INSITUTIONAL SURVEY

PLANNING FOR THE LIFECYCLE MANAGEMENT AND LONG-TERM PRESERVATION OF RESEARCH DATA: A FEDERATED APPROACH

Phase 1, Environmental Scan
June 18, 2013
IMLS 51-12-0695

UNIVERSITY OF KANSAS LIBRARIES & INFORMATION TECHNOLOGY
GREAT PLAINS NETWORK
GREATER WESTERN LIBRARY ALLIANCE
INSTITUTE OF MUSEUM AND LIBRARY SERVICES
# TABLE OF CONTENTS

**SURVEY QUESTIONS** ................................................................................................................................. 3

**RESULTS OF SURVEY OF GWLA & GPN INSTITUTIONS (PRELIMINARY AS OF 5/15/2013)** ........ 14

- **OVERVIEW OF SURVEY** ....................................................................................................................... 14
- **DETAILED REVIEW OF DATA** ................................................................................................................ 15
  - Response Rate ......................................................................................................................................... 15
  - Organizational Unit/Department Responding .......................................................................................... 15
- **SERVICES PROVIDED** ............................................................................................................................. 16
  - General Support and Services for Research Data .................................................................................. 16
  - Storage, Archiving, Preservation, and Sharing of Data .......................................................................... 17
  - Accessing and Using Research Data ........................................................................................................ 18
  - High Performance Computing ................................................................................................................ 20
  - Types of Digital and Non-digital Research Data that are Supported to Manage and Preserve Long Term Institutional Access .................................................................................................................. 22
- **RESEARCH DATA POLICIES** .................................................................................................................. 25
SURVEY QUESTIONS

IMLS - KU GWLA GPN Survey
Lifecycle Management of Research Data

This study surveys member institutions of the Great Plains Network and the Greater Western Library Alliance about services related to the management of research data. We are sending this survey to Deans and Directors of Libraries, to Chief Research Officers, and to Chief Information Officers. We welcome three individual responses from each member institution. Alternatively, one representative may be designated to respond on behalf of the institution.

The management of research data by institutions is an issue of growing interest and complexity for universities. Large-scale research instruments extend our observational power by many orders of magnitude but at the same time generate massive amounts of data. Researchers work feverishly to document and preserve changing or disappearing habitats, cultures, languages, and artifacts resulting in volumes of media in various formats. New software tools mine a growing universe of historical and modern texts and connect the dots in our semantic environment.

Research data management includes services that extend throughout the lifecycle of data such as organizing data; providing metadata; making data accessible for discovery and sharing; storing, archiving, and/or preserving data for long term retention and use; creation of data management plans; development of institutional policies for the appropriate management of data; training in best practices for managing data; and developing security or privacy provisions.

The purpose of the study is to assess what services GWLA and GPN institutions provide to help researchers manage research data and, within each institution, to discover which units or divisions are responsible for providing specific resources and services. We will use this information to inform an advisory council of university leaders from among the GWLA and GPN universities as they consider opportunities to leverage collective strengths and institute scalable and shared approaches to the management of research data.

This study is conducted and funded through a grant from the Institute for Museum and Library Studies and conducted on behalf of members of the Greater Western Library Alliance (GWLA), a consortium of 33 research libraries and the Great Plains Network, (GPN) a consortium of 20+ universities partnering to facilitate the use of advanced cyberinfrastructure. The University of Kansas serves as the home institution for the IMLS grant. Our project web site can be found at: http://imls.gwla.org.

This survey will take approximately 20 – 30 minutes of your time. We thank you in advance for helping us conduct this research by providing data from your institution. If you have any questions about this survey, please contact the principle investigator Deborah Ludwig (dludwig@ku.edu) or investigator Scott McEathron, (macmap68.ku.edu).

If you have any questions or concerns regarding this survey, please email dludwig@ku.edu.
IMLS - KU GWLA GPN Survey
Lifecycle Management of Research Data

If you have any questions or concerns regarding this survey, please email dludwig@ku.edu.

University Name

Person completing the survey (name and position)

Organizational Unit / Department

Email address of person completing the survey

For which are you filling out the survey?

- I am responding on behalf of my organizational unit
- I am responding on behalf of my entire institution

If you have any questions or concerns regarding this survey, please email dludwig@ku.edu.
Some universities provide services to support the use and management of research data. Please indicate by checking any or all boxes that apply what research data services are available at your institution and who is providing that service.

### General Support and Services for Research Data

<table>
<thead>
<tr>
<th>Service Description</th>
<th>University Library</th>
<th>Central IT</th>
<th>University Office of Research</th>
<th>Department or Research Center</th>
<th>Other</th>
<th>Not Offered</th>
<th>Don't Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.01) Provides assistance to researchers in developing data management plans for research proposals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.02) Offers formal training (e.g. workshops or courses) for researchers on data management practices</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.03) Consults with researchers on data management best practices (data archiving, data sharing, metadata, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.04) Dedicates funding resources (staffing, facilities, and services) that support long-term management of research data generated by campus researchers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.05) Ensures university compliance for research data in accordance with commercial licenses, government regulations, and funding agency mandates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.06) Consults with researchers on options for data licensing agreements for open or restricted access</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If you have any questions or concerns regarding this survey, please email dludwig@ku.edu.
### Storage, Archiving, Preservation, and Sharing of Data

<table>
<thead>
<tr>
<th>2.07</th>
<th>Helps researchers decide which data are important to preserve for long-term access.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.08</td>
<td>Advises researchers on data and or metadata standards for research data or datasets</td>
</tr>
<tr>
<td>2.09</td>
<td>Prepares metadata for researchers to enhance the discovery of their research data.</td>
</tr>
<tr>
<td>2.10</td>
<td>Provides standards and methods for de-identifying sensitive data</td>
</tr>
<tr>
<td>2.11</td>
<td>Provides short term networked data storage for researchers (5 years or less)</td>
</tr>
<tr>
<td>2.12</td>
<td>Provides long term networked data storage for researchers (more than 5 years)</td>
</tr>
<tr>
<td>2.13</td>
<td>Provides a repository on site to store metadata and data together</td>
</tr>
<tr>
<td>2.14</td>
<td>Provides a repository for sharing data with appropriate access controls</td>
</tr>
<tr>
<td>2.15</td>
<td>Provides assistance with identifying national or international data repositories for archiving research data</td>
</tr>
<tr>
<td>2.16</td>
<td>Provides guidance to researchers on offsite repositories for data and metadata deposit</td>
</tr>
<tr>
<td>2.17</td>
<td>Provides ongoing support for discovery, citation, and usability of data over the long term</td>
</tr>
<tr>
<td>2.18</td>
<td>Helps prepare data / data sets for deposit into a repository</td>
</tr>
</tbody>
</table>

If you have any questions or concerns regarding this survey, please email dludwig@ku.edu.
Some universities provide services to support the use and management of research data. Please indicate by checking any or all boxes that apply what research data services are available at your institution and who is providing that service.

### Accessing and Using Research Data

<table>
<thead>
<tr>
<th></th>
<th>University Library</th>
<th>Central IT</th>
<th>University Office of Research</th>
<th>Department or Research Center</th>
<th>Other</th>
<th>Not Offered</th>
<th>Don’t Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.19) Helps researchers locate sources of data</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2.20) Supports the assignment of persistent identifiers to research data sets. [Digital Object Identifiers, or DOIs are one example.]</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2.21) Supports the assignment of persistent identifiers for researchers. [ORCID ID’s are one example]</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2.22) Supports linking research data to research publications based on the research data</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2.23) Provides support for the analysis of data</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2.24) Provides support for visualization of data (for example, simulations, geographic information systems or GIS, or statistical visualization)</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

If you have any questions or concerns regarding this survey, please email dludwig@ku.edu.
Some universities provide services to support the use and management of research data. Please indicate by checking any or all boxes that apply what research data services are available at your institution and who is providing that service.

### High Performance Computing

<table>
<thead>
<tr>
<th>Service Description</th>
<th>University Library</th>
<th>Central IT</th>
<th>University Office of Research</th>
<th>Department or Research Center</th>
<th>Other</th>
<th>Not Offered</th>
<th>Don't Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.25) Provides a secure data facility for research data with access controls, backup and restore facilities meeting regulatory standards such as HIPAA</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2.26) Provides computing facilities for computationally intensive research, e.g. massively parallel technology</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2.27) Provides computing facilities for in-place analysis of extremely large research datasets</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2.28) Provides assistance in securing advanced or high performance computing resources located off-campus.</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

If you indicated above that "Other" entities provided one or more of these services, please list those other entities here.

---

If you have any questions or concerns regarding this survey, please email dludwig@ku.edu.
A helpful model for thinking about the different kinds of information and services needed for managing data is the Johns Hopkins University stack model for Data management described by Sayeed Choudhury. In this hierarchical model, layers depend on the layers listed below them. Preservation includes ensuring that there is enough information about the creation of the data that the data can be interpreted and reused without communication with its creators. Curation activities may include developing enhanced ability to find elements of the data or link data to other collections.

Data Management Layers

<table>
<thead>
<tr>
<th>Layers</th>
<th>Characteristics</th>
<th>Sample Actions</th>
<th>Researcher Implications</th>
<th>NSF Implications</th>
</tr>
</thead>
</table>
| Curation | Active and ongoing management of data through its lifecycle of interest and usefulness | • Provide ongoing bibliographic control for data.  
• Link research data to publications based on the data.  
• Provide tools for further analysis.  
• Harvest metadata for the data to share with external search engines. | • Feature Extraction  
• New query capabilities  
• Cross-disciplinary accessibility | • Offers competitive advantage  
• New opportunities for data use |
| Preservation | Ensures that archived data can be fully used and interpreted over time | • Add information to maintain the viability, render-ability, and understandability of data long term.  
• Monitor format obsolescence; migrate data to new digital formats as need.  
• Preserve tools and/or documentation for using the data. | • Ability to use own data in the future (e.g. 5 years)  
• Data sharing with others | • Satisfies NSF needs across directorates |
| Archiving | Data protection is applied to stored data, including fixity checking, and assignment of data identifiers | • Check for viruses as data is deposited  
• Establish checksum snapshots over time to ensure data has not change.  
• Assign a persistent identifier such as a DOI or handle.  
• Link metadata to data | • Data is better protected.  
• Provides persistent identifiers for locating, sharing, and referencing data. | • Could satisfy most NSF requirements |
| Storage | Bits on disk, tape, cloud etc. Backup and restore. | • Place data in networked storage  
• Invoke backups.  
• Set access protocols. | Responsibilities for  
• Restore  
• Sharing  
• Staffing | • Could be enough for now, but not near-term future |

Adapted from the "stack model" of Data Management Layers under development by John Hopkins University and based on the definition of data curation advanced by the University of Illinois Graduate School of Library and Information Science. See: http://www.clir.org/initiatives-partnerships/data-curation

If you have any questions or concerns regarding this survey, please email dludwig@ku.edu.
For which of the following types of digital and non-digital research data are support services in place to manage and preserve them for long-term access on your campus? (please check all levels of the stack model that apply)

<table>
<thead>
<tr>
<th>4.02) Digital texts or digital copies of texts and manuscripts</th>
<th>Curation</th>
<th>Preservation</th>
<th>Archiving</th>
<th>Storage</th>
<th>None</th>
<th>Don’t know</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.05) Digital images or digital copies of images</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.07) Digital audio recordings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.09) Digital video recordings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.10) Spreadsheets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.11) Digital Databases [or Data Sets] (e.g. surveys, census data, government statistics, etc)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.12) Computer code</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.13) Hardware or research equipment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.15) Spatial data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.16) Digital gene sequences or similar digital renderings of biological/organic/inorganic samples or specimens</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.17) Artistic products</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.18) Other (Please Specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Please describe the other type of data selected above:

---

If you have any questions or concerns regarding this survey, please email dludwig@ku.edu.
Policies For Research Data Management

Does your university have a general policy on ownership of research data?

- Yes
- No
- Don’t Know

If yes, please provide a reference to that policy.

Has your university established a policy for funding to cover the costs of research data management and storage

….. for externally funded research?

- Costs are primarily expected to be folded into the direct costs of those grants and contracts
- Costs are primarily paid for by the researcher’s department or center
- Costs are primarily paid for by the University from research overhead (F&A funds) or other sources
- Other (Please Specify Below)

- The University has no policy
- Don’t know

... for research NOT supported by grants, contracts, or other external sources of funding?

- Costs are primarily paid for by the individual researcher or research team
- Costs are primarily paid for by the researcher's department or center
- Costs are primarily paid for by the University from overhead (F&A funds) or other sources
- Other (Please Specify Below)

- The University has no policy
- Don’t know
... once external funding has expired?

- Costs are primarily paid for by the individual researcher or research team
- Costs are primarily paid for by the researcher’s department or center
- Costs are primarily paid for by the University from overhead (F&A funds) or other sources
- Other (Please Specify Below)
- The University has no policy
- Don’t know

What are the main institutional challenges for working with research data that your organization has identified?

[Blank field]

What services and/or future plans has your organization developed to meet the challenges listed above?

[Blank field]

If you have any questions or concerns regarding this survey, please email dludwig@ku.edu.
IMLS - KU GWLA GPN Survey
Lifecycle Management of Research Data

We thank you for your time spent taking this survey.
Your response has been recorded.

If you have any questions or concerns regarding this survey, please email dludwig@ku.edu.
RESULTS OF SURVEY OF GWLA & GPN INSTITUTIONS  
(PRELIMINARY AS OF 5/15/2013)

Overview of Survey
Amalia Monroe-Gulick and Deborah Ludwig

The survey questions are found in Appendix G and were developed by Research Team B members, named in Appendix A. The survey was sent to contacts representing the highest level of administration for libraries, information technology, and offices of research.

18 unique institutions out of 47 surveyed provided responses from at least one contact. The response rate was highest from libraries, with 13 of the 23 responses. Responding institutions all offers some level of data management services.

Given a fairly low response rate with responses skewed toward respondents from libraries, it is not possible to gain a comprehensive picture of research data management services in GWLA and GPN institutions at this time; however there are some generalizations we can make from these results and additional comparative information may be gleaned from looking at other national studies such as the recently-conducted SPEC Kit Survey by IMLS.

In general, all institutions responded that they provide some level of data management services. Information Technology organizations seems more likely to provide short term or long term storage for research data and to be involved in high performance computing initiatives. Libraries seem more likely to be involved in helping researchers develop data management plans, locate repositories for data deposit, or help with deposit data in an institutional repository. Research centers are most likely to provide data analysis and visualization services. Formal education and training services in research data management are not common.

With respect to data services for storage, archiving, preservation, and sharing, the most commonly offered services were storage and help identifying repositories where researchers could deposit data. The services most frequently not offered are local institutional repositories. When repositories are provided, the libraries are the most common provider of that service.
Detailed Review of Data

Response Rate

<table>
<thead>
<tr>
<th>Total Number of Surveys Sent Out</th>
<th>134</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Number of Universities Contacted</td>
<td>47</td>
</tr>
<tr>
<td>Total Responses</td>
<td>23</td>
</tr>
<tr>
<td># of Total Universities Responding</td>
<td>18</td>
</tr>
<tr>
<td>University-Level Response Rate</td>
<td>38%</td>
</tr>
<tr>
<td>Total Survey Response Rate</td>
<td>18%</td>
</tr>
</tbody>
</table>

- Out of the 47 individual universities contacted, there was at least one respondent from 18 (38%) of the universities.

- Three individuals were contacted at each university (libraries, IT, research administration), with a total of 134 surveys distributed. The resulting overall individual response rate was 18%.

- One university returned surveys from all of the departments/units contacted and two universities returned surveys from two of three departments/units contacted.

- Respondents also indicated if they were completing the survey on behalf of their organizational unit or on behalf of their institution. The majority of respondents indicated they completed the survey for their organizational unit (74%).

Organizational Unit/Department Responding

<table>
<thead>
<tr>
<th>Organizational Unit/Department</th>
<th># Responders</th>
<th>% of Responders</th>
</tr>
</thead>
<tbody>
<tr>
<td>Library/Library Unit</td>
<td>13</td>
<td>57%</td>
</tr>
<tr>
<td>Information Technology/IT Department</td>
<td>2</td>
<td>9%</td>
</tr>
<tr>
<td>Research Studies</td>
<td>6</td>
<td>26%</td>
</tr>
<tr>
<td>Provost</td>
<td>1</td>
<td>4%</td>
</tr>
<tr>
<td>Digital Services</td>
<td>1</td>
<td>4%</td>
</tr>
</tbody>
</table>

- Over half of the surveys were completed by those affiliated with a university library or a library unit. The second largest group represented was research studies.

- The overall individual response rate was not high. When this is considered with the majority of respondents reporting information on their organizational units, this will limit the overall analysis of understanding data services at a university-wide level because not all organizational units may be aware of services, policies, and challenges associated with other university units. In addition, with over half of the respondents representing libraries, the results may also be skewed toward knowledge of library-based data services rather than campus-wide services.
Services Provided
In the service provision section of the survey, five main categories were addressed:

1. General Support and Services for Research Data
2. Storage, Archiving, Preservation, and Sharing of Data
3. Accessing and Using Research Data
4. High Performance Computing
5. Support services management, preservation, and access to digital and non-digital research data for long-term access to campuses

General Support and Services for Research Data
In this section, questions were asked about support and service the following areas:

1. Development of data management plans
2. Consultation on data management
3. Consultation on options for data licensing agreements for open or restricted access
4. Dedicating funding resources that support long-term management of research data
5. Formal training on for researchers on data management planning
6. Ensuring university compliance for research data in accordance with commercial licenses, government regulations, and funding agency mandates

Overall, general support and services for data management is present at the responding universities. The average number of positive responses for the entire category of multiple questions was 27.

Support for the development of data management plans received the largest number of responses, (38). Formal training for data management planning received the fewest responses (19). This result is interesting because of the federally mandated inclusion of data management or sharing plans for NSF and NIH grants. It appears that the responding universities are offering assistance with developing data management plans, but do not have formal programs in place. There is potential for further research in this area to identify the nature of data management consultation services and if they are decentralized and on an “ad hoc” basis.

The libraries had the highest frequency of responses regarding general data services, with an average of 11. The make-up of the survey respondents may have influenced this potential skewing towards the high rate for the library selection. However, the question on assistance with university compliance with commercial licenses, government regulations, and funding agency mandates received 20 overall responses, with offices of research receiving 13 of those 20 (54%). Offices of research were the second largest respondent group.

1 The following discussion uses data that excluded the “not offered” choice.
All questions had responses that indicated a service was not offered, with the two highest being data management best practices (6) and dedicated funding resources (6).

Storage, Archiving, Preservation, and Sharing of Data

In this section of the survey, respondents were asked about services related to storing, archiving, preserving and sharing data. The average positive response rate was lower than the general services section, 18 compared with 27, indicating that this is an area that the responding universities have not addressed as much as general services, which focus more on issues related to data management.
The library and central IT options were most frequently selected in this category of questions. The library received the largest number of responses with the questions related to identification of repositories, metadata standards, and long-term discoverability and usability of data.

According to the results, central IT provides short-term networked data storage (5 years or less) with 18 responses out of a total 29 positive responses (62%). Respondents also indicated that central IT provides long-term data storage, but a smaller frequency (55%). This result is interesting, in part, because there were only two respondents from IT organizational units, but respondents representing different organizational units recognized the role of central IT in both long and short-term data storage.

The options of offices of research or research centers were rarely selected in questions related to storage, archiving, preservation, and storage. The office of research option was not selected for metadata standards, long-term support for discovery, citation, and usability of data, or providing a repository for sharing data.

However, in all questions, respondents indicated that the specific service was not offered through their unit or institution. Out of 26 total responses to the question “provides a repository for sharing data with appropriate access controls,” there were 12 (46%) responses of “not offered.” Also of note is the high rate of “not offered” responses to the question “provides a repository on site to store metadata and data together.” Out of the 27 total responses to this question, there were 11 (41%) “not offered” responses. Low rates of indicating these services are present in a unit or institution to these questions could indicate that universities are not yet offering centralized data repositories, for either short or long term. This is not an unexpected result because of the complicated nature of the issue.

**Accessing and Using Research Data**

The section of the survey addressing accessing and using research data can be divided into two sections:

1. Support for locating and using/analyzing data

2. Support for linking and assigning unique identifiers to researchers and research data.

There was an evident gap between positive response rates for the question in the two sub-sections of the category. The average response rate (i.e. excluding “not offered” choice) was 21 for locating/analyzing data and 9 for linking/assigning unique identifiers. This is not a surprising result because supporting researchers with finding and using data has historically been the focus of data service programs at university libraries. However, linking research and research data to publications and assigning unique identifiers is an emerging area across universities, and university department, therefore, not yet represented by formal service programs.
The responding universities indicated that their libraries most frequently assist researchers with finding data, but campus and research centers most frequently offer data analysis and visualization support. However, “not offered” is the second most frequent response to data visualization support, demonstrating while services are being offered, this is still an area of potential growth for data support services at universities. Offices of research and central IT do not play a large role in this category, which is not necessarily surprising, and potentially indicates a trend of decentralized data services across campuses.
Within the sub-section of the category of accessing and using data, the question related to linking and assignment of unique identifiers to researchers and research data sets, there was low response rate from the survey respondents. Respondents did indicate that 8 libraries offer assistance with linking research data to publications and assigning persistent identifiers to research data sets. However, this was matched by the same number of respondents selecting “not offered,” indicating that this is a service not yet frequently offered. The trend within this survey could be indicating that libraries may be the emerging leaders with these types of support services, but significant conclusions cannot be reached from this sample.

High Performance Computing

Central IT was the dominant organizational unit for providing services related to high performance computing, which is not a surprising result because of the inherent role of IT in computing. A potential area for further research is the role of the university offices of research in high-performance computing. Except for providing secure data facility with access controls, backup and restore facilities meeting regulatory standards, offices of research had an average selection frequency of 7, compared with IT’s average frequency of 11. There were six open-ended responses to elaborate on an “other” selection. These indicated that some institutions use
off-campus services, including other universities high performance computing facilities and other external vendors or partners.

High Performance Computing Services
by Organizational Unit

- Provides a secure data facility with access controls, backup and restore facilities meeting regulatory standards
- Provides computing facilities for computationally intensive research
- Provides computing facilities for in-place analysis of extremely large research datasets
- Provides assistance in securing advanced or high performance computing resources located off-campus.

Legend:
- University Library
- Central IT
- University Office of Research
- Department or Research Center
- Other
- Not Offered
Types of Digital and Non-digital Research Data that are Supported to Manage and Preserve Long Term Institutional Access

Overall, there was high indication of services present at the responding institutions in the category of supporting long term institutional access to both digital and non-digital data. Storage was the most common support service, which is not unexpected since it is considered the most basic approach to providing access. However, archiving was the second most frequent response indicating that the responding universities are acknowledging and taking action on the necessary service of archiving data, which goes beyond preservation. However, curation and preservation, which take many more resources, were less frequently selected.

The long-term access to non-digital data had a much lower response rate than the digital data options. The area of most concern is the lack of support services for maintaining long term access to hardware and research equipment from the responding universities. With technology quickly changing and becoming obsolete, maintaining hardware may become an issue because of data that might potentially be unusable if it is not (or cannot) be put in a format for newer technologies. However, it is possible that the high response rate from library units potentially skewed this section because libraries have not traditionally played a role in hardware management and preservation.
Support services for long-term access to digital data are more prevalent, at least among the responding universities, than support services for non-digital data. The average number of total responses indicating support among all digital types (excluding “other”) was 30. The most commonly supported digital data formats are:

1. Digital texts (57)
2. Digital images (56)
3. Digital audio recordings (46)

All three of these digital formats only received one “not offered” option. Since access to the above data types is commonly supported by libraries, this is not an unexpected result. As indicated by the survey respondents, the two least supported digital data formats for long term access are computer code (10) and digital sequences gene sequences or similar digital renderings of biological/organic/inorganic samples or specimens (16). One explanation could be that access to these data types are not generally supported by the units surveyed, and therefore, support services are unknown by respondents.

---

2 Total responses w/out “not offered” option
Frequency of Preservation/Management Support Services by Digital Research Data Type

- Curation
- Preservation
- Archiving
- Storage
- None
Research Data Policies

The research data policy section of the survey investigated four types of potential institutional policies:

1. General research data ownership policy
2. General policy for externally funded research
3. General policy for research NOT supported by grants, contracts, or other external sources of funding
4. General policy once external funding has expired

Out of 17 responses, 12 (71%) respondents indicated that their institutions have a general data ownership policy. The remaining policy areas had much closer results. A slight majority indicated that their institutions have policies on externally funded research. The remaining policy areas had slightly more negative responses, indicating areas for further research at the institutional level. However, the results of this section could once again been influenced by the internal-institutional affiliation of the respondents. Since libraries were highly represented, knowledge of specific types of funded research policies may not be known.
For policies related to externally funded research projects, 30% of respondents indicated that costs are primarily folded into the direct costs of the grants. According to the survey results, research departments or university-level funding are rarely responsible for financial contributions. However, 45% indicated that their institutions do not have specific policies on externally funded research.

![Specific Externally Funded Research Policies (%)](image-url)
Over half of the respondents indicated that their institutions do not have a policy specifically related to non-externally funded projects (53%). There were three open-ended responses to the “other” choice. These indicated that at one institution a policy is being drafted, a second institution stated all of the choices were a part of the institution’s policy, and a third institution wrote that they will store research data from this type of project for free on a short-term basis.

![Specific Non-External Funded Policies (%)](chart.png)
For research projects that have expired, 41% of the respondents indicated that their institution has no specific policy. The next highest response was that costs were paid by the researcher’s department or center. No respondent indicated that an individual researcher or research team took on the costs. There were no open-ended answers to the two “other” selections.