A Whirlwind Tour of Science on the NRP

Invited Presentation

Seventh National Research Platform Conference (7NRP)

UC San Diego

May 7, 2026

Dr. Larry Smarr

Founding Director Emeritus, California Institute for Telecommunications and Information Technology;

Distinguished Professor Emeritus, Dept. of Computer Science and Engineering

Jacobs School of Engineering, UCSD

http://lsmarr.net

NRP NATIONAL RESEARCH PLATFORM

NRP Nautilus

I Will Not Be Covering the NRP Science Applications I Reported on at 4NRP, 5NRP, and 6NRP

- **Particle Physics**
  - LHC (UCSD)
- **Astronomy**
  - LIGO (Caltech/MIT)
  - IceCube (U Wisconsin)
  - CASPER Radio Astronomy (UC Berkeley)
- **Earth Sciences**
  - WIFIRE and SmokeyNet (UCSD)
- **Chemistry**
  - OpenForceField Quantum Chemistry (OFF Initiative)
- **Botany**
  - Redwood Core Analysis (CSU Humboldt)
- **Biomedical**
  - Braingeneers (UCSC)
  - Brain Imaging with Electron Microscopes (UCSD)
  - TemPredict COVID (UCSD)

These Continue to be Major NRP Resource Users

Cal Poly San Luis Obispo Professor Lauren Chan
Evolutionary Biology and Conservation Computational Genomics

- **Genomic resources for non-model organisms allow us to:**
  - **Reconstruct the evolutionary history of groups**
  - **Test for evolutionary adaptation and population connectivity**
- **The inferences we can make can have important implications for conservation and management**
- **A challenge is that reptile and amphibian genomes can be huge:**
  - **Lizards: 2 billion base pairs (Gb), Frogs: 4-10 Gb, Salamanders: estimated 30-90+ Gb!**



Lauren Chan, Assoc. Prof.
Cal Poly SLO







In 2025: 253 GPU-hrs on namespace: cal-poly-chan-lab

Source: Lauren Chan

NRP NATIONAL RESEARCH PLATFORM

NRP Nautilus

Amphibian and Reptile Conservation Genetics

- Two main tasks studying the Federally Endangered Dunes Sagebrush Lizard
  - Assembly of a reference genome for this non-model organism
  - Whole genome resequencing for 100 individuals to estimate connectivity, inbreeding, and adaptation across the range
- Using the NRP!
  - High memory reference genome decontamination steps
  - GPUs to call variants across entire genome for 100 individuals
  - Post-calling filtering and data processing
  - Comparative genomics



Kat Arrizon, UG
Cal Poly SLO

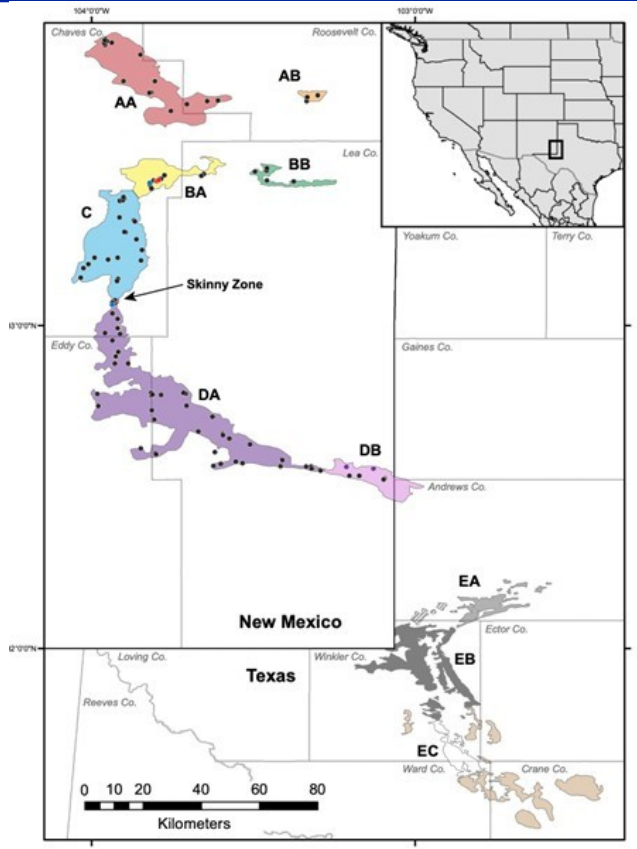


Emi Degembe, UG
Cal Poly SLO



Sydney Gutierrez, UG
Cal Poly SLO



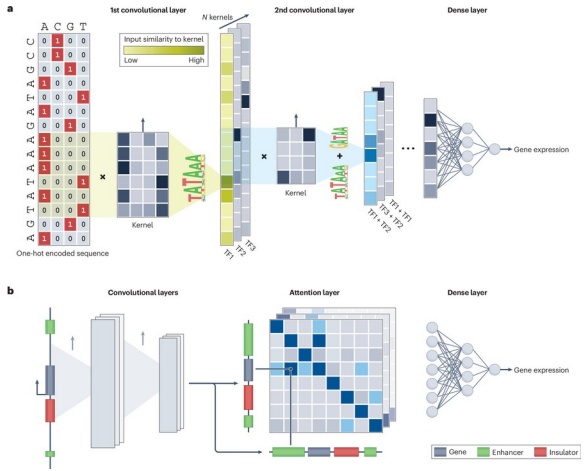


Source: Lauren Chan

NRP NATIONAL RESEARCH PLATFORM

NRP Nautilus

Stanford University Professor Anshul Kundaje Uses Deep Learning to Decode Regulatory Genomics



Deep learning architecture for predicting gene expression from DNA



Anshul Kundaje, Assoc. Prof. — Stanford (Genetics & CS)

- Deep learning models for regulatory DNA sequences
- ATAConvolutional neural networks map raw DNA sequence to TF binding, accessibility, and gene expression
- Interpretation tools (DeepLIFT, TF-MoDISco) recover regulatory motifs and predict variant effects
- Interpreting regulatory genomics with neural networks (keratinocyte differentiation)

NRP NATIONAL RESEARCH PLATFORM

In 2025: 129,544 GPU-hrs on namespace: kundajelab

NRP Nautilus

CSU Fullerton Graduate Student Akshat Desai
Used Machine Learning for Brain MRI Image Analysis and Classification

"My research under Professor Rakesh Mahto focuses on applying machine learning and radiomics analysis to brain MRI data for Alzheimer's disease classification.

Using MRI scans from the ADNI dataset (1,229 subjects), I developed a pipeline to extract high-dimensional texture and shape features from hippocampal and amygdala subfields.

These features were used in a two-stage machine learning framework (LASSO feature selection and XGBoost classification) to distinguish Alzheimer's Disease, Mild Cognitive Impairment, and Cognitively Normal subjects."
-Akshat Desai, CSU Fullerton



Graduate student Akshat Desai (left) with faculty research adviser Rakesh Mahto

NRP NATIONAL RESEARCH PLATFORM

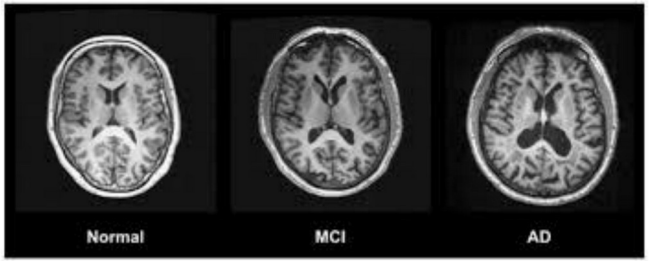
In 2025: Desai and Mahto Used 4500 GPU-hrs on namespace *jupyterlab*

NRP Nautilus

Large Dataset of Human Brain MRIs Are Analyzed with Machine Learning For Biological Signatures of Aging

"CENIC AIR Resources
Enabled Efficient Processing
of 1,229 MRI Scans,
including Automated Brain Segmentation,
Radiomics Feature Extraction,
and Machine Learning Analysis."

- **Dataset Source:** Alzheimer's Disease Neuroimaging Initiative (ADNI) - Large-scale neuroimaging dataset used for Alzheimer's disease research.
- **Subjects**
  - Total: **1,229 participants**
  - **Cognitively Normal (CN):** 672
  - **Mild Cognitive Impairment (MCI):** 282
  - **Alzheimer's Disease (AD):** 275
- MRI scans provide structural brain information for regions involved in memory and cognition.



Key Results:

1. Identification of Micro-Structural Imaging Biomarkers from Hippocampal and Amygdala Subfields
2. Evidence that Radiomic Features Capture Biological Signals Beyond Traditional Volumetric Measures

NRP NATIONAL RESEARCH PLATFORM

NRP Nautilus

Their Prior Published Research was on Breast Cancer Subtypes, Acknowledging NRP

Journal of Imaging

MDPI

Article

Multi-Class Classification of Breast Cancer Subtypes Using ResNet Architectures on Histopathological Images

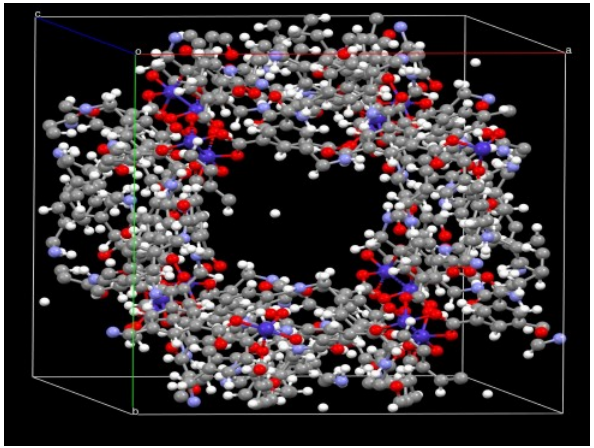
Akshat Desai [1] **and Rakeshkumar Mahto** [2,*]

[1] Department of Computer Science, California State University, Fullerton, CA 92831, USA; akshatdesai@csu.fullerton.edu
[2] Department of Electrical and Computer Engineering, California State University, Fullerton, CA 92831, USA

J. Imaging **2025**, *11*, 284

"ResNet training and testing were conducted on a high-performance workstation equipped with an NVIDIA A100 GPU (40 GB of VRAM), an 8-core CPU, and 32 GB of RAM, provided through the National Research Platform (NRP) program."

CSU San Bernardino Professor Kimberley Cousins
Quantum Molecular Dynamics for Atomic-Scale Materials Modeling





NRP GPU Acceleration Enables
Nanosecond-Scale Simulations
of Hundreds of Atoms
Running the
Vienna Ab initio Simulation Package
(VASP)

Welcome to Cal State San Bernardino JupyterHub

This JupyterHub is dedicated to Advanced Functional Materials research using The Vienna Ab initio Simulation Package (VASP)

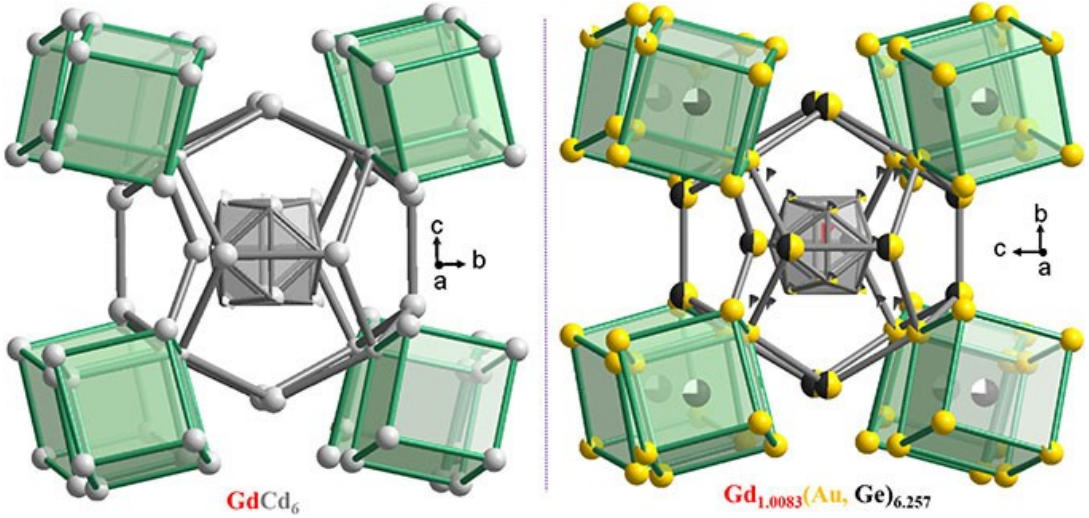
NRP NATIONAL RESEARCH PLATFORM

In 2025: 47,643 GPU-hrs on namespace: csusb-chaseci

NRP Nautilus

CSU San Bernardino Associate Professor Joyce Pham
Discovering New Inorganic Extended and Organic-Inorganic Hybrid Solids





Combining Experiment With Computational Tools
Provides A More Holistic Study of Their Atomic Arrangements

American Museum of Natural History Postdoc Joshua Fagin
Using Machine Learning to Model Quasar Variability



American Museum of Natural History

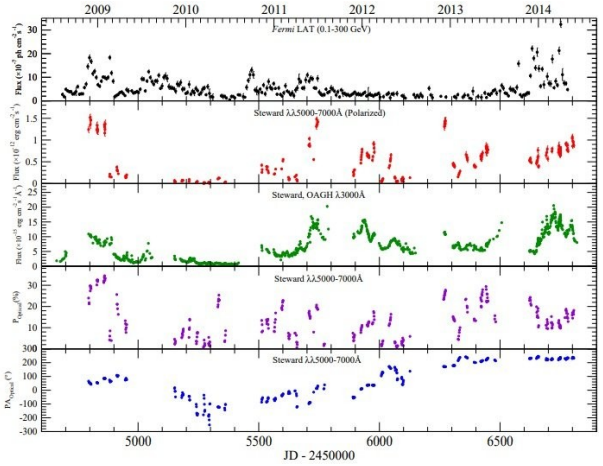


NRP Namespace PI:

Michael Benedetto, AMNH CISO and Deputy CIO



Joshua Fagin
AMNH Astrophysics
+
ML postdoc researcher at CUNY/AMNH



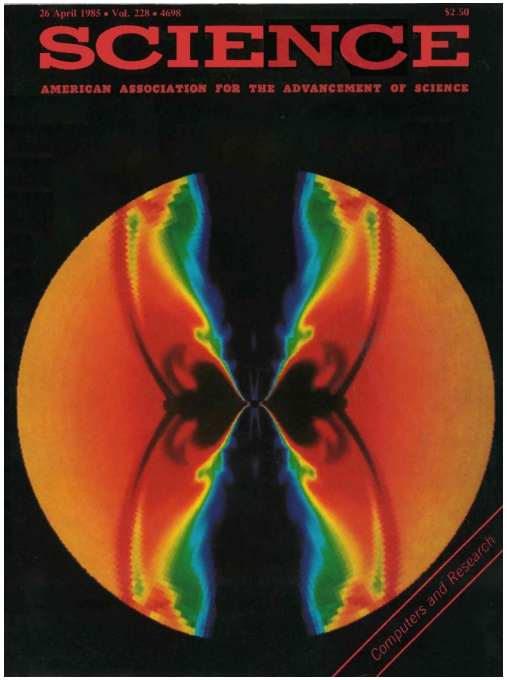
Quasar light curves: Irregularly Sampled, Noisy, Multivariate

In 2025: 17,516 GPU-hrs on NRP namespace: amnh-astro-jfagin

NRP NATIONAL RESEARCH PLATFORM

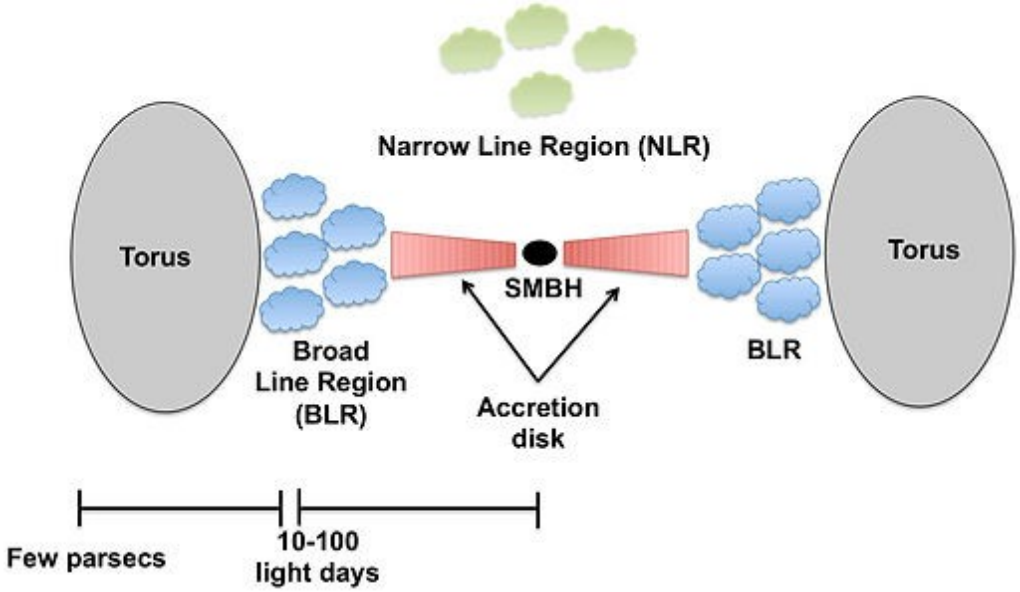
NRP Nautilus

Quasars are Bright Active Galactic Nuclei Powered by Accretion Of Matter Around Supermassive Black Holes At The Center Of Galaxies



April 26, 1985
www.science.org/toc/science/228/4698
An Approach to Complexity: Numerical Computations
Computation by John Hawley & Larry Smarr

Based on Forty Years of Computational Science



AGN schematic: central black hole, accretion disk, broad-line region (R. Ricci, U. Genève www.claudioricci.eu/agn-in-the-x-ray-band)

NRP NATIONAL RESEARCH PLATFORM

NRP Nautilus

Using Machine Learning on NRP
To Model Time Series of Quasar Variability

- Multiband time delays trace size and temperature of the accretion disk
- Simulation-based inference recovers black-hole mass, disk geometry, and reprocessing
- Scales to thousands of AGN from LSST/Rubin Observatory cadences

- Latent stochastic differential equations model intrinsic accretion-disk variability
- Neural networks handle irregular sampling, gaps, & heteroscedastic photometric noise
- End-to-end training on simulations enables fast inference on real survey data

NRP Usage Led to Joshua's 2025 Preprint

Joint Modeling of Quasar Variability and Accretion Disk Reprocessing Using Latent Stochastic Differential Equations

Joshua Fagin,[1,2,3] James Hung-Hsu Chan,[2,3] Henry Best,[1,2,3,4] Matthew O'Dowd,[1,2,3] K. E. Saavik Ford,[1,2,5,6] Matthew J. Graham,[7] Ji Won Park,[8,9] and V. Ashley Villar[10]

[1] *The Graduate Center of the City University of New York, 365 Fifth Avenue, New York, NY 10016, USA*

[2] *Department of Astrophysics, American Museum of Natural History, Central Park West and 79th Street, NY 10024-5192, USA*

[3] *Department of Physics and Astronomy, Lehman College of the CUNY, Bronx, NY 10468, USA*

[4] *Department of Theoretical Physics and Astrophysics, Faculty of Science, Masaryk University, Kotlářská 2, CZ-611 37 Brno, Czech Republic*

[5] *Department of Science, CUNY Borough of Manhattan Community College, 199 Chambers St, New York, NY 10007, USA*

[6] *Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA*

[7] *Department of Physics, Maths and Astronomy, California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, USA*

[8] *SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*

[9] *Prescient Design, Genentech, South San Francisco, CA 94080, USA*

[10] *Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138-1516, USA*

Draft version July 16, 2025
https://arxiv.org/pdf/2410.18423

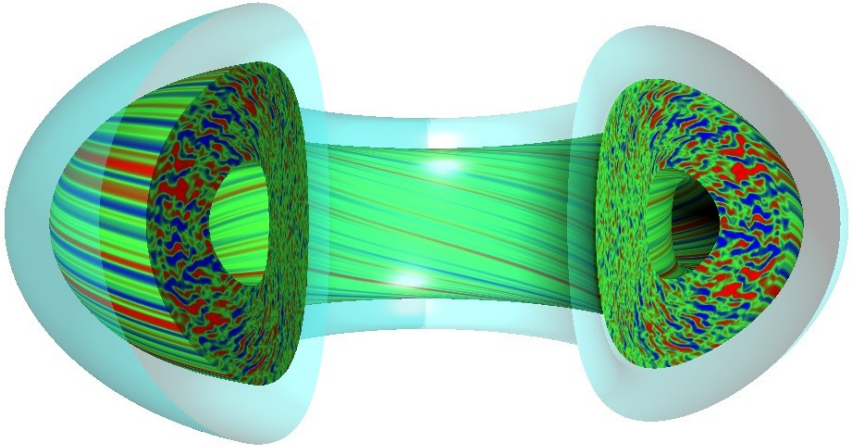
"This work used resources available
through the National Research Platform (NRP)
at the University of California, San Diego."

NRP NATIONAL RESEARCH PLATFORM

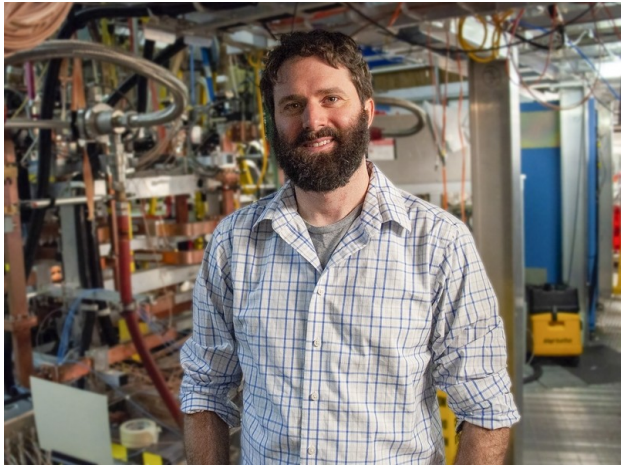
NRP Nautilus

MIT Nathan Howard
Unlocking the Secrets of Fusion's Core with AI-Enhanced Simulations

Gyrokinetic Plasma-Turbulence Simulation with CGYRO



Gyrokinetic turbulence in tokamak geometry — eddies drive cross-field transport (PPPL Hammett group viz)



Nathan Howard — Principal Research Scientist at the MIT Plasma Science and Fusion Center (PSFC)

- CGYRO solves the 5D gyrokinetic equation for ion- and electron-scale turbulence
- Resolves drift-wave eddies that set transport across magnetic flux surfaces
- Relies on many-node GPU resources.
- GPU-parallel runs make multi-scale, electromagnetic regimes tractable

In 2025: 1,516 GPU-hrs on namespace: fusion-psfc

NRP NATIONAL RESEARCH PLATFORM

NRP Nautilus

UC Riverside Mingxun Wang Lab
Computational Mass Spectrometry to Spot Hidden Environmental Pollutants

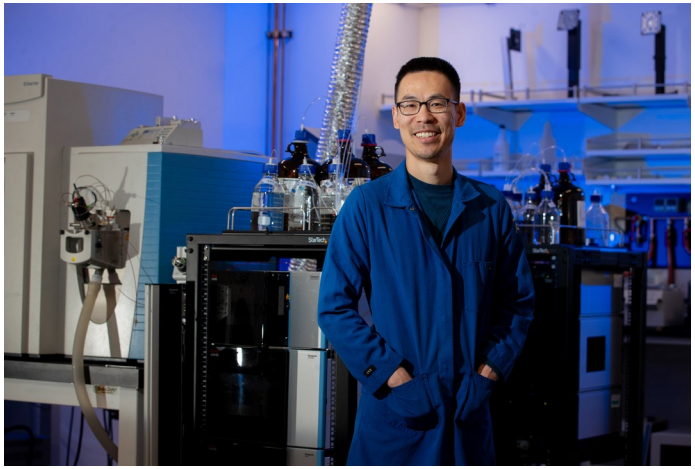
nature methods

Article

https://doi.org/10.1038/s41592-025-02660-z

A universal language for finding mass spectrometry data patterns

Published online: 12 May 2025



UCR Assistant Professor Mingxun Wang (CSE)

In 2025: 6,997 GPU-hrs on namespace: wang-bioinf

NRP NATIONAL RESEARCH PLATFORM

NRP Nautilus

User-Friendly Programming
Mass Query Language (MassQL)

- **Functions as a Search Engine for Mass Spectrometry Data, Enabling Researchers to Find Patterns That Would Otherwise Require Advanced Programming Skills**

- **Helps Spot Hidden Pollutants in Massive Chemical Datasets**

- **Connects GPU-Enabled Data Science with a Worldwide Network of Environmental Chemistry and Public-Health Surveillance**

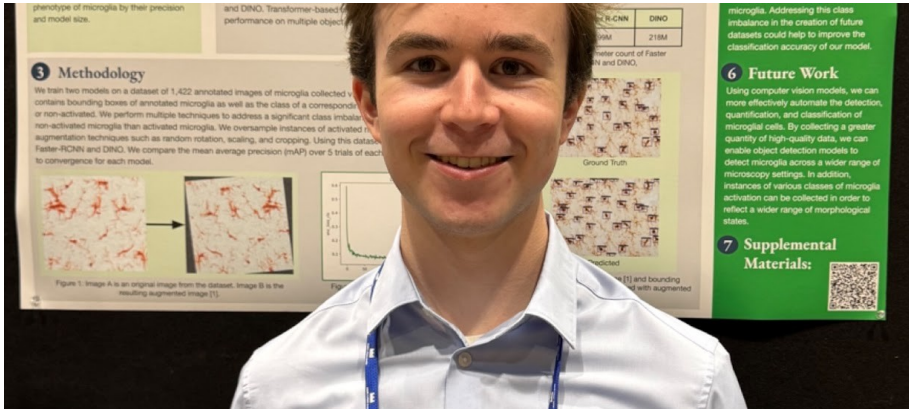
Cal Poly Humboldt Undergraduate John Gerving Trained Computer Vision to Identify Brain Microglia from Light Microscope Images



Kamila Larripa
Associate Professor
DEPARTMENT OF MATHEMATICS & DATA SCIENCE
Cal Poly Humboldt.

MAA
MATHEMATICAL ASSOCIATION OF AMERICA

MAA MATHFEST 2025 | SACRAMENTO, CA



John Won the 2025 MAA Janet L. Andersen Award for Undergraduate Research in Mathematical or Computational Biology

John Carried Out His Microglia Research
Under Cal Poly Humboldt's Professor Kamila Larripa.

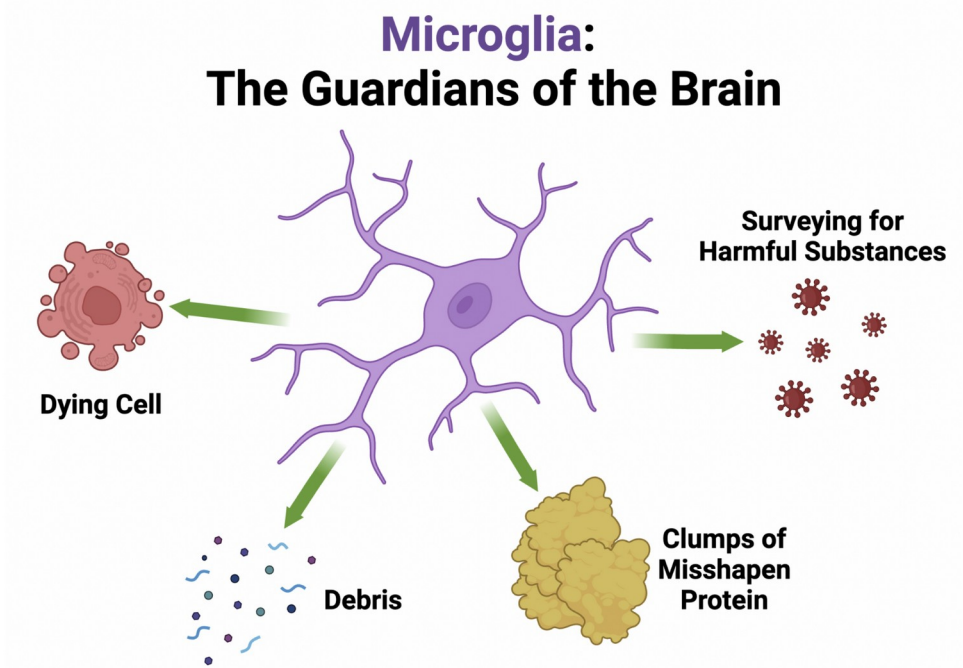
He is Currently an Undergraduate
at UC Berkeley in Applied Mathematics

NRP NATIONAL RESEARCH PLATFORM

In 2025: 2,282 GPU-hrs on namespace: cal-poly-humboldt-microglia

NRP Nautilus

Microglia are Immune Cells That Live Only in the Brain



An overview of what Microglia do in the brain. Figure made by : Rana Abdelhalim using BioRender.
www.ataxia.org/scasourceposts/snapshot-what-are-microglia/

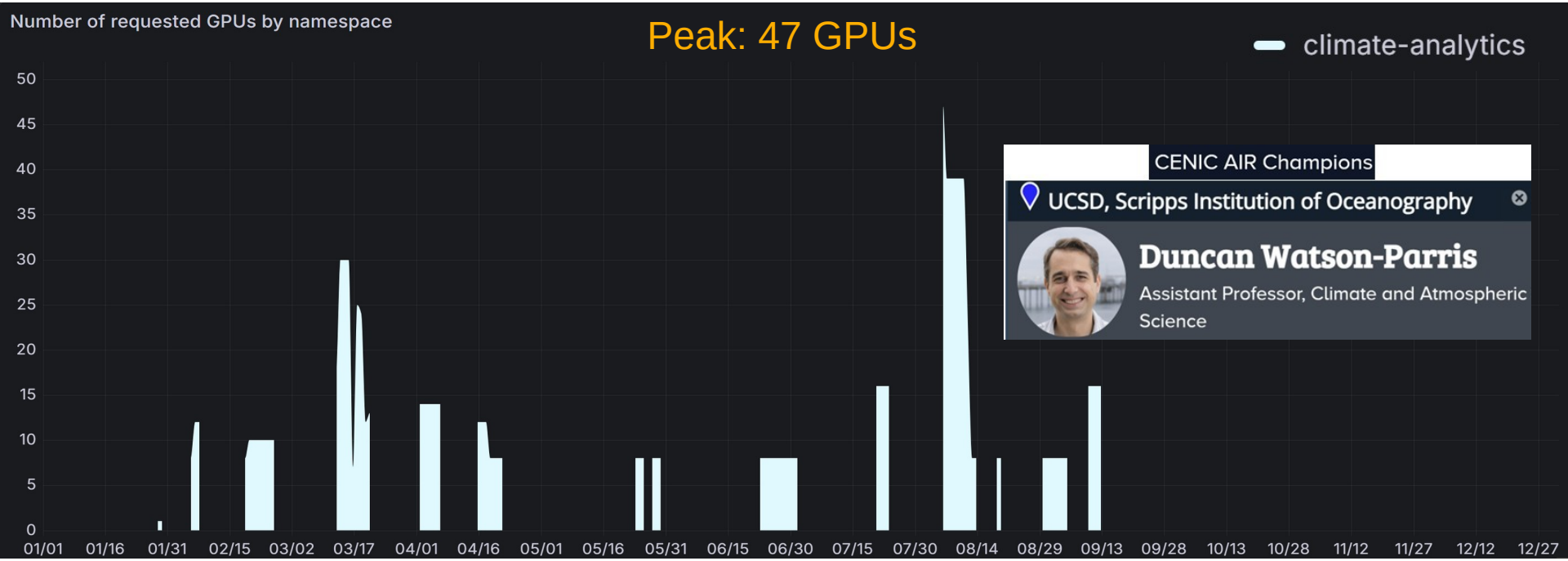
Microglia Have Multiple Phenotypes,
Whose Classification
Can Provide Insights
on the Overall Health of the Brain.

John Found the Optimal Computer Vision Technique to Classify Microglia Cells From Light Microscopy Images.

NRP NATIONAL RESEARCH PLATFORM

NRP Nautilus

UC San Diego Professor Duncan Watson-Parris
Using Machine Learning to Better Understand Clouds and Climate Interactions



In 2025: 7,754 GPU-hrs on namespace: climate-analytics

NRP NATIONAL RESEARCH PLATFORM

NRP Nautilus