

2018 Oak Ridge Associated Universities (ORAU) Visiting Faculty Travel Grant

**Project:** Investigating Data Management and Developing Collaborations across ORAU

**Name:** Plato L. Smith II, Ph.D.

**Email:** [plato.smith@ufl.edu](mailto:plato.smith@ufl.edu)

**Title:** Data Management Librarian

**College/Unit:** George A. Smathers Libraries

**Department:** Technology Support Services (TSS)

**Campus Address:** George A. Smathers Libraries, 521 Library West, PO Box 117000, Gainesville, FL 32611-7000

**Total funding request: \$763.49**

This ORAU Visiting Faculty Travel Grant proposal seeks to connect the University of Florida (UF) Libraries Data Management Librarian, Dr. Plato Smith, to an Oak Ridge National Laboratory Distributed Active Archive Center ([DAAC](#)) scientist through the University of Tennessee Libraries Data Curation Librarian, Christopher Eaker, for purposes of developing a new collaboration between UF and ORNL DAAC. The primary goal of this new collaboration is to advance the development of *socio-technical (people, policies, technologies, communities)* data management collaborations across multiple communities of practice at Oak Ridge Associated Universities (ORAU) for the purpose of “addressing challenges in securing research funding and attracting students into scientific and technical fields necessary to compete in the global arena<sup>1</sup>” (OARU, 2018). The key goal for this proposal is to explore “initial exploratory activities toward the creation of social and technical infrastructure solutions”<sup>2</sup> that further OARU’s commitment for promoting opportunities for collaborative partnerships in scientific research and education.

**Research questions:**

1. What are some of the challenges ORAU researchers face in securing research funding?
2. How can organizing ORAU stakeholders (e.g. discipline experts, data repository managers, and data appraisal experts) address challenges in securing future research funding?

**Budget Sheet:**

Cost Type	Total
Flight	\$416.00
Hotel for 2 nights (\$93 night) + tax (4%)	\$193.44
Ground transportation (cab/shuttle) from airport to hotel (\$20 x 2 trips)	\$40.00
Car parking at Jacksonville Airport (\$20 per day x 2 days)	\$40.00
Gas mileage (drive to/from JAX airport) 83.2 miles x 2 = 166.4	\$74.05
<b>Total Cost</b>	<b>\$763.49</b>
<b>Total Budget Request</b>	<b>\$763.49</b>

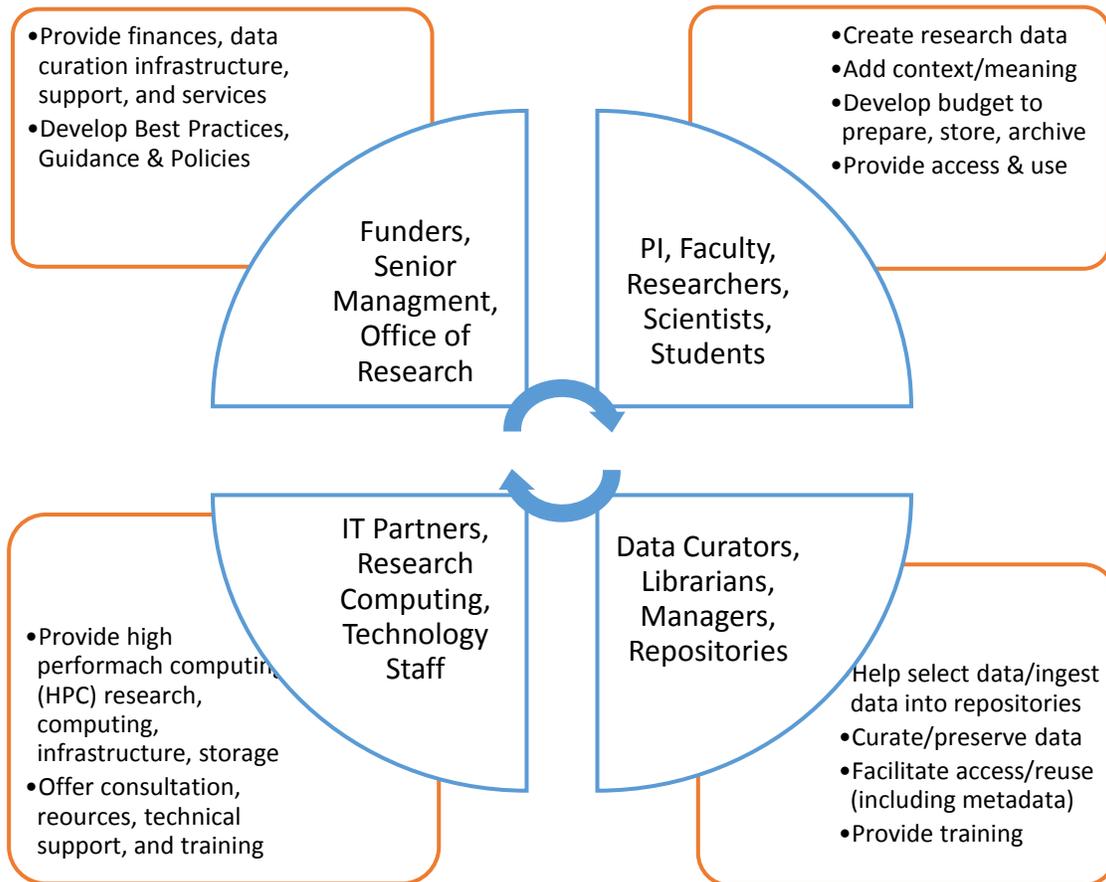
The budget for this proposal follows the University of Florida travel guide costs ([http://www.che.ufl.edu/PDF/Travel/Travel\\_Guide.pdf](http://www.che.ufl.edu/PDF/Travel/Travel_Guide.pdf)) and estimates based on flight priced on trip advisor on 5/9/2018, Comfort Inn Oak Ridge – Knoxville priced on Choice Hotels on 5/9/18.

<sup>1</sup> Oak Ridge Associated Universities (ORAU). (2018). University Partnerships. Retrieved May 10, 2018 from <https://www.orau.org/university-partnerships/default.aspx>.

<sup>2</sup> NSF 18-060. (2018). Dear Colleague Letter: Advancing Long-term Reuse of Scientific Data. Retrieved May 10, 2018 from [https://www.nsf.gov/pubs/2018/nsf18060/nsf18060.jsp?WT.mc\\_id=USNSF\\_25&WT.mc\\_ev=click](https://www.nsf.gov/pubs/2018/nsf18060/nsf18060.jsp?WT.mc_id=USNSF_25&WT.mc_ev=click).

**Introduction**

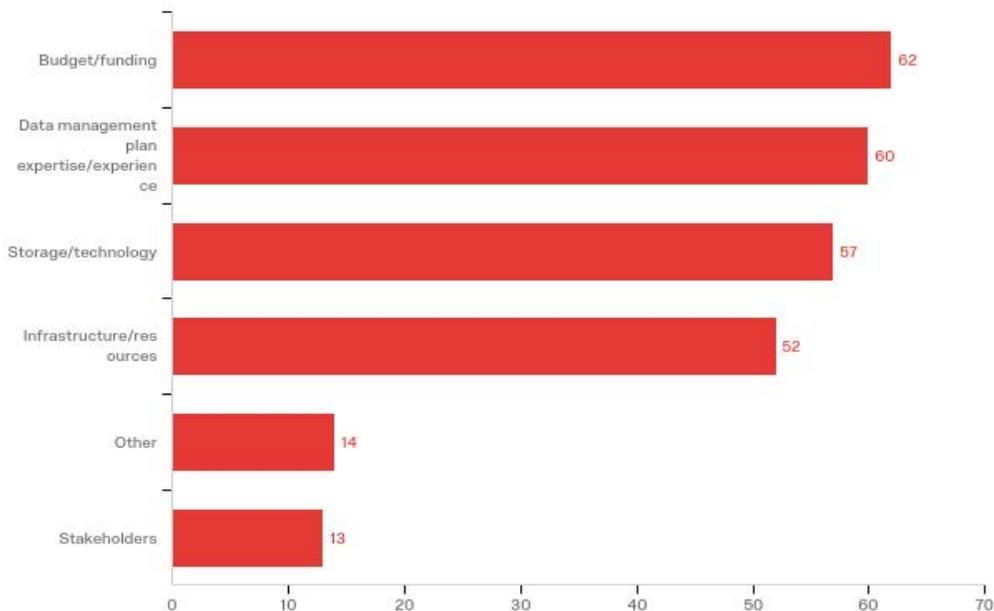
Good data management requires data skills, organization, technology, and resources in collaboration with multiple stakeholders (see Fig. 1). The sustainability of informatics infrastructure required to support good data management rests on financial support and the flexibility needed to store, manage, and preserve data for current and future use throughout the data curation lifecycle [1], science data lifecycle [2], and research cycle (pre, research, post) [3].



**Figure 1: Stakeholders’ roles aligned with Data Curation Lifecycle Responsibilities<sup>3</sup>**

Some of the challenges in securing research funding include: (1) not submitting a good data management plan, (2) failing to explain data lifecycle management within a funded project, and (3) not fulfilling funding agencies data sharing requirements, including long-term preservation. Results from a January 2017 – April 2017 data survey at the University of Florida UFIRB#201602303 revealed some of the barriers in managing and storing data for researchers that contribute to challenges in securing research funding. For example, results from data survey Question 15 (See Fig. 2) illustrate some common barriers to data management articulated by UF researchers, faculty, postdocs, staff, and students that participated in the survey. These barriers can contribute to non-competitive grant proposals that could affect securing research funding.

<sup>3</sup> DAF Implementation Guide. (2009). (Adapted) [http://www.data-audit.eu/docs/DAF\\_Implementation\\_Guide.pdf](http://www.data-audit.eu/docs/DAF_Implementation_Guide.pdf)



**Fig. 2: Q15 - What are some barriers to you with regard to managing and storing your research data?**

Data management and storage issues are not discipline or institution-specific. The survey results from Question 15 (See Fig. 2) may also be relevant to researchers at ORNL, the University of Tennessee, and other OARU. Using the UFIRB #201602303 data survey results as a frame of reference, this study seeks to explore data management and storage at ORNL DAAC and University of Tennessee at Knoxville (UTK) to gather data to share with UF researchers for improvement in future data management planning, training, and other activities to secure research funding.

One goal of the visit is to explore the idea for a conference proposal and community workshop on best practices for data management and planning for researchers and students in efforts to better develop relevant and timely data management education, support services, and training for faculty, staff, and students in ORAU communities of practice. Communities of practice are groups of people (scientists, working groups, external partners, and the human network of scientific domain collaborators) who share a concern, a set of problems, or a passion about a practice and who deepen their knowledge and expertise by interacting on an ongoing basis [4, 5].

ORAU	ORAU-EPSCoR	ORAU-HBCU/MEI	ORAU-SECU	ORAU-iSchool
140	34	23	10	11

**Table 1: ORAU, EPSCoR, HBCU/MEI, SECU, and iSchool**

Table 1 represents the total number of ORAU, ORAU and Historically Black Colleges and Universities/Minority Education Institutions ([HBCU/MEI](#)), ORAU in Established Program to Stimulate Competitive Research ([EPSCoR](#)) states, ORAU with an Information School ([iSchool](#)), and ORAU and Southeastern Conference Universities ([SECU](#)). Another goal for the visit is to discuss which ORAU category from Table 1 based on discussion with ORNL DAAC to initialize exploratory activities for more investigation, research, funding (e.g. [NSF CRII](#), [CyberTraining](#)).

**2018 Oak Ridge Associated Universities (ORAU) Visiting Faculty Travel Grant**

The ORNL DAAC has a good online resource on data management that includes (1) Best Practices for Data Management, (2) Writing a Data Management Plan, and (3) How to's and Resources (<https://daac.ornl.gov/datamanagement/>). The U.S. Department of Energy Office of Scientific and Technical Information has a good resource that includes (1) Data ID Services (digital object identifier (DOI) service), (2) API Documentation, (3) DOE MARC Records System, and (4) Open Archives Initiatives (OAI) (<https://www.osti.gov/data-services-developer-tools>). The USGS has a very good online resource on data management that includes (1) data management checklist, (2) data lifecycle processes, and (3) Fundamental Science Practices (FSP) FAQs: Release of Scientific Data (<https://www2.usgs.gov/datamanagement/index.php>). These resources are fundamental for researchers/stakeholders (e.g. discipline experts, data repository managers, and data appraisal experts) responsible for the aggregation, representation, dissemination, preservation, and destruction of data across all ORAU. This proposal seeks to initiate stakeholders to explore (a) community-specific agreements, (b) common data types, (c) data repository Findability, Accessibility, Interoperability, and Reuse (FAIR)<sup>4</sup>, (d) minimal metadata to make data FAIR, and (e) best practices associated with data management plans.

**Location and potential partners to be visited:**

1. University of Tennessee at Knoxville (UTK)
2. ORNL Distributed Active Archive Center (DAAC)

Institution: University of Tennessee at Knoxville

Name: **Christopher Eaker**

Title: Assistant Professor & Data Curation Librarian

College/Unit: University of Tennessee Libraries

Department: Public Services

Office: (865) 974-4404

Email: [ceaker@utk.edu](mailto:ceaker@utk.edu)

Institution: Oak Ridge National Laboratory (ORNL)

Name: **Leslie Hook**

Profile: <https://www.ornl.gov/staff-profile/leslie-hook>

Office: (865) 241-4846

Email: [hookla@ornl.gov](mailto:hookla@ornl.gov)

**Purpose or objective of visit:**

- Discuss leveraging DMP resources and developing data management support and training workshops into research collaborations partnerships and senior stakeholders support

**History of contact between two parties:**

- **University of Tennessee at Knoxville (UTK)** – There is very limited history of contact with this party. I met Chris Eaker at the DataOne Users Group ([DUG](#)) Annual Meeting in Asilomar, CA in 2015. He was co-Chair of the DataOne Users Group and completed his term. I was selected as co-Chair of the DataOne Users Group from 2015 – 2017. I did not

---

<sup>4</sup> Wilkinson, M. D., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3, Article number: 160018. doi: 10.1038/sdata.2016.18.

## 2018 Oak Ridge Associated Universities (ORAU) Visiting Faculty Travel Grant

speak to him again until I sent him an email on 5/4/18 requesting an ORNL contact for the ORAU Travel Grant application. Chris then forwarded my request to Leslie Hook at ORNL on 5/7/18. Leslie then responded to Chris that she would forward his request to an ORNL DAAC Scientist during the latter part of week of 5/7/18. Thus, the development of a new contact with an ORNL DAAC Scientist is in process as of today, 5/10/18.

- **Oak Ridge National Laboratory (ORNL)** – There is no history of direct contact with this party. However, indirect contact with ORNL party via UTK party occurred on 5/7/18.
- **ORNL Distributed Active Archive Center (DAAC)** – There is no history of direct contact with this party. However, indirect contact with ORNL DAAC party via ORNL party is allegedly set to occur during the end of week of 5/7/18.

### Anticipated outcomes of the travel:

- Develop new collaborative partnerships with the UTK, ORNL, and ORNL DAAC
- Explore the data management resources and training available at these facilities
- Discuss ideas for addressing any gaps in these resources for broader impact and reach
- Investigate resources and tools that enable capacity, reproducibility, and sustainability
- Develop a ‘**research data science**’ assessment survey for ORAU and SECU
- Identify funding agency (e.g. [NSF CAREER](#)) to further investigate ORAU and SECU

‘**Research Data Science**’ is defined by CODATA-RDA as an ensemble of (1) Open Science principles and practices ([FAIR](#)) and research data management and curation skills, (2) the use of a range of data platforms and infrastructures, (3) large scale analysis, (4) statistics, (5) visualization and modeling techniques, (6) software development and annotation, and (7) more[6].

### References:

1. Digital Curation Centre. (2008). DCC Curation Lifecycle Model | Digital Curation Centre. Available from <http://www.dcc.ac.uk/resources/curation-lifecycle-model>.
2. Faundeen, J. L., Burley, T. E., Carlino, J. A., Govoni, D. L., Henkel, H. S., Holl, S. L., ... Zolly, L. S. (2013). The United States Geological Survey Science Data Lifecycle Model. U.S. Geological Survey Open-File Report 2013–1265, 4p. Available from <http://tinyurl.com/y7fzftsf>.
3. Australian National Data Service (ANDS). (2018). Data management overview. Retrieved May 10, 2018 from <http://www.ands.org.au/working-with-data/data-management/overview>.
4. Macklin, A. S. (2007). Communities of Practice In G. Bodner & M. Orgill (Eds.), *Theoretical Frameworks for Research in Chemistry/Science Education* (pp. 204-227). Upper Saddle River, NJ: Pearson.
5. United States Geological Survey. (2015). U.S. Geological Survey Community for Data Integration (CDI) Science Support Framework (SSF) The CDI Science Support Framework (SSF) USGS Data Assets Flow through the CDI Science Support Framework. Retrieved from <https://www2.usgs.gov/cdi/cdi-ssf/cdi-ssf-components.pdf>.
6. CODATA. (2018). Committee on Data of the International Council for Science. CODATA-RDA School of Research Data Science. Retrieved May 10, 2018 from <http://www.codata.org/working-groups/research-data-science-summer-schools>.