Is it worth it? Weighing the benefits of research data management infrastructure for the library

FREE 60-minute webinar, April 20 - 14:00 ET

Amy Gay, Assistant Head of Digital Initiatives, Binghamton University Libraries

Mark Hahnel, Founder and CEO, Figshare

Jonathan Petters, Assistant Director of Data Management and Curation Services, Virginia Tech
Is it worth it?
Weighing the benefits of research data management infrastructure for the library

Mark Hahnel
CEO - Figshare
what is figshare?

Figshare is a leading repository platform for storing, accessing, and citing research outputs including data, papers, preprints, theses, teaching materials, conference outputs, and more…
Track your outputs’ reuse, engagement, and impact

- Turn-key SaaS
- Researcher or administrator-led deposit workflows
- ISO 27001 certified
- WCAG 2.1 AA compliant
- Use Figshare-provided AWS storage or any compatible storage provider
- Integrate with your institution’s SSO for quick and easy access
- API and FTP uploader for programmatic uploads

- All items are assigned a digital object identifier (DOI) or handle
- Views, downloads, citations, and altmetrics are displayed with click-throughs to Dimensions and Altmetric for a deep-dive into the impact of your data
- Access to a statistics dashboard for your administrative staff
- Integrate with CRIS or RIMS systems like Symplectic Elements and Pure to measure the impact of your institution’s research outputs
Open Research is good for Equity and Transparency

1. **Accessibility**: Open Research makes research more accessible to everyone, regardless of their background, financial status, or location.

1. **Reproducibility**: Open Research encourages transparency and reproducibility in research. This means that scientific findings can be verified and replicated by other researchers.

1. **Community engagement**: Open Research can promote engagement with communities outside of academia. This can help to ensure that research is more relevant and responsive to the needs of society.
What do we mean in by ‘Research data management infrastructure’?

- software,
- services,
- people,
- hardware the library may be responsible for
Who Researchers would be willing to receive support from to help in reviewing, curating and preparing their data for public release

41% relied upon Publishers

38% upon their own Institution (including research offices, peers and librarians)
<table>
<thead>
<tr>
<th>Concern</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concerns about misuse of data</td>
<td>36%</td>
<td>37%</td>
<td>38%</td>
<td>43%</td>
</tr>
<tr>
<td>Not receiving appropriate credit or acknowledgement</td>
<td>33%</td>
<td>32%</td>
<td>31%</td>
<td>39%</td>
</tr>
<tr>
<td>Unsure about copyright and data licensing</td>
<td>35%</td>
<td>33%</td>
<td>30%</td>
<td>35%</td>
</tr>
<tr>
<td>Contains sensitive information or requires consent</td>
<td>29%</td>
<td>22%</td>
<td>31%</td>
<td>30%</td>
</tr>
<tr>
<td>I am unsure I have permission from my funder or institute</td>
<td>32%</td>
<td>26%</td>
<td>26%</td>
<td>25%</td>
</tr>
<tr>
<td>Organising data in a presentable and usable way</td>
<td>32%</td>
<td>25%</td>
<td>23%</td>
<td>26%</td>
</tr>
<tr>
<td>Costs of sharing data</td>
<td>19%</td>
<td>26%</td>
<td>24%</td>
<td>26%</td>
</tr>
<tr>
<td>Lack of time to deposit data</td>
<td>21%</td>
<td>24%</td>
<td>20%</td>
<td>21%</td>
</tr>
</tbody>
</table>
NIH issues a seismic mandate: share data publicly

The data-sharing policy could set a global standard for biomedical research, scientists say, but they have questions about logistics and equity.
52 funders listed on Sherpa Juliet require data archiving as a condition of funding, while a further 34 encourage it.

More than two-thirds of respondents of the State of Open Data 2022 are supportive “to some extent” of a national mandate for making research data openly available. This number has been declining since 2019.
<table>
<thead>
<tr>
<th>funder_country</th>
<th>funder_type</th>
<th>funder_name</th>
<th>author_contribution</th>
<th>code_availability</th>
<th>competing_interest</th>
<th>data_availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singapore</td>
<td>Government</td>
<td>Agency for Science, Technology</td>
<td>61</td>
<td>4</td>
<td>82</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ministry of Education (Singapore)</td>
<td>64</td>
<td>7</td>
<td>82</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>National Medical Research Institute</td>
<td>60</td>
<td>4</td>
<td>80</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>National Research Foundation</td>
<td>63</td>
<td>5</td>
<td>86</td>
<td>62</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>trust_markers / pct</th>
</tr>
</thead>
<tbody>
<tr>
<td>funder_country</td>
</tr>
<tr>
<td>Singapore</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Binghamton University Libraries + Research Data Management: Chapter One

Amy Gay

Assistant Head of Digital Initiatives for Digital Scholarship

Binghamton University Libraries
About Binghamton University

- State University of New York (SUNY) public school
- R1 research institution (since 2018 / 2019)
- Made up of 7 different colleges and schools (and growing)
- About 15,000 undergraduate students and about 4,000 graduate students
- Plans to hire more research faculty over the next five years
Why the libraries are talking about RDM

- The shifts to more open data management and sharing policies by funders
- SUNY-wide task forces and discussion groups
- Already offer workshops related to RDM practices
- Continued increase in consultation requests
What we currently do

- Assist with drafting data management plans
- Offer consultations
- Organize and lead workshops
- Provide joint workshops and information sessions with the Division of Research
- Manage an institutional repository
What we envision as the libraries role going forward

- **Educators**
  - Gathering, organizing, cleaning, visualizing, analyzing, preserving and sharing data

- **Consultants**
  - Share our expertise on the above and bring in our metadata experts

- **Guides**
  - Point researchers to the people in the libraries and outside the libraries for different stages of support
  - And tools!

- **Content sharers**
  - Promotion of research output
  - Repository recommendations
Benefits / challenges of the libraries being involved

Benefits

● Being in the know of where people are sharing their research
● Provide guidance related to our expertise
● Build stronger relationships with other research service points
● Be more connected with research happening at Binghamton and across SUNYs
Benefits / challenges of the libraries being involved

Challenges

- Current bandwidth of people involved
- Variety of needs and research output
- Staying current on where people are sharing their research
- Being overwhelmed by the constant change
- Recommendations of “best” places to share research
- Funding!
Who we hope to collaborate with

In the Libraries:
- Digital scholarship team
- Our metadata experts
- Subject librarians

Institutionally:
- Division of Research
- Information Technology Services
- The Graduate School
- Human Resources

Across SUNY:
- Other librarians and research support teams -- build a community of practice
What’s to come in the year ahead

- Campus environmental scan
- More collaboration with the Division of Research
- Create learning materials and additional resources for finding and sharing research content
- Help with building research profiles
- Expand collaboration with other librarians across SUNY
Virginia Tech Data Services

Services designed to “boost the value and impact of research data created by researchers at Virginia Tech. Our expertise creates efficiencies for researchers across the research lifecycle”
Virginia Tech Data Services

We help with:
- Data Management
- Planning and Access
- Metadata Schema
- Informatics (e.g. Analysis, Cleaning, Visualization)
- Data Curation
- Publishing Research Data

Staff of 15
Virginia Tech Data Repository

To highlight, preserve, and provide access to research products (e.g. datasets) of the Virginia Tech community, and in doing so help to disseminate and make usable the intellectual output of the university in its land-grant mission

2016 to 2021 - Samvera platform
2021 to present - figshare for institutions platform
Barriers to Service Provision

• These are new services for Libraries
  ○ OUTREACH - The library: it’s not just for books!
  ○ Needs assessments (geospatial data survey)
Barriers to Service Provision

• These are new services for Libraries
  ○ OUTREACH - The library: it’s not just for books!
  ○ Needs assessments (geospatial data survey)

• Buying vs. Building repository infrastructure
  ○ What support is/will be available locally?
    Externally? (our repository choice process)
  ○ Data preservation needs <-> credibility
Future plans

On campus

• Outreach around campus, now that services are more mature
Future plans

On campus

• Outreach around campus, now that services are more mature
• Coordination of campus research data services
Future plans

On campus

• Outreach around campus, now that services are more mature
• Coordination of campus research data services
• Digital collection coordination in Libraries (?)
Future plans

On campus
- Outreach around campus, now that services are more mature
- Coordination of campus research data services
- Digital collection coordination in Libraries (?)

Coordinating efforts w/others (e.g. institutions, professional societies, businesses)