Workforce Needs in the Age of Blended Data

John Towns
Executive Associate Director for Engagement
National Center for Supercomputing Applications (NCSA)
University of Illinois
jtowns@ncsa.illinois.edu

ILLINOIS
NCSA | National Center for Supercomputing Applications

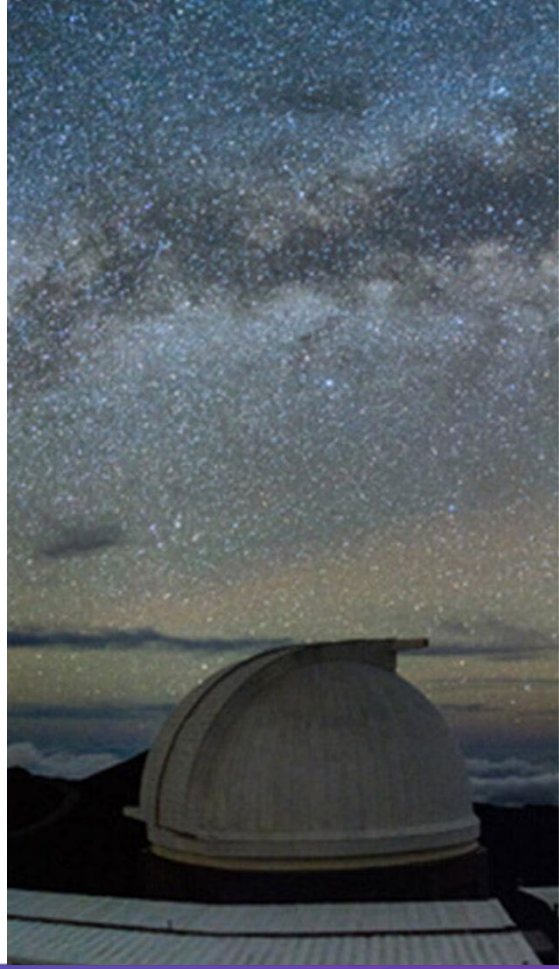
Why a Discussion of Workforce?

- Effective use of blended data, particularly at scale, will require an advanced research computing and data resources and services ecosystem
- Computing hardware is the easy problem
    - yes, there are financial challenges
- A skilled technical workforce is a critical element of this ecosystem
    - serious workforce issues are on the horizon
    - not specific to endeavours in the use of blended data!
    - affects everyone making use of advanced research computing and data resources and services
- Demand for skilled workforce is evident
- This is still an emerging profession with associated challenges

Blended Data is not a New Concept

- The research community has been making use of this for some time
- Many efforts over the years to produce “science ready data sets”
  - metadata schemas, storage formats, databases, data management systems,...
  - even within domains, datasets not managed consistently
  - using multiple data sources (blending) requires significant effort in “data wrangling”
- More recently, interdisciplinary teams are drawing on various data sources that are used together to advance science

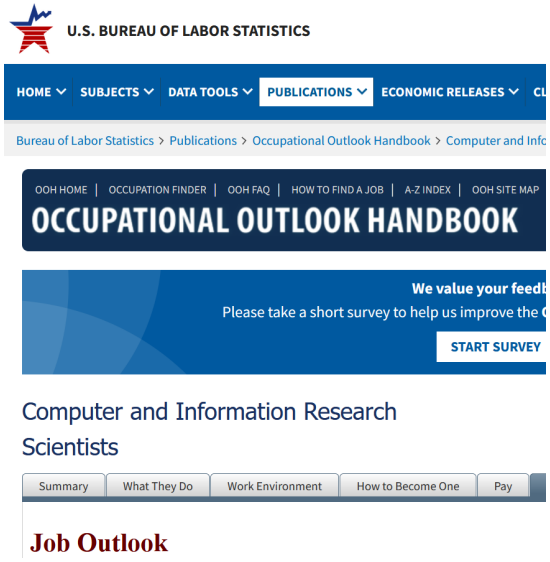
And this is just about preparing the data!



Demand for Research Computing and Data Professionals: BLS

Bureau of Labor Statistics (BLS) predicts:

- employment of … research computing professionals, is projected to grow 15 percent from 2020 to 2030
- much faster than the average for all occupations



https://www.bls.gov/ooh/computer-and-information-technology/computer-and-information-research-scientists.htm#tab-6

Demand for Research Computing and Data Professionals: NSF

National Science Foundation (NSF) predicts:

- increasing demand for research computing professionals in all scientific fields over the next decade
- the number of researchers who rely on high-performance computing (HPC) will double over the next five years
- projected growth rate of 7-10 percent per year through 2025

https://www.nsf.gov/about/performance/strategic_plan.jsp

There is no Professional Association

(that may not be the right answer anyhow)

- Some historical efforts in the community (mostly HPC oriented)
  - IEEE
  - ACM: SIGHPC, SIGAPP
  - Society of HPC Professionals: small organization in Houston, TX
- More recently
  - Campus Champions: expanded to regional, state, domain
  - LCI, Round 2: workshops only at this point
  - ACI-REF/CaRCC, RCD Nexus
  - Regional organizations: RMACC, CADRE, HPC2, GPN,…
  - PEARC Conference series
- This is encouraging!
  - need to figure out how to bring these together
  - get past “let a thousand flowers bloom?”

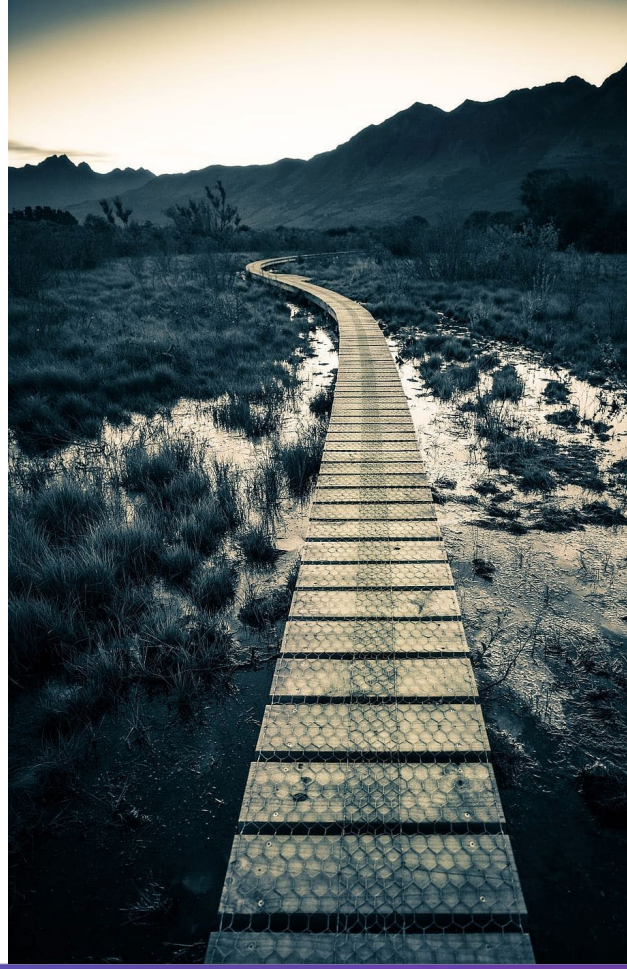
BUT…. these are only pieces of the solution

- Where will these technical professionals come from?
  - we lack certificate and degree programs
  - we lack scalable training and workforce development to create the needed workforce
- Some things that we need
  - Education and training opportunities
    - degrees and certifications
    - ongoing professional development
  - Advocacy for the emerging profession
    - perception of second-class status
  - Definition of the profession
    - titles and career paths
  - and so much more….

We must fight the tendency to reward technical prowess with management responsibility. Too often we fail to prepare technical staff we promote and put them at risk!

So what do we do?

- Focus on the outcomes we want for this community
  - a professional organization is perhaps only one way to accomplish them
- Many efforts emerging in the community to attempt to address these needs
  - efforts are not aligned; we need to organize the organizations
- We must develop coherence in the community
  - create a space for doing this



Q&A

ILLINOIS NCSA