DSC Report: 2024

Tim Dennis

2024-05-22

Table of contents

Email Draft

Dear Todd,

Thank you for your and Athena's response and for considering our budget request. Thanks for the opportunity to provide more detailed information about the Data Science Center (DSC).

The Data Science Center (DSC) supports the UCLA research community by offering a range of computational and data scholarship services and infrastructure, including:

Service Provision

Consultations: We offer one-on-one and group consultations on the full data/research lifecycle: acquisition, preparation, analysis, statistical modeling, data visualization, management, sharing, and archiving. These sessions are usually short-term problem-solving assistance for specific phases in larger research projects. Our goal is to help improve research practices, advocate for emerging standards like Open Science, and teach researchers how to do something rather than just doing it for them.

Instruction: We regularly offer Carpentries workshops on data science topics, including Python, R, Jupyter, Git, machine learning, and data cleaning. Additionally, we lead and program UC Carpentries instruction, UC GIS Week, and UC Love Data Week. We also provide traditional class-based instruction as requested. Lead the UCLA Carpentries program and provide curricular development as part of the Carpentries community as lesson authors or maintainers.

Infrastructure Services

Administration and provision of essential data science infrastructure, such as:

- **UCLA Dataverse**: A data repository for storing and sharing research data.
- **Redivis**: A data platform for large datasets that also might need extra security protocols (HIPPA, etc.).
- **PointClouds via Potree**: A platform for visualizing and interacting with large point cloud (3D) datasets.
- **ESRI/ArcGIS Online**: A suite of tools for creating, sharing, and analyzing geographic information. The Spatial Data Science Librarian has led recent improvements in access and licensing and works across the UC system to improve access to GIS tools.
- **Deep Learning Machine**: High-performance computing resources for deep learning and other advanced data science applications. This is a mid-tier computing service for researchers who first need more computing than a laptop can perform but are not ready to jump to HPC.

Roles and Responsibilities

- **Leigh Phan, Data Scientist**: Provides consulting on computation and data projects, leads ML/AI support for the DSC, administers and mediates research access to the Deep Learning Machine, Carpentries instructor, and leads programming for UC Love Data Week.
- **Jamie Jamison, Data Collections Manager**: Manages and administers UCLA Dataverse and Redivis, provides data consultations with focus on data management, sharing, and curation, serves as the point of contact with the Dataverse consortia, and is a Carpentries instructor.
- **Dr. Zhiyuan Yao, Spatial Data Science Librarian**: Leads GIS services, provides spatial analysis/GIS consulting, is a Carpentries instructor and local coordinator, serves as the campus ESRI/ArcGIS Online representative and provisioner, and is the point person for UC GIS Week and related activities.
- **Doug Daniels, Emerging Technologies Librarian**: Provides consulting and lifecycle support for 3D data, manages LuxLab operations (3D printing, scanning, etc.), identifies new emerging technologies and services, and is a Carpentries instructor and teacher of emerging tech tools and practices, administers Potree server for 3D in-browser visualization.

Metrics

```r
# Calculate the number of unique departments for workshops since 2017
num_depts_wkshp <- ucla_workshops %>% drop_na(standardized_department) %>% filter(institutio

# Calculate the number of unique departments for workshops in 2023
num_depts_wkshp_2023 <- ucla_workshops %>% filter(date >= "2023-01-01" & date <= "2023-12-31

# Calculate the number of unique departments for consultations since 2019
num_depts <- dsc_consult %>% drop_na(department) %>% distinct(department) %>% nrow()

# Calculate the number of unique departments for consultations in FY 2020-21
num_depts_20_21 <- dsc_consult %>% filter(year(start_date_time) == 2020) %>% drop_na(departm

# Data for sparklines
workshop_attendance <- ucla_workshops %>% count(year = year(date)) %>% arrange(year) %>% pul
consults_by_year <- dsc_consult %>% filter(!is.na(start_date_time)) %>% count(year = year(sta

# Generate sparklines
workshop_sparkline <- sparkline(workshop_attendance, type = "line", height = "20px", width =
consult_sparkline <- sparkline(consults_by_year, type = "line", height = "20px", width = "10

# Print the metrics with HTML for sparklines
cat(sprintf('
<p><strong>Overall Workshop Attendance:</strong> Since 2017, our workshops have been attende
<p><strong>Top Attending Departments:</strong> The top attending departments include Geograp
<p><strong>Workshop Growth:</strong> The attendance at DSC events has shown consistent growt
<p><strong>Consultation Services:</strong> The number of consultations has grown consistentl
<p><strong>Affiliation of Attendees:</strong> The top statuses or affiliations of attendees
<p><strong>Consultation Topics:</strong> Top consultation topics frequently include coding/p
  num_depts_wkshp, num_depts_wkshp_2023, workshop_sparkline, workshop_sparkline, num_depts, n
```


<p>Overall Workshop Attendance: Since 2017, our workshops have been attende
<p>Top Attending Departments: The top attending departments include Geograp
<p>Workshop Growth: The attendance at DSC events has shown consistent growt
<p>Consultation Services: The number of consultations has grown consistentl
<p>Affiliation of Attendees: The top statuses or affiliations of attendees
<p>Consultation Topics: Top consultation topics frequently include coding/p
<p>Overall Workshop Attendance: Since 2017, our workshops have been attende
<p>Top Attending Departments: The top attending departments include Geograp
<p>Workshop Growth: The attendance at DSC events has shown consistent growt
<p>Consultation Services: The number of consultations has grown consistentl
<p>Affiliation of Attendees: The top statuses or affiliations of attendees

```
<p><strong>Consultation Topics:</strong> Top consultation topics frequently include coding/p
<p><strong>Overall Workshop Attendance:</strong> Since 2017, our workshops have been attende
<p><strong>Top Attending Departments:</strong> The top attending departments include Geograp
<p><strong>Workshop Growth:</strong> The attendance at DSC events has shown consistent growt
<p><strong>Consultation Services:</strong> The number of consultations has grown consistentl
<p><strong>Affiliation of Attendees:</strong> The top statuses or affiliations of attendees
<p><strong>Consultation Topics:</strong> Top consultation topics frequently include coding/p
<p><strong>Overall Workshop Attendance:</strong> Since 2017, our workshops have been attende
<p><strong>Top Attending Departments:</strong> The top attending departments include Geograp
<p><strong>Workshop Growth:</strong> The attendance at DSC events has shown consistent growt
<p><strong>Consultation Services:</strong> The number of consultations has grown consistentl
<p><strong>Affiliation of Attendees:</strong> The top statuses or affiliations of attendees
<p><strong>Consultation Topics:</strong> Top consultation topics frequently include coding/p
<p><strong>Overall Workshop Attendance:</strong> Since 2017, our workshops have been attende
<p><strong>Top Attending Departments:</strong> The top attending departments include Geograp
<p><strong>Workshop Growth:</strong> The attendance at DSC events has shown consistent growt
<p><strong>Consultation Services:</strong> The number of consultations has grown consistentl
<p><strong>Affiliation of Attendees:</strong> The top statuses or affiliations of attendees
<p><strong>Consultation Topics:</strong> Top consultation topics frequently include coding/p
<p><strong>Overall Workshop Attendance:</strong> Since 2017, our workshops have been attende
<p><strong>Top Attending Departments:</strong> The top attending departments include Geograp
<p><strong>Workshop Growth:</strong> The attendance at DSC events has shown consistent growt
<p><strong>Consultation Services:</strong> The number of consultations has grown consistentl
<p><strong>Affiliation of Attendees:</strong> The top statuses or affiliations of attendees
<p><strong>Consultation Topics:</strong> Top consultation topics frequently include coding/p
<p><strong>Overall Workshop Attendance:</strong> Since 2017, our workshops have been attende
<p><strong>Top Attending Departments:</strong> The top attending departments include Geograp
<p><strong>Workshop Growth:</strong> The attendance at DSC events has shown consistent growt
<p><strong>Consultation Services:</strong> The number of consultations has grown consistentl
<p><strong>Affiliation of Attendees:</strong> The top statuses or affiliations of attendees
<p><strong>Consultation Topics:</strong> Top consultation topics frequently include coding/p
<p><strong>Overall Workshop Attendance:</strong> Since 2017, our workshops have been attende
<p><strong>Top Attending Departments:</strong> The top attending departments include Geograp
<p><strong>Workshop Growth:</strong> The attendance at DSC events has shown consistent growt
<p><strong>Consultation Services:</strong> The number of consultations has grown consistentl
<p><strong>Affiliation of Attendees:</strong> The top statuses or affiliations of attendees
<p><strong>Consultation Topics:</strong> Top consultation topics frequently include coding/p
```

Overall Workshop Attendance: Since 2017, our workshops have been attended by 148 different departments, schools, centers, or units from UCLA. In 2023, workshops engaged 86 different departments, schools, centers, or units from UCLA.

Top Attending Departments: The top attending departments include Geography, Information

Studies, Library, Social Science, and Environment and Sustainability.

Workshop Growth: The attendance at DSC events has shown consistent growth over time, highlighting the increasing interest and participation from the UCLA community.

Consultation Services: The number of consultations has grown consistently, reflecting the increased demand for one-on-one research support. Since 2019, consultations have come from 108 different departments. In FY 2020-21, data services were provided to 0 campus departments.

Affiliation of Attendees: The top statuses or affiliations of attendees include Graduate Students, Undergraduates, and Faculty.

Consultation Topics: Top consultation topics frequently include coding/programming, data analysis, and GIS-related inquiries, illustrating the importance of these skills in data-intensive research.

I hope this provides a clearer picture of our service portfolio and what we are engaged in.

Thanks,

Tim