




Inclusion in Data: Incorporating DEIA Components into Research Data Services Practices

Chelsea H. Barrett, Seton Hall University, South Orange, NJ, US, chelsea.barrett@shu.edu 

Abstract

The RDAP Summit 2022 theme of “Envisioning an Inclusive Data Future” reeled in various talks which provided ample avenues for incorporating DEIA components into data management strategies as well as outreach to diverse communities. This commentary will focus on a few of the strategies and initiatives discussed in presentations and talks at the conference that could and potentially should be integrated into established Research Data Services programs.

Received: July 1, 2022 **Accepted:** November 11, 2022 **Published:** February 15, 2023

Competing Interests: The author declares that they have no competing interests.

The *Journal of eScience Librarianship* is a peer-reviewed open access journal. © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

See <http://creativecommons.org/licenses/by/4.0>.

∞ OPEN ACCESS

Introduction

The prevalence and dependence upon Research Data Services (RDS) in universities and institutions worldwide has given rise to innovative and timely research from a variety of disciplines and perspectives. In many institutions, librarians are integrated into the RDS team, aiding with research as well as guiding researchers to appropriate avenues to data collection, data management, and data preservation. As more ideas are utilizing the copious amounts of data available, there are many instances where researchers are unaware of areas where there may be diversity, equity, inclusion and/or accessibility (DEIA) risks that need to be addressed. As a member of a RDS team, it is imperative to identify areas where DEIA is present to ensure that information is accessible, and that projects and research conducted are inclusive to all populations.

The Research Data Access and Preservation (RDAP) 2022 Summit with the theme of “Envisioning an Inclusive Data Future” had 307 attendees from more than 4 continents (Research Data Access and Preservation Association 2022), all with diverse backgrounds and perspectives on handling, collecting, visualizing, and distributing data. When looking from the lens of RDS, many of the presentations provided new and interesting viewpoints and initiatives on what can be done to incorporate DEIA into overall RDS processes and practices.

Data Training and DEIA

Working with data and incorporating DEIA components is not always at the forefront of a data management plan. Researchers who are working with data are typically concerned with the proper protocol of collecting and handling data which unfortunately does not always relate to DEIA issues and concerns. Those who serve on a RDS team may need an avenue to prepare researchers to think about DEIA issues throughout the research process.

Peace Ossom Williamson, the Associate Director of the National Library of Medicine’s (NNLM) National Center for Data Services gave a presentation on a three-year internship which “provides practical experiences for graduate students that include the soft and hard skills for data librarian positions, including working on a team, building a network, understanding the data lifecycle, and working with data” (Ossom-Williamson 2022). Within this internship, graduate students learn proper data cleaning, data structuring, analysis, and visualization protocols. This initiative aims to assist students in incorporating DEIA components into the data process and will hopefully serve as a recruitment tool for people of color seeking to enter data librarianship. Colby Witherup Wood from Research Computing Services at Northwestern University proposed a related idea that is a bit more condensed than the previous initiative and provides a more collaborative approach to training called “Bring Your Own Data (BYOD) Working Groups” (Wood 2022). These working groups run for about eight weeks with 60-minute roundtable sessions with 3-7 participants. The roundtables run around a selected theme led by a member of the team. An appealing aspect of this setup is that participants can speak with the other participants in an open forum to get feedback on data management practices and several aspects of the research process as it pertains to their individual projects.

From a RDS perspective, having the opportunity to train researchers would combat many of the common mistakes that come up in the research process. The internship initiative is a bit more structured than the working groups, however dependent upon the modality preferred, either avenue is appropriate for implementation and can still be fruitful for any RDS team.

DEIA and Data Visualization

An important part of the research process is the presentation of the data. The receiving audience must be able to understand and interpret the findings of the research. Data visualizations enable researchers to create digestible visuals for the audience to get a clear and concise narrative of the data and analysis of that data. However, how often do researchers think about the implications of DEIA of the data?

Kristen Adams from the Miami University Libraries presented on an initiative revised to meet DEIA requirements. A group of librarians created online self-paced, asynchronous modules on data management, data curation, data visualization and other data-related topics to be either integrated in courses or as a standalone for students to take on their own time (Adams 2022). However, upon reviewing the content of the modules, the modules were not accessible to most populations, as the original target audience were those in the STEM field at a certain proficiency level. The modules were revised to include learners of all abilities, different disciplines and learning levels, creating a more inclusive opportunity for different populations. On a similar vein of making sure practices incorporate DEIA components, Negeen Aghassibake from the University of Washington Libraries spoke on data scientists doing a review on a bigger scale. Aghassibake urged data professionals to evaluate the “best” data practices that have been routine and instead make them inclusive. When working with data and creating data visualizations, there needs to be understanding that data is not neutral and has biases that can be harmful to various populations. An example used in Aghassibake’s presentation referred to the practice of identifying American Indians/Alaska Natives with Pacific Islanders and Asian Americans. By doing this and consolidating differing racial groups into one category, it erases the distinction between the two which is bad practice (Aghassibake 2022). It is important to remember there are many ways where working with data can be harmful, but in learning these issues, data can also become inclusive. It all depends on the intentionality of the person working with the data, the knowledge of best practice and the overall data collection process.

DEIA with Grants Research

There are many who may work in tandem or closely with an office that specializes in grants and research. This may involve the RDS team encountering various research ideas, Institutional Review Board (IRB) materials, and overall questions about the research process. When assisting researchers on grants and overall research, it may be of use to preemptively address potential issues before they arise.

Dessi Kirilova from the Qualitative Data Repository (QDR) spoke about the DataPro Tool, which would allow the researchers to input information about their research, data and the overall project and then receive

feedback, protocol recommendations and data management steps to take throughout the data lifecycle (Kirilova 2022). One of the greatest aspects of this tool is that the information obtained by the tool will advise the researcher against potential risks to human participants before the data collection or research has even begun. This allows for conversation beforehand to address issues and produce solutions so that the research does not do harm. Any RDS team that works closely with grants or research at their institution may find significant use in a questionnaire like this, even if it is an iteration of this idea. This form can also provide data and insights on prevalent issues that researchers may be running into which may prompt the creation of related workshops or further training to combat these issues before they arise.

Conclusion

The implications of DEIA in data are not always obvious though still need to be addressed and adhered to in the data research process. Those tasked with assisting researchers in this capacity have a responsibility to identify issues and risks with potentially affected populations and rectify those issues prior to data collection, analysis, and/or visualization. The talks given at the RDAP 2022 Summit provided many ways to analyze and correct existing practices to truly be inclusive and as fair as possible when it comes to data. By implementing one or more of these initiatives, RDS teams will train researchers to utilize and employ the best and most inclusive practices possible.

Acknowledgements

The author thanks Tina Griffin from the University of Illinois Chicago for being a wonderful mentor in the process of writing this commentary.

The content of this article is based on the presentations at RDAP Summit 2022, available from [Open Science Framework: RDAP Summit 2022 Public materials](#).

Competing Interests

The author declares that they have no competing interests.

References

Adams, Kristen. 2022. "The Impact of 'Academic Capitalism' on 'Belongingness:' Institutional Data's Impact on Diversity, Equity, and Inclusion." Recorded March 15, 2022 at Research Data Access and Preservation (RDAP) 2022 Summit, 29:20.

<https://www.youtube.com/watch?v=LQQeogbMHEo&t=1758s>.

Aghassibake, Negeen. 2022. "From 'Best' Practices to Inclusive Practices: Critical Approaches to Data Structures." Presented March 15, 2022 at Research Data Access and Preservation (RDAP) 2022 Summit.

<https://osf.io/ftkgw>.

Kirilova, Dessi. 2022. "DataPro-A Tool for Characterizing and Communicating Research Data Risks and Responses." Recorded March 15, 2022 at Research Data Access and Preservation (RDAP) 2022 Summit, 0:01. <https://www.youtube.com/watch?v=6J9SKYeNNIM>.

Ossom-Williamson, Peace. 2022. "Developing Diverse Data Librarians: Construction of a National Internship Program." Recorded March 15, 2022 at Research Data Access and Preservation (RDAP) 2022 Summit, 10:04. <https://www.youtube.com/watch?v=LQQeogbMHEo&t=604s>.

Research Data Access and Preservation Association. 2022. "RDAP Summit 2022: Opening, Keynote, and Presentations Session 1." Recorded March 15, 2022 at Research Data Access and Preservation (RDAP) 2022 Summit, 0:28. https://www.youtube.com/watch?v=CTSKYahg0Wg&list=PLm34yOXj7cB8yJCuW5v6_PkKubhZ0XWpL&index=1.

Wood, Colby Witherup. 2022. "Bring Your Own Data (BYOD) Working Groups: A New Service to Multiply Staff Impact and Create Community." Recorded March 15, 2022 at Research Data Access and Preservation (RDAP) 2022 Summit, 18:26. <https://youtu.be/7CU6cbDLwgg?t=1106>.