Developing Data Management Services

What Support do Researchers Need?

Christine Kollen, Data Curation Librarian,
Office of Digital Innovation & Stewardship

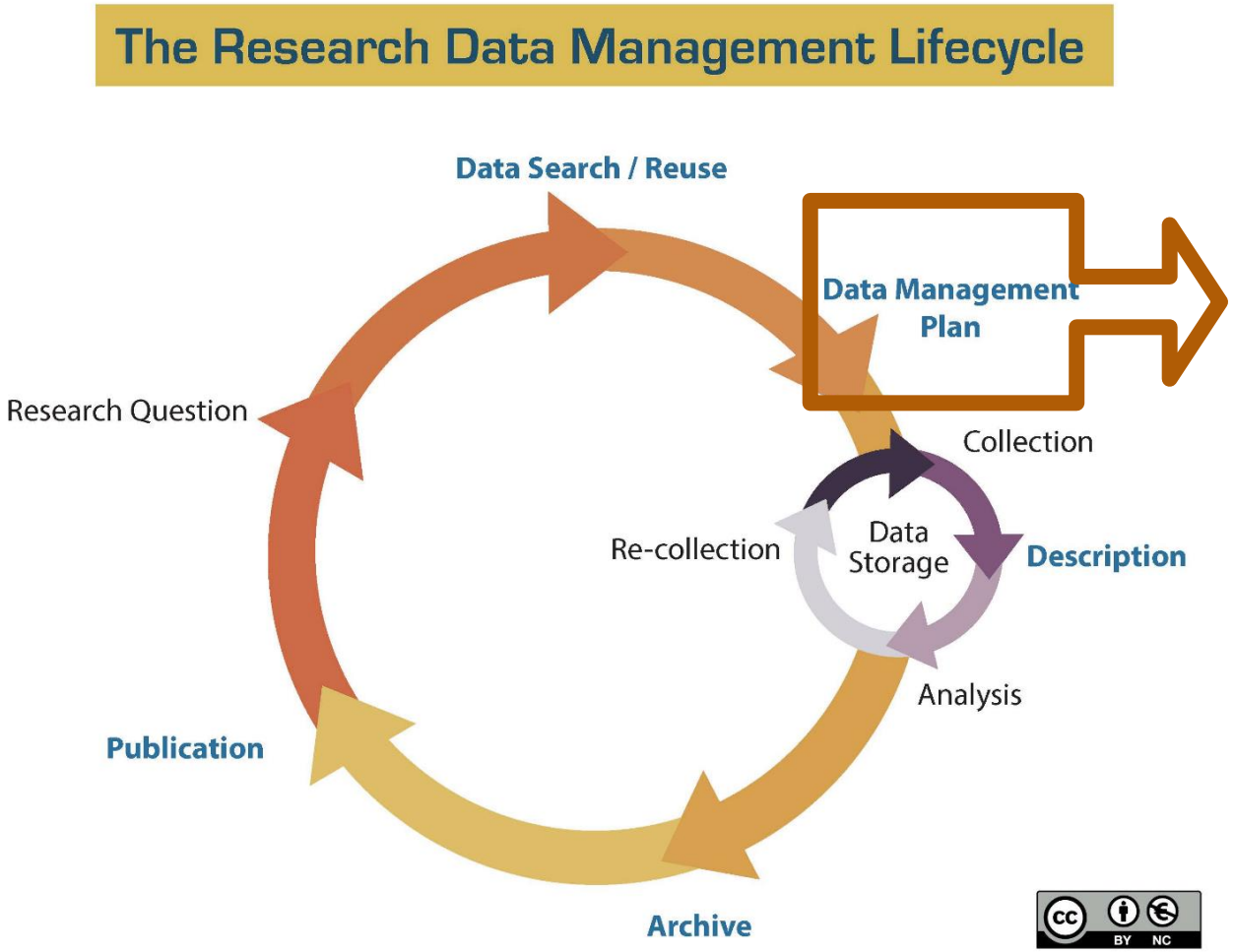
Providing Access to Research Results

- To increase visibility, usage, and impact – and remove barriers – important to provide open access to research results, publications and data
- In 1999, Office of Management & Budget (OMB) amended Circular A-110 to require research data produced with federal funding be made publicly available
- In 2003, NIH required Data Sharing plans (some grant proposals)
- In 2011, NSF required data management plan

Providing Access to Research Results (cont.)

- In 2013, Office of Science and Technology Policy released memo directing federal agencies to develop plans for making research results freely available (publications and data)
- Applied to 19 federal agencies
- Most agencies have come out with new policies to address the memo

Data Life Cycle



Documents the lifecycle of data and provides details on data collection for storage, access, sharing, and reproducibility.

This can ensure the availability and accessibility of research results after the project is complete.

Enhances the value of research, increases the visibility of the institution, encourages collaboration.

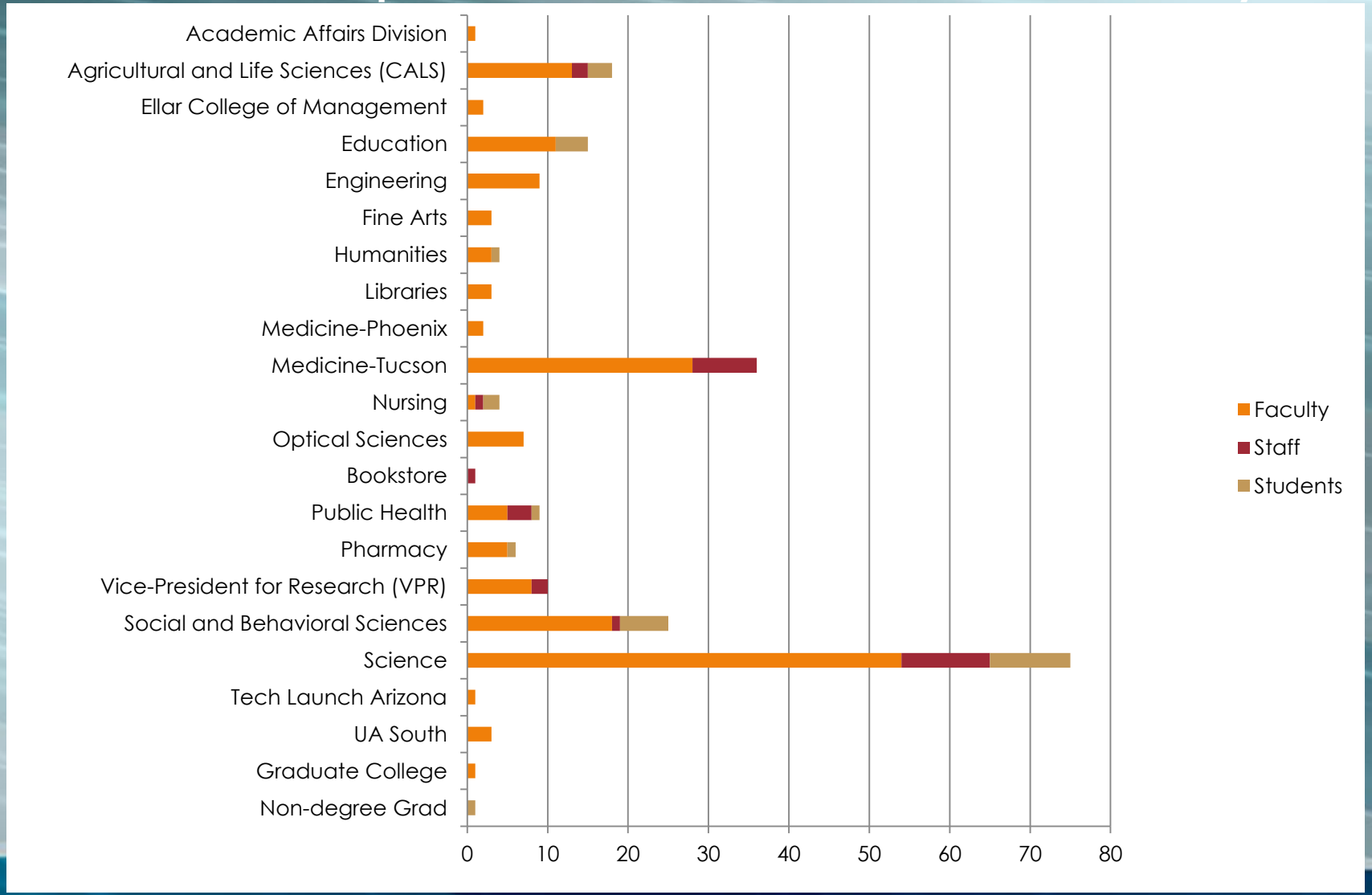
Current Data Management Services

- Data Management Resources website
- Workshops on data management and curation
- Consult on data management plans and metadata
- Customized DMP Tool
- Researchers and graduate students can make research data available in the Campus Repository and geospatial data available in the Spatial Data Explorer
- Other options for sharing data include disciplinary repositories

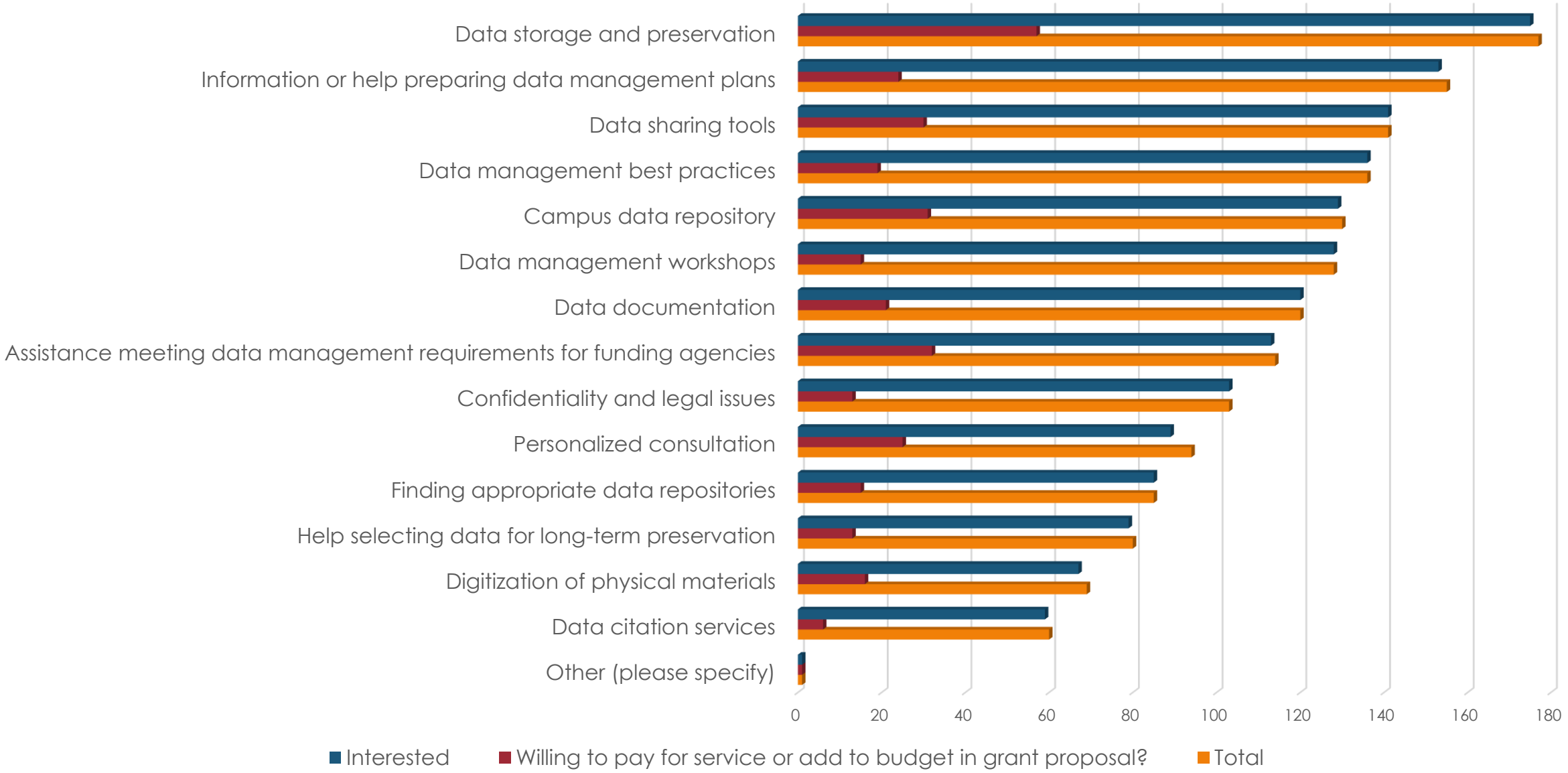
Research Data Management Survey

- Survey administered during spring semester 2014 by Libraries and RCGC Campus Data Management Subcommittee
- Purpose -- discover research data management and curation needs of researchers
  - How is research data managed, stored, shared and reused across campus?
  - What are researcher's attitude toward data creation, sharing and preservation?
  - What is the demand for existing services and what new services do researchers need?

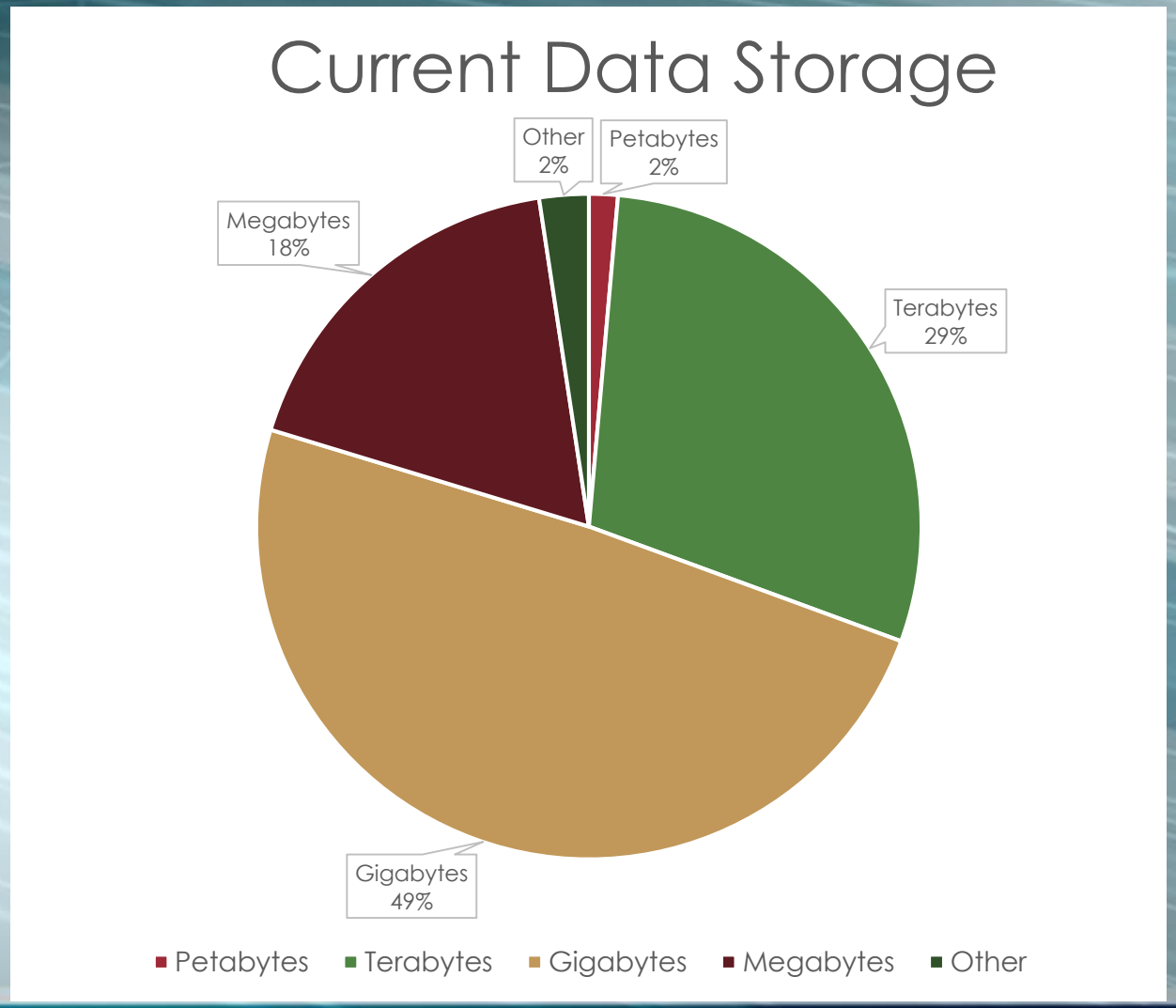
Who Responded to the Survey?

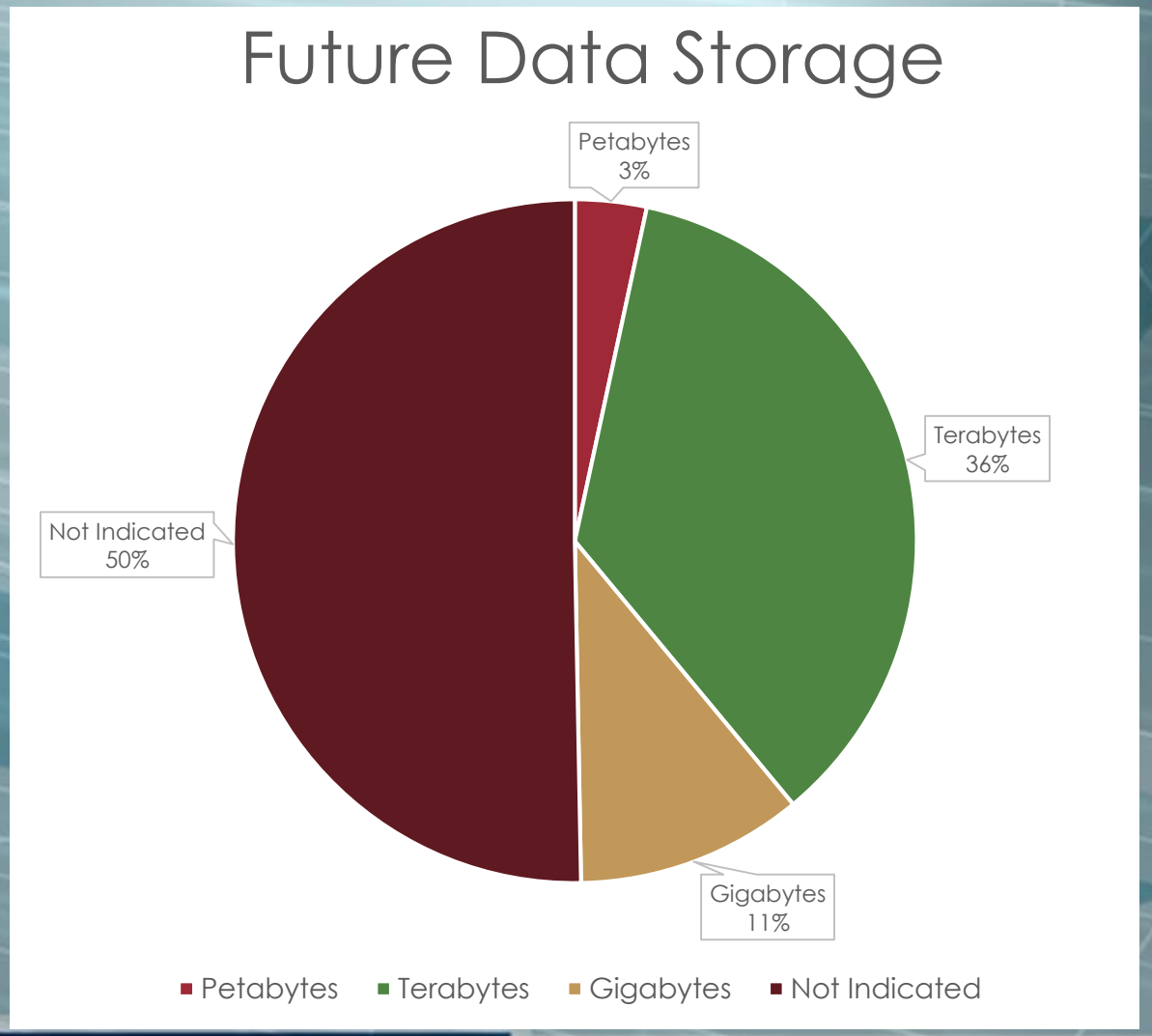


What Data Management Services do Researchers Need?

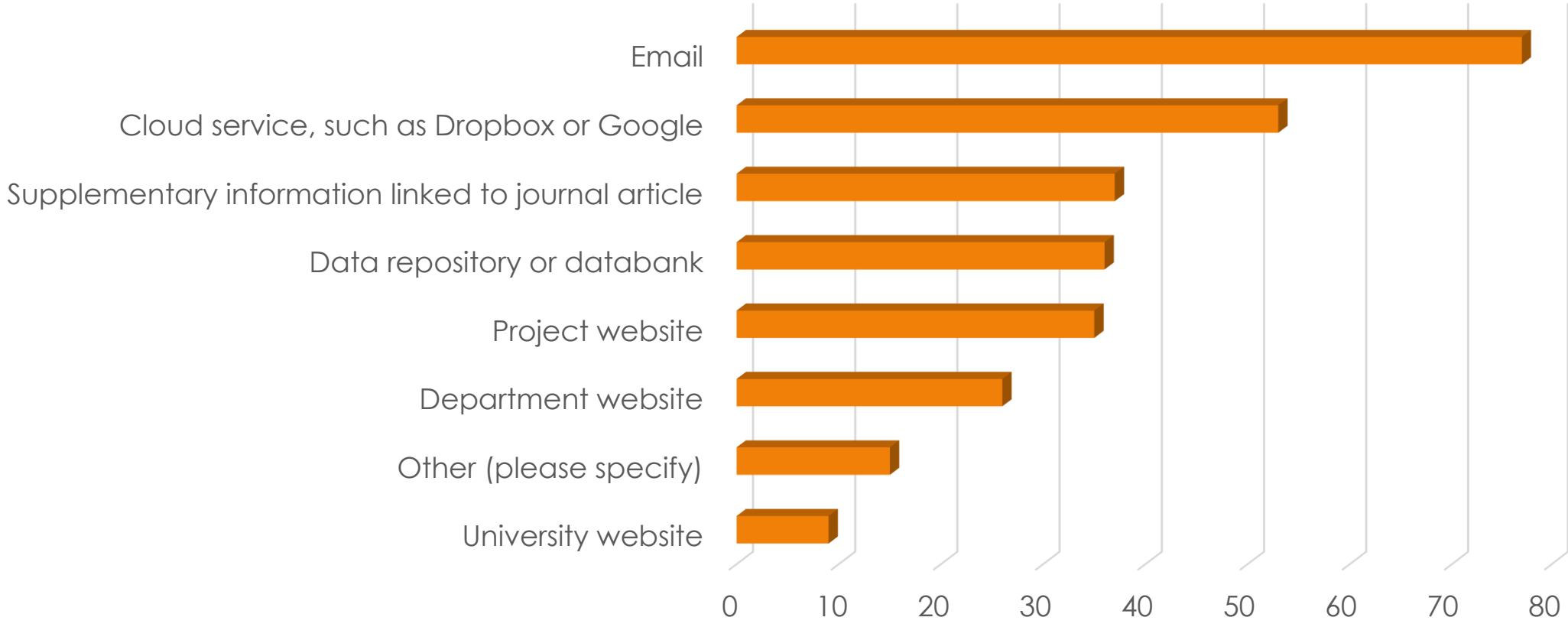


Researcher's Current and Future Data Storage Needs?

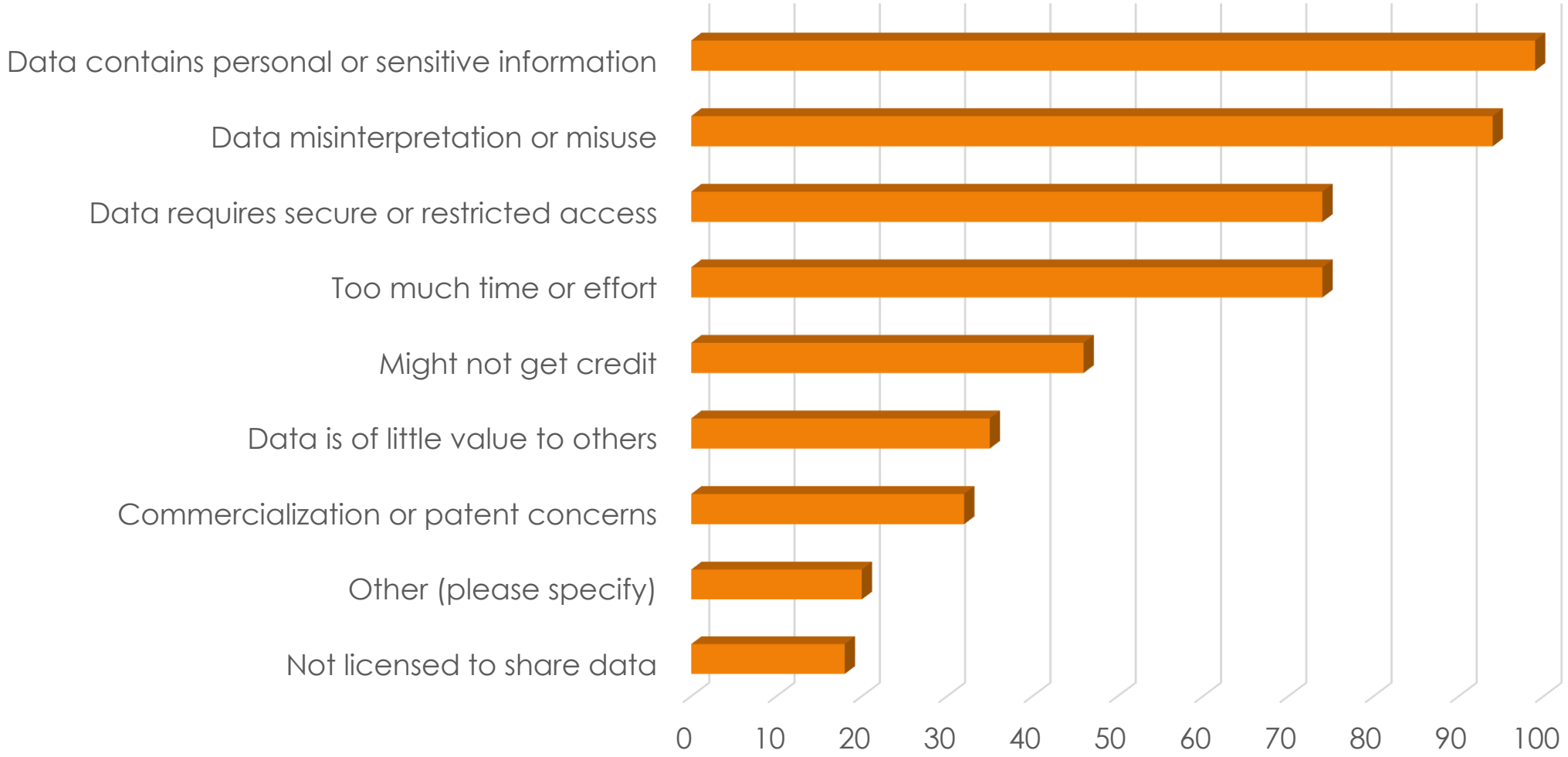




Methods of Sharing Data with Others



Reasons for Not Sharing Data



Recommendations - Data Storage

box.com implemented for campus

- Share files with collaborators
- Anticipate usage during active stage
- Will help address some current data storage needs
- Does not provide preservation or curation services, not setup for researchers to share data beyond immediate collaborators
- Use for confidential information (e.g. FERPA), but not for regulated information (e.g. HIPAA, ITAR)

Recommendations – Data Repository

- UA Campus Repository – facilitates access to research, creative works, publications, and teaching materials of the UA
  - Can handle data files, although not an ideal solution
  - There is a limit on file size
- Disciplinary repositories
  - Some funding agencies require that data is deposited in specific data repository
  - Researchers may prefer to make data available in a disciplinary repository
    - Do not cover all disciplines, some have limitations, what are the repository's long-term viability
    - Some UA researchers want their data identified as from UA

Data Management and Data Curation (DMDC) Pilot

- Work with 3-5 research projects through the data lifecycle to evaluate data management services
- Pilot will:
  - Develop data management services needed by each project
  - Evaluate – what worked well, what didn't work well
  - Training needs
  - Feasibility of implementing DM services for campus

DMDC Pilot (continued)

- Funding – proposal requested funding for servers, staff support, configure iRODS; hire data management staff
- Developed application process and opened to campus through the Office of Research and Discovery

DMDC Initial Results

- Six research projects selected last spring
- Are providing the following services and recommendations:
  - Help map out data workflow
  - Assist with data documentation and implementing best practices
  - Implement iRODS for data storage
  - Provide referrals - campus workshops or to another unit
  - Recommendations:
    - where to publish research
    - metadata standards to use
    - archiving older data

DMDC Participant Issues

- Several of the projects have privacy and security concerns – can some of their data be made publicly available? How to de-identify their data?
- What data should be archived – not necessarily all data collected
- How to deal with older data in proprietary formats?

Questions?

- Contact – Chris Kollen kollen@email.arizona.edu or 520-305-0495