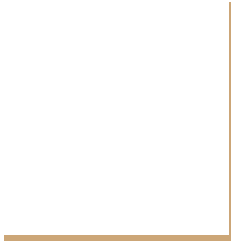




An Example of Campus Champion Facilitation in the Long Tail of Science

Asya Shklyar
July 4 2020
PEARC



About

20+ years of IT work (since 1995)

Startups, corporations, consulting

And now higher ed

~3 years at Pomona College

Director of HPC

Supporting research and teaching

Founded a technology exploration lab

hpc.pomona.edu | @pomonahpc



<https://www.pomona.edu/museum/collections/james-turrell-skyspace>

My Other Identities

XSEDE Campus Champion

EMPOWER program mentor

PEARC mentor

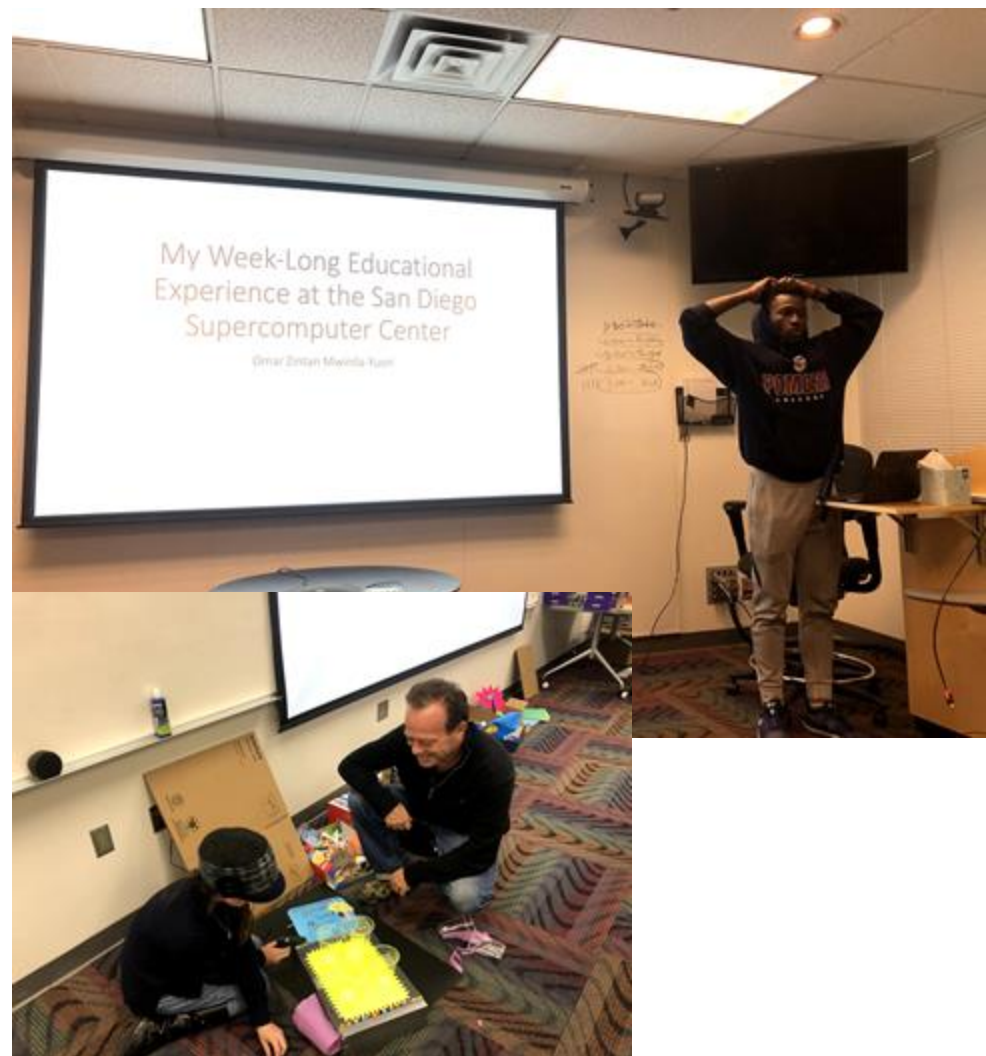
Snap The Gap event last year

Big Data and Data Science MeetUps

Mentoring coworkers

Helping artists in SoCal learn about technology

Speaking at SCaLE



How did I decide to become a Campus Champion

I learned about XSEDE while working as a consultant and was intrigued but I was not working at a college

There is this notion that if you have on-prem infrastructure you are all set

Not true, not only you are not all set, it is a never ending battle for more resources

Where are the resources?

Cloud

Right. \$\$\$\$\$\$ after 3-4 months and \$\$\$\$ getting your data back (and it is even more complex than local infra)

Aha! XSEDE is perfect

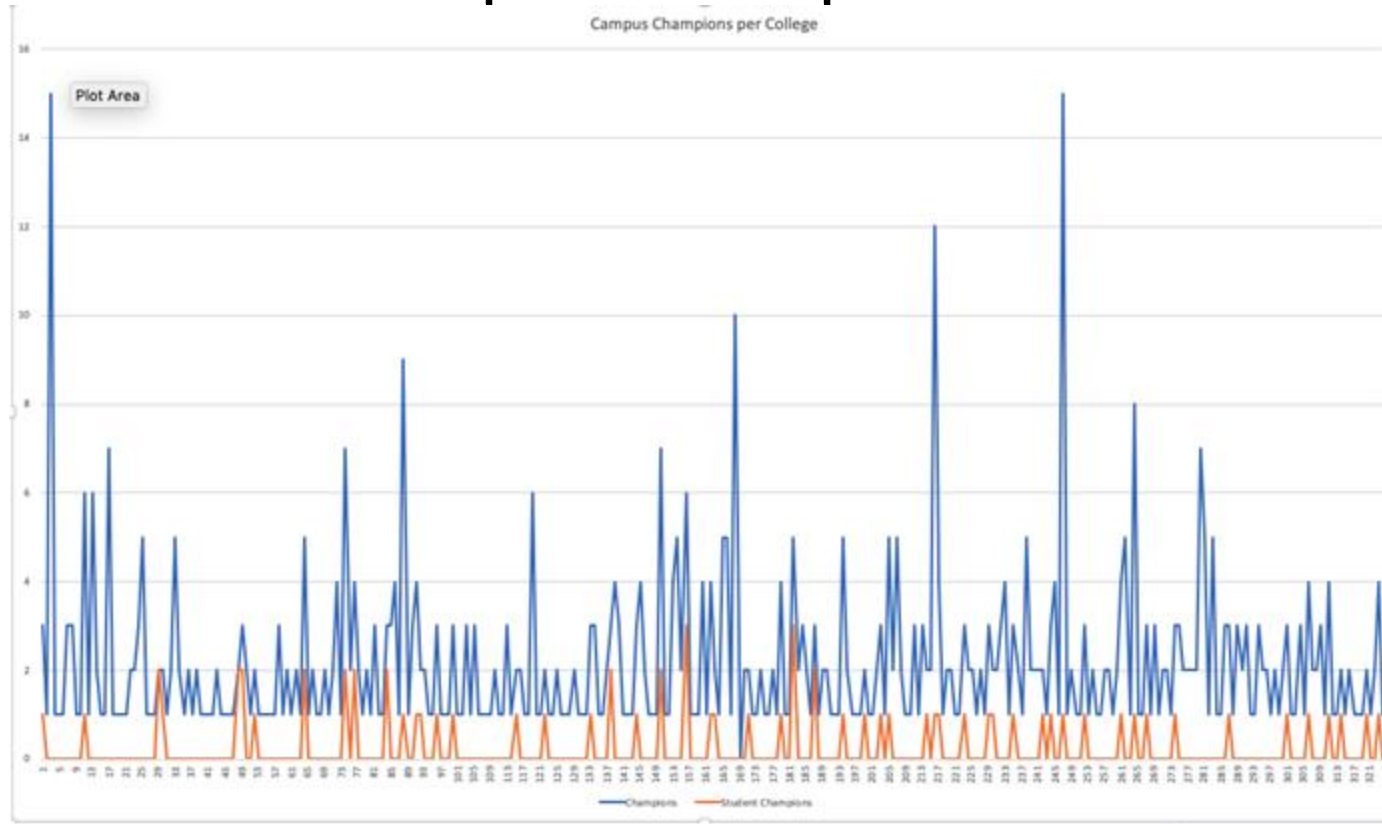
The school does research, the researchers have no idea they can have access

Access to so much information, discussions, networking, AND compute+storage! Opportunities for students!

Student Campus Champions

Pomona College

Asya Shklyar, Andrew Crawford, Omar Zintan Mwinila-Yuori (student), Samuel Millette (student), Sanghyun Jeon, Nathaniel Getachew (student)



Some Data Science!

Only 2 institutions with a total over 10 champions!

Only 3 institutions with close to 10 champions!

Majority are in 1-4 range

Only 2 institutions with 3 student campus champions!

Majority of the institutions have both!

Surprised that there aren't more student campus champions at the institutions with most campus champions...

Infrastructure

1000 cores local, plus XSEDE, plus cloud

32 core Epyc, 28 core Skylake

V100 and P100

Expensive Flash storage and Minio

Slurm and Kubernetes

R/RStudio and Jupyter/Hub/Lab ecosystem

Python, Scala, Julia



Interesting Projects

right now*

Psychology: suicide prevention, Python+HDDM (hierarchical Bayesian parameter estimation of the Drift Diffusion Model) and Netlogo for opioid epidemic modeling

Economics: NLP, Topic Modeling, LDA, Sentiment Analysis, Twitter API, OCR, visualization, very iterative

Politics: GIS, ESRI, Pix4D, QGIS, geocoding, Stata, public datasets, social media, orthomosaics

Physics: Nuclear fusion simulations and quantum irreversibility

Computer Science: vGPUs, ML/AI, pytorch, GitLab, HCI

Environmental Analysis: sensor data, queues, NOSQL, graph databases, user interfaces, data viz

Museum: Augmented Reality and Virtual Reality for remote access

Other Projects

WebMO for Gaussian, MOPAC, Gamess

JupyterHub for Environmental Analysis

Data Wrangling workflows

StarFish for finding all the data, dating it, establishing ownership, automating workflows

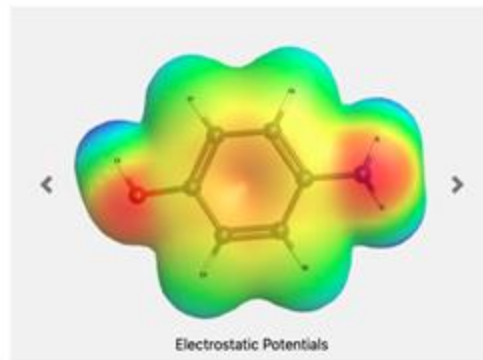
CentOS 7 to 8 migration

Converting workflows to containers where appropriate

What is WebMO?

WebMO is a web-based interface to computational chemistry packages.

- WebMO allows students and non-specialists to run state-of-the-art computational chemistry programs from the web-browser on their computer
- Intuitive for novices: point-and-click molecular editor, pre-defined job types, reasonable default values
- Flexible for experts: advanced job options, full access to input and output files



A multi-user version of the notebook designed for companies, classrooms and research labs



Visibility into Your Data

- Instant Search/Find
- Duplicate Deletion
- Aging
- Utilization/Ownership
- Cross Volume Search/Reporting



Rapid File Movement

- On-Premise & Cloud
- Archive
- Backup
- File Migration



Policy-Based Action

- Hash, Sync & Verify
- Automated Metadata Tagging & Extraction
- Run Your Own Code

Future

Teaching online

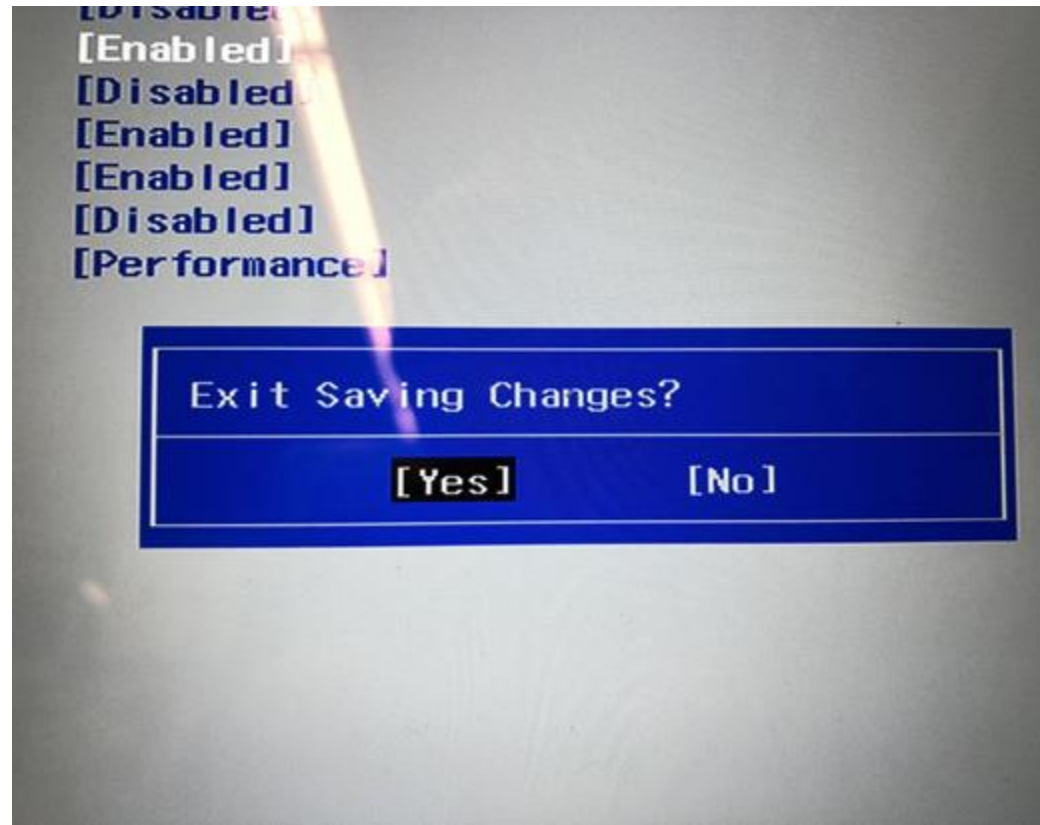
Raspberry Pi clusters (64 bit!!)

128 cores in a 2 socket server.. Optane..

Science Gateway for Data Science

Training undergraduate students

Participating in the HPC community
diversity efforts





Students programming DJI Robomaster



HPC Pondr team

"What is Human?" class



In The Know Lab



Supported by Pomona College and Shodor/EMPOWER

App Store Preview



HPC Pondr 4+

Quiz for HPC Terminology

Pomona College

★★★★★ 5.0, 9 Ratings

Free

iPhone Screenshots

