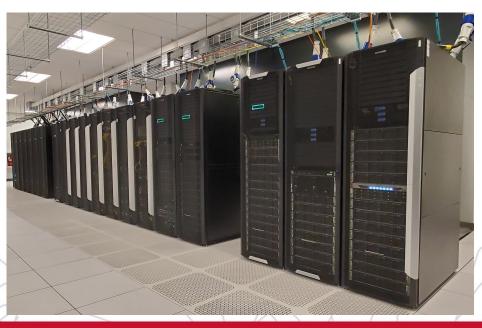
Opuntia



1,860 CPU cores
4 Nvidia K40 GPU nodes



600 TB shared storage





Research Computing Data Core

- Clusters utilization:
 - Serving ~ 100 research groups from 18 Departments

Cluster	Awarded SUs	Awarded Gb
Carya	58,994,064	504 TB
Sabine	23,449,046	511 TB
Opuntia	5,300,000	367 TB

• 121 Allocation renewals for the compute cycle September 2021 to August 2022



New Data Visualization Lab AERB building, room 201

- Flexible lab, designed for different layouts
- Groups stations
- Instructional
- Research
- Seminars





New Data Visualization Lab AERB building, room 201





Seminars and Workshops

- Look for FEATURED EVENTS in our website
- One webinar every week
- Featuring cougars and building UH's data science community
- Promoting partnerships with industry (HPE, NVIDIA, ConocoPhillips, Humana Institute, Mark III Systems, many more) and facilitating opportunities for students and faculty to network
- Partnership with professional organizations
 - Geophysical Society
 - Society of HPC Professionals
- Hands-on workshops: STATA, Python, Matlab, GPUs

FEATURED EVENTS

View All Events >



Intro to Deep Learning

Brought to you by the University of Houston HPE Data Science Institute and Mark III



HPC and FEM

Brought to you by the University of Houston HPE Data Science institute and the Society of HPC Professionals



Research Areas

- Faculty Interest Groups
 - Visualization
 - Image Analysis
 - Robotics
 - Scientific Computing
 - ML/AI
 - Natural Language Processing



Current HPE DSI Collaboration Projects

- Data Science Projects (Ishita Sharma):
 - DSI-College of Medicine: Medical Legal Planning Project(MLP)
 - This project revolves around the evaluation of patients suffering with HHLN(Health Harming Legal Needs) is associated on various factors such as chronic diseases/number of hospital visits and so on
 - The factors considered for the effective implementation of MLP by patients
 - My contribution involves the creation of different databases for the data being collected through patient surveys
 - DSI-College of Education Research Department Project
 - The creation of database management system of the assessment data for flexible data extraction by researcher scientists
 - DSI-Data Science Community
 - Providing data science skills consultation and solutions to the undergraduate researchers

Ishita is available to consult on Data Science projects. Contact us if interested.





Communication

We continue building our brand!

 A lot of new sections in our website, hpedsi.uh.edu

 Featured Data Science short stories, covering success events, interviews, webinars and more



Using Bayesian Techniques to Speed Up HPC Solutions

Bayesian inference interprets probability as a measure of believability or confidence that an indivi...



Data, Society and Justice

UH Team Empowers Re-entry Programs Across Harris County

Source: HPE Data Science Institut...



Communication

- Our newsletter keeps you up to date!
- Articles, research, opportunities and upcoming events relevant for all data science
- Are you not getting The Front End by email? Contact Tim Holt at cjholt@central.uh.edu

Front End

The FrontEnd August Newsletter

The <u>Hewlett Packard Enterprise Data Science Institute</u> champions research and academic programs to strengthen the University of Houston's position within the developing data landscape.



Privacy Versus Personalization

Taking inspiration from justice theory, proposed in 1965 by psychologist John Stacey Adams, the dean of the C. T. Bauer College of Business and his colleagues sought to add fairness to privacy policy agreements. They devised an IT-enabled artificial intelligence system that links tiers of customer information with levels of personalization.

Read more..



Smartphone-Based Rapid COVID-19 Test

"We live in a world where you can have all your basic necessities delivered to your home through an app, and yet this pandemic has exposed how far behind the diagnostics industry lags compared to consumer technology and the convenience economy," said Andrew Paterson, UH alumnus and Luminostics co-founder, of his company's newest technology.

Read more...



Education





Courses and Workshops

- FREE to CURRENT UH community (faculty, students, and staff)
 - Non-UH affiliates can attend for a fee
- Digital Badges offered upon completion
 - Open badge standard
 - Linkable to job social media sites
- 15 contact Hours (typically two 1.5 hrs meeting per week)
 - Online





Courses (Fall and Spring)

- Fall and Spring (2021)
 - 101 Introduction to Cluster Computing and Linux
 - 211 R for Data Science
 - 212 Scientific Programming with Python
 - 251 Data Visualization using Paraview and Tableau
 - 261 Principles of Data Management
 - 311 Introduction to Machine Learning
- Fall only (2021)
 - 221 Data Science and Scientific Computing using Julia
 - 271 Excel for Data Science
 - 301 Parallel Computing with OpenMP



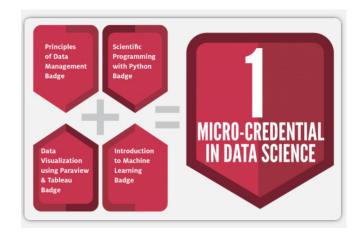
Enrollment

	Spring 2021	Summer 2021	Fall 2021
101 Cluster Computing	128	157	52
201 Scientific Programming with C++	61		
211 R for Data Science	58	72	39
212 Python	91	148	72
Bootcamp Python (Summer Only)		199	
221 Data Science with Julia	22		22
251 Data Visualization	133	115	58
261 Principles of Data Management	48	85	50
271 Excel for Data Science			58
301 Parallel Computing with OpenMP			15
303 GPU Programming	28		
311 Machine Learning	90	102	42



Micro-credentials

- Micro-credentials now being offered
 - Micro-credential in Data Science
 - FREE to CURRENT UH students, faculty and staff
 - Others can participate for a fee
 - Requires a pass or badge in 4 data science courses
 - 212 Scientific Programming with Python
 - 251 Data Visualization using Paraview and Tableau
 - 261 Principles of Data Management
 - 311 Introduction to Machine Learning





Partnerships





Engaged Data Science – Data & Society Minor

- Partnership with Honor College
- 15-hour program (open Honors) launched fall 2020
- Declared minors 22; Fall '21 enrollment ~60
- Courses & Curriculum: project-based learning across multiple semesters; role of data in specific fields (e.g., criminal justice, public health)
- Approach: Bring humanities questions to data science to create career pathways for people who need to communicate with data
- Cross-course evaluation: individual trajectories & project-based learning outcomes
- Expanding affiliated faculty: unified around undergraduate education, project-based learning, evolving understanding of data



Engaged Data Science – Undergraduate Research

- Pharis Fellowship (summer '21)
 - 11 undergraduate students 10wk, full-time, online \$4K stipend
 - Project areas: health, education, criminal Justice
- Data Analytics in Students Hands (DASH)
 - ~20 undergraduate students ongoing academic year projects
 - Project areas: food insecurity, incarceration & reentry, disaster relief, covid-19 & frontline healthcare workers, equity in scientific funding





Engaged Data Science – K-12 Outreach

Project Engagement Encouraging Rising Students (PEERS)

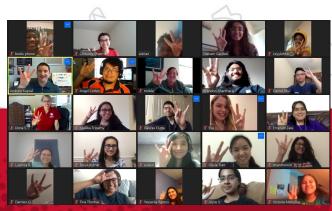
- Local HS students mentored by UH undergraduates (near-peer)
 - **2018-21:** ~220 participants | **2021-22 (target):** ~100 participants
- City-wide capstone competition (May) 11 teams
- UH student org: Honors in Community Health (HICH)
- Multiple external conference presentations (APHA, TEACH-S, TSOPHE)

HPE DSI Girls Coding Camp

- 6-8th grade: Scratch, HTML, Game Design, Python Programming
- 9-12th grade: Python Programming and Game Design







Engaged Data Science – Community Partnerships

Community Health Worker Initiative (CHWI)

- Training and supporting CHWs (~300 since 2018)
- Rethinking potential for data to address social determinants of health
- \$2.2M external/partnership funding
 - Vaccine access, pediatric neurological health, rare type diabetes
- Data, Society, and Justice Initiative
 - Restoring Justice wholistic indigent defense
 - Harris County Justice Administrative Department
 - Public Defenders Office







Bauer College of Business

- Collaboration on AI in Business
 - Kevin Hong, Professor
 - Yumei He, Postdoc
 - Meng Li, Associate Professor
- Research support
- Teaching in the HPE DSI
 - Advanced Excel for Data Science

HPE DSI Student Labs

- Undergraduate research
 - External organization
 - Dual mentoring
 - Data science focus
- HPE
- Humana Inc./Humana Institute
- Harris County Public Defender's Office/Prof. Grigorenko
- PCIC/Prof. Buck
- Montrose Center/IRWGS/Prof. Gregory

