Iowa Research Online
The University of Iowa's Institutional Repository

University of Iowa Libraries Staff Publications

12-12-2012

# Report on the University Libraries' Data Management Needs Survey

Shawn Averkamp
*University of Iowa*

Xiaomei Gu

Copyright © 2014, The University of Iowa.

Hosted by Iowa Research Online. For more information please contact: lib-ir@uiowa.edu.

# Report on the University Libraries' Data Management Needs Survey

December 12, 2012

Submitted by Shawn Averkamp and Xiaomei Gu

## EXECUTIVE SUMMARY

As research becomes more data-intensive and collaborative, the University of Iowa is seeing a growing demand for data management skills and resources. Funding agencies are increasingly requiring that researchers include data management plans in new grant proposals, and many now require that research data be archived and shared in digital repositories.

Using an online questionnaire, this study aimed to provide starting points to understand University of Iowa researchers' needs and concerns regarding data management. The All Faculty & Staff listserv was used to send a survey link to the target survey population, faculty and staff members at the University of Iowa who are directly involved in research (the exact number of target population is unknown). The survey was open from June 26, 2012 to September 18, 2012. A total of 784 responses, both partial and complete, to the survey, were received. Because participants were self-selecting and the exact target population is unknown, this total should not be interpreted as a representative sample of the University research community.

The 13-question survey included questions on types of data, sensitive or confidential data, funding agencies, data management plan requirement, data storage or sharing, and where researchers get assistance and guidance for data management and organization. The survey results illustrate a range of needs across the University's disciplinary spectrum as well as a view into current services, both internal and external, that researchers consult to meet those needs. Key findings and observations include:

- Researchers generate data in a wide range of formats, and most work with two or more.
- Nearly one quarter of respondents report writing data management plans as part of their funding requirements.
- Over one third of respondents report not receiving assistance with data management. Of those respondents, however, roughly half expressed interest in finding help. Among those who receive assistance, most find help within their departments or colleges.
- Availability of storage is a concern for many researchers.

In order to better identify gaps in service or opportunities for promoting and expanding existing services, we plan to conduct a combination of focus group and one-on-one interviews with faculty and Professional & Science (P&S) researchers and their research assistants (if applicable). These interviews will explore each area of data management in depth to learn how researchers find assistance, where they would prefer to receive assistance, their level of satisfaction with services they receive, and their awareness of existing available services. Results will be summarized in a report planned to be released by the end of June 2013.

## Table of Contents

EXECUTIVE SUMMARY ..... 1
INTRODUCTION ..... 3
Survey Population ..... 3
Survey Design and Pretesting ..... 3
Survey Distribution ..... 4
Survey Responses ..... 4
FINDINGS ..... 4
Participant Demographic Information ..... 4
Willingness to be Interviewed ..... 6
Data ..... 6
Funding Sources ..... 9
Data Management Plans ..... 10
Data Organization ..... 12
Storage ..... 14
Comments and Concerns ..... 17
NEXT STEPS ..... 20
Interview Design ..... 20
Survey of Existing Services ..... 20
Timeline ..... 21
REFERENCES ..... 22
APPENDIX 1. SURVEY INSTRUMENT ..... 23
APPENDIX 2. SURVEY RESULTS ..... 29
APPENDIX 3. LIST OF SELECTED REPOSITORIES ..... 35

# INTRODUCTION

As research becomes more data-intensive and collaborative, the University of Iowa is seeing a growing demand for skills and resources to deal with capturing, organizing, storing, publishing, archiving, and sharing data (i.e. data management). Funding agencies, such as National Science Foundation (NSF), National Institute of Health (NIH), and National Endowment for the Humanities (NEH), are increasingly requiring that researchers include data management plans in new grant proposals, and many now require that research data be archived and shared in digital repositories.

Currently, the Digital Research and Publishing unit in the University of Iowa Libraries offers dataset archiving (50GB or less) in its institutional repository and data management plan consultation services. In addition, ITS - Research Services and several other research centers on campus, such as the Clinical Trials Statistical and Data Management Center (CTSDMC), the University of Iowa Center for Health Effects of Environmental Contamination (CHEEC), the Institute for Clinical and Translational Science (ICTS), and the Iowa Social Science Research Center (ISRC), also offer data management services within certain research disciplines. The University of Iowa Libraries is interested in assessing the needs of researchers across campus for managing their data, in order to inform the coordination of services in the Libraries and across campus to support effective data stewardship throughout the research lifecycle.

The University of Iowa Libraries charged Shawn Averkamp, Data Services Librarian, and Xiaomei Gu, Clinical Education Librarian, to conduct a needs assessment using a mixed-mode survey methodology involving web-based questionnaires and face-to-face interviewing. The study protocol was approved by the Institutional Review Boards (IRB) of the University of Iowa. This report summarizes findings from the web-based survey. Survey results will guide the design of focus group and one-on-one interviews with researchers across campus.

## Survey Population

The target population for the survey was faculty and staff members at the University of Iowa who are directly involved in research. To reach this audience, survey links were sent by email to all University of Iowa faculty and staff via the ITS All Faculty & Staff listserv.

## Survey Design and Pretesting

A questionnaire of 13 questions was developed based on similar studies in other institutions (Provost's Task Force on the Stewardship of Digital Research Data, 2012; "LSE Research Data," 2012; Koralis, 2012) and conversations with individuals at Research Services at Information Technology Services (ITS), Iowa Social Sciences Research Center (ISRC), and Division of Sponsored Programs (OSP). In addition to questions on demographic information and willingness to be further interviewed, the survey included questions on sensitive or confidential data, types of data, funding agencies, data management plan requirement, data storage and sharing, and where researchers get assistance and guidance for data management and organization. A complete list of questions can be found in Appendix A.

The questionnaire was programmed and administered using Qualtrics survey software (http://uiowa.qualtrics.com). Participants were allowed to skip questions. The survey was pretested before distribution by several researchers from multiple disciplines.

## Survey Distribution

Before the survey was disturbed, Linda Walton, Director of Hardin Library and UI Branch Libraries & Associate University Librarian, requested the deans on campus encourage their faculty and staff researchers to respond the survey. Pre-notification email messages were also sent to contacts in ISRC and OSP.

One single survey URL was disturbed via ITS All Faculty & Staff listserv on Tuesday, June 26, 2012. An email reminder with a specific deadline was sent out on Tuesday, August 28, 2012. The survey was closed on Tuesday, September 18, 2012.

## Survey Responses

There were 784 responses to the survey. A response was counted only when a respondent answered at least one question and submitted the survey, so this total includes both partial and complete responses. We also acknowledge it is possible that some respondents responded multiple times since one single survey URL was distributed via a listserv. Since the exact number of the target population, faculty and staff involved in research at the University, is unknown, it is not possible to calculate the survey response rate.

The response rate for each question (i.e. question response rate) ranged from 14% to 99%. Only two questions had a response rate lower than 61%: Questions 8, which was not prompted for those who answered "no" in Question 7, and Question 13, which was an optional, open-ended question intended to capture any additional information. So, their low response rates do not necessarily indicate the questions themselves are irrelevant or poorly phrased.

# FINDINGS

Survey results excluding free-text responses can be found in Appendix 2.

## Participant Demographic Information

Three questions in the survey relate to demographic information.

***Question 1.*** **Which one of the following best describes your primary research area?**

- **Interdisciplinary. Please specify: _____**
- **Arts and Humanities**
- **Business**
- **Computer and information science**
- **Education**

- **Engineering**
- **Health sciences**
- **Life sciences**
- **Law**
- **Physical sciences**
- **Psychology**
- **Social sciences**
- **Other. Please specify: _____**
- **Not applicable. I do not conduct any research.**

This was a single-choice question, and a total of 780 responded. 574 selected one of the research areas given (Figure 1). 174 selected "Not applicable. I do not conduct any research" and were directed to the end of the questionnaire, where they had the option to describe their department/unit (Question 12) and enter any additional comments (Question 13). 32 chose "Other" to specify their primary research areas. Some of these free-text responses could fall under the options given. For example, "pharmacy" and "medical research" would fall into the Health Sciences category. Others such as "university housing & services" and "media relations" seemed like department or unit names rather than research areas or disciplines.

| Answer | | Response |
|---|---|---|
| Law | | 6 |
| Business | | 12 |
| Computer and information science | | 13 |
| Education | | 14 |
| Engineering | | 19 |
| Psychology | | 19 |
| Other. Please specify: | | 32 |
| Physical sciences | | 32 |
| Life sciences | | 42 |
| Interdisciplinary. Please specify: | | 50 |
| Social sciences | | 51 |
| Arts and Humanities | | 75 |
| Not applicable. I do not conduct any research. | | 174 |
| Health sciences | | 241 |
| Total | | 780 |

**Figure 1.** Responses to *Question 1*, "Which one of the following best describes your primary research area?

***Question 3.*** **Which one of the following best describes your position?**

- **Faculty researcher**
- **P & S researcher**
- **Other. Please specify: _____**

This was also a single-choice question, and a total of 561 responded. 241 reported to be faculty researchers, and 184 reported to be staff researchers. 136 chose "Other" to specify their positions. 11 of these 136 respondents reported to be graduate students, which fell outside of the target population. Many graduate students are included in the ITS All Faculty & Staff listserv because of their research assistantships or teaching assistantships, and they may have responded from either perspective. 7 of the 136 participants in the "Other" category reported to work in the capacity of information technology services. Other free-text responses given could fall under either faculty researchers or staff researchers.

***Question 12.*** **What is your department or unit? List your primary affiliation if you have a dual appointment. (Optional)**

This is an optional free-text question placed toward the end of the survey. A total of 477 respondents from various departments listed their department/unit information in varying degrees of detail.

## Willingness to be Interviewed

In Question 11, a total of 183 respondents indicated they are willing to be interviewed at a later time. Most are faculty members.

## Data

***Question 2:*** **"Data" is any recorded material necessary to validate your research findings. This is often numeric, tabular data, but it can also be textual data, images, audio/video files, or even artifacts. Do you collect, generate, or use data in your research?**

573 out of 735 participants responded to Question 2 that they do collect data. 162 answered that they do not generate or collect data in their research and were taken to the end of the survey. 49 did not respond to this question and were allowed to continue with the survey.

A cross-tabulation of Question 2 against Questions 1 and 3 would indicate that not all of the respondents fell into the target demographic population. Of the 49 respondents who did not respond to Question 2, 38 answered Question 1 about their primary research area as "Not applicable. I do not conduct any research." Additionally, 18 respondents of the 573 who said in Question 2 that they do collect data answered in Question 1 that they do not conduct research. Of these 56 respondents who claimed to not conduct research, 25 did not self-identify with any role in Question 3, but 3 identified as P&S researchers, and 30 identified as "Other," with roles varying from "Resident" to "ITS Staff" to "Research Support Specialist" to "Administrative support," to name a few.

We cannot deduce the precise reasons for these respondents' choices to continue completing a survey directed at campus members producing data in the course of research, but from the roles represented, it would seem that at least some of the respondents completed the survey because they either support research in their position and do not feel comfortable claiming ownership over it, or they do not conduct research but produce or manage data in the course of their job (ex., managing student data or Human Resources data) and have concerns about the state of data management on campus, research or otherwise.

## Types of Data

***Question 5:*** **What kinds of data do you collect, generate, or use in the course of your research? Check all that apply.**

- **Artifacts, samples, and/or specimens**
- **Tabular or relational data (e.g. spreadsheet, database)**
- **Geospatial data**
- **Images**
- **Audio**
- **Video**
- **Computer programs or code**
- **Genomic data**
- **Survey / questionnaire data**
- **Other. Please specify: _____**

558 participants responded to Question 5, "What kinds of data do you collect, generate, or use in the course of your research? Check all that apply." Most participants reported use of tabular/relational data (404 responses) or textual data (391 responses). Figure 2 lists the full distribution of responses to Question 5.

42 respondents reported collecting, generating, or using other types of data than the options listed. Some of these free-text responses, such as "physical measurements", "interviews and field notes," and "time series data," represented refinements of the options given, while others such as "instrument or telescope data" and academic and medical records were types that did not fall cleanly into any of the options offered. Seven respondents noted primary and secondary sources from archives or database searches as "data" that they use in the course of their research.

Evaluating the types of data used by researchers could possibly help us predict data management needs. For instance, creators of tabular data may need assistance with data interoperability or metadata standards. Creators of image, audio, video or other large data files will likely have issues with storage. Also, certain media, such as audio and video, are frequently rendered in proprietary or short-lived file formats. Creators of these types of data may need guidance or assistance in format migration.

| Answer | | Response |
|---|---|---|
| Other. please specify: | | 42 |
| Geospatial data | | 76 |
| Genomic data | | 96 |
| Audio | | 102 |
| Computer programs or code | | 128 |
| Video | | 130 |
| Artifacts, samples, and/or specimens | | 188 |
| Survey / questionnaire data | | 258 |
| Images | | 276 |
| Textual data (e.g. text files, Word documents, PDF) | | 391 |
| Tabular or relational data (e.g. spreadsheet, database) | | 404 |

**Figure 2.** Responses to *Question 5*, "What kinds of data do you collect, generate or use in the course of your research? Check all that apply." 558 participants responded.

## Sensitive and Confidential Data

***Question 4:*** **Do you work with sensitive or confidential data?**

382 of 562 respondents reported in Question 4 that they work with sensitive or confidential data, 170 said they do not, and 10 respondents said they do not know (Figure 3). Approximately half of respondents who chose "Yes" also chose Health Sciences as their primary research area in Question 1 (185 respondents). Additionally, based on the free-text responses to Question 1, at least 13 of those who chose "Interdisciplinary" and 8 of those who chose "Other" could be considered as working in Health Sciences, bringing that total to 206. Of the remaining respondents who chose "Yes" to Question 4, 33 chose Social Sciences, while the remaining research areas from Question 1 claimed 15 or fewer respondents. "Not applicable, I do not conduct research" had 29 respondents claiming to work with sensitive or confidential data. Based on the free-text responses given for Question 12 ("Please list your department or unit"), it appears that many of these respondents work with data in non-research capacity, such as collecting and managing patient, students, or human resources data.

| Answer | | Response |
|---|---|---|
| I don't know | | 10 |
| No | | 170 |
| Yes | | 382 |
| Total | | 562 |

**Figure 3.** Responses to *Question 4*, "Do you work with sensitive or confidential data?"

## Funding Sources

***Question 6.*** **Where do you receive funding for your research? Check all that apply.**

- **Department of Defense (DoD)**
- **Department of Education (ED)**
- **Department of Energy (DOE)**
- **National Institutes of Health (NIH)**
- **National Science Foundation (NSF)**
- **National Endowment for the Humanities (NEH)**
- **National Aeronautics and Space Administration (NASA)**
- **Privately funded**
- **Not applicable. I don't receive funding for my research**
- **Other. Please specify: _____**

This was a multiple-choice question, and a total of 547 responded (Figure 4). Among the sources specified by respondents, some could fall into the options given, and others include the University (or its colleges or departments), the State of Iowa, and federal agencies that are not given. Several additional federal funding agencies were listed by respondents:

- Agency for Healthcare Research and Quality (AHRQ)
- Centers for Disease Control and Prevention (CDC)
- Department of Administration (VA)
- Department of Health and Human Services (HHS)
- Department of Transportation (DOT)
- Department of Justice National Institute for Occupational Safety and Health (NIOSH)
- Federal Highway Administration (FHWA)
- Health Resources Services Administration (HRSA)
- National Endowment for the Arts (NEA)
- National Institute of Nursing Research (NINR)
- Substance Abuse and Mental Health Services Administration (SAMHSA)
- US Food and Drug Administration (FDA)
- U.S. Department of Agriculture (USDA)
- U.S. Environmental Protection Agency (EPA)
- U.S. Department of Transportation (DOT)

| Answer | | Response |
|---|---|---|
| Department of Energy (DOE) | | 11 |
| National Aeronautics and Space Administration (NASA) | | 12 |
| Department of Education (ED) | | 13 |
| National Endowment for the Humanities (NEH) | | 16 |
| Department of Defense (DoD) | | 25 |
| National Science Foundation (NSF) | | 82 |
| Other. Please specify: | | 122 |
| Privately funded | | 147 |
| Not applicable. I don't receive funding for my research | | 160 |
| National Institutes of Health (NIH) | | 229 |

**Figure 4.** Responses to *Question 6*, "Where do you receive funding for your research? Check all that apply." 547 participants responded.

## Data Management Plans

***Question 7.*** **Some funding agencies require a data management plan as part of the grant proposal. Are you required to write data management plans?**

This was also a single-choice question. Of 551 respondents, 186 reported they are required to write data management plans and were then prompted with another question on data management plans (Question 8). 99 selected "I don't know." This group was also prompted with Question 8. 266 of the 551 reported that they are not required to write data management plans, and were then taken to Question 9.

***Question 8.*** **Where do you find assistance for writing data management plans? Check all that apply.**

- **My department or college**
- **UIowa research center or institute. Please specify: _____**
- **Division of Sponsored Programs**
- **University Libraries**
- **Information Technology Services (ITS)**
- **Institute for Clinical and Translational Science (ICTS)**
- **I don't receive assistance, but I would like help**
- **I don't need assistance writing data management plans**
- **Other. Please specify: _____**

This was multiple-choice question. Since two subgroups of respondents in the previous question skipped this question, there were only 273 total responses (Figure 5). It should be noted that 17 selected both " I don't receive assistance but would like help" and at least one of the other options given. Similarly, 13

selected both “I don’t need assistance writing data management plans” and at least one of the other options given.

When respondents selected “UIowa Research Center or Institute”, they had the option to specify which center or institute using a directory from the Office of the VP for Research (OVPR) (http://research.uiowa.edu/directory-centers-and-institutes). Responses chosen from the OVPR list included IIHR - Hydroscience & Engineering, Clinical Trials Statistical & Data Management Center (CTSDMC), Iowa Social Science Research Center, Preventive Intervention Center, and Health Effectiveness Research Center (HERCe).

A small number of respondents (36 of 273) chose ‘Other’ to specify where they find assistance. Several listed colleagues, professional societies, National Aeronautics and Space Administration (NASA) guidelines, and online sources.

| Answer | | Response |
|---|---|---|
| University Libraries | | 7 |
| UIowa research center or institute. Please specify: | | 17 |
| Information Technology Services (ITS) | | 21 |
| Institute for Clinical and Translational Science (ICTS) | | 23 |
| Division of Sponsored Programs | | 26 |
| Other. Please specify: | | 36 |
| I don’t need assistance writing data management plans | | 53 |
| I don’t receive assistance, but I would like help. | | 76 |
| My department or college | | 109 |

**Figure 5.** Responses to **Question 8**, “Where do you find assistance for writing data management plans? Check all that apply.” 273 participants responded.

***Question 13.*** **Please add any additional comments or concerns related to data management needs and services on campus. (Optional)**

Data management plans were addressed in only one response to Question 13. The respondent recounted an experience in pursuing help with a data management plan:

> *“When I first had to write a data management plan, I contacted Sponsored Programs for help. They did not have any examples they could provide, or even any resources to go to for help. Therefore, I went online and found some wonderful online examples and resources at other*

*universities. It seems that they should be more on top of new requirements for grants, as other universities appear to be."*

## Data Organization

***Question 9.*** **Where do you find guidance for data organization? (e.g., creating metadata, file naming, file versioning, data storage) Check all that apply.**

- **My department or college**
- **UIowa research center or institute. Please specify: _____**
- **Division of Sponsored Programs**
- **University Libraries**
- **Information Technology Services (ITS)**
- **Institute for Clinical and Translational Science (ICTS)**
- **I don't receive assistance, but I would like help**
- **I don't need guidance in organizing my data**
- **Other. Please specify: _____**

A total of 535 responded to this multiple-choice question (Figure 6). It should be noted that 36 selected both "I don't receive guidance but would like help" and at least one of the other options given. Similarly, 30 selected both "I don't need guidance in organizing data" and at least of the other options given. When respondents selected "UIowa research center or institute", they again had the option to specify which center or institute using a directory from the Office of the VP for Research (OVPR) (http://research.uiowa.edu/directory-centers-and-institutes). Free-text responses chosen from the OVPR list included IIHR - Hydroscience & Engineering, Clinical Trials Statistical & Data Management Center (CTSDMC), Iowa Social Science Research Center, Preventive Intervention Center, Center for Health Effects of Environmental Contamination (CHEEC), and Health Effectiveness Research Center (HERCe).

A small number of respondents (57 of 535) chose "Other" to specify where they find guidance, and several listed colleagues and collaborators, National Aeronautics and Space Administration (NASA) guidelines, professional societies, vendors, and online sources.

| Answer | | Response |
|---|---|---|
| Division of Sponsored Programs | | 13 |
| UIowa research center or institute. Please specify: | | 23 |
| Institute for Clinical and Translational Science (ICTS) | | 28 |
| University Libraries | | 34 |
| Information Technology Services (ITS) | | 54 |
| Other. Please specify: | | 57 |
| I don't receive assistance, but I would like help. | | 141 |
| I don't need guidance in organizing my data. | | 155 |
| My department or college | | 192 |

**Figure 6.** Responses to *Question 9*, "Where do you find guidance for data organization? (e.g., creating metadata, file naming, file versioning, data storage) Check all that apply." 535 participants responded.

***Question 13.* Please add any additional comments or concerns related to data management needs and services on campus. (Optional)**

Data organization was addressed in 25 of the free-text responses to Question 13 regarding additional comments and concerns. Some of these responses spoke to a general need for assistance in data organization while many specifically addressed the topics of training, tools, and issues with analog data. Areas of data organization where respondents indicated a need for assistance or guidance included: implementation of standards and best practices, description or organization of legacy data, training and tools, and data entry/migration from analog to digital.

## Training and Tools

15 respondents left comments related to desired training in software, including Microsoft Excel, Microsoft Access, and SAS, as well as training in database design and coding XML, XSLT, HTML, and Javascript. Two respondents perceived a lack in training opportunities for software and tools used in qualitative research and in the humanities:

> *"I am not really clear about "data management" which I assume is more applicable to large-scale studies, generally quantitative. I have seen the term used in relation to NSF announcements, for example. I plan to work on applied research projects, although mostly (but not entirely) with qualitative data - but here even though the numbers may appear small, the actual data (e.g., transcripts for in-depth interviews, ethnographic notes and observationsm, etc.) become voluminous. I would appreciate having opportunities to learn to use software like QSR NVivo or Ethnograph or NUD*ST for qualitative data management."*

*"It would be nice to have someone who could help with Humanities data management software!!! IT'S is all STRM/social sciences oriented--never even heard of Zotero!"*

### Analog data

Five respondents commented on issues related to migrating data from analog to digital. Two respondents, both existing repositories (Iowa Lakeside Laboratory, Paleontology Repository/University of Iowa Museum of Natural History), spoke to the lack of resources to fund digitization initiatives. Two other respondents expressed needs for assistance in data organization and compliance with standards:

*"My research group has mountains of data and many samples collected over many years in many formats on various types of media from many different/changing instruments and computer systems. We have tried to keep things organized as best we can, but it is far from ideal and things have become non-uniform over time. Unfortunately, we do not have resources to go back in and reorganize."*

*"My department/unit is a registered drug manufacturer with the FDA. Our data management needs are for policies, procedures, and batch records for our registered drug products. All documentation and data management systems need to meet cGMP compliance and be compliant with 21CFR11 (part 11). In order to meet this requirement currently we are using a paper and file system for all formal documentation."*

## Storage

***Question 10:*** **Where do you store or share your data? Check all that apply:**

- **On my computer**
- **On a shared drive or server on campus**
- **On external media (e.g., external hard drive, memory stick, CD)**
- **With a commercial online storage service (e.g. Dropbox, Google Docs)**
- **In a disciplinary repository (e.g. Dryad, ICPSR) or institutional repository (e.g. Iowa Research Online). Please specify:**
- **On paper (e.g. lab notebooks, printouts)**
- **I don't know**
- **Other. Please specify:**

541 participants answered the question on storage, choosing one or more responses. The majority of respondents reported storing data on their personal computer and/or a shared drive or server on campus. Most respondents selected more than one method of storage or sharing, though we cannot determine from the results if additional storage methods are backup for a primary storage method or if respondents use different methods for different purposes or types of data.

| Answer | | Response |
|---|---|---|
| I don't know | | 6 |
| In a disciplinary repository (e.g. Dryad, ICPSR) or institutional repository (e.g. Iowa Research Online). Please specify: | | 34 |
| Other. Please specify: | | 37 |
| With a commercial online storage service (e.g. Dropbox, Google Docs) | | 94 |
| On paper (e.g. lab notebooks, printouts) | | 257 |
| On external media (e.g., external hard drive, memory stick, CD) | | 280 |
| On my computer | | 375 |
| On a shared drive or server on campus | | 388 |

**Figure 7.** Responses to *Question 10*, "Where do you store or share your data? Check all that apply." 541 participants responded.

### Disciplinary and institutional repositories

34 respondents answered that they store or share data in a disciplinary or institutional repository, though only 18 listed a specific repository. Repositories listed included:

- CHILDES (Child Language Data Exchange System)
- Dryad
- GenBank
- EarthChem
- NASA data centers: NSSDC (National Space Science Data Center), CDAWeb (Coordinated Data Analysis Web), SSCWeb (Satellite Situation Center Web), OMNIWeb
- NGDC NOAA SPIDR (Space Physics Interactive Data Resource)
- ESA CAA (European Space Agency Cluster Active Archive)
- Virtual observatory websites
- Planetary Science Archive, Planetary Data System
- Office of the State Archaeologist
- Iowa Department of Public Health
- FDA eLEXNET (Electronic Laboratory Exchange Network)

Half of the 18 text responses, however, listed storage or database management systems, such as Xythos and REDCap, which may indicate that some researchers do not share a common definition of

"repository." Because of this, we cannot tell how many of the 16 respondents who did not list a specific repository actually share information in institutional or disciplinary repositories.

### Other methods

37 respondents chose "Other" and specified other methods for storing or sharing data beyond the options listed. Some of these responses, such as "in my office" and "on back-up media" were redundant of the listed options. A number of respondents listed lab servers, so the option "shared drive or campus server" may have been too vague or perhaps not applicable to the researcher's situation. Nine respondents said they store or share their data on external servers--offsite backup, collaborators' university servers, or sponsors' servers. Two respondents mentioned using online survey software, Qualtrics, and three noted using a clinical trials database or REDCap. (The University holds campus-wide licenses to both of these products). Two responses, "email" and "FTP site", imply means of sharing data, though we did not ask respondents to specify which purpose, storage or sharing, they were addressing.

### Comments and concerns

Storage and/or sharing were addressed in 26 of 111 responses to Question 13 regarding additional comments or concerns, more than any other topic represented. Many spoke to the issue of storage costs and availability, citing departmental storage limitations (particularly for large files), excessive costs, and alternatives to University server storage upon hitting the limit:

> *"Data storage, just like IT support, should be paid for out of indirect costs. In the modern age of research, data is no different from electricity; an essential part of getting science done."*
>
> *"The cost of data storage is currently too high for widespread use by investigators that have very limited resources given the current funding climate."*
>
> *"Institutional servers have had inadequate memory for my data storage needs, which forced me to use external hard drives. I maintain these hard drives with dual backups in a fire proof safe in the lab."*
>
> *"We are chronically running up against data storage limitations on departmental servers."*

A few respondents were satisfied with the level of service they receive with regards to storage:

> *"Our biggest "data" storage needs continue to be banking the DNA, cell, and tissue samples we collect from patients, but we are moving into massively parallel DNA sequencing. This has generated computer hard drive space needs that are many times the total drive space we had ever used up to this point, and promises to continue to expand rapidly. Thus far, ITS has been able to accommodate these needs very well."*
>
> *"...Data has been stored using a central repository model over the 24 year period on various campus servers. We currently use an ITS NAS service that is super reliable with redundant backup, but is expensive. We also use an ITS server for archiving data that has less robust*

*backup, but is relatively inexpensive. We believe that we have a relatively robust system in place, but are always looking for ways to reduce our costs."*

*"I guess, all I really would like to know, what storage opportunities are available at the University. We do have a Administrator who takes care on storage, but just in case he does not have all the information, it would be good to know the opportunities."*

*"I find the hard drive space supplied by my department/college to be adequate for my needs."*

Some respondents shared comments and concerns related to sharing data. A few offered suggestions, such as a Google Drive-like system (or a site license to Google Drive) or central repositories for software, codes, data, and metadata:

*"A program that closely mimics Google Drive would be good, or just purchasing a 'site license' for campus use of google drive would be fine. There are few improvements to this system that I could recommend."*

*"I very strongly support a central repository for codes, data, and metadata related to the research activities of IIHR. This is a critical need. I'd be happy to be a part of any effort to provide such a service (e.g., committee to define needs, evaluate solutions, test implementations)."*

*"Networked research file space and code repositories (i.e. svn) would be a wonderful service that would drastically improve my work."*

## Comments and Concerns

***Question 13: Please add any additional comments or concerns related to data management needs and services on campus. (Optional)***

111 respondents entered responses for this question. Responses were categorized based on recurring themes, including:

- storage
- organization (comments or concerns around organization of data)
- training/tools
- analog data
- data management plans
- confidentiality
- profile (comments further describing respondents' job position or data management practices)
- services (comments describing existing services used or self-identifying as service providers)
- libraries (comments related to University Libraries or libraries in general)
- non-research (comments about non-research data management)
- referrals (suggestions or offers for collaboration on this study)

Some of the responses fell in to more than one category. Others were completely unrelated to the study or were unintelligible and are not discussed here.

Some of these comments and concerns have already been discussed in previous sections of this report. In addition, respondents shared comments, suggestions, and concerns about existing services and issues with non-research data, while others used the opportunity to comment on general library services or the role libraries should play in data management services. A few respondents used this space to refer us to other stakeholders on campus they thought would be of help to the study.

### Existing services

Seven respondents named services they currently use or identified themselves as service providers of some aspect of data management assistance.

Services used include:

- Biology department ITS
- College of Public Health

Self-identified service providers include:

- Holden Comprehensive Cancer Center (HCCC) Population Research Core
- a Health Care Information Systems (HCSIS) storage area network (SAN) administrator
- a College of Nursing staff member who prepares faculty research data for statistical analysis
- Carver College of Medicine (CCM) Flow Cytometry Facility
- Facilities Management Utilities and Energy Management ("used by Civil and Environmental Engineering for Research relating to Environmental issues and concerns")

### Non-Research Data Management

A few respondents expressed the need for training/tools, guidance, and storage for data not generated in the course of research. While this study did not intend to address the data management needs of this contingent, it should be noted that some solutions adopted by the University to assist researchers could potentially aid these campus units:

> *"I do most/all of the data management for the Honors Program. This includes working with Access Databases and the Data Warehouse, storing and organizing photos with Picasa, and collecting information needed for smooth operation of a student staff using Google Docs, and conducting queries using Qualtrics. The majority of my co-workers do not know how to work with any of these programs, so I am having to help them consistently in their job responsibilities when these components are needed. Database management is not my primary responsibility, and we do not have anyone in our department that focuses on the proper procedures for data management and storage."*

*"We generate several terabytes of photo and video data every year and would love to keep it if we had the means to do so. Currently we delete this material some of which is irreplaceable due to storage constraints. It's sad to see some fine photos and video disappear forever due to a lack of centralized archival storage."*

## Libraries

12 Respondents left comments related to the University Libraries services or to the role of libraries in data management services. Most of the praise and complaints about University Libraries services were unrelated to data management, prompted most likely by a misunderstanding of the purpose of the survey or a perceived opportunity to offer general feedback on library services.

Four respondents voiced concern about the role the University Libraries should play in offering data management services. Two of these feared a reallocation of library resources that would decrease the quality of library services they presumably use, namely purchasing materials for the collection:

*"My primary concern is the library resources will be *diverted* to data management at the *expense* of its core mission, which is to maintain books and journals to support teaching and research - something that has been in a steady state of decay in recent years."*

*"I worry that the expenses of data collection and management is swamping the need to collect old fashioned books."*

The other two respondents spoke to the role they feel the Libraries should play in data management:

*"I feel strongly that the University Library ought to strive to maintain its traditional and valued function to support the research and intellectual needs of the faculty and students. The impulse to change into a quantitative, "information" and "data-driven" institution seems to me, and numerous of my colleagues, not only to be trendy and short-sighted, but to undermine the deepest values of the library in the liberal arts and sciences."*

*"Please don't try and take over. Our storage systems and database methods are reliable and cost effective. I'd be happy to collaborate with the UI libraries, but our data is not going to be as important to them as it is to us."*

## Referrals

Finally, five respondents suggested we contact or collaborate with other stakeholders or themselves, including ITS, systems administrators, IT directors, and faculty researchers.

# NEXT STEPS

## Interview Design

The survey results illustrate a range of needs across the University's disciplinary spectrum as well as a view into current services, both internal and external, that researchers consult to meet those needs. In order to better identify gaps in service or opportunities for promoting and expanding existing services, we plan to conduct a combination of focus group and one-on-one interviews with faculty and P&S researchers and their research assistants (if applicable) to learn more about data management needs and practices within different disciplines across campus.

Our researcher interviews will focus on these areas of data management:

- data management planning (e.g., planning for a research project, compiling a written plan as part of a grant proposal)
- data organization (e.g., applying metadata standards, database design, tracking data files, data entry, analog to digital format conversion, file migration, selecting and learning tools for analysis or organization)
- storage (e.g., choosing temporary or long-term storage, backing up data)
- sharing (e.g., selecting an institutional or disciplinary repository, preparing data for deposit, applying licenses)
- confidential data (e.g., data anonymization, compliance with state and federal regulations)

Within each area, we hope to generate discussion that helps to answer the following questions:

- Where do researchers find assistance?
- How satisfied are they with the services they receive?
- Where would they prefer to find assistance?
- How aware are they of existing available services?

Research cultures also vary widely by discipline, so we plan to draw on the domain expertise of subject liaison librarians within the Libraries to design appropriate interview questions and compositions of focus groups. We hope to create an environment where participants feel comfortable speaking freely, so liaison input will help us determine which participants to interview in a focus group or one-on-one. Participants will be selected from the subset of survey respondents who specified in Question 11 that they are willing to be contacted to answer further questions. We will also solicit recommendations from contacts within various disciplines on campus.

## Survey of Existing Services

Before conducting researcher interviews, we plan to first meet with existing service providers on campus to better understand what services are currently available. We hope to learn more about the following topics:

- What services are provided? To whom? What are the costs for researchers?
- How do service providers view their role in research planning?
- How are providers funded? Does this have an impact on the range and level of services they are able to provide?
- How do providers promote their services?
- What services would providers like to offer? What are the obstacles to providing these services?
- Where do providers' clients find assistance with areas of data management not addressed by the providers?

The survey results revealed many groups on campus that researchers turn to for data management assistance. We plan to meet with the following service providers most frequently named in the survey results:

- Information Technology Services (ITS)
- Iowa Social Science Research Center (ISRC)
- Health Care Information Services (HCIS)
- Institute for Clinical and Translational Science (ICTS)
- Clinical Trials Statistical and Data Management Center (CTSDMC)
- Office of Sponsored Programs (OSP)
- IIHR--Hydroscience and Engineering

## Timeline

In order to conduct focus group and one-on-one interviews by early Spring semester 2013, we plan to compile questions and meet with service providers by the end of January. After completing necessary Institutional Review Board requirements, we hope to schedule interviews for February and March 2013, with analysis and reporting completed by the end of June 2013.

## REFERENCES

Koralis, S. (2012, February 20). *Survey on Data Management in Academic Libraries*. Retrieved February 25, 2012 from http://hastac.org/blogs/sdkeralis/2012/02/20/survey-data-management-academic-libraries

*LSE Research Data Preservation Survey. (2012).* Retrieved November 21, 212, from http://lsedice.wordpress.com/2012/03/13/research-data-preservation-survey-for-lse-researchers/

Provost's Task Force on the Stewardship of Digital Research Data. (2012). *Research Data Stewardship at UNC.* Retrieved November 21, 2012, from http://sils.unc.edu/sites/default/files/general/research/UNC_Research_Data_Stewardship_Report.pdf

## APPENDIX 1. SURVEY INSTRUMENT

You are invited to participate in a brief survey being conducted by the University of Iowa Libraries.

As research becomes more data-intensive and collaborative, The University of Iowa is seeing a growing demand for skills and resources to deal with capturing, organizing, storing, publishing, archiving, and sharing data (i.e. data management). Funding agencies (including NSF, NIH, and NEH) are increasingly requiring that researchers include data management plans in new grant proposals, and many now require that research data be archived and shared in digital repositories.

We are interested in hearing from faculty and staff researchers from across campus including the scientific, social sciences, and humanities communities to learn what kind of data is being generated and how researchers are finding help in managing it. Your participation will help us assess data management needs at the University of Iowa and will inform the coordination of services across campus to support effective data stewardship throughout the research life cycle.

The survey should take approximately 10 minutes to complete. Click on the following link to access the survey: https://uiowa.qualtrics.com/SE/?SID=SV_eIKUbETb5AVPOZK

If you have any questions or comments, please contact Shawn Averkamp, Data Services Librarian, at shawn-averkamp@uiowa.edu or Xiaomei Gu, Clinical Education Librarian, at xiaomei-gu@uiowa.edu.

If you have questions about the rights of research subjects, please contact the Human Subjects Office, 105 Hardin Library for the Health Sciences, 600 Newton Rd, The University of Iowa, Iowa City, IA 52242-1098, (319) 335-6564, or e-mail irb@uiowa.edu.

/**********
Distribution of this message was approved by the Provost and the VP for Human Resources. Neither your name nor e-mail address was released to the sender. The policy and guidelines for the UI Mass Mail service, including information on how to filter messages, are available at: http://its.uiowa.edu/apps2/support/massmail.

**********/

**Data Management at UIowa**

We invite you to participate in a research study being conducted by investigators from The University of Iowa Libraries. The information collected in this study will be used to assess data management needs at The University of Iowa and will inform the coordination of services across campus to support effective data stewardship throughout the research life cycle.

If you agree to participate, we would like you complete the 12-question survey about your primary research area and your collection of study data. You will also be asked to provide your primary department or unit. You are free to skip any questions that you prefer not to answer. It will take approximately ten minutes to complete the survey. If you would like to elaborate on any of your responses, please feel free to use the comment form at the end of the survey.

If you collect, generate, or use data in your research, you will be asked to provide your email address if you agree to be contacted about a follow-up interview procedure. You may participate in this part of the study without providing your email address.

You will be asked to provide information over the Internet. It is possible that your responses could be viewed by persons who have access to the computers hosting the web site or by unauthorized persons who gain access to the web site computers. We will use a secure web site and computers to collect and store the study information. We will separate your email address from your survey responses and assign your survey responses a study code number. The study code number will be linked to your email address only if you agree to participate in an in-depth interview regarding data management. The list linking your email address and your study identification code will be stored in a separate location that is accessible only to the researchers.

Taking part in this research study is completely voluntary. If you do not wish to participate in this study, close your web browser window without submitting your survey responses.

If you have any questions or comments, please contact Shawn Averkamp, Data Services Librarian, at shawn-averkamp@uiowa.edu or Xiaomei Gu, Clinical Education Librarian, at xiaomei-gu@uiowa.edu.

If you have questions about the rights of research subjects, please contact the Human Subjects Office, 105 Hardin Library for the Health Sciences, 600 Newton Rd, The University of Iowa, Iowa City, IA 52242-1098, (319) 335-6564, or e-mail irb@uiowa.edu.

Thank you very much for your consideration of this research study.

If you agree to participate in this study, click the NEXT button to go to the first survey question. If you wish to keep a copy of this information page, please save or print the page before going to the next page.

***Question 1: Which one of the following best describes your primary research area?***

- ❍ Interdisciplinary. Please specify: _____
- ❍ Arts and Humanities
- ❍ Business
- ❍ Computer and information science
- ❍ Education
- ❍ Engineering
- ❍ Health sciences
- ❍ Life sciences
- ❍ Law
- ❍ Physical sciences
- ❍ Psychology
- ❍ Social sciences
- ❍ Other. Please specify: _____
- ❍ Not applicable. I do not conduct any research.

***Question 2: "Data" is any recorded material necessary to validate your research findings. This is often numeric, tabular data, but it can also be textual data, images, audio/video files, or even artifacts. Do you collect, generate, or use data in your research?***

- ❍ Yes
- ❍ No

If No is selected, then skip to "What is your department or unit?"

***Question 3: Which one of the following best describes your position?***

- ❍ Faculty researcher
- ❍ P & S researcher
- ❍ Other. Please specify: _____

***Question 4: Do you work with sensitive or confidential data?***

- ❍ Yes
- ❍ No
- ❍ I don't know

***Question 5: What kinds of data do you collect, generate, or use in the course of your research? Check all that apply.***

- ❑ Artifacts, samples, and/or specimens
- ❑ Tabular or relational data (e.g. spreadsheet, database)
- ❑ Textual data (e.g. text files, Word documents, PDF)
- ❑ Geospatial data
- ❑ Images
- ❑ Audio
- ❑ Video
- ❑ Computer programs or code
- ❑ Genomic data
- ❑ Survey / questionnaire data
- ❑ Other. please specify: _____

***Question 6: Where do you receive funding for your research? Check all that apply.***

- ❑ Department of Defense (DoD)
- ❑ Department of Education (ED)
- ❑ Department of Energy (DOE)
- ❑ National Institutes of Health (NIH)
- ❑ National Science Foundation (NSF)
- ❑ National Endowment for the Humanities (NEH)
- ❑ National Aeronautics and Space Administration (NASA)
- ❑ Privately funded
- ❑ Not applicable. I don't receive funding for my research
- ❑ Other. Please specify: _____

***Question 7: Some funding agencies require a data management plan as part of the grant proposal. Are you required to write data management plans?***

- ❍ Yes
- ❍ No
- ❍ I don't know

If No is selected, then skip to "Where do you find guidance for data ..."

***Question 8: Where do you find assistance for writing data management plans? Check all that apply.***

- ❑ My department or college
- ❑ UIowa research center or institute. Please specify: _____
- ❑ Division of Sponsored Programs
- ❑ University Libraries
- ❑ Information Technology Services (ITS)
- ❑ Institute for Clinical and Translational Science (ICTS)
- ❑ I don't receive assistance, but I would like help.
- ❑ I don't need assistance writing data management plans
- ❑ Other. Please specify: _____

***Question 9: Where do you find guidance for data organization? (e.g., creating metadata, file naming, file versioning, data storage) Check all that apply.***

- ❑ My department or college
- ❑ UIowa research center or institute. Please specify: _____
- ❑ Division of Sponsored Programs
- ❑ University Libraries
- ❑ Information Technology Services (ITS)
- ❑ Institute for Clinical and Translational Science (ICTS)
- ❑ I don't receive assistance, but I would like help.
- ❑ I don't need guidance in organizing my data.
- ❑ Other. Please specify: _____

***Question 10: Where do you store and/or share your data? Check all that apply.***

- ❑ On my computer
- ❑ On a shared drive or server on campus
- ❑ On external media (e.g., external hard drive, memory stick, CD)
- ❑ With a commercial online storage service (e.g. Dropbox, Google Docs)
- ❑ In a disciplinary repository (e.g. Dryad, ICPSR) or institutional repository (e.g. Iowa Research Online). Please specify: _____
- ❑ On paper (e.g. lab notebooks, printouts)
- ❑ I don't know
- ❑ Other. Please specify: _____

***Question 11: Are you willing to be contacted to answer further questions about your data management practices and needs?***

- ❍ Yes. Please enter your email address. _____
- ❍ No

***Question 12: What is your department or unit? List your primary affiliation if you have a dual appointment. (Optional)***

***Question 13: Please add any additional comments or concerns related to data management needs and services on campus. (Optional)***

## APPENDIX 2. SURVEY RESULTS

1. Which one of the following best describes your primary research area?

| Answer | | Response |
|---|---|---|
| Law | | 6 |
| Business | | 12 |
| Computer and information science | | 13 |
| Education | | 14 |
| Engineering | | 19 |
| Psychology | | 19 |
| Other. Please specify: | | 32 |
| Physical sciences | | 32 |
| Life sciences | | 42 |
| Interdisciplinary. Please specify: | | 50 |
| Social sciences | | 51 |
| Arts and Humanities | | 75 |
| Not applicable. I do not conduct any research. | | 174 |
| Health sciences | | 241 |
| Total | | 780 |

2. "Data" is any recorded material necessary to validate your research findings. This is often numeric, tabular data, but it can also be textual data, images, audio/video files, or even artifacts. Do you collect, generate, or use data in your research?

| Total Responses | 735 |
|---|---|

| Answer | | Response | % |
|---|---|---|---|
| No | | 162 | 22% |
| Yes | | 573 | 78% |
| Total | | 735 | 100% |

3. Which one of the following best describes your position?

| Answer | | Response |
|---|---|---|
| Other. Please specify: | | 136 |
| P & S researcher | | 184 |
| Faculty researcher | | 241 |
| Total | | 561 |

4. Do you work with sensitive or confidential data?

| Answer | | Response |
|---|---|---|
| I don't know | | 10 |
| No | | 170 |
| Yes | | 382 |
| Total | | 562 |

5. What kinds of data do you collect, generate, or use in the course of your research? Check all that apply.

| Total Responses | 558 |
|---|---|

| Answer | | Response |
|---|---|---|
| Other. please specify: | | 42 |
| Geospatial data | | 76 |
| Genomic data | | 96 |
| Audio | | 102 |
| Computer programs or code | | 128 |
| Video | | 130 |
| Artifacts, samples, and/or specimens | | 188 |
| Survey / questionnaire data | | 258 |
| Images | | 276 |
| Textual data (e.g. text files, Word documents, PDF) | | 391 |
| Tabular or relational data (e.g. spreadsheet, database) | | 404 |

6. Where do you receive funding for your research? Check all that apply.

| Total Responses | 547 |
|---|---|

| Answer | | Response |
|---|---|---|
| Department of Energy (DOE) | | 11 |
| National Aeronautics and Space Administration (NASA) | | 12 |
| Department of Education (ED) | | 13 |
| National Endowment for the Humanities (NEH) | | 16 |
| Department of Defense (DoD) | | 25 |
| National Science Foundation (NSF) | | 82 |
| Other. Please specify: | | 122 |
| Privately funded | | 147 |
| Not applicable. I don't receive funding for my research | | 160 |
| National Institutes of Health (NIH) | | 229 |

7. Some funding agencies require a data management plan as part of the grant proposal. Are you required to write data management plans?

| Answer | | Response |
|---|---|---|
| I don't know | | 99 |
| Yes | | 186 |
| No | | 266 |
| Total | | 551 |

8. Where do you find assistance for writing data management plans? Check all that apply.

| Total Responses | 273 |
|---|---|

| Answer | | Response |
|---|---|---|
| University Libraries | | 7 |
| UIowa research center or institute. Please specify: | | 17 |
| Information Technology Services (ITS) | | 21 |
| Institute for Clinical and Translational Science (ICTS) | | 23 |
| Division of Sponsored Programs | | 26 |
| Other. Please specify: | | 36 |
| I don't need assistance writing data management plans | | 53 |
| I don't receive assistance, but I would like help. | | 76 |
| My department or college | | 109 |

9. Where do you find guidance for data organization? (e.g., creating metadata, file naming, file versioning, data storage) Check all that apply.

| Total Responses | 535 |
|---|---|

| Answer | | Response |
|---|---|---|
| Division of Sponsored Programs | | 13 |
| UIowa research center or institute. Please specify: | | 23 |
| Institute for Clinical and Translational Science (ICTS) | | 28 |
| University Libraries | | 34 |
| Information Technology Services (ITS) | | 54 |
| Other. Please specify: | | 57 |
| I don't receive assistance, but I would like help. | | 141 |
| I don't need guidance in organizing my data. | | 155 |
| My department or college | | 192 |

10. Where do you store and/or share your data? Check all that apply.

| Total Responses | 541 |
|---|---|

| Answer | | Response |
|---|---|---|
| I don't know | | 6 |
| In a disciplinary repository (e.g. Dryad, ICPSR) or institutional repository (e.g. Iowa Research Online). Please specify: | | 34 |
| Other. Please specify: | | 37 |
| With a commercial online storage service (e.g. Dropbox, Google Docs) | | 94 |
| On paper (e.g. lab notebooks, printouts) | | 257 |
| On external media (e.g., external hard drive, memory stick, CD) | | 280 |
| On my computer | | 375 |
| On a shared drive or server on campus | | 388 |

11. Are you willing to be contacted to answer further questions about your data management practices and needs?

| Total Responses | 523 |
|---|---|

12. What is your department or unit? List your primary affiliation if you have a dual appointment. (Optional)

| Total Responses | 477 |
|---|---|

13. Please add any additional comments or concerns related to data management needs and services on campus. (Optional)

| Total Responses | 110 |
|---|---|

# APPENDIX 3. LIST OF SELECTED REPOSITORIES

This appendix lists URLs of disciplinary and institutional repositories mentioned in this report.

- CDAWeb (Coordinated Data Analysis Web)
  http://cdaweb.gsfc.nasa.gov/
- CHILDES (Child Language Data Exchange System)
  http://childes.psy.cmu.edu/
- Dryad
  http://datadryad.org/
- EarthChem
  http://www.earthchem.org/
- ESA CAA (European Space Agency Cluster Active Archive)
  http://caa.estec.esa.int/caa/home.xml
- FDA eLEXNET (Electronic Laboratory Exchange Network)
  https://www.elexnet.com/elex/
- GenBank
  http://www.ncbi.nlm.nih.gov/genbank/
- The Interuniversity Consortium for Political and Social Research (ICPSR)
  www.icpsr.umich.edu/
- Iowa Research Online
  http://ir.uiowa.edu/
- NASA data centers: NSSDC (National Space Science Data Center)
  http://science.nasa.gov/planetary-science/planetary-science-data/national-space-science-data-center-nssdc/
- NGDC NOAA SPIDR (Space Physics Interactive Data Resource)
  http://spidr.ngdc.noaa.gov/spidr/
- OMNIWeb
  http://omniweb.gsfc.nasa.gov/ow.html
- SSCWeb (Satellite Situation Center Web)
  http://sscweb.gsfc.nasa.gov/