Building Data Science Support Capacity through Graduate Fellowship Programs

Jeffrey C. Oliver, Maliaca Oxnam, Susan J. Miller, Vignesh Subbian
University of Arizona
Data science relies on collaboration

Jeff Oliver
UA Libraries
@jcoliverAZ

Maliaca Oxnam
Data Science Institute
@maliaca

Vignesh Subbian
Engineering
@V_Subbian

Susan Miller
Data Science Institute
@sjmiller_az
But what is “data science”
Data science is Venn diagrams
Interest in data science is growing

Frequency of the search “what is data science” for past 10 years on Google
Data science requires subject matter expertise
We can’t hire someone from *every* domain

Christopher Marlowe credited as one of Shakespeare’s co-writers of *Henry VI*

October 23, 2016

*Crisis Text Line* is developing predictive analyses to identify threats to participants

February 9, 2015

eBird uses machine learning to predict habitat suitability for bird migration

April 30, 2016
What do we need to support data science on campus?

- **Experience**
  - Applying modern computational and statistical approaches to address domain-specific questions

- **Interest**
  - Data science education
Graduate student population has significant capacity

- **Experience**
  - Applying modern computational and statistical approaches to address domain-specific questions

- **Interest**
  - Data science education
The Data Science Ambassadors support the campus

The Data Science Ambassadors (DSA) program capitalizes on domain knowledge and data science expertise in the UA graduate student population to assist researchers with their data science needs.

https://datascience.arizona.edu/ambassadors/data-science-ambassadors-program
Ambassadors are trained in pedagogy

All ambassadors complete two-day Carpentries Instructor Training

- Pedagogy
- Motivation
- Imposter syndrome
- Stereotype threat
- Growth mindset

https://carpentries.github.io/instructor-training/
Ambassadors develop custom engagement plans

- A roadmap of expectations
- Developed by each individual ambassador
- Tailored to the constituents they serve (i.e. scholars in their college)
- Iterative feedback from directors and fellow ambassadors
Ambassadors curate living document of resources

A living document provides a one-stop-shop for ambassadors to use when navigating resources

---

Data Science Resource Guide

Data science resources & contacts

Campus units
- Data Science Institute
- BIO5 Science Promotion
- Research Bazaar
  - Python, R, and HPC support
  - CyVerse @CyVerseOrg
Everybody loves cold hard cash

- $500 stipend per semester
- Expectation of 15 hours of work each semester
Time is a limited resource

- Directors support ambassadors’ guarding time
- “Why would I tell my student to do this?”
  - Benefits to student
    - Networking
    - Skills development
  - Benefits to research group
    - Best data practices
We give colleges agency to increase buy-in

- Switched from self-nomination to system where colleges nominate candidates
- Colleges provide half the stipend; can provide more if they want more than one ambassador
After two years, a doubling in domain coverage

- **Year 1:**
  - 17 applications
  - 8 ambassadors in
  - 4 colleges

- **Year 2:**
  - 13 applications
  - 9 ambassadors in
  - 8 colleges

- 100+ consults
- 4 Software Carpentry workshops
- 12+ stand-alone workshops (e.g. R, Python, Orange, GIS)
Ambassadors’ efforts dwarf single-individual support

- Survey of data science needs in the College of Agriculture & Life Sciences
- Introducing data science to the humanities
- Workshop on the Brain Imaging Data Structure (BIDS) Ecosystem
- Introduction to time series machine-learning with PyTorch
- Introduction to deep learning for astronomers
Ambassadors are driving data science education efforts

Two Ambassadors co-organized the inaugural UA Women in Data Science event in 2019

DS Ambassadors are leading organizational efforts for 2020 UA Research Bazaar

DS Ambassadors help organize R Ladies and Code for Venezuela events
Due date for nominations was May 1, 2020
Interest in program remains strong, despite challenges

- 13 nominations
- Five colleges
- Multiple colleges nominated more than one student
What questions do you have?

Jeff Oliver
@jcoliverAZ

Vignesh Subbian
@V_Subbian

Maliaca Oxnam
@maliaca

Susan Miller
@sjmiller_az

Slides available on Google Drive at
https://bit.ly/3f6frT4