

Is Open Science FAIR for Stakeholders?

Twenty-First International Conference on Grey Literature

Plato L. Smith II, University of Florida
11:00am, Session 1, October 22, 2019
German National Library of Science and Technology
Leibnizhaus, Holzmarkt 5
Hannover, Germany

University of Florida



- Established 1853
- Carnegie classification – R1: Doctoral Universities – Highest Research Activity
- Public land grant university on 2,000 acre campus
- 50,000+ students; 4,000+ faculty
- Located in Gainesville, Florida, USA
- State University System of Florida

Table of Contents

1. Open Science, FAIR, and Stakeholders
2. NIH NIEHS P42 Data Management and Analysis Core (DMAC) 2018 requirement – NEW CORE!
3. Proposal for Development of a UF DMAC
4. Developing socio-technical data management collaborations with stakeholders
5. References

Open Science, FAIR, Stakeholders

- “**Open Science** is the practice of science in such a way that others can **collaborate and contribute**, where **research data, lab notes** and other **research processes** are freely available, under terms that **enable reuse, redistribution and reproduction** of the research and its underlying **data and methods**.” – FOSTER
- **FAIR – (Findable, Accessible, Interoperable, Reusable)** – See: [FAIR- 15 Principles](#)
 - **F1.** (Meta)data are assigned a globally unique and persistent (1 of 4 elements).
 - **A1.** (Meta)data are retrievable by their identifier using a standardized communications protocol (1 of 2 elements with 1st element having two sub elements).
 - **I1.** (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation (1 of 3 elements).
 - **R1.** Meta(data) are richly described with a plurality of accurate and relevant attributes (1 of 4 elements). – GO FAIR/Joint Declaration of Data Citation Principles – [Final](#) (2014)
- **Stakeholders** – “While the term *stakeholder* may be essentially a management term, it points to an **extremely broad range of actors** who are attended to by a broad range of subject matter disciplines relevant to management... Note that ‘**key stakeholders**’ also include ‘**insiders**’ such as public managers and employees, as well as ‘**outsiders**’ such as political overseers and **funders**.” – Bryson (2007)

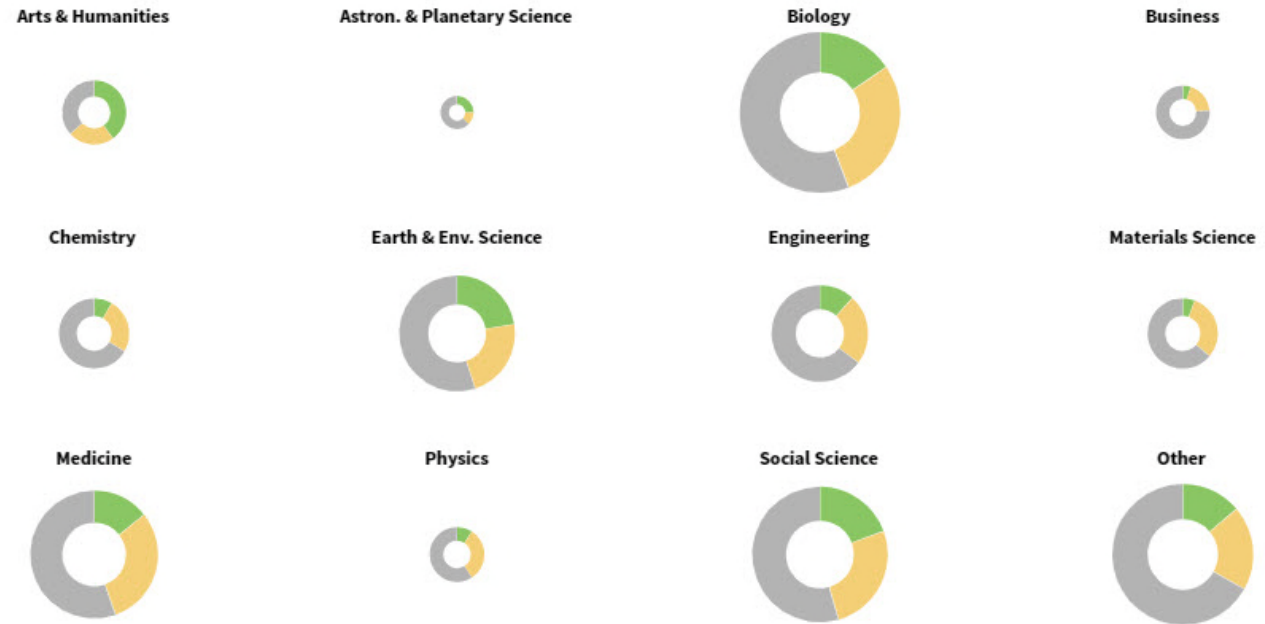
Open Science, FAIR, Stakeholders

“A **diverse set of stakeholders**—representing academia, industry, funding agencies, and scholarly publishers—have come together **to design** and **jointly endorse** a concise and measureable **set of principles** that we refer to as **the FAIR Data Principles.**” (Wilkinson et al., 2016)

Familiarity with FAIR principles

The majority of researchers surveyed as part of a recent study on open data had never heard of FAIR, regardless of their field. Of the 748 researchers that responded to this question, 144 said they were familiar with the principles. Circles are sized by number of respondents.

■ I am familiar with the FAIR principles ■ I have previously heard of the FAIR principles but I'm not familiar with them ■ I've never heard of the FAIR principles before now



Source: [State of Open Data](#)

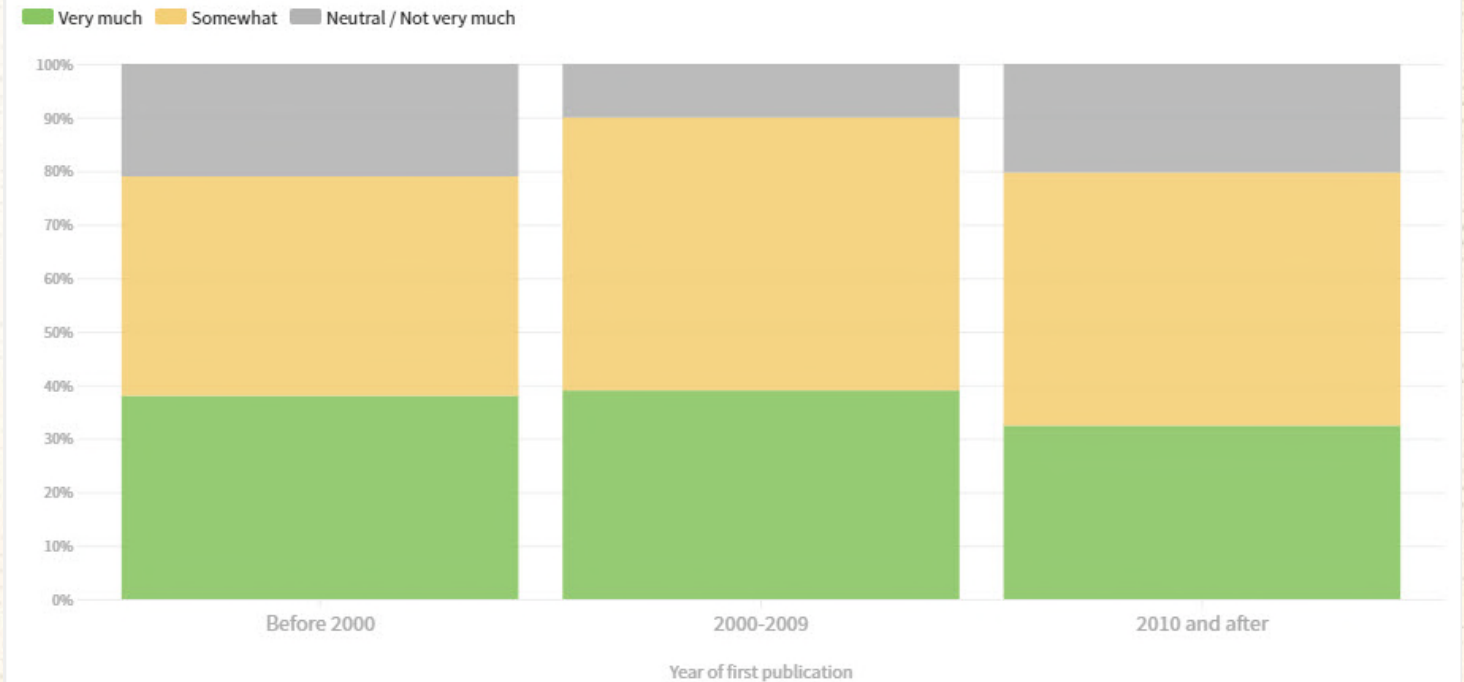
Fig. 1 – Familiarity with FAIR principles (Brock, 2019)

Open Science, FAIR, Stakeholders

“There are **numerous and diverse stakeholders who stand to benefit** from overcoming these obstacles: **researchers wanting to share, get credit**, and reuse each other’s data and interpretations; **professional data publishers** offering their services; **software and tool-builders** providing data analysis and processing services such as reusable workflows; **funding agencies** (private and public) increasingly concerned with **long-term data stewardship**; and a **data science community mining**, integrating and analysing new and existing data to **advance discovery**.” (Wilkinson et al., 2016)

Compliance with FAIR principles

Of the participants who were familiar with FAIR, about a third said that their data management practices were very compliant with the principles. That proportion is similar across scientists at different stages of their career.



Source: [State of Open Data](#) • Data shown for respondents familiar with FAIR principles

Made with Flourish

Fig. 2 – Compliance with FAIR principles (Brock, 2019)

NIH NIEHS P42 Data Management and Analysis Core (DMAC)

2018 requirement – **NEW CORE!**

Purposes

- Support the management and integration of data assets across the Center, irrespective of dataset size
- Establish, coordinate, and monitor processes for data analysis
- Work with project/core leaders to ensure high data quality through lifecycle of data
- Identify opportunities for integrating project/core-generated data with other existing datasets
- Foster and enable interoperability of data between BMR and ESE projects to accelerate impact of Center's research
- Promote best principles so data is **Findable, Accessible, Interoperable, and Reusable (FAIR)**

Comprehensive data management plan

- 1) coordination with projects and cores
- 2) fostering data sharing and interoperability
- 3) data quality assurance and quality control.

Provides Points of Contact (in Relation to Overall Center)

- Data Collection Tool
- SRP-hosted conference calls/webinars data management and analysis

- May also include additional functionalities of biostatistics, bioinformatics, geographical information systems, and computational modeling, etc.
- Note: Not required to create a repository for the Center's data

DMAC does not have set budget

Data sharing policies: <https://grants.nih.gov/policy/sharing.htm>; <https://fairsharing.org/>

SRP Data Collection Tool: <https://tools.niehs.nih.gov/srp/rtc/index.cfm>

Proposal for development of a UF DMAC

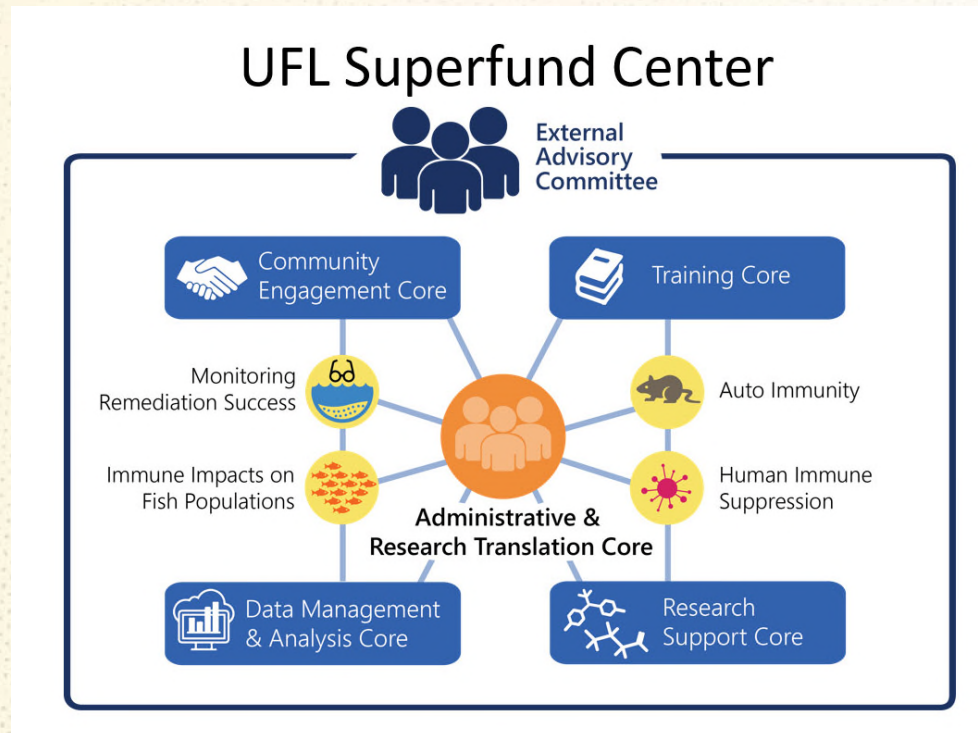


Fig. 3 – Proposal for UFL Superfund Center (Vulpe et al., 2018)

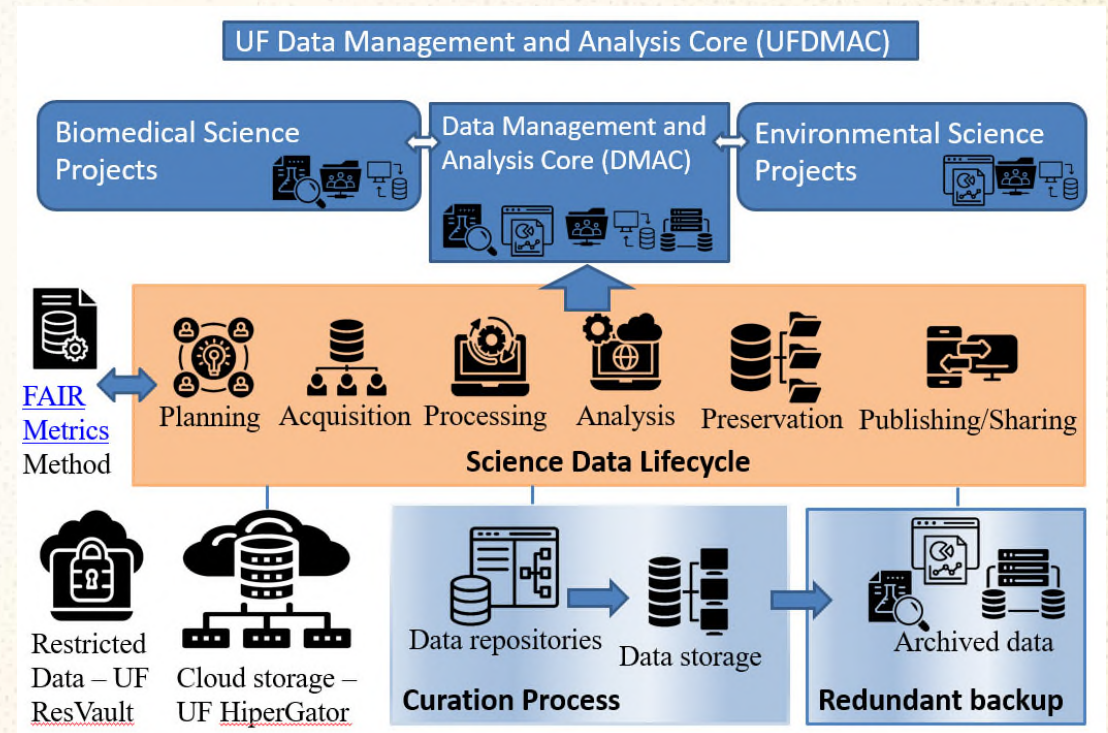


Fig. 4 – Proposal for UF DMAC (Smith and Barbazuk, 2018) – modeled after Northeastern Univ.'s [PROTECT](#)

Developing socio-technical data management collaborations with stakeholders

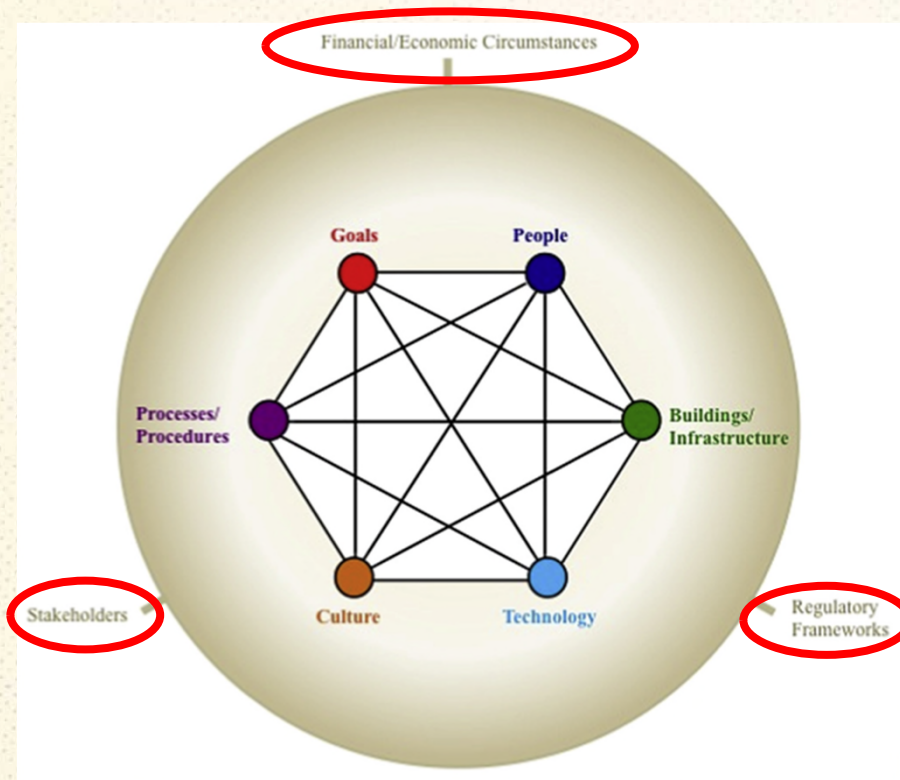
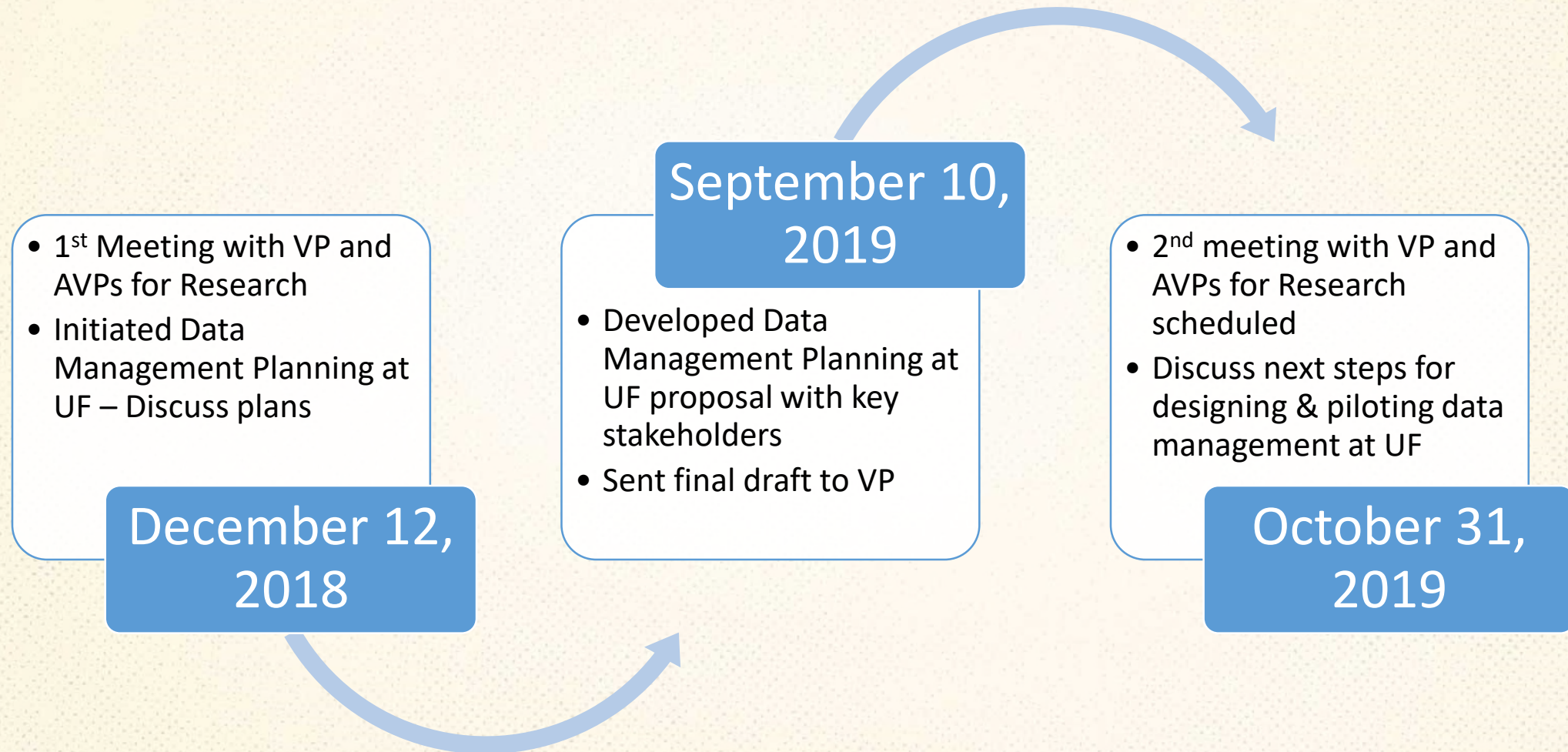


Fig. 5 – Social-technical system, illustrating the interrelated nature of an organizational system, embedded within an external environment.
(Davis et al., 2014, p. 173)

- **UF Office of Research**
- **UF George A. Smathers Libraries**
- **UF Research Computing**
- **UF Clinical and Translational Science – IT (CTS-IT)**
- **UF Informatics Institute (UFII)**
- **UF Center for Environmental and Human Toxicology (CEHT)**
- **UF Interdisciplinary Center for Biotechnology Research (ICBR)**

Developing socio-technical data management collaborations with stakeholders



Acknowledgements

-
- Dr. David Norton, UF Office for Research Vice President
 - Dr. Sobha Jaishankar, UF Office for Research Assistant Vice President
 - Judith Russell, UF George A. Smathers Libraries Dean
 - Patrick Reakes, UF George A. Smathers Libraries Senior Associate Dean
 - Dr. Eric Deumens, Director, UF Research Computing
 - Christopher Barnes, Director, UF Clinical Translation Science IT
 - Dr. Christopher Vulpe, Professor, UF Center for Environmental Health and Toxicology (CEHT)
 - Dr. William “Brad” Barbazuk, Director, UF Interdisciplinary Center for Biotechnology Research (ICBR) Informatics
 - Colleagues, coworkers, and members of UF ARCS, UF DMCWG, and UFII/Carpentries

1. Brock, J. (2019). “A love letter to your future self”: What scientists need to know about FAIR data. Nature Index 360. Accessed online October 16, 2019 from <https://go.nature.com/2tl0x4y>.
2. Bryson, J. M. (2007). What to do when stakeholders matter. *Public Management Review*. v.6, (1). doi.org/10.1080/14719030410001675722.
3. 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>.
4. Data Citation Synthesis Group: joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014. <https://bit.ly/31oCLmM>.
5. FOSTER. (nd). Open Science Definition. Accessed online October 16, 2019 from <https://bit.ly/2IJEnTO>.
6. GO FAIR. (nd). FAIR Principles. Accessed online October 16, 2019 from <https://www.go-fair.org/fair-principles/>.
7. HATII, University of Glasgow, Digital Curation Centre, and University of London computing Centre (ULCC). (2013). CARIOD RDM Matrix. Accessed online October 16, 2019 from <https://bit.ly/2IUIX0U>.
8. NIH NIEHS. (2018). P42 Superfund Hazardous Substance Research Training Program. Accessed online October 16, 2019 from <https://bit.ly/2MK1q05>.
9. NIH NIEHS. (2019). Superfund Research Program. Currently Funded Multiproject Research Centers (P42). Accessed online October 16, 2019 from <https://tools.niehs.nih.gov/srp/programs/index267.cfm>.