



Data Science as a Research Support Service and the Role of the Libraries: UF's CNI Executive Roundtable Experience

Authors: Plato L. Smith II, Erik Deumens, & Christopher Barnes

Conference title: 24th International Conference on Grey Literature

Acronym: ICGL

Dates: 5 December 2022

Place: Virtual

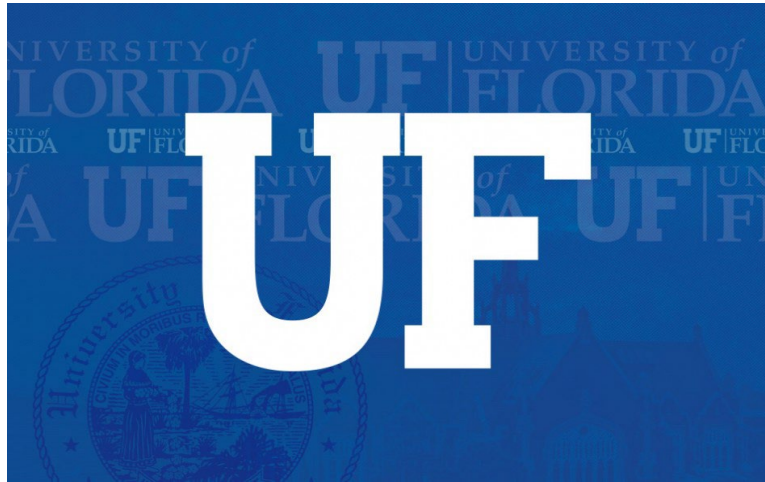
Website: <https://textrelease.com/gl2022conference/program.html>

Session: Plenary Session Two

Part: Stakeholders and Policies Influencing Grey Literature



University of Florida



- Established 1853
- Carnegie classification – R1: Doctoral Universities – Highest Research Activity
- Public land grant university on 2,000 acre campus
- 54,000+ students; 5000+ faculty
- No. 5 for Public Universities in U.S. (U.S. News & World Report Ranking)
- Located in Gainesville, Florida USA
- State University System of Florida



Framing the Context

“Contemporary research – particularly when addressing the most significant, inter-disciplinary research challenges – cannot effectively be done without a range of skills relating to data, including ***the principles and practice of Open Science and research data management and curation, the use of a range of data platforms and infrastructures, large scale analysis, statistics, visualization and modelling techniques, software development and annotation***, etc., etc.” – CODATA-RDA Schools of Research Data Science (2022)



Table of Contents

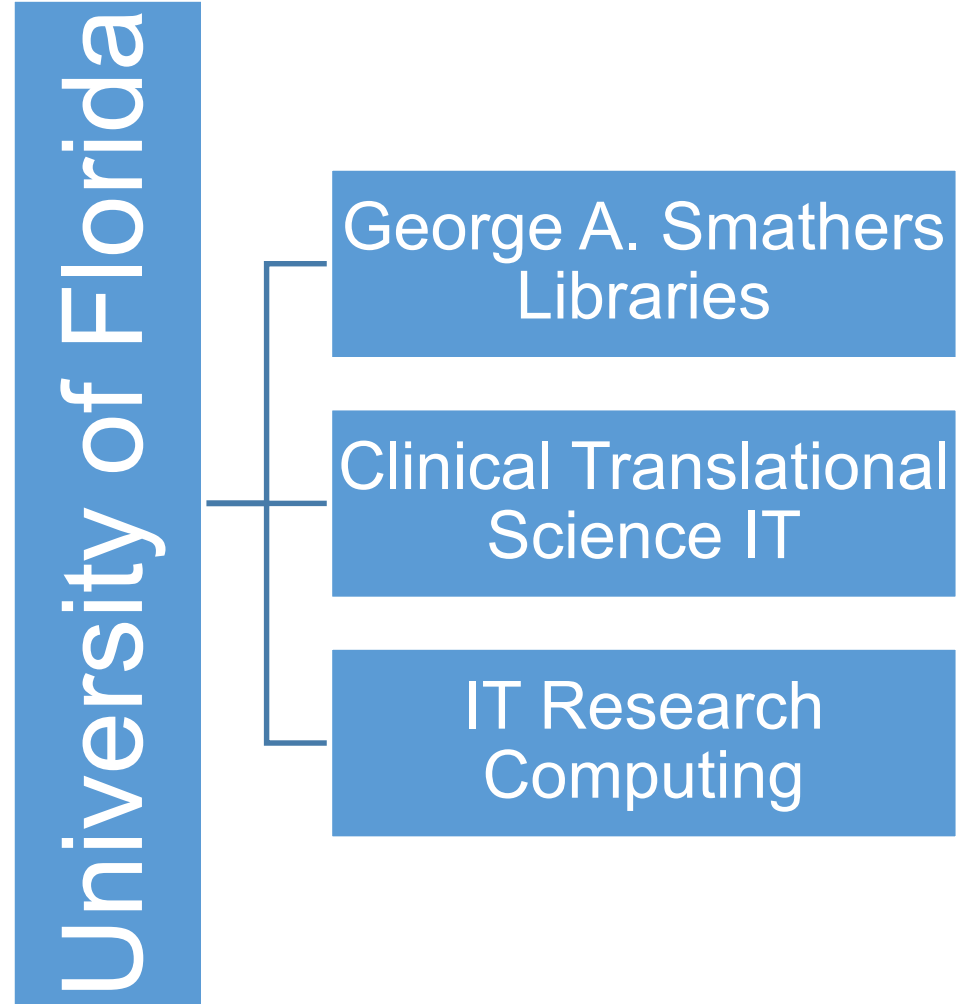
1. University of Florida Strategic Planning and Stakeholders
2. CNI Executive Roundtable: Call of Expressions of Interest – Data Science as a Research Support Service and the Role of Libraries
3. CNI Executive Roundtable Discussion Questions
4. Some key CNI Executive Roundtable responses from the University of Florida
5. Research support drivers
6. Recommendations: Program Evaluation, Equality, Equity
7. Current and future projects related to data support services initiated and led by the libraries at the University of Florida
8. The future of data science support services model
9. Horizon Opportunities – Learning from international examples (Macneil, 2022))
10. References

University of Florida Strategic Planning and Stakeholders



UF Strategic Planning alignment

- **Goal 4 of 7** – Growth in research and scholarship that enhances fundamental knowledge and improves the lives of world’s citizens.
- **Objective 4 of 4.** Processes and systems that facilitate excellence in research and scholarship - <https://president.ufl.edu/initiatives/uf-strategic-planning/>





CNI Executive Roundtable: Call of Expressions of Interest – Data Science as a Research Support Service and the Role of Libraries

CNI

Coalition for Networked Information (CNI) Washington, DC USA

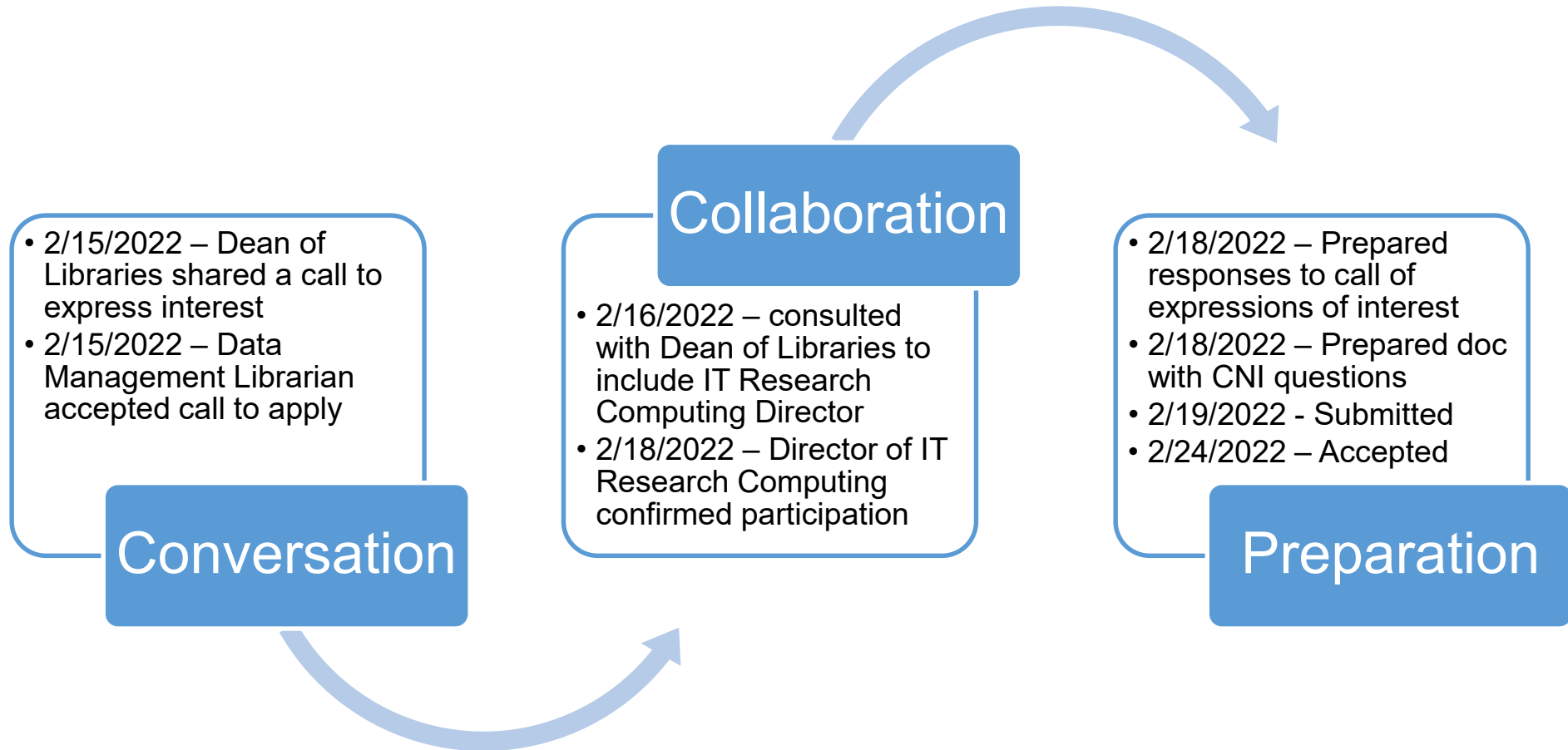
- “The Coalition for Networked Information (CNI) is dedicated to supporting the transformative promise of digital information technology for the advancement of scholarly communication and the enrichment of intellectual productivity. Over 200 institutions representing higher education, publishing, information technology, scholarly and professional organizations, foundations, and libraries and library organizations make up CNI’s members; CNI is entirely funded through membership dues. Semi-annual membership meetings bring together representatives of CNI’s constituencies to discuss ongoing and new projects and to plan for future initiatives.” – About CNI <https://www.cni.org/about-cni>

CNI

Executive Roundtable – Data Science as a Research Support
Service and the Role of Libraries – March 16, 2022, 4pm EST USA

- “The goal of the Executive Roundtable was to gain insight into how data science expertise is being provided to the university community in the broader context of research support services at key CNI member institutions. CNI was particularly interested in what roles, if any, the university library is playing in research support. To enable Roundtable participants to fully understand the context at each institution, participating institutional teams prepared to speak to at least most of the following issues” covered in the ten (10) questions provided to the participating institutions in the roundtable discussion. The University of Florida was one of 13 key CNI member institutions selected to participate in this executive roundtable discussion.

CNI Executive Roundtable: Call of Expressions of Interest – Data Science as a Research Support Service and the Role of Libraries





CNI Executive Roundtable Discussion Questions (1 of 2)

1. How are research support services organized, staffed, supported and funded at your institution? What is included in the portfolio of research support services? How is support services funding related to grant proposals and other sources of funding that directly support researchers? Are costs for institutional support services written into grant budgets?
2. What is the scope of data science research support services?
3. What is the library's role in research support service provision? Which services does it provide or support? Who are the other players?
4. How are you staffing this support function? What are the qualifications, backgrounds, and training for support providers? How much disciplinary expertise (and in which disciplines) is expected along with generalized data science expertise? How large is the staff and how are they organized?
5. How do you define the boundaries, differences, and relationships between research data management and research data science support at your institution? Note that at some organizations data science and research data management are treated together and considered as a continuum of activities; in others they are very distinct.



CNI Executive Roundtable Discussion Questions (2 of 2)

6. Who are the primary users of these research support services? Faculty only? Graduate students? Undergraduates?
7. How much demand is there for these research support services and how are you managing or rationing this demand? What are the limits on institutional general purpose research support services in this area? In other words, when do research programs need to shift to dedicated, self-funded data science support staff or collaborators?
8. What disciplines are making use of, perhaps relying upon, these research support services? Are there other areas (life sciences, for example) where these functions are being handled through funded research collaborations outside the scope of institutional research support services?
9. How are academic programs at the institution organized, and how do they relate to research support services (for example, participants in the academic programs might supply expertise to research support services capstone projects, clinics, or other mechanisms)?
10. What training or education in data science is being offered as part of research support services (other than formal academic program courses), and who is offering it?



Some key CNI Executive Roundtable responses from the University of Florida



Large scale computing and data storage infrastructure costs are covered by startup funds and grants



All faculty, students, and staff and collaborators are supported. Multidisciplinary teams are supported across colleges & depts.



There is now more demand than capacity to serve. We are trying to develop a business model that will scale to the current demand.



The libraries and IT offer short training opportunities that complement semester-long courses taught by the colleges.

Research support drivers (training future researchers)

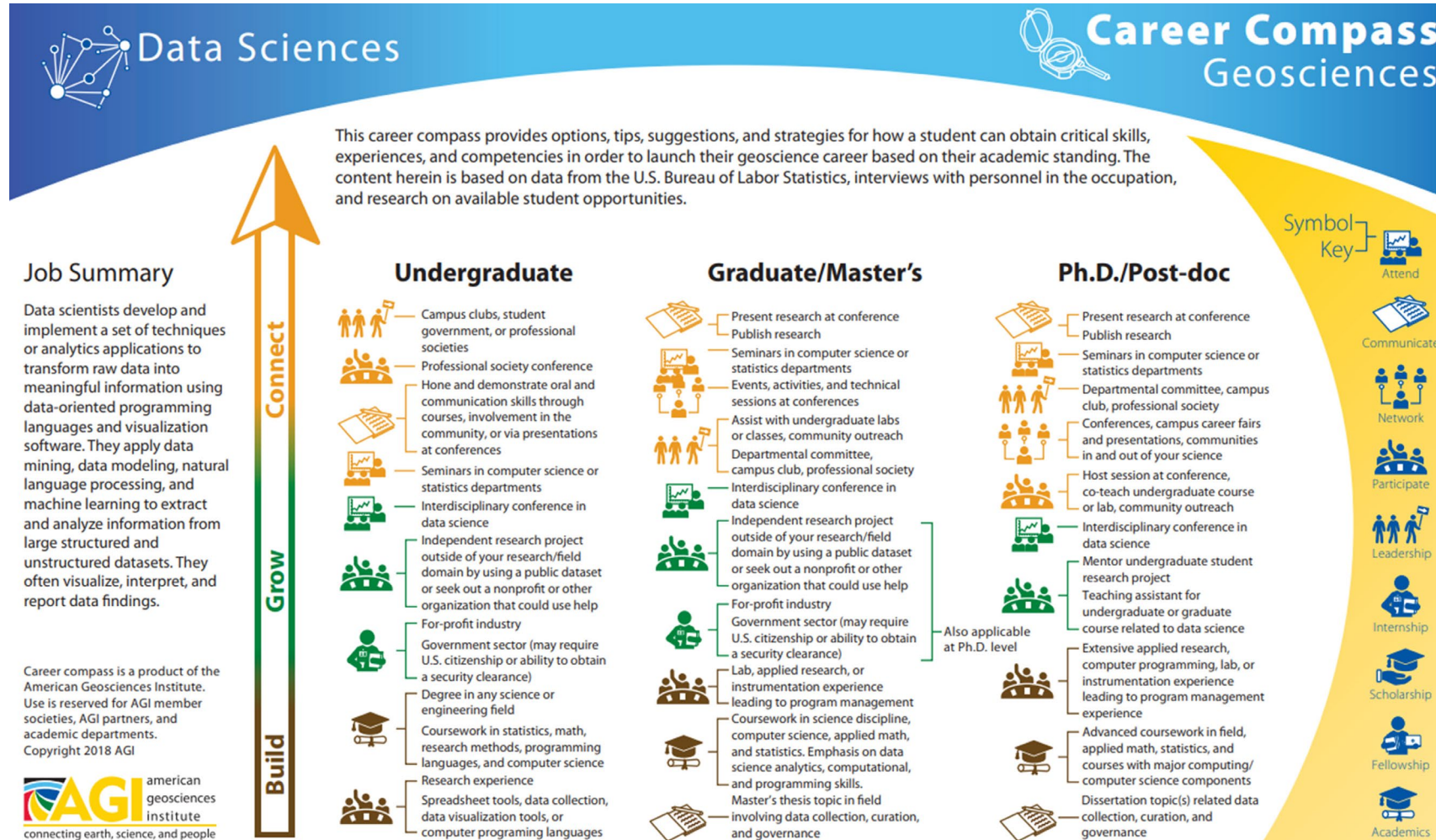


Figure 1: American Geosciences Institute, Data Sciences Career Compass



Research support drivers (Research Data Management Support Services (RDMS))



Figure 2: Research Data Management Services (Jones et al., 2013)

Recommendations: Program Evaluation, Equality, Equity

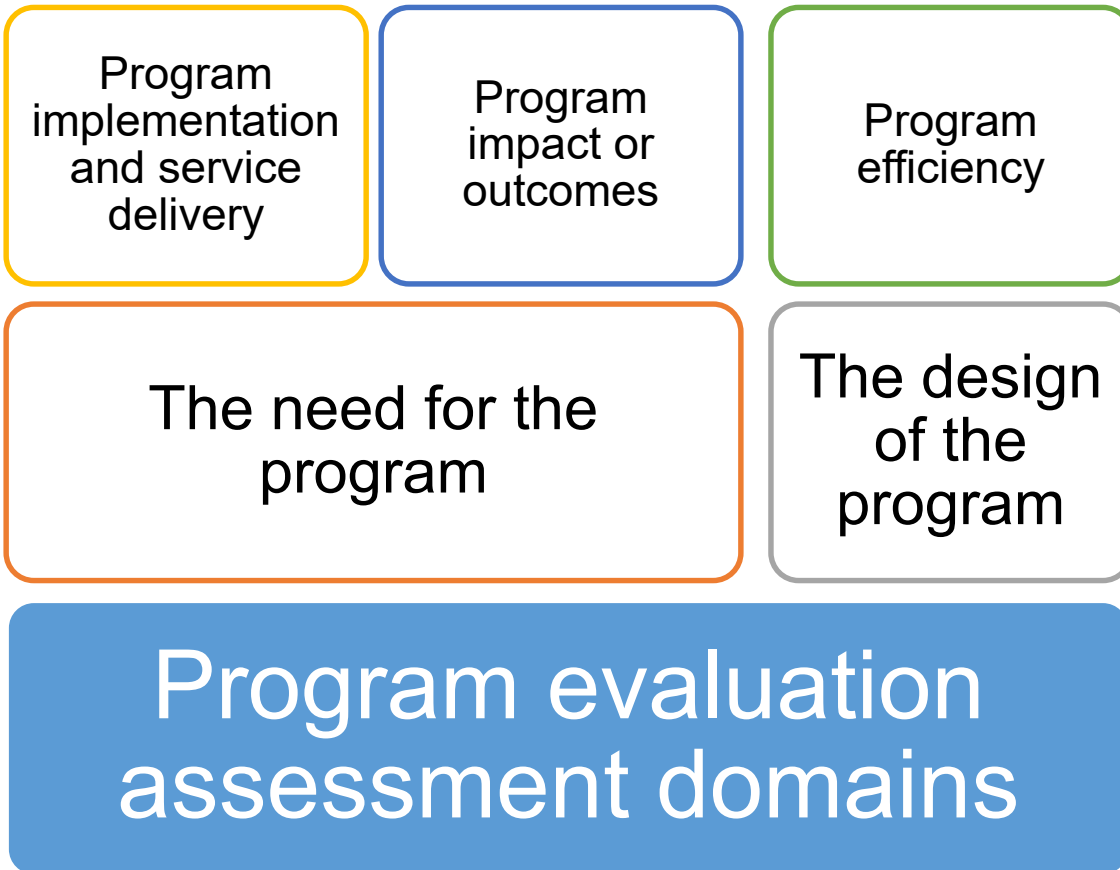


Figure 3: Program evaluation assessment of one or more of five program domains (Rossi et al., 2004, p.29)

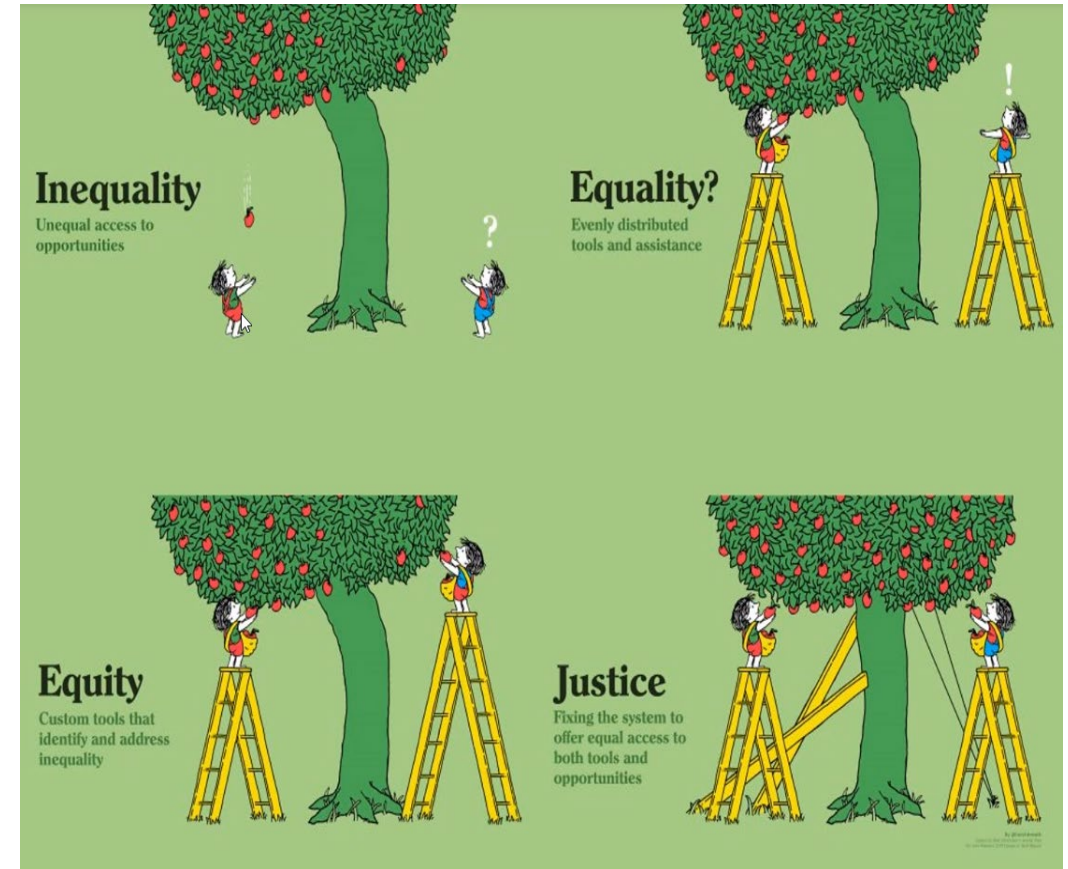


Figure 4: Equality, Equity, and Justice



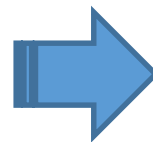
Current and future projects related to data support services initiated and led by the libraries at the University of Florida

Data Services Continuing Professional Education (DSCPE) Capstone Partners – Fall 2022 (Current)

- DSCPE Capstone Project - Review of research data management programs at select peer institutions

Ithaka S+R Coordinating Research Data Support Services Across Campus: 2023 – 2025 (Future)

- **Launch of a New Cohort-Based Research and Consulting Project includes:** Brandeis University, Carnegie Mellon University, Clemson University, Harvard University, Indiana University, Johns Hopkins University, Montclair State University, San Diego State University, University of Chicago, University of Pittsburgh, **University of Florida**, University of Virginia, and Yale University.



















Data Services Continuing Professional Education

About Capstone Partners Cohort Members Contact Pilot Program Team

Capstone Partners

A special thank you to all of our Fall 2022 Capstone hosts!

 The University of Alabama at Birmingham	 BROWN University Library	 George A. Smathers Libraries UNIVERSITY of FLORIDA
 HARVARD Dataverse	 HARVARD LIBRARY	 UNIVERSITY of MARYLAND Health Sciences and Human Services Library
 TAUBMAN HEALTH SCIENCES LIBRARY UNIVERSITY OF MICHIGAN		
 NIH NLM	 Northeastern University Library	 NYU HEALTH SCIENCES LIBRARY National Center for Data Services
 Oregon State University Libraries and Press	 Penn Libraries Leon Levy Dental Medicine Library	 Stanford MEDICINE Lane Medical Library
 Lamar Soutter Library UMass Chan Medical Center	 Utah State University MERRILL-CAZIER LIBRARY	 UVA Health CLAUDE MOORE HEALTH SCIENCES LIBRARY



The future of data science support services model: Building a Machine-Actionable Pipeline for Research Data (University College London, Supercomputing 2022 Poster)

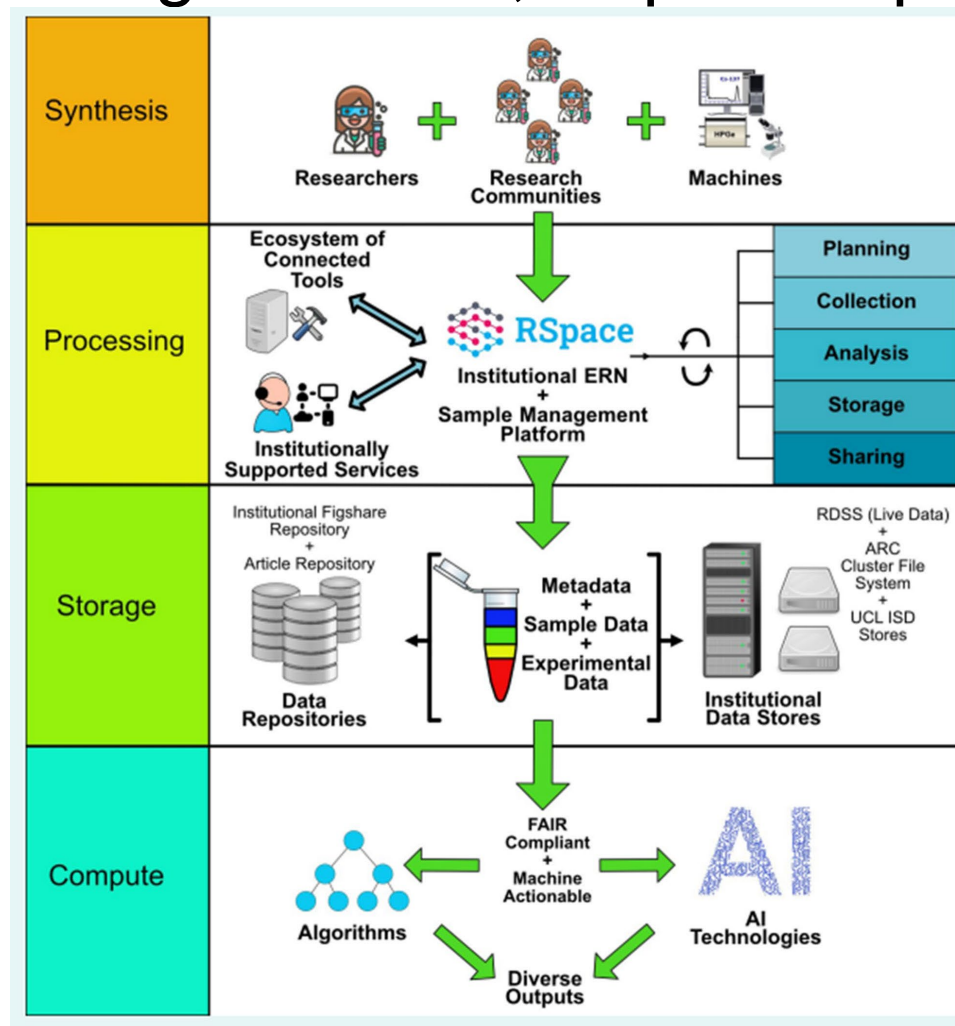


Figure 5: Building a Machine...Ahmed, M. et al., 2022 (used with permission)

Horizon Opportunities – Learning from international examples (Macneil, 2022)

- **National provision of research infrastructure**
 - **Australia** – Australia Research Data Commons
 - National Data Assets - <https://ardc.edu.au/program/national-data-assets/>
 - **Netherlands** -- SURF service provision to Dutch universities - <https://www.surf.nl/en>
 - Cyberinfrastructure including storage
 - Metadata management via iRODS
 - Authentication
 - **Canada** -- **Digital Research Alliance service provision to Canadian Universities**
 - Cyberinfrastructure including storage
 - Data repository (Dataverse)
 - Data management plan (DMP Assistant)
- **Global Open Research Commons**
 - Concept: Suite of interconnected research infrastructure and services
 - Possible providers: Supranational, national or domain organization
 - Lead investigator: RDA GORC International Model Working Group
- **Institution-level innovation**
 - University College London (UCL) example

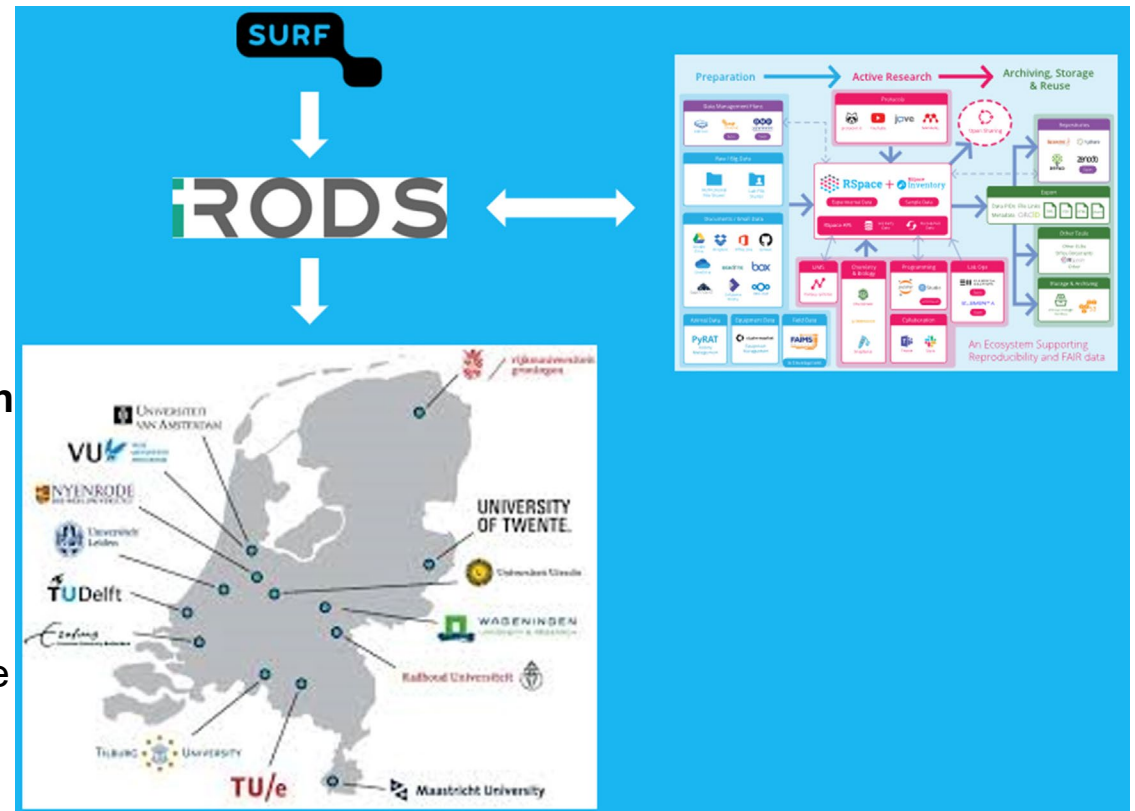


Figure 6: Working with universities to provide integrated research infrastructure (Wilson and Macneil, 2022)



Thank you



Figure 7: 21st International Conference on Grey Literature.
TIB Leibniz Information Centre for Science and Technology
University Library. Hannover, Germany. October 22-23, 2019.



References

- CODATA-RDA Schools of Research Data Science. (2022). <https://codata.org/initiatives/data-skills/research-data-science-summer-schools/>.
- CNI. (2022). <https://www.cni.org/>.
- Data Services Continuing Professional Education (DSCPE). (2022). <https://dscpe.github.io/>.
- Davis, M. C., Challenger, R., Jayewardene, D. N. W., & Clegg, C. W. (2014). Advancing socio-technical systems thinking: A call for bravery. *Applied Ergonomics*, 45(2 Part A), 171–180. <https://doi.org/10.1016/j.apergo.2013.02.009>.
- Ithaka S+R. (2022). Coordinating Research Data Support Services Across Campus. <https://sr.ithaka.org/blog/coordinating-research-data-support-services-across-campus/>.
- Jones, S., Pryor, G. & Whyte, A. (2013). ‘How to Develop Research Data Management Services - a guide for HEIs’. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: <https://www.dcc.ac.uk/guidance/how-guides>.
- Leeds University Business School. (2022). Socio-technical systems theory. <https://tinyurl.com/vfjt87y>.
- National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. 2005. Facilitating Interdisciplinary Research. Washington, DC: The National Academies Press. <https://doi.org/10.17226/11153>.
- National Science Foundation (NSF). (2022). Leading the World in Discovery and Innovation, STEM Talent Development and the Delivery of Benefits from Research – NSF Strategic Plan for Fiscal Years (FY) 2022 – 2026. https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf22068.
- National Science Foundation (NSF). (2022). NSF CISE Vision 2030: A Vision for NSF CISE Opportunities and Challenges Over the Next Decade. https://nsf.gov/cise/advisory/CISE_Vision_2030.pdf.
- National Science Foundation (NSF). (2021). Transforming Science Through Cyberinfrastructure: NSF’s Blueprint for a National Cyberinfrastructure Ecosystem for Science and Engineering in the 21st Century. <https://www.nsf.gov/cise/oac/vision/blueprint-2019/CI-LWD.pdf>.
- Rossi, P., Lipsey, M. W., Freeman, H. E. (2004). Evaluation: A Systematic Approach, 7th Edition. Sage Publications, Inc. Thousand Oaks, CA.
- Wilson, J. and Macneil, R. (2022). Towards an integrated approach to Research data management: UCL and Leiden/SURF examples. EUDAT 2022.