Open Data, Diversity Data: Diversity Scholars and Campus Library Research Data Services

Woodbrook, Rachel; De Vera, Emma; Segoshi, Megan
https://dx.doi.org/10.7302/878
https://hdl.handle.net/2027.42/167203
http://creativecommons.org/licenses/by-nc/4.0/

Downloaded from Deep Blue, University of Michigan's institutional repository
Open Data, Diversity Data: Diversity Scholars and the Campus Library

Rachel Woodbrook (woodbr@umich.edu)
Emma De Vera (edevera@umich.edu)
Megan Segoshi (diversityscholars@umich.edu)
https://tinyurl.com/OpenDataDEIA
Agenda

1. Research
2. Toolkit
3. Partnership & (Potential) Applications
Research

- Problem statement
- Support
- Research
  - Partnership & Research sample
  - Mixed methods
- Toolkit
Research question

“What implications do Diversity, Equity, Inclusion and Accessibility considerations have for best practices in each step of the data lifecycle?”
The data lifecycle is depicted as eight stages in a circle leading into each other. The first two stages, which take place before starting a research project, are "Finding data (for secondary research)" then "Data planning." The next three stages take place during the project: "Data collection," "Data processing/analysis," then "Active data management." The final three stages are undertaken after the project is completed: "Data curation," "Data sharing," then "Data archiving/preservation." This final stage may lead back into finding data.
Framework for Diversity Scholarship

Diversity Scholarship is research and scholarship that broadly seeks to:

- **inform** understanding of historical and contemporary issues of social inequality across societal contexts and life domains
- **illuminate** the challenges and opportunities that arise when individuals from different backgrounds and frames of reference come together in significant societal contexts
- **inform** our understanding of systems of power and privilege and their interactions with groups historically underrepresented and marginalized based on identities
- **highlight** the experiences of disenfranchised populations, whose narratives have traditionally been relegated to the outer periphery of intellectual inquiry and academic scholarship
- **foreground** the knowledge systems, assets and resources, and cultural strengths of members of historically marginalized communities
National Center for Institutional Diversity

- Supports programming and resources for scholars committed to diversity and diversifying the academy, either through their research, teaching, or service
- Existing working relationship; partnership borne out of recognition of need for resources for diversity scholars
- Partnership as model for possible collaborations to address real needs
- Provided access to sample and expertise of NCID staff
Diversity Scholars Network (DSN)

- Scholarly community of over 900 “diversity scholars” across the country and some international
- Consists primarily of assistant professors, but also includes a number of senior scholars, postdoctoral fellows, faculty of practice, and research faculty/staff. Members are required to be affiliated with an institution at the time of their application.
- Represent a wide variety of disciplines and methodologies
Diversity Scholars Network (DSN) (cont.)

- DSN members are supported through community building and networking opportunities, having their scholarship highlighted and promoted publicly by NCID, programming and workshops, and publication and grant opportunities.
Survey Findings: Question 1

● No correlation found between demographics and likelihood of using a toolkit:
  ○ 131/140 respondents “somewhat likely” or “very likely” to use

● Biggest anticipated barrier:
  ○ Time and resources (83/140)
Survey Findings: Question 2

Important data life cycle stages: Comfortable vs. Uncomfortable

Data lifecycle stage:

- Data collection
- Finding existing data
- Data planning
- Active data management
- Data curation
- Data sharing
- Data archiving/processing
- Prefer not to say
- NA

Number of respondents

Comfortable
Uncomfortable
What toolkit resources would diversity researchers find most useful?

<table>
<thead>
<tr>
<th>Potential resource</th>
<th>No. Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examples of community engagement</td>
<td>100</td>
</tr>
<tr>
<td>Checklist of questions for making data decisions</td>
<td>97</td>
</tr>
<tr>
<td>Templates for one-page data applications or use agreements</td>
<td>95</td>
</tr>
<tr>
<td>Consent form language for data sharing</td>
<td>90</td>
</tr>
<tr>
<td>Resources on hidden metadata</td>
<td>84</td>
</tr>
<tr>
<td>Resources on data accessibility</td>
<td>75</td>
</tr>
<tr>
<td>Case studies of data shared back to participants</td>
<td>75</td>
</tr>
<tr>
<td>Resources on participant rights</td>
<td>70</td>
</tr>
<tr>
<td>Bibliography/Reading material</td>
<td>69</td>
</tr>
<tr>
<td>Resources on de-identification</td>
<td>66</td>
</tr>
</tbody>
</table>
Examples of Useful Resources for Diversity Scholars

On this page you will find resources that may be useful to diversity scholars whether or not you are working with data. The resources are listed in no specific order. Additionally, you are welcome to complete the example checklist that is provided.

<table>
<thead>
<tr>
<th>Example</th>
<th>Description</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consent form language/templates</td>
<td>Consent form language/templates for sharing (raw data, de-identified data, sharing with other researchers, publicly, etc.)</td>
<td>See for example</td>
</tr>
<tr>
<td>Case Studies</td>
<td>Case studies of underlying data shared back to participants in useful ways.</td>
<td>--</td>
</tr>
<tr>
<td>Checklist of Questions</td>
<td>A checklist of questions for making data decisions, including how far to open data, and to whom.</td>
<td>See for example</td>
</tr>
</tbody>
</table>

- **Box:** Secure cloud-based file storage. Retiring—will no longer available as of December 1, 2021.
- **Canvas:** Learning Management System
- **Deep Blue Data:** Repository for publicly sharing & preserving completed research data
- **Desktop Backup:** Automatic backup for Windows and Macs
- **Dropbox:** Cloud file storage with automatic back up, readily accessible from any device
- **Data Den:** Archive storage for large files
  - For research only
  - Large files only
Applications/Implications

- Context change in the middle of project timeline (COVID-19, Movement for Black Lives)
- Resources can help frame and model just data practices for future and current research inside and outside of academic settings
- Important to consider specific needs of data from vulnerable populations
Research Data Life Cycle

Text description of image: The data lifecycle is depicted as eight stages in a circle leading into each other. The first two stages, which take place before starting a research project, are “Finding data (for secondary research)” then “Data planning.” The next three stages take place during the project: “Data collection,” “Data processing/analysis,” then “Active data management.” The final three stages are undertaken after the project is completed: “Data curation,” “Data sharing,” then “Data archiving/preservation.” This final stage may lead back into finding data.
Proposed resources

Chicago Beyond’s *Why am I always being researched?*
- Use during data planning
- Focuses on approaching data ownership, use, and co-creation

Open Data Institute’s *Data Ethics Canvas*
- Use to prepare for data sharing
- Focuses on context and intention behind data collection and use for new audiences
Planned dissemination

- Sharing research, toolkit
- Talking about how best to reach intended audience
- Feedback and closing the loop
Further information

- **Project site:**
- **LYRASIS White paper:** http://dx.doi.org/10.7302/10
- Forthcoming: toolkit, publications, data, bibliography
- **Contact us:**
  - woodbr@umich.edu / DEIA.Data.Toolkit@umich.edu
Thank you!
## Important data lifecycle stages: comfortable vs. uncomfortable

<table>
<thead>
<tr>
<th>Stage</th>
<th>No. Comfortable</th>
<th>No. Uncomfortable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data planning</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Finding existing data</td>
<td>29</td>
<td>4</td>
</tr>
<tr>
<td>Data collection</td>
<td>53</td>
<td>5</td>
</tr>
<tr>
<td>Active data management</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Data analysis/ processing</td>
<td>33</td>
<td>18</td>
</tr>
<tr>
<td>Data curation</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>Data sharing</td>
<td>4</td>
<td>32</td>
</tr>
<tr>
<td>Data archiving/ preservation</td>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>NA</td>
<td>1</td>
<td>15</td>
</tr>
</tbody>
</table>