President's Emerging Leaders | University of Minnesota Driven to Discover

Implementing Research Cyberinfrastructure for the 21st Century

A report to the University of Minnesota's Research Cyberinfrastructure Alliance and the PEL Project Sponsors by the

2008-2009 President's Emerging Leaders

Tracy Anderson, College of Biological Sciences
Craig Gjerdingen, Carlson School of Management
Bryan Herrmann, Office of Admissions, Morris Campus
Katherine Himes, Office of the Senior Vice President for Academic Affairs and Provost
Lisa Johnston, University Libraries

Acknowledgments:

Thanks to our sponsors Ann Hill Duin and Bernard Gulachek, our project leads Kemal Badur and John Sonnack, our advisors from the University of Minnesota Research Cyberinfrastructure Alliance, and Dave Dorman from the PEL Program.

Table of Contents

I. Executive Summary

Project Charge

To develop recommendations for a strategic approach to Research Cyberinfrastructure that addresses the growing need on campus and leverages partnerships across the University. Currently, research technology and emerging global partnerships are generating data with computational and storage needs that outpace the information infrastructure available.

What is Research Cyberinfrastructure?

Research Cyberinfrastructure includes the instruments, sensors, high performance computational systems, massive storage systems, data resources, and visualization facilities, tied together by high speed networks and made to work together by advanced software to accomplish goals that would not be possible by any single information technology system. It also includes the people, processes, training, security, policies, and capabilities to sustain the systems and networks over time. Implementing Research Cyberinfrastructure requires a high level of coordination and collaboration between researchers and an information technology workforce with expertise in scientific computing.

Methodology

We conducted an online user-needs survey to 8,424 University faculty, research staff, and students asking them to report the current state of Research Cyberinfrastructure support at the University, as well as to assess their future needs. The survey was comprised of 130 questions on the following Research Cyberinfrastructure trends: data storage, data management, and networking infrastructure; collaboration with other researchers; tools and applications; high performance computing; and learning and workforce development, as well as future trends within each of these areas. We had 780 successful responses (9.2% response rate). These data formed the basis for our trends and recommendations. In addition, we performed a meta-analysis of recent University reports and surveys on this topic to enhance our trends and recommendations.

Recommendations

1. Develop enterprise-wide, integrated cyberinfrastructure to align with national cyberinfrastructure.
2. Better position services around the end-user.
3. Improve and expand the University's ability to handle data.
4. Re-brand Research Cyberinfrastructure at the University to improve accessibility for all disciplines.
5. Create incentives for virtual collaborative relationships.

What are the Benefits?

Continued focus on Research Cyberinfrastructure will help the University become a top three public research university through accelerating research-enabled discovery; increasing partnerships, research collaborations, and communication within the University; increasing the University's competitiveness for funding from federal agencies; and, finally, sparking new types of research and Research Cyberinfrastructure collaborations.

II. Introduction

> *"The 2008-2009 President's Emerging Leaders (PEL) Project was commissioned to lead this effort toward Implementing Cyberinfrastructure for 21st Century Research."*

In support of the University of Minnesota's goal to become one of the top three public research universities in the world, a strategic alliance to support research cyberinfrastructure was developed in Fall 2007 with the vision of facilitating access to high-end, high quality, efficient and effective research computing systems and services, enhancing interdisciplinary research and allowing researchers to explore radically new concepts, approaches, and tools.

The Research Cyberinfrastructure Alliance (RCA) functions as a virtual organization of research support and information technology service providers throughout the university. Founding members of the University's RCA included individuals from college-level research computing units, the Minnesota Supercomputing Institute, and key leaders from the University's Office of Information Technology (OIT), the Office of the Vice President for Research (OVPR), and the University Libraries.

A key step for the RCA was to develop strategies that enable researchers to easily find and engage with relevant research computing resources and support. The 2008-2009 President's Emerging Leaders (PEL) Project was commissioned to lead this effort toward Implementing Cyberinfrastructure for 21st Century Research. See Appendix A, PEL Project Proposal, for details.

The primary role for the President's Emerging Leaders RCA project was to work with key stakeholders (faculty, staff, students, collegiate IT directors, OIT staff, and the RCA group) to develop a strategic approach to cyberinfrastructure that leverages partnerships across the University. In completion of this goal, we hope that this report serves as the springboard for stakeholders to implement our plan to: help the University become a leader in academic research; accelerate research-enabled discovery; increase partnerships, research collaborations, and communication within the University; increase the University's competitiveness for funding from federal agencies; and, finally, spark new types of research and CI collaborations.

Research cyberinfrastructure (CI) is a key ingredient in fostering interdisciplinary research, garnering national funding, and transforming the University of Minnesota into a top three public research university. Increasingly, research technology and emerging global partnerships are generating data that have computational and storage needs that outpace the information infrastructure currently available. Our project addresses the growing need for CI on campus and strategically lays out ways that the RCA and the University partners might respond. Specifically, in this report we:

1. Provide a background by examining local, national, and international CI trends from an academic setting;
2. Report the results of our survey assessment of researchers' CI needs and describe the current CI available through a meta-analysis of resources across the University;

3. Present our recommendations for action through our combined data analysis and aligned priorities.

III. Background of National and International Cyberinfrastructure Initiatives

Dramatic advances in information technology make way for exciting new opportunities for research. New technologies allow projects to span disciplines and institutions, enabling researchers to seek new answers to critical questions in ways that were impossible just years or even months ago. Indeed, research computing is a top priority for leading universities and research institutions around the globe; furthermore, cyberinfrastructure is seen as a key factor in securing research funding and attracting and retaining top faculty and students.

The expanded research agendas in many disciplines are now outpacing the computing resources available to individual researchers, departments, or even institutions. Enabling this research requires large-scale investments in high-performance computing, storage and networking, as well as the development of cyberinfrastructure to integrate these components into a meaningful whole.

Cyberinfrastructure includes the instruments, sensors, high performance computational systems, massive storage systems, data resources, and visualization facilities, tied together by high speed networks and made to work together by advanced software to accomplish goals that would not be possible by any single information technology system. It also includes the people, processes, training, security, policies, and capabilities to sustain the systems and networks over time. Implementing cyberinfrastructure requires a high level of coordination and collaboration between researchers and an information technology workforce with expertise in scientific computing.

The following aspects of cyberinfrastructure are investigated at the national and international level. By providing exemplary agendas and initiatives as a background to this report, we will help shape and support our recommendations for action in the local cyberinfrastructure here on campus. The aspects focused on here are: (1) Policy leaders, (2) Cyberinfrastructure visionaries, and (3) disciplinary subcultures.

1. Policy Leaders

These organizations are paving the way for institutions to develop cyberinfrastructure for the 21st century:

a. **National Science Foundation (NSF) Vision for Cyberinfrastructure-Cyberinfrastructure for the 21st Century**

The National Science Foundation (NSF) is building a comprehensive infrastructure to capitalize on dramatic advances in information technology for computation-enabled discovery. NSF's mission for cyberinfrastructure (CI) is to:

- Develop a human-centered CI that is driven by science and engineering research and education opportunities;
- Provide the science and engineering communities with access to world-class CI tools and services, including those focused on: high performance computing; data, data analysis and

visualization; networked resources and virtual organizations; and learning and workforce development;
- Promote a CI that serves as an agent for broadening participation and strengthening the nation's workforce in all areas of science and engineering;
- Provide a sustainable CI that is secure, efficient, reliable, accessible, usable, and interoperable, and that evolves as an essential national infrastructure for conducting science and engineering research and education; and
- Create a stable but extensible CI environment that enables the research and education communities to contribute to the agency's statutory mission.
- For more information about NSF's investments in cyberinfrastructure, see Appendix B.

b. National Institutes of Health (NIH)

Read alone the NIH statement on data sharing seems more like an attempt to discourage or at least de-emphasize the sharing of research data. However, the extensive number of resources devoted to implementing the data sharing policy suggests that NIH is attempting to find balance between legitimate reasons for limiting access while also promoting sharing as a general value within its grantee community. There are no firm deadlines for sharing data nor required public databases as is the case with the Department of Energy or the CSREES of the USDA. However, NIH does require that a data sharing plan be included in grant applications and that the plan will be followed if funds are granted. NIH also makes a point of noting that they typically award grants to institutions rather than individuals which puts the awards in a slightly different liability context. NIH also will provide funds above the main grant request for the support of data sharing if the request is included in the application. NIH does provide some general guidelines on deadlines in the FAQ in which they say "While NIH also understands that an institution's desire to exercise its intellectual property rights may justify a need to delay disclosure of research findings, a delay of 30 to 60 days is generally viewed as a reasonable period for such activity." In the question that immediately follows, NIH also makes it clear that, provided no confidentiality issues arise, unpublished data should also be shared (prior to publication if feasible). Additionally, in their guidance document NIH recommends, where appropriate, depositing data in data archives."

c. Microsoft's Report on 2020 Science

Microsoft's report makes it clear that science has the potential to have an unprecedented impact on our world in the 21st Century, from how long we live, to what we know about ourselves, our planet and the universe, to understanding how to control and eradicate disease, to how to protect the entire life-support systems of the earth. As a consequence, it is difficult to overestimate how profound is the scientific revolution now under way. CI will accelerate the pace of scientific discovery now that we are unconstrained by many of the restrictions imposed by print. To counteract this evolution, it proposes the following recommendations:

a. Establish science and science-based innovation at the top of the political agenda.
b. Rethink policies. Tomorrow's scientists will be amongst the most valuable assets that any nation will have. As well as being required to be scientifically and mathematically literate, tomorrow's scientists

will also need to be computationally literate. Achieving this urgently requires a re-think of education policies now, not just at the undergraduate and postgraduate training level, but also at the school level since today's children are tomorrow's scientists.

c. Get the word out. The science community needs to find new ways to raise the importance of science both in the public arena and in terms of its importance for the political agenda.
d. Re-think science funding and science policy structures to make the connection between funding and scientific output, and how we even measure scientific output. Focus more on research and collaborations to facilitate such research, on funding longer timescale projects so people work on solving problems.
e. Create new kinds of research institutes that are highly interdisciplinary, combine teaching, training and research, and focused on 'grand challenges' rather than 'grand disciplines' and simply producing papers.
f. Call to action the development of new conceptual and technological tools.
g. Develop innovative public private partnerships to accelerate science-based innovation. Governments, universities and businesses need to find new kinds of ways to work together - public-private partnerships (PPPs) are needed in order to really accelerate science and science-based innovation.
h. Find better mechanisms to create value from intellectual property, for both universities and industry (and probably governments) to find new and better ways to generate value from science-based intellectual property (IP).

For more information about the Microsoft 2020 report, see Appendix C.

d. National Academies of Science's Committee on Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age

What began as the impetus for this committee were complaints from journal editors and publishers over the recurring improper manipulation of images accompanying article submissions. This led to the broad evaluation of current standards and practices in data integrity and transparency in their report "Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age" (2009, http://books.nap.edu/openbook.php?record_id=12615&page=R1)

The report was written by researchers, law and policy professionals, and publishing representatives and takes a very broad look at what must happen for the future sustainability of data in the digital age. Specifically it makes ten recommendations that ensure:

- research data are managed (researchers obtain training, data professionals are involved),
- data are shared (publicly accessible in a timely way),
- data are retained for future use (curated, archived).

It was noted that these recommendations not be set as "unfunded mandates" from the funding agencies, but included in the research infrastructure and research environment as determined by the researchers, institutions, and related enterprises in each field.

e. Committee on Institutional Cooperation (CIC)

The CIC, founded over fifty years ago, is a consortium comprised of the Big Ten universities plus the University of Chicago. Leveraging the collective impact of twelve institutions, the CIC collaborates on many projects, including library initiatives, course sharing, leadership development, and technology development. To facilitate collaborations and advance initiatives on 'home' campuses, university officers meet multiple times per year. For example, CIC provosts meet semiannually, CIC chief information officers (CIOs) meet quarterly, and CIC graduate deans meet semiannually. In addition, professional technology groups (e.g., IT wireless, IT networking, IT security) representing various levels of management meet frequently.

In 2004, the CIOs of CIC universities launched OmniPoP, a fiberoptic network collaboration tool, with the goals of connecting the twelve institutions to each other and major research hubs, providing additional bandwidth through Internet 2.0, and reducing costs. By purchasing this resource collectively, CIC universities saved an initial $25 million, and yearly cost savings are projected to be $600,000 per institution. This unique, innovative resource is available as a research and teaching technology. Currently, CIOs are working to advance security protocols for OmniPoP.

f. EDUCAUSE IT Directors Survey (Appendix D)

This EDUCAUSE report outlines the findings from their survey of 369 IT administrators of academic institutions. It concludes that an atmosphere of collaboration, resource sharing, and active pursuit of economies of scale are all associated with more effective integration of CI resources. However one of the more interesting findings in this study related to the terminology of Cyberinfrastructure, for example:

> "As one of our advisers remarked, the very term cyberinfrastructure is part of the problem; it means everything and it means nothing. It's not that the National Science Foundation (NSF) was unclear when it coined the term in 2001, but its subsequent usage by the higher education community has been inconsistent, sometimes broadening and sometimes restricting its original scope. As we conducted this study, we heard many times from respondents to our quantitative survey and in our qualitative interviews that the term has become a buzzword and lacks enduring meaning. Several respondents preferred the equivalent European terms e-science and the somewhat less restrictive e-research. Although various questions in our survey asked about research outside science and engineering and inquired about CI technologies used in creative activities and teaching and learning, the bulk of our findings did reflect science and engineering research use."

2. *Cyberinfrastructure Visionaries*

These national and international institutions have made headway in various areas of cyberinfrastructure providing a framework to build from:

a. The **University of California- San Diego's** Supercomputing Institute has recommended several protocols in the management and storage, and preservation of large scientific data sets. The Director, Fran Bergman provides a nice broad summary of these techniques (Berman, 2008), however UCSD provides a specific example of cyberinfrastructure planning with their Ocean Observatories Initiative (OOI) (Ramsey and Cowles, 2009). OOI will be composed of a network of ocean observation components and infrastructure, allowing scientists to study ocean processes regionally, globally, and in coastal areas.

- UC San Diego in collaboration with the Consortium for Ocean Leadership will receive $32 million from NSF and the American Recovery and Reinvestment Act of 2009 over the next five years to develop and construct the networked cyberinfrastructure.
- This will allow for the first time scientists to interact with some of the oceans' most extreme environments. Interest in expanding international capabilities exists in countries including Canada, Japan, Europe, Australia, China and Korea.
- In 2010 the program will add an Education and Public Engagement team to use data and infrastructure to develop new and innovative approaches to outreach. Initial data flow is expected early 2013.

b. The **University of California Humanities Research Institute** (UCHRI), located at UCI, is a multicampus research unit of the UC Office of the President. Founded in 1987, UCHRI's distinctive mission is to foster intellectual community, research, and public programs across campus boundaries; mobilize the strength of the University of California humanities faculty as a whole; and promote innovative collaborative and interdisciplinary research among humanities scholars and researchers in disciplines across the social sciences, sciences, technology, and medicine. UCHRI has two projects of note:
   - The institute leads a cyberinfrastructure initiative to strengthen research support for the humanities, arts, and social sciences called the Humanities, Arts and Social Science (HASS) Grid.
   - Also led the development of the Humanities, Arts, Science, and Technology Advanced Collaboratory (HASTAC), a consortium of top-level humanities and science institutions in the U.S. and abroad dedicated to new and productive partnerships across disciplines in order to create tools and databases that will allow for the production of new kinds of knowledge.

c. The **University of Washington' s eScience Institute** was established in 2008 to address the changing methods in which researchers interact with data, instruments and emerging technologies. Their 2009 Preliminary report (Fox et. al, 2009) outlines the significant changes in user attitudes toward technologies and data use as well as their preliminary solutions to these problems. For more information, visit the institute's Web site: http://escience.washington.edu.

d. At **Penn State University** a sub-unit of their IT department was recently created called the **Research Computing and Cyberinfrastructure group**. This unit works with campus scholars to address cyberinfrastructure problems and provide support solutions for areas such as high performance computing, visual computing technologies, and interdisciplinary projects. To see a full list of their services, visit http://rcc.its.psu.edu.

e. **Indiana University's Cyberinfrastructure Task Force** was established in 2004-2005 to determine how the University should invest in research IT. Their goal was to consider IU's needs for shared cyberinfrastructure investments. Specifically, the group asked scholars to focus on needs that could help support a doubling of IU's externally funded research by 2010-2011. The report to the Vice President for Research & Information Technology conveys ten specific recommendations. View the report: https://scholarworks.iu.edu/dspace/handle/2022/469

f. **Massachusetts Institute of Technology's Research computing task force** and the University Libraries offer a workshop titled "Managing Research Data 101" which is designed to train data producers on best practices to support the latest cyberinfrastructure challenges. The course covered topics such as data management planning for grant applications, metadata, security/backups, file naming, and best practices for archiving.

g. The **University of Texas at Austin** has developed the UT Grid (Boisseau, 2006). The UT Grid is a comprehensive campus cyber-infrastructure project to integrate the numerous and diverse computational, visualization, storage, data and information, and instrument/device resources of The University of Texas at Austin (UT). This joint project between UT Austin and IBM has a focus and approach with important fundamental differences from multi-institution grids and discipline-specific grids. These distinctions, coupled with new locally-developed software for providing both portal and shell-based user interfaces to numerous grid software technologies, will facilitate rapid deployment, adoption, and evolution of UT Grid, while enabling it to serve as a platform for both production computing (for research and education) and grid computing research. The first stages of UT grid are well under way after only two months: the construction of grid user portals and grid user nodes as interfaces, and the integration of serial and parallel computing resources for high-throughput computing.

   At their coordinate campus, the **University of Texas at El Paso** has started the Cyber-shARE Center of Excellence. Cyber-ShARE will bring together experts in computer science, mathematics, and earth and environmental science to develop software applications, services, and other digital tools for gathering and computing data over the Internet for use in scientific research. The center will be an important part of a major National Science Foundation initiative to improve the performance of the United States' cyberinfrastructure through the immense amounts of useful data and high-performance computing power that can be shared by researchers over the Internet.

h. **The Research Data Working Group in Canada** is a multidisciplinary collaborative effort of universities, institutes, libraries, granting agencies, and individual researchers to address challenges and issues surrounding data access and preservation. Their goal is to focus on the necessary actions, next steps and leadership roles for researchers and institutions to ensure Canada's research data is accessible and usable for current and future generations of researchers. The group recently created a Gateway to Scientific Data in addition to numerous resources on data management and curation. This work stems from a 2008 "Gap Analysis" reporting the strengths and problem areas of data curation at the national research level which can be found at http://data-donnees.gc.ca/eng/about/achievements.html. Also mentioned on this Web site is information about the National Consultation on Access to Scientific Research Data (NCASRD), a partnership initiative of the National Research Council Canada, the Canada Foundation for Innovation, Canadian Institutes of Health Research and Natural Sciences and Engineering Research Council (NSERC).

i. The **University of British Columbia** has had many recent developments at this major Canadian university since cyberinfrastructure was first discussed in December 2005. Interestingly, it appears as though the University is leading many national efforts, rather than reacting to them, including:
   - WestGrid, one of seven high performance computing consortia in Canada, part of a national organization called Compute/Calcul Canada. Thirteen of the fourteen partners are educational institutions, and five network partners have joined the effort, which spans the four western Canadian provinces. The goal is to build a user community across

- Canada in disciplines ranging from the sciences and engineering to arts and humanities. Visit: http://www.westgrid.ca/
  Canadian Biodiversity Cyberinfrastructure Network: The Herbarium, Spencer Entomological Collection and Botanical Garden collaborated with a group of university museums and collections from across Canada on a successful proposal to fund a Canadian Biodiversity cyber-infrastructure network. More information: http://www.cbin.ec.gc.ca/
- A University e-strategy to enable students, faculty and staff to "excel in one of the world's leading universities by enhancing learning, research and community through leading-edge technology initiatives. Our activities include project development, information and resource sharing, and setting long term goals for technology at UBC." The purpose is to align technology initiatives with the University's strategic goals. Strategies and tactics include:
  - e-town hall meetings
  - e-strategy governance
  - e-research work groups
- The Incorporation of cyberinfrastructure into several classes and library offerings.

j. In **the United Kingdom, the Digital Curation Center (DCC)** has for years been a leader in government-based eResearch assessment and solutions. The DCC works with federal funding agencies, like the JISC, and helps create mandates and best practices for digital research management. The DCC report *Dealing with Data; Roles, Rights, Responsibilities and Relationships: Consultancy Report* (Lyon, 2007) outlines the priorities and future steps of this UK initiative and their website provides a number of relevant reports including the results of their National Data Infrastructure study: http://www.jisc.ac.uk.

k. In Australia, the cyberinfrastructure effort is headed by a government-based approach to create a **National Data Service**. The ANDS serves to direct national policy on data management issues, outline best practice for data curation, and create a unified collection of unique research data resources from around Australia. Read about the ANDS from the Technical Working Groups proposal "*Towards the Australian Data Commons: A proposal for an Australian National Data Service"* (2007) or more information is available: http://www.ands.org.au.

3. ***Disciplinary subcultures***

An exploration of resources available in the discipline of neuroscience and the broad field of humanities reveals the minutia of achievements and directions that many disciplines and areas of study (i.e., disciplines with similar epistemologies, ontologies, etc.) have undertaken. One question that these examples raise is: What is the balance between disciplinary subcultures around cyberinfrastructure vs. the notion of creating multiple cottage industries? That is, while disciplines have specific cyberinfrastructure needs, is it a good use of resources to create disciplinary solutions? Should national publications, NSF and NIH initiatives, Web sites provide solutions for all disciplines related to cyberinfrastructure? How does cyberinfrastructure balance the rapid growth of 'cottage industries'?

a. Neuroscience (a multidisciplinary example)

Neuroscience is a young, multidisciplinary field, and major funding agencies have launched specific cyberinfrastructure initiatives focused on neuroscience. In addition, many cyberinfrastructure organizations specific to neuroscience have been created in recent years. These organizations offer conferences, funding, and virtual communities, and range from the national to international level:

- ***Neuroinformatics* Journal:** A newly launched publication that is a disciplinary resource for cyberinfrastructure-related topics. *Neuroinformatics* publishes original and review articles, emphasizing data structure and software tools related to data analysis, modeling, integration, and sharing across all areas of neuroscience.
- ***Collaborative Research in Computational Neuroscience (CRCNS):*** This Web site provides a forum for data sharing within the neuroscience community. In addition, the project supplies yearly awards are administered by this section of NSF, which is a joint initiative between NSF and NIH. More information: http://www.nsf.gov/crcns
- ***Cyber-Enabled Discovery and Innovation (CDI):*** NSF's five-year initiative to create revolutionary science and engineering research outcomes made possible by innovations and advances in computational thinking. Results of this effort include building virtual organizations. More information: http://www.nsf.gov/crssprgm/cdi
- ***International Neuroinformatics Coordinating Facility (INCF):*** Fourteen countries established the INCF in 2005, with the purpose of coordinating and fostering international activities surrounding discovering and innovation in neuroscience. The INCF consists of a secretaritat (central facility) and national nodes. In addition to hosting workshops, the INCF also hosts the Congress of Neuroinformatics, focused exclusively on informatics within the neurosciences, such as high performance computing, grid infrastructure, and multi-dimensional data analyses. Visit: www.incf.org
- ***Code Analysis Repository and Modeling for e-Neuroscience (CARMEN):*** An e-Science pilot project funded by the Engineering and Physical Sciences Research Council in the United Kingdom. This project functions as a virtual laboratory for data sharing, and is focused on neurophysiology. Learn more: www.carmen.org.uk
- ***Biomedical Informatics Research Network (BIRN):*** Launched in 2001 as part of the NIH National Center for Research Resources as a collaborative environment for biomedical research and clinical information management. The BIRN is configured as a central coordinating center plus hubs around the country. The University of Minnesota is one such hub. See www.nbirn.net
- ***Neuroimaging Informatics Tools and Resources Discovery (NITRC):*** This initiative is part of the NIH Blueprint for Neuroscience Research, and includes software tools for research, an online community, and funding opportunities. Visit: www.nitrc.org
- ***Neuroscience Information Framework (NIF):*** Part of the NeuroLex community, working in partnership with INCF. (A good example of collapsing the cottage industry.) Visit: nif.nih.gov

b. Humanities (a non-science example)

Cyberinfrastructure is an integral part of digital scholarship in the humanities. Several new initiatives have emerged in this area to help shape and define the future needs of this community. For example:

- The report by the **American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences** titled "Our Cultural Commonwealth" reviews the cyberinfrastructure needs and strategic priorities that pertain to the advancement of the humanities. The Commission recommends eight strategies to overcome the constraints and possible barriers that humanities and social science researchers face in the digital landscape. In order to preserve a cultural heritage, it recommends:

1. Address cyberinfrastructure at the university and college level as a matter of priority through federal funding and academic investment;
2. Develop policies that foster openness and access;
3. Promote cooperation between public and private sectors (i.e. Universities and internet-focused companies);
4. Cultivate leadership in support of cyberinfrastructure;
5. Encourage digital scholarship through training of researchers;
6. Develop new university and national computing centers dedicated to the humanities and social sciences;
7. Develop and maintain community-based standards and tools;
8. Create extensive and reusable digital collections through scholarly-driven digitalization projects.

The Commission's report calls on not only researchers to implement these strategies, but for a strong collaboration among researchers that also includes scholarly societies, university leadership, librarians and archivists, and the commercial and government sector to take action. Access the report at http://www.acls.org/uploadedFiles/Publications/Programs/Our_Cultural_Commonwealth.pdf

- **Mass digitalization efforts** have already begun to impact the scholarship of the liberal arts communities. Google Books is a notable example, but other non-profit initiatives like the Open Content Alliance's (OCA) "Internet Archive" and multi-institutional collaborations such as HathiTrust, are actively harvesting print collections for the cyber-based world presenting humanities researchers with new research opportunities such as text-mining.
- The **Data Curation Education Program (DCEP)** at the Graduate School of Library and Information Science held a Summer Institute for Humanities Data Curation at the University of Illinois at Urbana-Champaign in May 2009. The Center for Informatics Research in Science and Scholarship (CIRSS) hosted the Institute. Data curation is the active and on-going management of research data through its lifecycle of interest and usefulness to scholarship, science, and education. http://cirss.lis.uiuc.edu/CollMeta/dcep/SummerInstituteHumanities.htm
- **Project Bamboo**, developed by researchers from the University of Chicago and the University of California-Berkeley, is attempting to create a distributed system of shared technology services to support the digital humanities (both research and teaching). By looking at how researchers and students engage in digital scholarship, the project will build a user-based support system to promote the use of digital technologies in the humanities. http://projectbamboo.org/
- The **Institute for Advanced Technology in the Humanities** (IATH) is a research center at the University of Virginia that develops tools and cyber-environments for scholarly humanities research. The research projects include a number of digitization efforts.

IV. Cyberinfrastructure at the University of Minnesota

This review provides a background on the recent initiatives at the university which are relevant for developing capacities and plotting a future direction for cyberinfrastructure support on campus. For a visual description of intersecting cyberinfrastructure priorities on campus, see our brainstorming document attached (Appendix E).

1. Transforming the U

In 2006 three reports were created as part of the University of Minnesota's Transforming the U initiative that was connected to Research Cyberinfrastructure. The reports, The Task Force on Research Infrastructure, the Task Force on Collaborative Research and the Administrative Service and Productivity Task Forces and Steering Committee, made several important recommendations, including:

- Ensuring that the research and support aspects of cyberinfrastructure are optimized and integrated across the University for research application as well as production capacity
- Planning for cyberinfrastructure at the University as soon as possible. The Task Force on Research Infrastructure recommended that key stakeholders should be convened immediately to identify key needs and directions.
- (Regarding infrastructure support for collaborative research, which applies to cyberinfrastructure.) Considering how to create information technology required to support collaborative activities, which often is more complex than for non-collaborative activities. IT must provide data sharing to multiple users. Technology could help build collaborative research networks by identifying researchers with expertise or similar research interests.
- Enhancing technology goes beyond equipment and requires a professional, well-trained staff to deliver knowledge and assistance.
- Transforming the "centralized vs. decentralized" administrative structure into a new model of administrative support to the academic enterprise to best use University resources
- Recognizing the University of Minnesota, its campuses, colleges, departments, and units as a single enterprise. Establishing uniform standards and systems to reduce duplication of administrative processes and their associated support structures.

2. Research Cyberinfrastructure Alliance

In late 2007, the Office for the Vice President for Research (OVPR), Office of Information Technology (OIT), and the University Libraries created the Research Cyberinfrastructure Alliance (RCA), a virtual organization intended to foster collaboration between and among these enterprise services and various college-level research computing units. A chief goal of the RCA is to "enable researchers to work together with technologists and professionals with expertise in specialized scientific applications, image processing, life-cycle data management, operation of clusters and parallel systems and algorithm development to plan, coordinate and conduct research projects˝.

Project scope and direction for the RCA was established in a report by consultant Eric Celeste, who interviewed four groups of researchers at the University in order to assess needs and desires for centralized research computing services. These preliminary interviews indicated a great demand for engineering expertise and technology.

- Move toward a layered approach focusing on local service, but more developed in scope than the PEL survey. For example, local service providers should be familiar with project/disciplinary categories and possibly funded off of grants. This creates a 'node- like' configuration. Local staff need to be sufficiently nimble, such that 'pilot' projects can be tested prior to a large launch. However, move from the pay-as-you-go model. IT people need to be involved during the life of the project. Can build CI into grants in a predictable manner, and seed projects should be explored.
- University investment in cyberinfrastructure is necessary to be top three in this area.

- Most research groups do not have their own support staff to support data storage, software, backup, etc.
- Researchers that generate large amounts of data are having to sacrifice data storage needs, because University servers cannot keep up with the amount of data generated. This is not helpful for the scientific process/data management plan, as important data may be lost.

3. Office of the Vice President for Research (OVPR) 6-year Research Cyberinfrastructure Plan (Senate Research Committee, 2009)

In 2009 the OVPR released a 6-year cyberinfrastructure plan that aimed to align current and future cyberinfrastructure investments with the university's goals in order to implement a long term research infrastructure planning process. This report derived from assessments of campus research assets and through listening session with faculty and found:

1. An increasing need for computing and imaging;
2. An increasing integration across disciplines;
3. A need for a global view of research;
4. Priorities should be for the greatest good rather than return on investment;
5. Research clusters could be used. An example is spatial analysis. This allows better coordination, broad researcher involvement, and big picture.

Principles defined by the report included:

1. Promote shared infrastructure over individual set-ups;
2. Increase capacity for sharing remotely and use across all campuses, as well as with other research facilities and universities;
3. Adopt a portfolio approach to provide diverse types of funding;
4. Prioritize infrastructure that is of system-wide value and requires system-wide financing;
5. Develop criteria for inclusion of infrastructure priorities.

Data and ppt available here http://www.research.umn.edu/infrastructure/

4. Office of Information Technology (OIT)

In 2008 OIT unveiled a 6-year its six-year information technology plan to identify and invest in those technology projects and services that are truly transformative to the University while maintaining and leveraging the infrastructure necessary to support them. This plan used the University's seven criteria for decision-making to articulate enterprise technology investment principles, the idea of six-year governance groups, and priority projects. While this report was very high-level and not mired in detail, ideas outlined in the plan are very much in line with the other University reports listed in this section.

5. University Libraries E-science and Data Services Collaborative

In June 2008, the libraries created an E-science and data services collaborative that addressed the needs identified in the sciences assessment. The group had just released a report "Data stewardship at the University of Minnesota" which identifies the near and long-term strategies of the libraries cyberinfrastructure efforts and supplemental support services. Follow-up assessment on data needs found here: http://www.lib.umn.edu/about/scieval/Sci%20Eval%20DataResources.pdf

The libraries completed a user-needs study of researchers in 2006 titled "Understanding Scientists: A University Libraries Sciences Assessment" (Marcus, 2007). The report describes:

1. Interdisciplinary & collaborative research obstacles, including: a lack of access to shared resources between collaborators; a wide variety of preferred communication methods; and lack of knowledge on how to use technology-based communication (teleconferencing, etc.);
2. Data storage needs such as: ease of access; solutions for shared data across disciplines; lack of education on database development and data sharing; discipline-based sharing restrictions/protection arisen out of competitive nature of research;
3. Data management misunderstandings. For example, when asked "How long do you need to keep your data" some believe that their data will not be relevant in 20 years others feel that all data must be preserved indefinitely.

V. Survey Methodology and Findings

During March 24th-April 8th, 2009, our team conducted an online user-needs survey to faculty, research staff and students asking them to report the current state of cyberinfrastructure support at the University as well as to assess their future needs. The survey comprised of 130 questions (see full survey instrument in Appendix F) on the following cyberinfrastructure topics: data storage, data management, and networking infrastructure; collaboration with other researchers; tools and applications; high performance computing; and learning and workforce development, as well as topics within each of these areas.

Following consultation with the Office of Measurement Services and a pilot survey tested on members of the Senate Committee on Information Technology, the survey was delivered by email (see Appendix F) to 8,424 researchers at four University of Minnesota campuses. Researchers were selected based on their primary role (i.e., Faculty, postdoc, researcher) and/or their involvement with cyberinfrastructure (i.e., PI on current grants). In an effort to increase the survey response rate, each respondent was entered into a prize drawing for one of five $50 VISA gift cards; early respondents within the first week were entered twice.

Respondent Demographics

After two weeks, our survey generated 780 successful responses out of a pool of 8,424 selected researchers (a 9.3% response rate) from a broad cross-section of affiliates representing various roles (Fig. 1), research disciplines (Fig. 2), environments (Fig 3), and campuses (Fig. 4).

Fig. 1: Primary Role

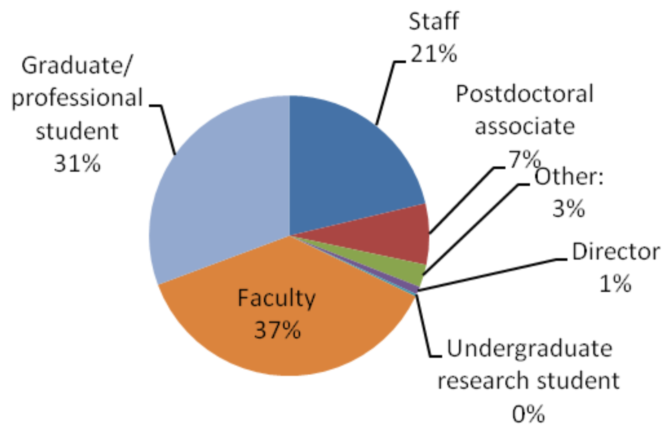


Fig. 2: Research Disiplines

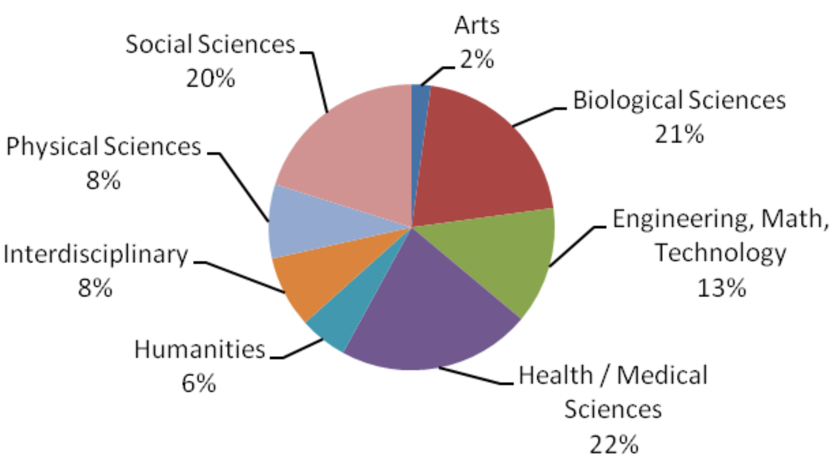


Fig. 3: Research Enviroment

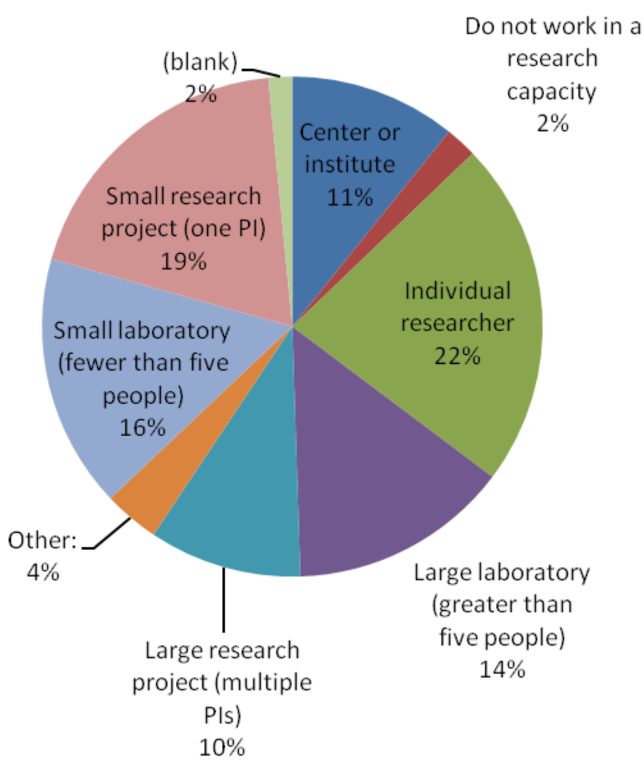


Fig. 4: Campus Affiliation

UMN Campus	Researchers Solicited	Successful Responses	Campus response rate
Twin Cities	7767	708	9.1%
Duluth	493	50	10.1%
Morris	102	9	8.8%
Crookston	41	6	14.6%

Survey Question Results

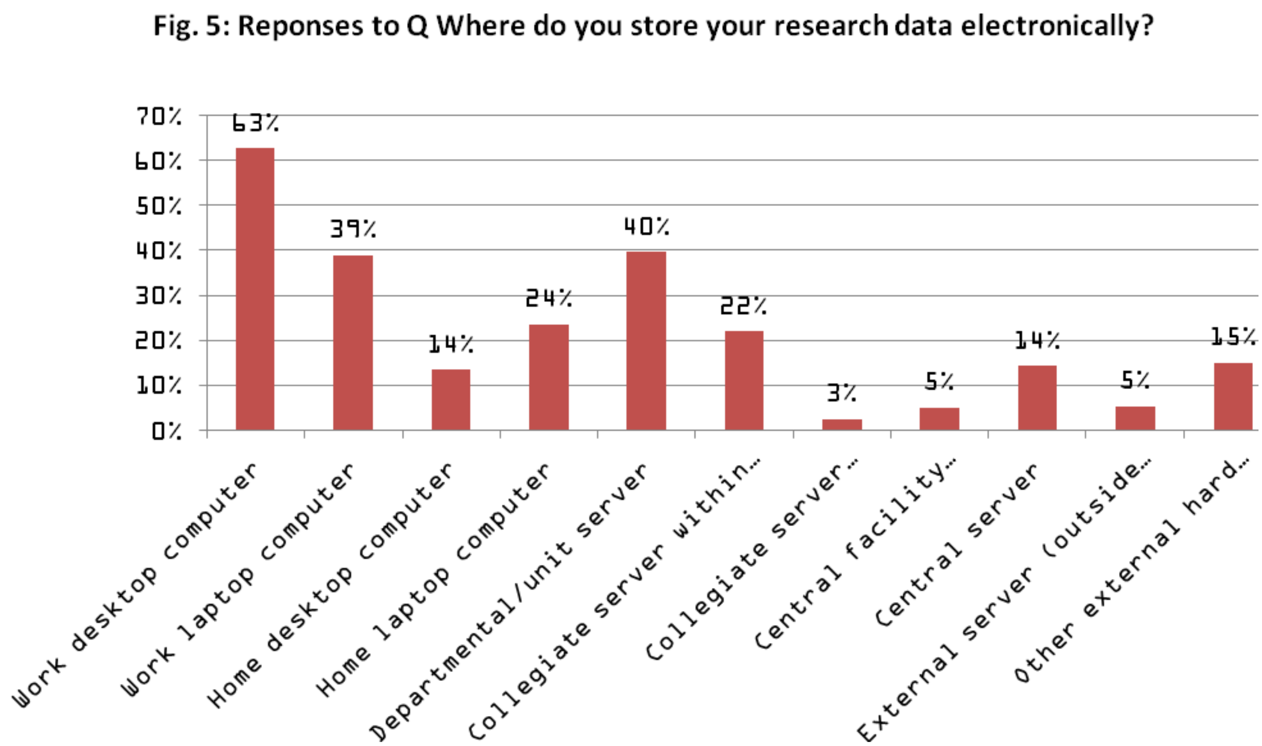
The results and comments received in the user-needs assessment (see all result data and accompanying graphs in Appendix G) demonstrated several trends which became the driving force of our recommendations. These trends, discussed in detail below, include: data storage and backup, local point of contact, data use, terminology, and interdisciplinary practices.

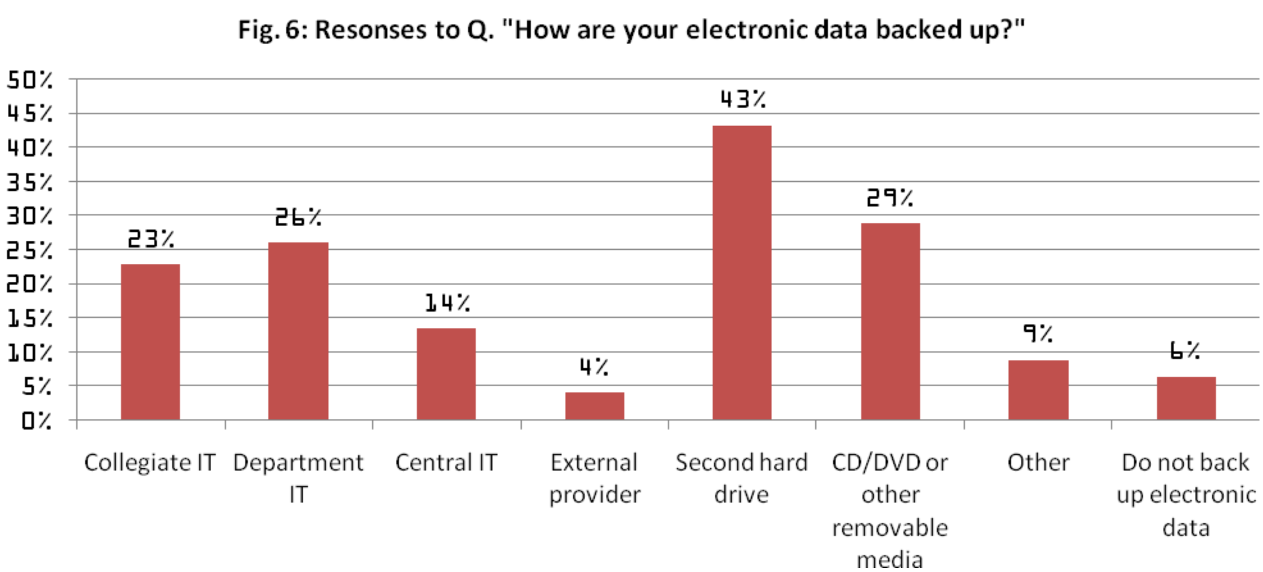
Trend 1: Storage and backup

Local data storage access (Fig. 5) and back-up support (Fig. 6) is more commonly used over University- or cloud-level storage. However comments reflect a need/desire for more central support.

What they are doing: Only 14% of respondents use central data storage solutions. More prevalent were work computer desktops and laptops (63% and 39% or users) and departmental servers (40%). External hard drives and other media were also noted by 15% of respondents. Back-up solutions show a preference for local and "low-tech" storage. Nearly half of respondents (43%) reported using a secondary hard drive to back-up their data, and 29% use CD/DVD or other removable media. Interestingly, 72% of researchers have never lost data due to the lack of a back-up (27% have).

Compared by Discipline: Storage solutions and backup were consistent across all disciplines but vary by user type. Grad students more likely to store data on personal computers. Faculty are more likely to store data on Central.





What they say:

> "In my research group, research data is mainly managed by the respective owner. We evaluated the file sharing service from central IT but found it too cumbersome to use. It works best if research groups can set up their own servers to have control over their data storage platform." -postdoc in the Institute of Technology.

> "There needs to be a mechanism for centralized storage of data with high security for individual research groups. This at least partially removes the responsibility for secure backups from the research group alone."
>
> -Ph.D. Research Assistant, Institute of Technology

From this trend grew PEL Recommendation 1: Develop enterprise-wide, integrated cyberinfrastructure to align with national efforts. (Discussed more fully in Section VI.)

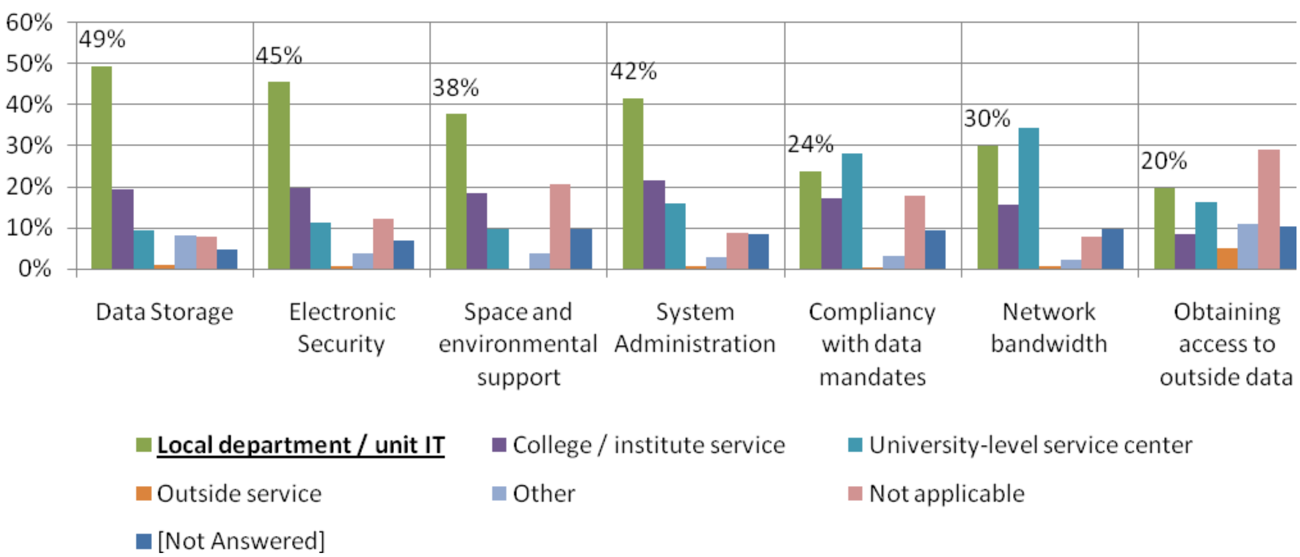
Trend 2: Local point of contact

Local IT department/unit contact is primary end-user resource (Fig. 7) for research needs including storage, electronic security, space & environment, and system administration. Central University level services are contacted for compliance & privacy issues, and bandwidth needs. Outside data services are not needed for many.

Compared by Discipline: Health Sciences are more likely to contact university-level service for support (15% vs~5%). Could be an artifact of how AHC views "university level". Outside data services are not needed particularly in the Health Sciences.

What they are doing: A majority of respondents access the following resources and support through their "local IT point of contact" based in their department/college/institute: Data storage (68%), electronic security (65%), space/environment (56%), and systems administration (64%). However local connections are much more split with University level service centers when it came to compliance & privacy (local, 41% vs. university-level, 28%), bandwidth (local, 46% vs. university-level, 34%), and obtaining outside data (local, 26% vs. university-level, 16%).

Fig. 7: Responses to Q: "Who is your primary contact for providing each of these services?"



The extent to obtain access for the following resources also indicated a preference (very large to moderate) for local support in the lab and departments for data storage (83% and 55% respectfully), virtual collaboration (64% and 41% respectfully), software and tools (85% and 55% respectfully), and, to a lesser extent, high performance computing (47% and 23%).

What they say:

> "It's hard to know who to contact in the tech area for some research tech support. The basic tech person assigned to the dept. is fine for routine computer and computer related software things, but beyond the basics it's not clear who to contact for what."
> -Faculty Member in the College of Liberal Arts

> "As far as I know, there is no inventory for data storage and we don't have a contact person for data storage. We just save our data in multiple places and on an external hard drive. There is no oversight for data storage."
> - postdoc, School of Dentistry

From this trend grew PEL Recommendation 2: Better position data storage, management, and networking services around the end-user. (Discussed more fully in Section VI.)

Trend 3: Data use: storage, access, and retention

The amount of data storage may be driven by a high demand from a small number of users from all disciplines. Frequency of data access is daily by most researchers and nearly all in the physical sciences. Most researchers want to keep their data forever.

What they are doing: The majority of researchers, 80%, reported generating less than 1 GB of data per week (Fig 8). Only 19% of respondents do generate greater than 1GB of data per week. Data are accessed (Fig. 9) daily by 64% of researchers, weekly by 25%, and monthly, quarterly, yearly at 5%, 3%, and 1% respectfully. Data are typically stored indefinitely by 70% of the population and are shared with departmental colleagues.

Fig. 8: Responses to the Q: "How much data do you generate each week?" answered by those who "generate more than 1GB/week."

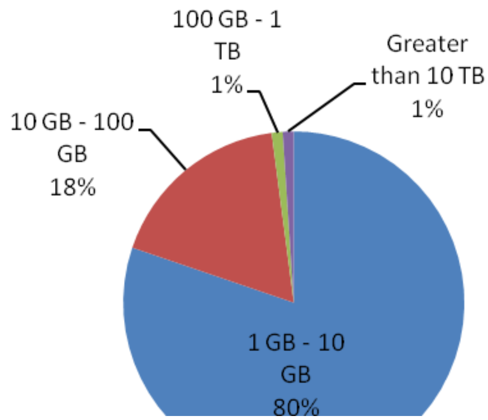
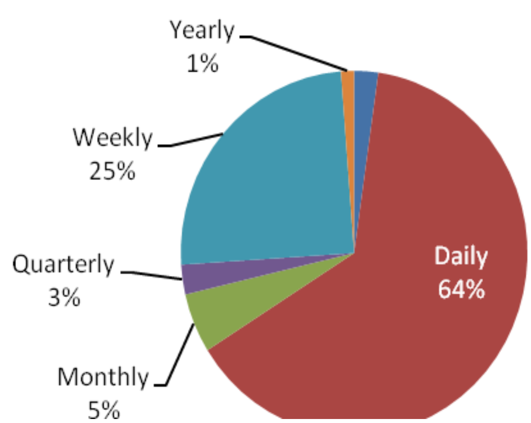


Fig. 9: Responses to the Q: "How often do you access your electronic data?"



Compared by Discipline: Of the 19% who do generate over 1 gigabyte of data per week, interestingly the rate of responses by discipline did not suggest a significant difference between Biological (21%), Physical (15%), Engineering (14%), Health (19%), Interdisciplinary (11%), and Social Sciences (15%), as each ranged between 11-21%. Humanities and Arts represented 2% respectively.

Humanities researchers access their data as frequently (75% daily) as Biological and Engineering Researchers. Daily access is most consistent in the Physical Sciences (89%), and by many in Arts (56%), Biological (75%), Engineering (71%), Health (59%), Humanities (75%), Interdisciplinary (62%). Social Sciences access their data daily (48%), weekly (33%), or monthly (10%)

What they say:

> "Once or twice a year I have to negotiate for storage space; I always feel as if my research is at odds with IT policy. I always feel as if I'm living on borrowed time, no confidence in having access to adequate data storage for research in the future."
> -Faculty Member in the College of Education & Human Development

> "Instead of looking at individual requirements for storage, management, software and security take a look at how CLA now manages its research services. There is one single point of contact , they provide space, security, management, software -- all you do is apply and you get an account WITH phone or email support!" -Faculty Member in the College of Liberal Arts

From this trend grew PEL Recommendation 3: Improve and expand the University's ability to handle data. (Discussed more fully in Section VI.)

Trend 4: Terminology

Humanities researchers surveyed felt that this survey and accompanying terminology (i.e. cyberinfrastrucuture) did not 'speak' to them; however, humanities have self-identified cyberinfrastructure needs. These trends were evident in the comment section of the survey (see Appendix F).

What they say:

> "This instrument seems to have been constructed using the sciences as its template. I find it hard to translate the terms into my research in the Humanities." - Faculty member in the College of Liberal Arts

> "This survey is largely geared to researchers outside the Humanities. Consequently, many if not most of the questions are stretch -- in terms of relevance. Please find a way to incorporate a genuine commitment to research in the humanities (without which no university with which we aspire to compete, either public or private, is ranked highly)." - Faculty member in the College of Liberal Arts

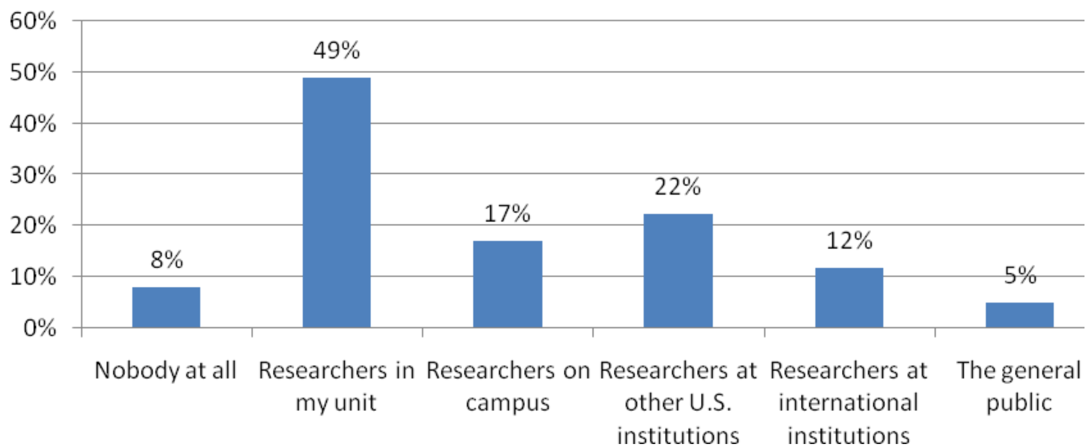
From this trend grew PEL Recommendation 4: Re-brand Research Cyberinfrastructure at the University to improve accessibility for all disciplines. (Discussed more fully in Section VI.)

Trend 5: Interdisciplinary practices

Data is shared by nearly all researchers, most share locally on campus. Department data sets and resources are not well documented. The majority of all disciplines did not have or know of any documented inventory of data sets.

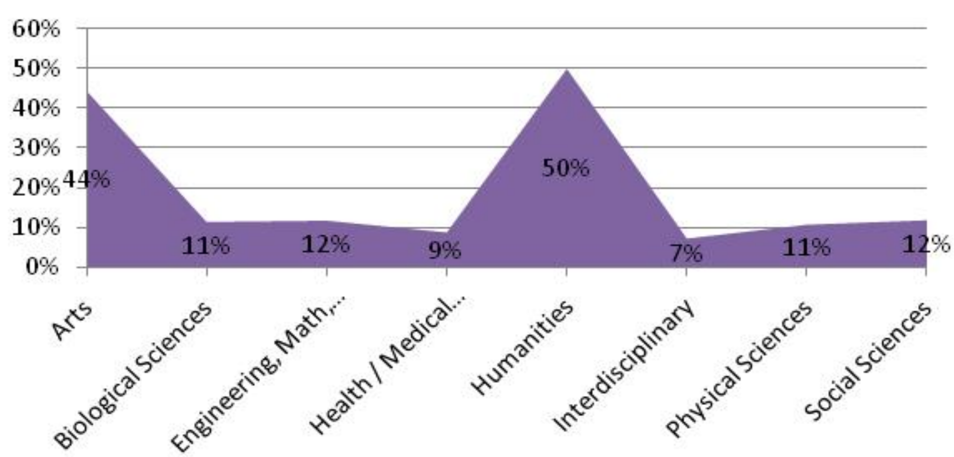
What they are doing: Data is shared by 92% of all researchers, but primarily with researchers in their own unit or on campus (51%, 18% respectively). Only 8% of researchers do not share their data at some level. Of those, 52% would share their data, 30% would not and 19% did not know.

Fig. 10: Responses to the Q: "With Whom Are You Sharing Your Data?"



Compared by Discipline: The majority of disciplines share their data in statistically balanced ways with Biological and Health Sciences leading the way with sharing within their unit. Of those researchers, 8% of respondents, who are not sharing their data, the arts and humanities appear to stand apart from the rest with 44% and 50% of respectively (fig. 11).

Fig. 11: Discipline breakdown of "Nobody at all" responses of Fig. 10.



What they say:

> "International collaboration will increase; how do we encourage & assure compliance with security standards from our collaborators?" -Faculty in the College of Education and Human Development

From this trend grew PEL Recommendation 5: Create incentives for virtual collaborative relationships. (Discussed more fully in Section VI).

VI. Recommendations for Action

A note on the recommendations: We encourage the RCA to prioritize these recommendations based on the criteria for decision-making outlined in the University's strategic positioning process:

1. Centrality to Mission
2. Quality, Productivity, and Impact
3. Uniqueness and Comparative Advantage
4. Enhancement of Academic Synergies
5. Demand and Resources
6. Efficiency and Effectiveness
7. Development and Leverage of Resources

Also, three strategic questions can help to guide the process of recommendation to action:

- What should be the University's strategy to develop research computing resources in the next five years?
- What are ways to forge a common cyberinfrastructure with central and distributed support? What are best practices for articulating central and distributed support? How can units across the University best work in partnership?

- How can the University compete at the national level for cyberinfrastructure grants and resources?

PEL RCA Project Recommendations

From the survey results and the trends outlined in the Methodology and results section we drafted the following recommendations for implementing cyberinfrastructure for the 21st century research. By implementing these recommendations we feel the University will be much more competitive in the research landscape.

PEL Recommendation 1: Develop enterprise-wide, integrated cyberinfrastructure to align with national efforts.

Accountability at the senior leader level is needed to ensure the University's competitiveness in the future (i.e., obtain grants, become a top three public research university, recruit, evolve research and establish new programs). At the same time, the University has the opportunity to leverage existing 'cottage industries' and to align internal services.

Strategies:

- Create partnerships between collegiate IT directors and end-users in their college to identify local contacts. Directors will then work with local contacts to inform them of new services and offerings.
- Assist end-users to help them understand better data storage solutions and collaborative opportunities through relationships with local points of contact.
- Refocus the role of the collegiate IT director into a local point-of-contact: concierge/liaison mode. Leverage end-user behavior (i.e., prefer to go to their local point-of-contact) to drive up-to-date information from central through the collegiate IT director, have the collegiate IT director hold consultations with end-users, etc.
- Leverage existing permanent cottage industries at the University. Discurage the transient cottage industries (i.e., a PI will fund someone's salary for two years off of a grant, purchase a large microscope, and then the research scope becomes obsolete and the salary funding dries up). Transient cottage industries may have their place in a limited capacity, but not in a widespread manner across the University system. EDUCAUSE's findings support this.
- Move from cottage industry structure to system-wide solutions. Integrate cyberinfrastructure needs with other services for end-users. Focus on the integration of multiple cottage industries that end-users need, rather than repairing only one (i.e., a systems approach to cottage industries). This has the potential to lead to significant cost-savings for the University.
- Develop system-wide University policy to streamline and direct the ongoing efforts of SPA, Collaborate, RCA, OVPR for benefit of the end-user.

PEL Recommendation 2: Better position services around the end-user.

Provide IT contacts with the tools to share and leverage existing resources across University service providers. Researchers do not care who offers the service; they just want the opportunity to take advantage of the offering.

Strategies:

- Focus on what the end-user needs from the time of initial grant planning forward.
    - As part of the new collaborative planning process, collegiate IT staff should work with researchers early in the process. EDUCAUSE findings support this, and indicate that IT experts want to collaborate with end-users early in their grant planning process.
    - Develop a promotional handout for IT staff: "Questions to ask researchers about their IT needs."
- Respond to clusters identified in the recent OVPR 6-year research infrastructure survey. Leverage the clusters as end-user service provider content areas. IT staff support should represent each of these areas.
- Consider a visual model for the current and future positioning of services for the end-user. Are end-users at the center of the new model or the edge, with the local IT contact at the center (See Appendix G).

PEL Recommendation 3: Improve and expand the University's ability to handle data.

In order to meet the growing needs of researchers, IT service providers should provide short-term data storage as needed (i.e., a bank model), utilize external cloud storage for long-term archiving, and provide high speed data connectivity and automatic back-up, with the appropriate compliance and privacy mechanisms.

Strategies:

- Create flexible storage, not individual end-user capacity, similar to as-needed money in a bank (i.e., banks do not have on hand the dollar amount of all bank investments; rather, they have sufficient cash to whether withdrawals).
- Create local servers that are funded per grant, provide instant data access to researchers, and allow researchers to have on-demand control/availability, with customizable interfaces. These local servers must connect to central cloud storage with high-speed, automatic back-up, necessary compliance/privacy, and that facilitates local sharing and long-term archiving.
- Utilize the cloud storage capacity currently available from the Committee on Institutional Cooperation (CIC).
- Investigate off-campus options from the private sector (i.e., Amazon) taking into account scalability, high-availability, low-latency (fast) at commodity costs.

PEL Recommendation 4: Re-brand Research Cyberinfrastructure at the University to improve accessibility for all disciplines.

In order to bridge the jargon gap among the research community, the IT service providers and University leaders should re-brand cyberinfrastructure to ensure efficient workforce training and awareness of technology needs.

Strategies:

- Eliminate the term cyberinfrastructure. An appropriate alternative is e-research or some of the British terms. This suggestion is justified by EDUCAUSE's survey results.

- Create dedicated IT points-of-contact within the disciplines or clusters. Currently, accessibility is a large problem, as no one point-of-contact exists, and establishing in-roads is very challenging.
- Create and maintain a glossary of commonly used cyberinfrastructure terms/jargon for campus end-users. This information should appear on a site with grant application information, along with a list of local points-of-contact. In addition, this glossary will help to bridge the gap between University end-users and cyberinfrastructure service providers.
- Educate and provide awareness on cyberinfrastructure to engage end-users and University leader. This could be done at New Faculty Orientation, graduate/professional student orientation, Responsible Conduct in Research Seminars, etc.

PEL Recommendation 5: Create incentives for virtual collaborative relationships.

Give researchers the necessary tools to help develop collaborative relationships and facilitate local and external data sharing in order to establish the University as a top research institution.

Strategies:

- Create a universal platform for data sharing within the University, and one that interfaces well with external audiences. This platform may need to be modified slightly within each major area of research (i.e., social sciences, humanities, science). End-user focus groups and conference calls with external audiences will be instrumental in developing this.
- Interface with NIH and NSF to learn whether they have developed such tools. If so, ask for the University to be a pilot in using such tools.
- Train researchers in the use of these tools (note: tools should be very easy to use and intuitive, thus requiring minimumal training). Without training, researchers likely will not use these tools.
- Provide internal incentives for collaboration (e.g., help from SPA, compliments in OVPR newsletter, compliment from dean). However, this may need to be a stick, rather than carrot, model. That is, this may need to become a mandatory part of grant applications through SPA.

VII. References

"ACLS Commission on Cyberinfrastructure." http://www.acls.org/programs/Default.aspx?id=644.

ANDS Technical Working Group. (2007). *Towards the Australian Data Commons: A proposal for an Australian National Data Service*, Canberra, October 2007, p24. http://www.pfc.org.au/twiki/pub/Main/Data/TowardstheAustralianDataCommons.pdf

Berman, F. (2008). Got data?: a guide to data preservation in the information age. *Communications of the ACM. 51* (12), pp. 50-56.

Blue Ribbon Task Force on Sustainable Digital Preservation and Access (2008). Sustaining the Digital Investment: Issues and Challenges of Economically Sustainable Digital Preservation, Interim Report. http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf

Boisseau, John R. (2004). "UT Grid: A Comprehensive Campus Cyberinfrastructure," *13th IEEE International Symposium on High Performance Distributed Computing (HPDC-13 '04),* pp. 274-275.

Butler, D. (2007). Agencies Join Forces to Share Data . *Nature*, *446* (7134), 354. doi: 10.1038/446354b.

Fox, Louis, Cara Lane, Ed Lazowksa, with Janice Fournier, Greg Koester, and William Washington. "Information Technologies for eScience: A Preliminary Report from the University of Washington" (Research Bulletin, Issue 6). Boulder, CO: EDUCAUSE Center for Applied Research, 2009, available from http://www.educause.edu/ecar.

Furlani, C., Greer, C., & Romine, C. (2009, January 29). Interagency Working Group on Digital Data. Washington, DC. Retrieved February 4, 2009, from http://sites.nationalacademies.org/xpedio/groups/pgasite/documents/webpage/pga_047703.pdf.

Lougee, W. et al. (2007). Agenda for Developing E-Science in Research Libraries: *Final Report and Recommendations to the Scholarly Communications Steering Committee, the Public Policies Affecting Research Libraries Steering Committee, and the Research, Teaching, and Learning Steering Committee.* ARL, November 2007. http://www.arl.org/bm~doc/ARL_EScience_final.pdf

Lyon, L. (2007). *Dealing with Data; Roles, Rights, Responsibilities and Relationships: Consultancy Report*. UKOLN, University of Bath. June 2007, page 54. http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/reports/dealing_with_data_report-final.pdf

Marcus, C. et. al. (2007). Understanding Research Behaviors, Information Resources, and Service Needs of Scientists and Graduate Students: A Study by the University of Minnesota Libraries. URL http://www.lib.umn.edu/about/scieval/Sci\%20Report\%20Final.pdf

University of California Press Release. (2009). Cyberinfrastructure for Ocean Observation Tap, 9/2/2009. Accessed from http://www.universityofcalifornia.edu/news/article/21839.

VIII. Appendices

A. PEL Project Proposal
B. National Science Foundation
C. Microsoft 2020 Science
D. EDUCAUSE Survey
E. UMN Intersecting Cyberinfrastructure Efforts Map
F. PEL Survey Instrument
G. PEL Survey Result Data, Graphs, Handouts, and Poster

Appendix A. PEL Project Proposal

President's Emerging Leaders Project Proposal Implementing Cyberinfrastructure for 21st Century Research 2008-2009

As a result of innovative partnerships between the Office of Information Technology, collegiate units, and other central units, outstanding technological support systems for administrative and academic needs have been developed to serve faculty, staff and students. Central and local IT support staffs maintain servers and desktop computers to ensure availability and security of information technology tools and resources. The University's Digital Media Center, along with academic technology support staffs in local units, provides a variety of services to instructors and students using technology in teaching and learning.

Dramatic advances in information technology make way for exciting new opportunities for research as well. New technologies allow projects to span disciplines and institutions, enabling researchers to seek new answers to critical questions in ways that were impossible just years or even months ago. Indeed, research computing is a top priority for leading universities and research institutions around the globe; furthermore, *cyberinfrastructure*[1] is seen as a key factor in securing research funding and attracting and retaining top faculty and students.

The expanded research agendas in many disciplines are now outpacing the computing resources available to individual researchers, departments, or even institutions. Enabling this research requires large-scale investments in high-performance computing, storage and networking, as well as the development of cyberinfrastructure to integrate these components into a meaningful whole.

Cyberinfrastructure includes the instruments, sensors, high performance computational systems, massive storage systems, data resources, and visualization facilities, tied together by high speed networks and made to work together by advanced software to accomplish goals that would not be possible by any single information technology system. It also includes the people, processes, training, security, policies, and capabilities to sustain the systems and networks over time. Implementing cyberinfrastructure requires a high level of coordination and collaboration between researchers and an information technology workforce with expertise in scientific computing.

In support of the University's goal to become one of the top three public research universities in the world, a strategic alliance to support research cyberinfrastructure was developed in Fall 2007 with the vision of facilitating access to high-end, high quality, efficient and effective research computing systems and services, enhancing interdisciplinary research and allowing researchers to explore radically new concepts, approaches, and tools. Initial sponsors of the University's Research Cyberinfrastructure Alliance (RCA) included the University's Office of Information Technology (OIT), the Office of the Vice President for Research (OVPR), and the University Libraries. Among the founding members were individuals from college-level research computing units, the Minnesota Supercomputing Institute, and key leaders from OIT.

As an alliance, the RCA brings key groups together to coordinate and align distributed strategic research cyberinfrastructure and related services to benefit researchers. The alliance enables researchers to work together with technologists and professionals with expertise in wide-ranging areas to plan, coordinate, and conduct research projects. By leveraging resources in this manner, we improve efficiency and effectiveness, and also guide future investments in the best manner possible for the entire University community.

The RCA functions as a virtual organization of research support and information technology service providers, working to identify a portfolio of research technology resources and support services throughout the University. A key next step for the RCA is to develop strategies that enable researchers to easily find and engage with relevant research computing resources and support.

Strategic Questions
The PEL team will work together with the RCA and other stakeholders to discover and recommend:

- What are the strengths of the University in research computing?
- How are colleges and research centers meeting the needs of their researchers?
- What issues are researchers dealing with in research computing?
- What should be the University's strategy to develop research computing resources in the next five years?
- What are ways to forge a common cyberinfrastructure with central and distributed support? What are best practices for articulating central and distributed support? How can units across the University best work in partnership?
- Who are key stakeholders across the campus that need to be engaged to spread the vision of CI?
- How can the University compete at the national level for cyberinfrastructure grants and resources?

Project Goals and Outcomes

The primary role for the President's Emerging Leaders will be to work with key stakeholders (faculty, staff, students, collegiate IT Directors, OIT staff and the RCA group) to develop a strategic approach to cyberinfrastructure that leverages partnerships across the University. We have identified the following goals for the project:

1. Investigate how research computing is currently supported at the University of Minnesota; compare and contrast these findings with other research institutions;
2. Collect information about the needs and priorities of researchers using or desiring to use research cyberinfrastructure;
3. Develop an outreach process to inform researchers of cyberinfrastructure services and resources that are available;
4. Identify cost factors, financial aspects of cyberinfrastructure services, and what makes the most financial sense for the University to optimize research productivity.

As the RCA, we look forward to partnering with the PEL group to refine these goals and identify additional strategic goals for the project.

At the end of the project, the following outcomes will be among the indicators of success:

- All faculty at the University are aware of research cyberinfrastructure resources and support services
- Major granting agencies (NSF, NIH, etc.) are aware of the University of Minnesota's cyberinfrastructure alliance efforts and thus increased strategic support of our researchers
- An increased number of grant applications that take advantage of cyberinfrastructure resources and services are submitted
- Metrics are identified and deployed to track how a coordinated, aligned research cyberinfrastructure is making an impact

Approach

Phase I, Assessment and Information Gathering: In this phase, project participants will survey University researchers, document local research computing resources and investigate cyberinfrastructure efforts at other institutions and at the national level
Phase II, Development: Working with RCA members, project participants will use the data gathered in Phase I to identify and make recommendations for actions, projects and tools to work towards the strategic goals identified above
Phase III, Implementation: If time permits, the PEL team will identify and make recommendations for initiation of pilot cyberinfrastructure projects based on the work in Phase II, and work with RCA members to develop long-term strategies and projects

Project Sponsors:

- Ann Hill Duin, Associate Vice President and Associate CIO
- Bernard Gulachek, Senior Director, Office of Strategy Management, OIT

Project Leads:

- Kemal Badur – College of Liberal Arts, Research Computing and Engineering Services
- John Sonnack – OIT Office of Strategy Management

Proposed Advisors:
Members of the 2008-09 University of Minnesota Research Cyberinfrastructure Alliance

---

Appendix B: National Science Foundation

Broadening Participation and Increasing Access **found at http://www.nsf.gov/pubs/2009/nsf09052/nsf09052.jsp?govDel=USNSF_25**

The NSF is building a comprehensive infrastructure to capitalize on dramatic advances in information technology for computation-enabled discovery. NSF's mission for cyberinfrastructure (CI) is to:

- Develop a human-centered CI that is driven by science and engineering research and education opportunities;
- Provide the science and engineering communities with access to world-class CI tools and services, including those focused on: high performance computing; data, data analysis and visualization; networked resources and virtual organizations; and learning and workforce development;
- Promote a CI that serves as an agent for broadening participation and strengthening the nation's workforce in all areas of science and engineering;
- Provide a sustainable CI that is secure, efficient, reliable, accessible, usable, and interoperable, and that evolves as an essential national infrastructure for conducting science and engineering research and education; and
- Create a stable but extensible CI environment that enables the research and education communities to contribute to the agency's statutory mission.

NSF's guiding principles are that investments will be science-driven, recognize the uniqueness of NSF's role, provide for inclusive strategic planning, enable U.S. leadership in science and engineering, promote partnerships and integration with investments made by others in all sectors, both national and international, and rely on strong merit review and on-going assessment, and a collaborative governance culture. Strong partnerships involving other federal agencies, universities, industry and state government also are critical to success.

NSF also will promote resource sharing between and among academic institutions to optimize the accessibility and use of high performance computing assets deployed and supported at the campus level. Further, resource sharing is a trend that is expected to see significant growth in the very near future as advances in sensors and sensor networks, high-throughput technologies, and instrumentation, automated data acquisition, computational modeling and simulation, and other methods and technologies materialize.

The anticipated growth in both the production and repurposing of digital data raises complex new issues of stewardship, curation and long-term access. Responding to the challenges and opportunities of a data-intensive world, NSF will pursue a vision in which science and engineering digital data are routinely deposited in well-documented form, are regularly and easily consulted and analyzed by specialist and non-specialist alike, are openly accessible while suitably protected, and are reliably preserved. At this time it has become clear that increasingly capable hardware is not the only requirement for computation-enabled discovery. Sophisticated software, visualization tools, middleware and scientific applications created and used by interdisciplinary teams are critical to turning flops, bytes and bits into scientific breakthroughs.

NSF has a set of goal and strategies that are guided by several foundational principles, intended to promote and foster CI maturity across the United States. These include developing CI community, promoting exchange-of-service, coordinating investments, broadening access to resources, supporting robust software development ecosystems, supporting emerging trends in virtual communities, fostering innovation among forefront data management and distribution systems, realizing the full potential of investments in tools and large facilities through use of CI, investing in the high-risk/high-gain research to ensure continue growth in CI forefront, and providing a comprehensive CI flexible, extendable architecture and framework that emphasizes interoperability and open standards.

NSF will focus on new frontiers in science and engineering through networked resources and virtual organizations (VOs) such as experimental facilities and field equipment, distributed instrumentation, sensor networks and arrays, mobile research platforms, HPC systems, data collections, sophisticated analysis and visualization facilities, and advanced simulation tools. Virtual Organizations (VOs) are revolutionizing science and engineering research and education.

Three elements of VOs are essential: the creation of a common technological framework that promotes seamless, secure integration across a wide range of shared, geographically-distributed resources; the establishment of an operational framework built on productive and accountable partnerships developed among system architects, developers, providers, operators, and end-users who span multiple communities; and the establishment of an effective assessment and evaluation plan that will inform NSF's ongoing investments in cyberinfrastructure for the foreseeable future.

Cyberinfrastructure is enabling powerful opportunities to collaborate, model and visualize complex scientific and engineering concepts, create and discover scientific and educational resources for use in a variety of settings, both formal and informal, assess learning gains, and personalize learning environments. These changes both demand and support a new level of technical competence in the science and engineering workforce and in citizenry at large.

Appendix C: Microsoft 2020 Science

The report also focuses largely, on the biological sciences broadly defined, from molecular biology to systems biology to organic biology and ecosystems science because it is in the natural sciences where the greatest impact of computer science will be felt.

It is clear that science has the potential to have an unprecedented impact on our world in the 21st Century, from how long we live, to how we live, to what we know about ourselves, our planet and the universe, to understanding how to control and eradicate disease, to how to protect the entire life-support systems of the earth. As a consequence, it is difficult to overestimate how profound is the scientific revolution now under way. CI will accelerate the pace of scientific discovery now that we are unconstrained by many of the restrictions imposed by print.

1. We are at a point where we are taking a leap from the application of computing to support scientists to ‘do’ science to the integration of computer science concepts, tools and theorems into the very fabric of science.
2. Conceptual and technological tools developed within computer science are starting to have wide-ranging applications in sciences investigating complex systems, most notably in biology and chemistry.
3. Computer science concepts and tools in science form a third, and vital component of enabling a ‘golden triangle’ to be formed with novel mathematical and statistical techniques in science, and scientific computing platforms and applications integrated into experimental and theoretical science. This combination is likely to accelerate key breakthroughs in science and benefits to society.
4. There is an immediate and important challenge to address end-to-end scientific data management, from data acquisition and data integration, to data treatment, provenance and persistence.

5. Scientific publishing is being affected. Even near-term developments in the computing infrastructure for science which links data, knowledge and scientists will lead to a transformation of the scientific communication paradigm.
6. Today's CI is the starting point for fundamental new developments in biology, biotechnology and medicine, but also for potentially profound developments in the future of computing (Reversal affect).
7. Finally, the education of tomorrow's scientists will need to be completely computationally and mathematically literate. Education policy and funding need to address this right now.

The principles of natural information processing methods are yet to be fully understood, but progress in the biosciences continually unveils more detail. What is known already informs the development of molecular computing concepts. The course of computation in nature's molecular computers is directly driven by the physicochemical properties of the materials that implement the computation. Through these largely stochastic processes, macromolecules can provide much more powerful components than conventional silicon architectures. The combinatorial building block principle for assembling specialized macromolecules offers an inexhaustible set of basic functions. The new computer revolution will complement and extend, not overthrow, established methods. Enabling a wide range of novel uses for information processing, will lead to materials, components, and devices capable of responding with life-like adaptability to their environment.

First, a new revolution is just beginning in science. The building blocks of this revolution are being transformed into revolutionary new conceptual and technological tools with wide-ranging applications in the sciences, especially sciences investigating complex systems, most notably the natural sciences and in particular the biological sciences. This represents nothing less than the emergence of 'new kinds' of science. Second, that this is a starting point for fundamental advances in biology, biotechnology, medicine, and understanding the life-support systems of the Earth upon which the planet's biota, including our own species, depends. The new conceptual and technological tools have the potential to accelerate a new era of 'science-based innovation' and a consequent new wave of economic growth. Economic growth from new health, medical, energy, environmental management, computing and engineering sectors, some of which are unimaginable today is not only entirely plausible, it is happening already. Third, is the importance and potentially profound impact of what is occurring already at the intersection of computing, computer science and the other sciences.

Recommendations

1. Establish science and science-based innovation at the top of the political agenda.
2. Tomorrow's scientists will be amongst the most valuable assets that any nation will have. As well as being required to be scientifically and mathematically literate, tomorrow's scientists will also need to be computationally literate. Achieving this urgently requires a re-think of education policies now, not just at the undergraduate and postgraduate training level, but also at the elementary school level since today's children are tomorrow's scientists.
3. The science community needs to find new ways to raise the importance of science both in the public arena and in terms of its importance for the political agenda.
4. Re-think science funding and science policy structures to make the connection between funding and scientific output, and how we even measure scientific output. Focus more on research and collaborations to facilitate such research, on funding longer timescale projects so people work on solving problems.

5. Create new kinds of research institutes that are highly interdisciplinary, combine teaching, training and research, and focus on 'grand challenges' rather than 'grand disciplines' and simply producing papers.
6. A call to action to develop new conceptual and technological tools.
7. Develop innovative public private partnerships to accelerate science-based innovation. Governments, universities and businesses need to find new kinds of ways to work together - public-private partnerships (PPPs) are needed in order to really accelerate science and science-based innovation.
8. Find better mechanisms to create value from intellectual property, for both universities and industry (and probably governments) to find new and better ways to generate value from science-based intellectual property (IP).

Appendix D: EDUCAUSE Survey

Methodology: Survey was of 369 IT administrators (mostly CIOs)

Findings and conclusions

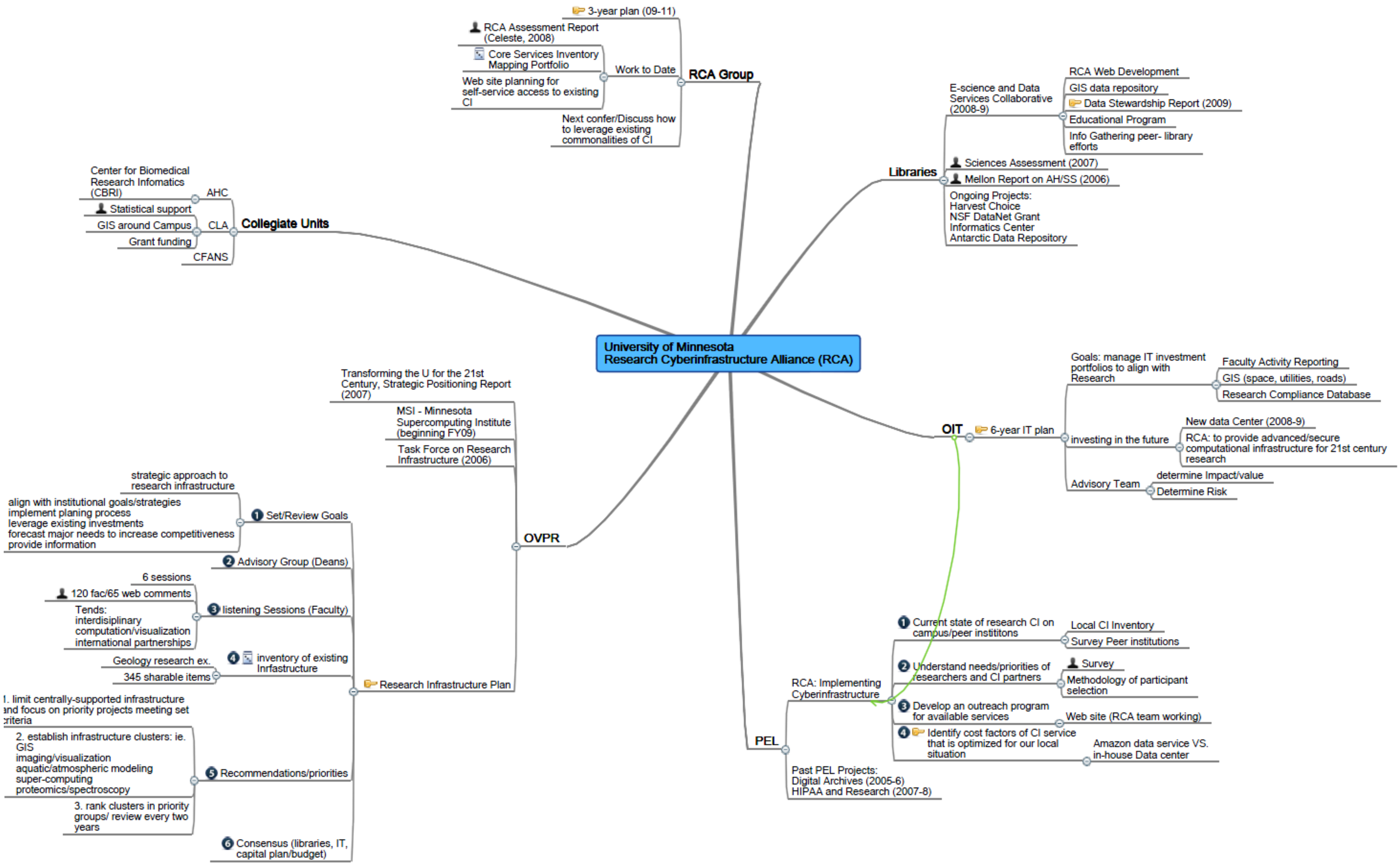
- Respondents rated overall knowledge of CI technologies high but their resources for meeting their research-related responsibilities low.
- CIOs are most knowledgeable about advanced network infrastructure resources; least knowledgeable about CI applications and tools.
- Rated research-related collaborative practices at their institutions low, and their own organizations' effectiveness at integrating CI technologies even lower.
- Under half are making use of high-performance computing resources or advanced network infrastructure resources for research. Due to high cost of these technologies?
- Researchers obtain access to most CI technologies mainly through the use of their own resources or those of their labs. Central IT is next most frequently used source. Central IT is primary source for advanced network infrastructure.
- A campus environment in which information about the use of CI technologies is openly shared between researcher and central IT is important to achieving the full potential of cyberinfrastructure.
- For all 5 CI technologies, more funding for central IT infrastructure and services would help them support more effective research use at their institutions. Needs for data storage & management, advanced NW infrastructure, and resources for collaboration within virtual communities appeared to be greatest. Respondents also identified better communication and outreach between researchers and central IT as something that would improve their support for research.
- Substantial numbers identified a greater role for central IT in developing budgets for the institution's research grants and contracts as one possible mechanism for improving the supply of funding for central IT's support of research.
- Achieving economies of scale in the use of specialized research IT resources relies in part on the researchers themselves offering some of the capacity of those resources to their colleagues under a variety of collaborative scenarios, and it relies on them as well to work with central IT where appropriate to ensure optimal management of shared or shareable resources.

- Respondents gave their institutions mediocre scores for collaboration and somewhat poorer scores for achieving economies of scale and for providing incentives to collaborative behavior. Higher scores were seen at institutions where central IT provides and funds various combinations of CI technologies, where the CIO is knowledgeable about them, and where inventories of CI resources are more complete.
- Institutions whose ratings for collaborative practices and institutional incentives were higher also said they did better at integrating CI resources.
- Research enterprise model vs. cottage-industry model. Cottage-industry: inefficiency and cost inflation. Research Enterprise: synergies among researchers and between researchers and central IT can improve research quality and more rapidly advance the frontiers of knowledge.
- Pennsylvania State University – leader in CI support, cited as model example by several interviewees.

Conclusion

Atmosphere of collaboration, resource sharing, and active pursuit of economies of scale are all associated with more effective integration of CI resources.

Appendix E. Intersecting Cyberinfrastructure Efforts Map



Appendix F. PEL Survey Instrument

Here is where we print the questions of our survey and results. The screenshots of survey in final form, along with emails, reminder emails.

Initial Email

President's Emerging Leaders | University of Minnesota Driven to Discover℠

Greetings,

The 2008-2009 President's Emerging Leaders (PEL) Project on Research Cyberinfrastructure at the University of Minnesota is asking for your help in voicing your needs and concerns regarding the current and future needs of research technology support across the University.

This project has been endorsed by senior leadership, and your contribution by completing the survey will help shape the future direction of research technology support at the University.

Survey participants will be entered into a drawing for one of five $50 VISA gift cards. You have until April 8, 2009 to complete this survey; however, if you complete the survey within one week, then your name will be entered twice in the drawing. Please allow approximately 30 minutes to complete the survey.

Take me to survey

Thank you for your time and contribution in advance.

2008-2009 RCA PEL Project Team
Tracy Anderson, College of Biological Sciences
Craig Gjerdingen, Carlson School of Management
Bryan Herrmann, Office of Admissions, Morris Campus
Katherine Himes, Office of the Senior Vice President for Academic Affairs and Provost
Lisa Johnston, University Libraries

This e-mail was sent to rosex001@umn.edu by the 2008-2009 RCA PEL Project Team.

Reminder Email:

President's Emerging Leaders | University of Minnesota Driven to Discover℠

Greetings,

Just two days remain to contribute your thoughts to the 2008-2009 President's Emerging Leaders (PEL) Project on Research Cyberinfrastructure at the University of Minnesota. Your contribution by completing the survey will help shape the future direction of research technology support at the University.

Survey participants will be entered into a drawing to win for one of five $50 VISA gift cards. But don't wait, the survey will close after Wednesday, April 8, 2009. Please allow approximately 30 minutes to complete the survey.

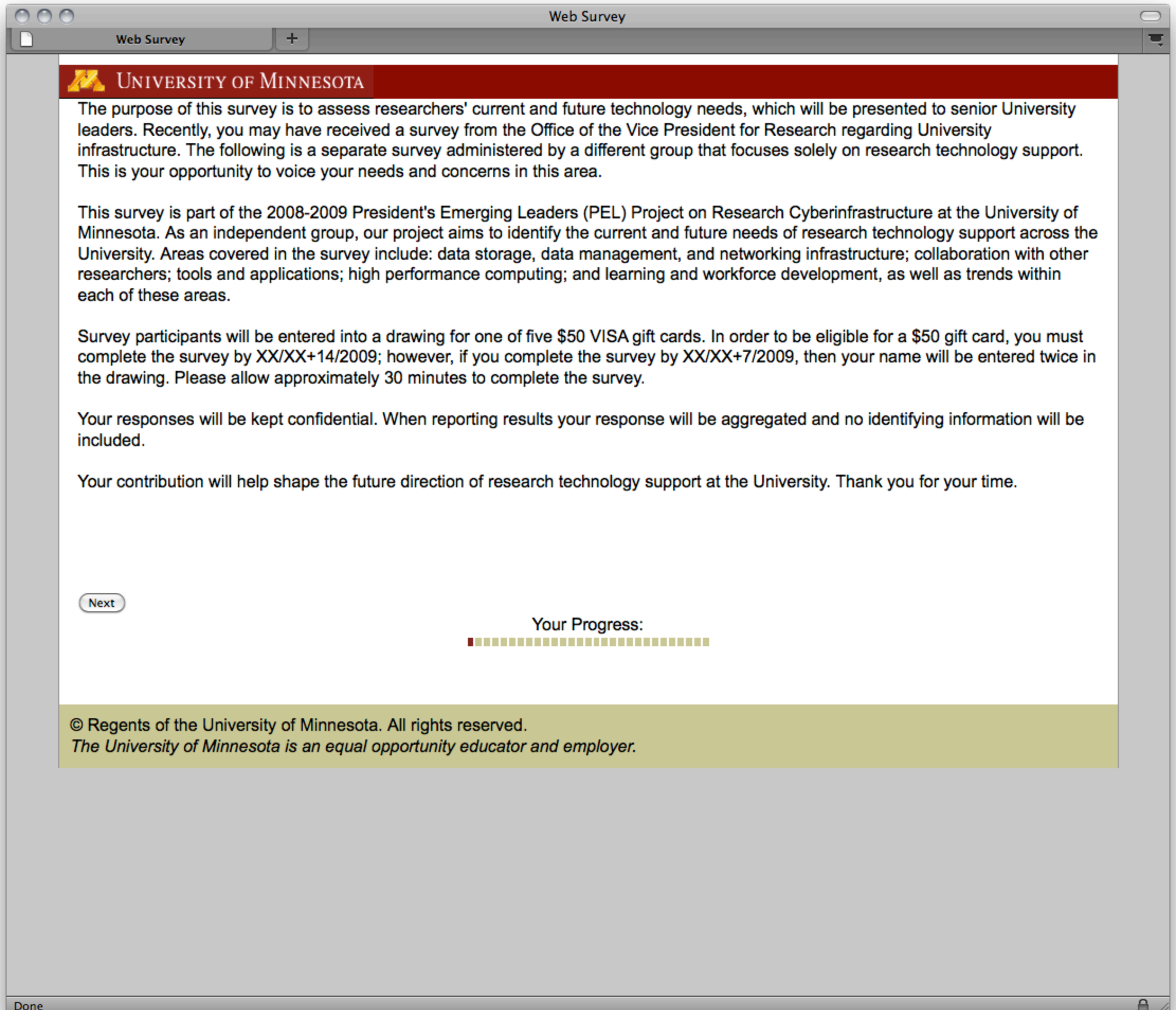
Take me to survey

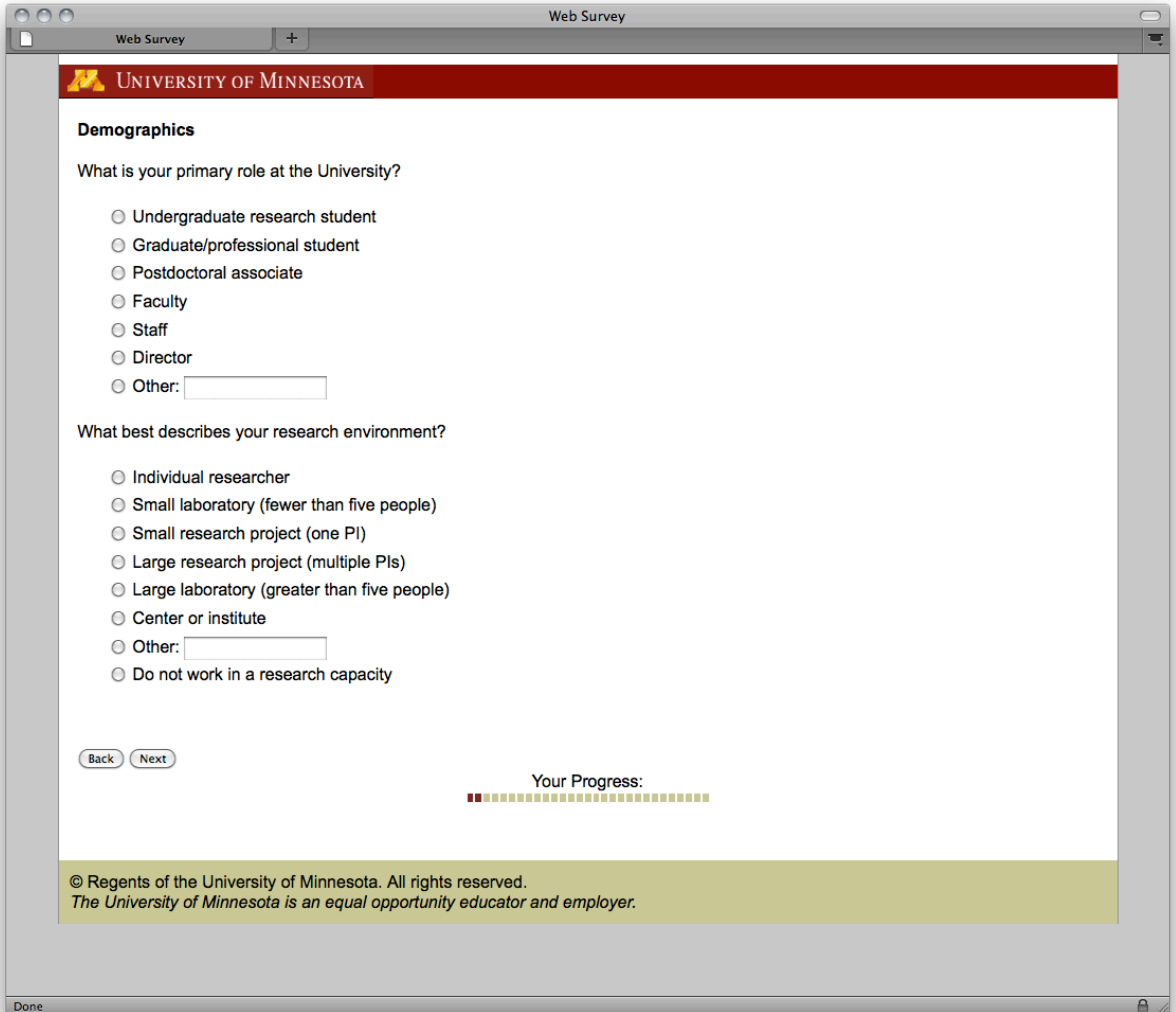
Thank you for your time and contribution in advance.

2008-2009 RCA PEL Project Team
Tracy Anderson, College of Biological Sciences
Craig Gjerdingen, Carlson School of Management
Bryan Herrmann, Office of Admissions, Morris Campus
Katherine Himes, Office of the Senior Vice President for Academic Affairs and Provost
Lisa Johnston, University Libraries

This e-mail was sent to white003@umn.edu by the 2008-2009 RCA PEL Project Team.

Survey Screenshots:




Web Survey

Web Survey

University of Minnesota

Demographics

What is your primary role at the University?

- ○ Undergraduate research student
- ○ Graduate/professional student
- ○ Postdoctoral associate
- ○ Faculty
- ○ Staff
- ○ Director
- ○ Other: ______

What best describes your research environment?

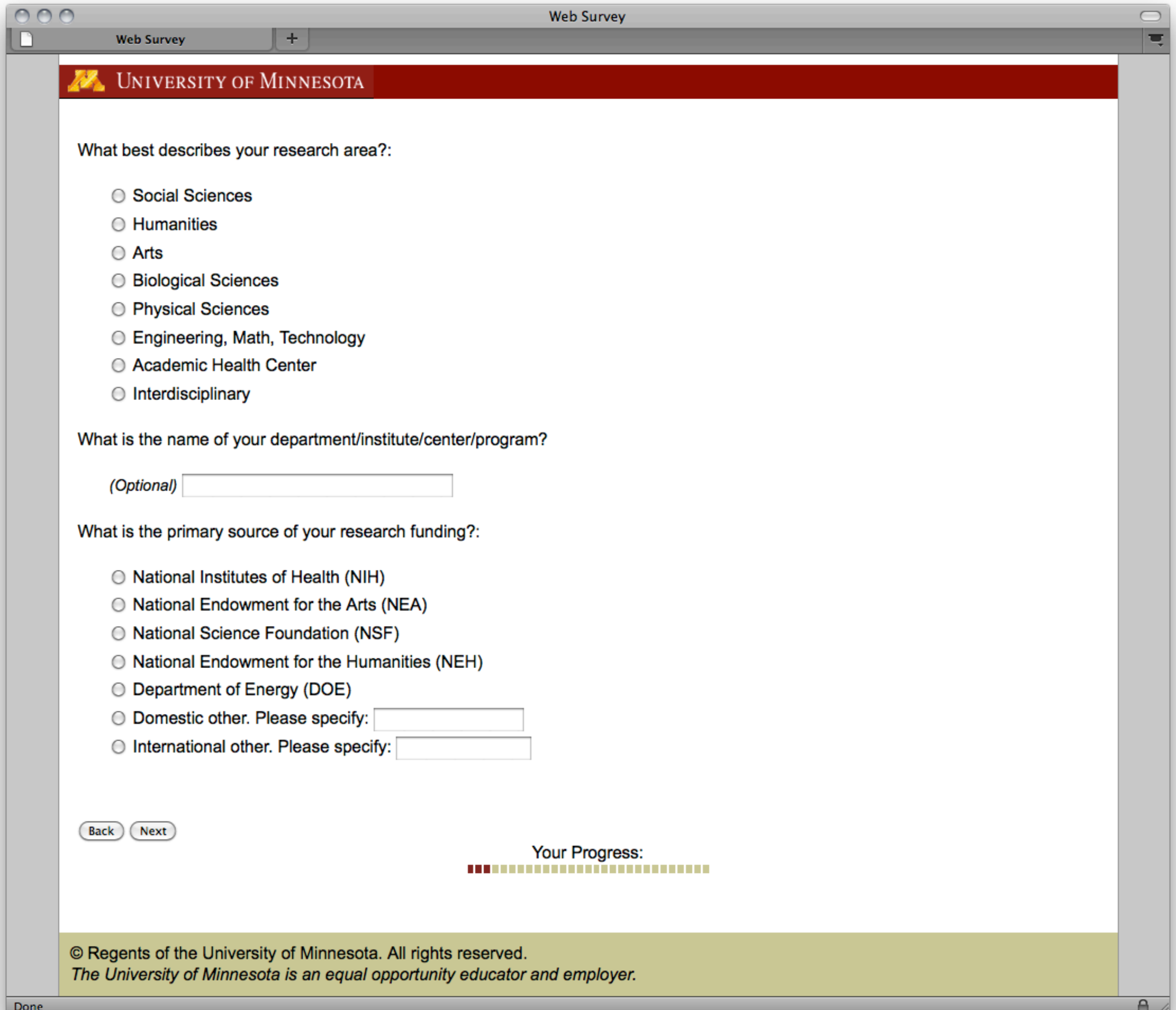
- ○ Individual researcher
- ○ Small laboratory (fewer than five people)
- ○ Small research project (one PI)
- ○ Large research project (multiple PIs)
- ○ Large laboratory (greater than five people)
- ○ Center or institute
- ○ Other: ______
- ○ Do not work in a research capacity

Back Next

Your Progress:

The University of Minnesota is an equal opportunity educator and employer.

Done


Web Survey

University of Minnesota

What best describes your research area?:

- Social Sciences
- Humanities
- Arts
- Biological Sciences
- Physical Sciences
- Engineering, Math, Technology
- Academic Health Center
- Interdisciplinary

What is the name of your department/institute/center/program?

(Optional)

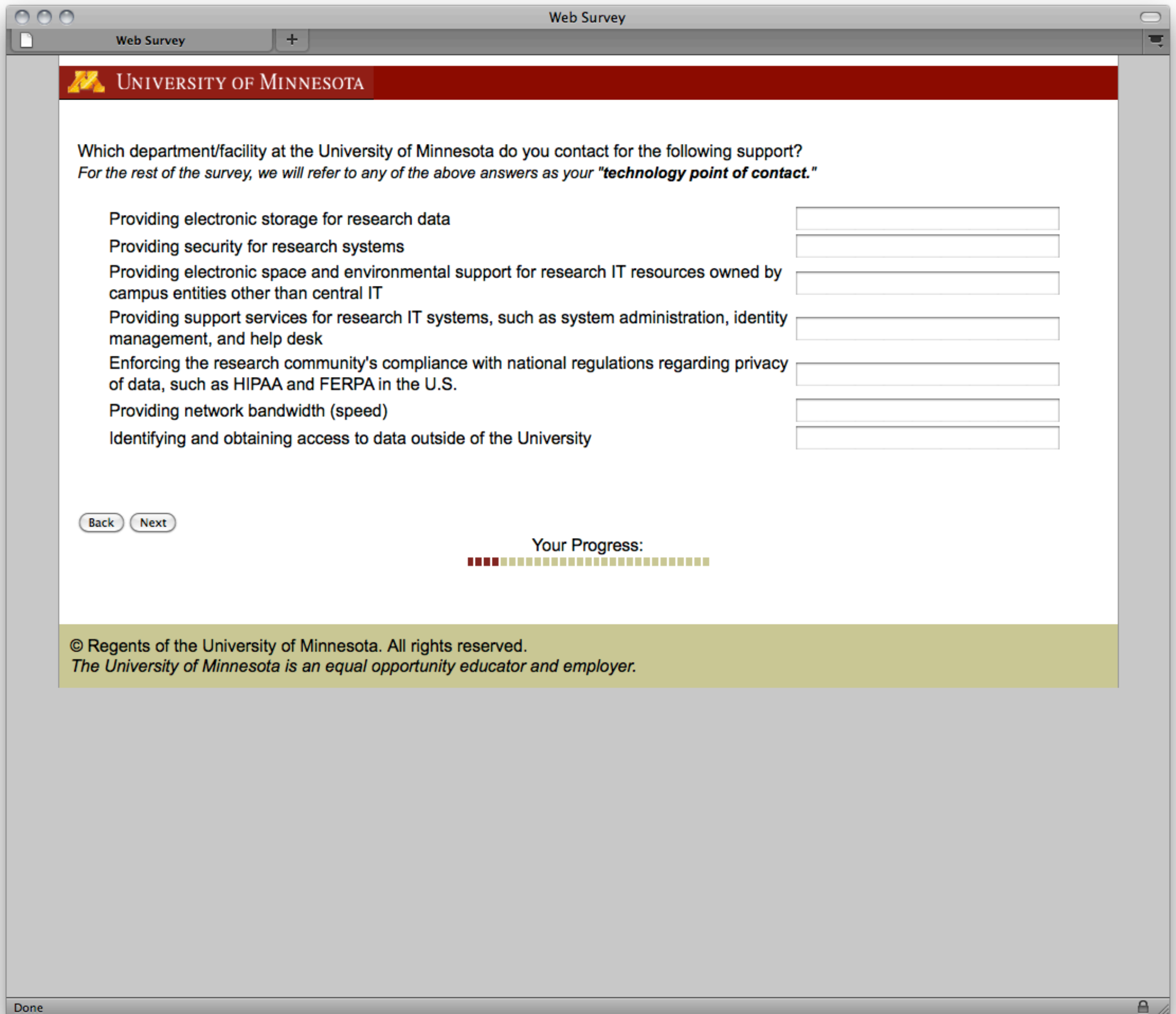
What is the primary source of your research funding?:

- National Institutes of Health (NIH)
- National Endowment for the Arts (NEA)
- National Science Foundation (NSF)
- National Endowment for the Humanities (NEH)
- Department of Energy (DOE)
- Domestic other. Please specify:
- International other. Please specify:

Back Next

Your Progress:

Done


Web Survey

Web Survey

UNIVERSITY OF MINNESOTA

Which department/facility at the University of Minnesota do you contact for the following support?
*For the rest of the survey, we will refer to any of the above answers as your "**technology point of contact.**"*

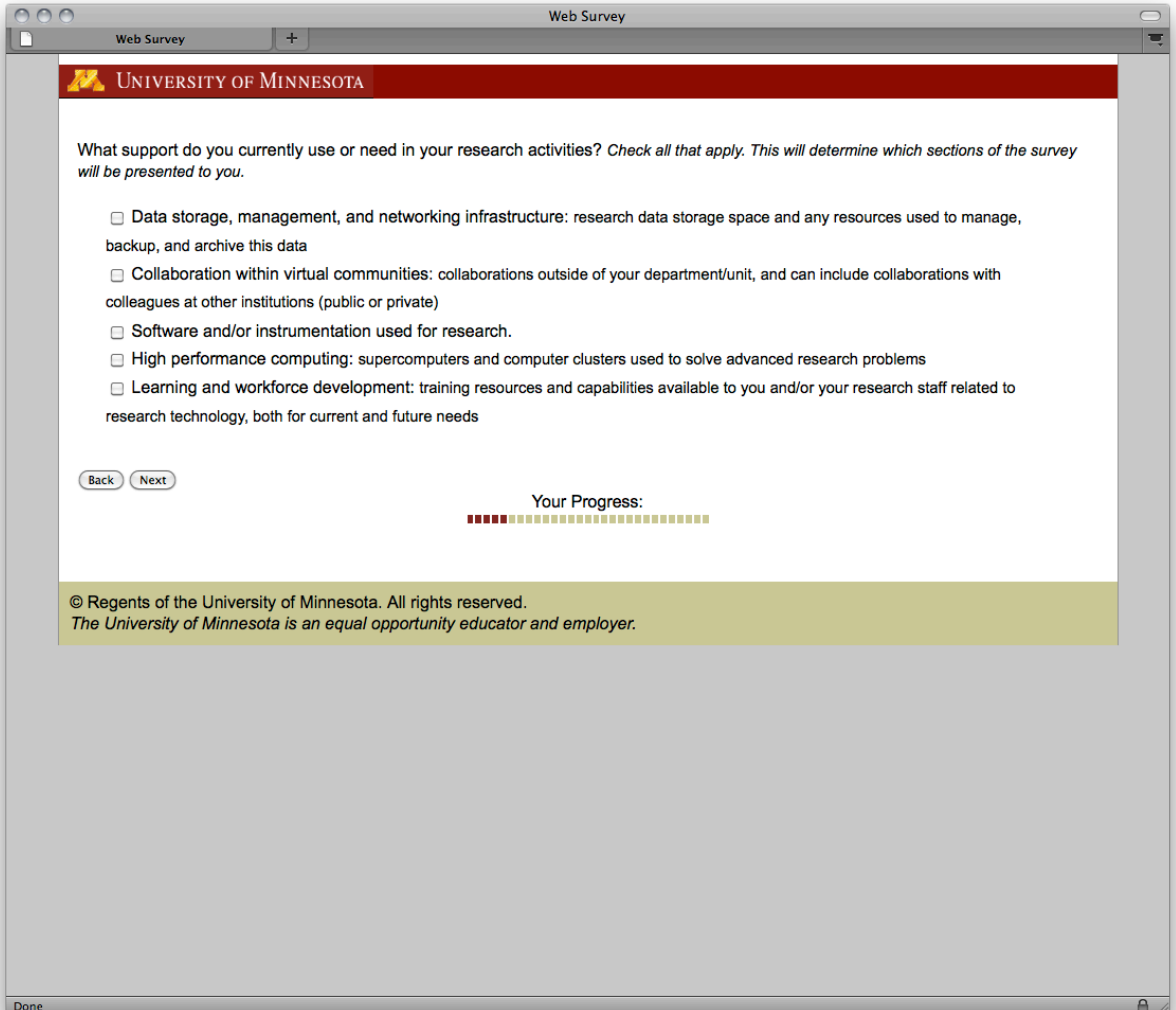
Providing electronic storage for research data	
Providing security for research systems	
Providing electronic space and environmental support for research IT resources owned by campus entities other than central IT	
Providing support services for research IT systems, such as system administration, identity management, and help desk	
Enforcing the research community's compliance with national regulations regarding privacy of data, such as HIPAA and FERPA in the U.S.	
Providing network bandwidth (speed)	
Identifying and obtaining access to data outside of the University	

Back Next

Your Progress:

© Regents of the University of Minnesota. All rights reserved.
The University of Minnesota is an equal opportunity educator and employer.

Done


Web Survey

UNIVERSITY OF MINNESOTA

What support do you currently use or need in your research activities? *Check all that apply. This will determine which sections of the survey will be presented to you.*

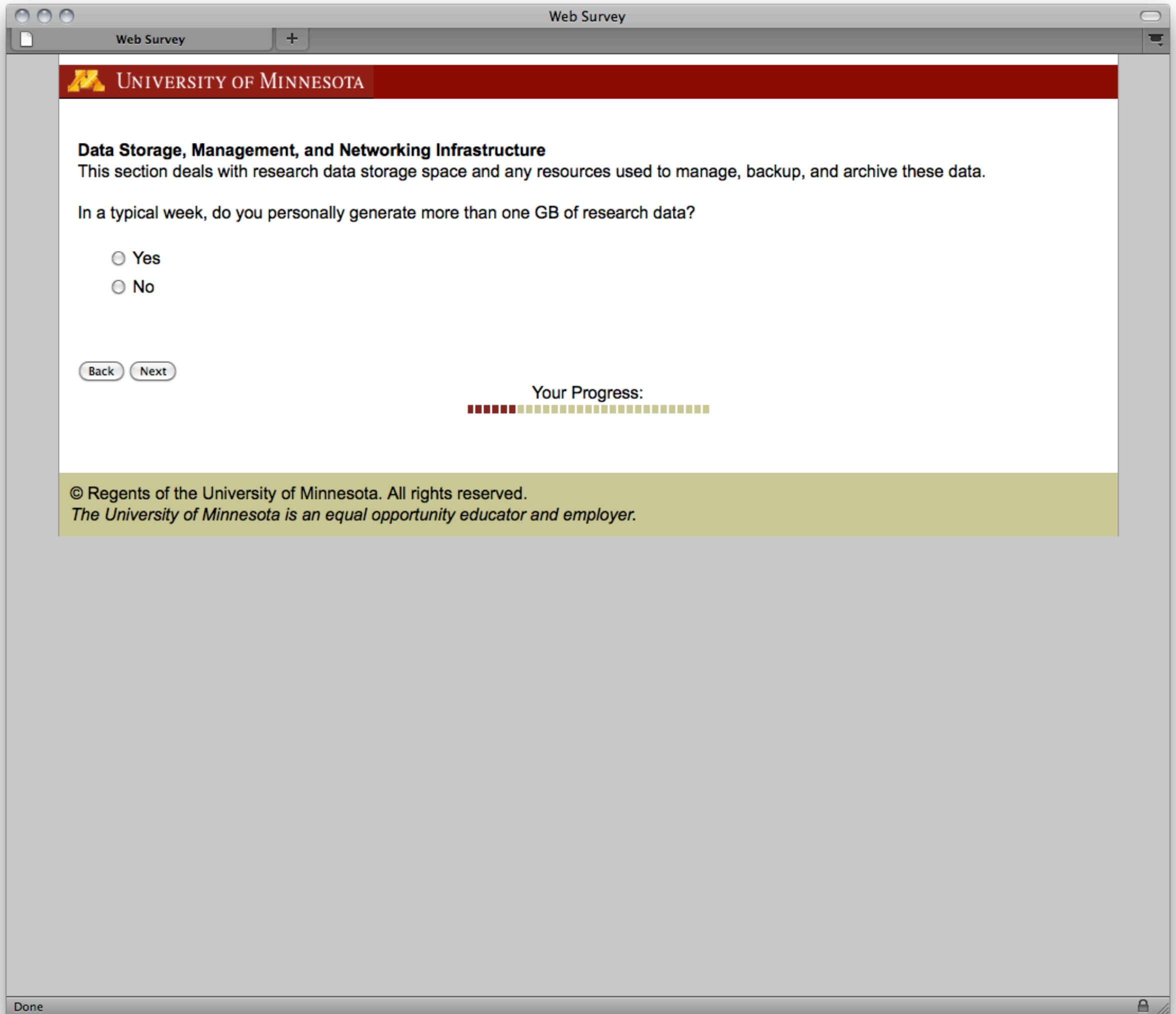
- ☐ Data storage, management, and networking infrastructure: research data storage space and any resources used to manage, backup, and archive this data
- ☐ Collaboration within virtual communities: collaborations outside of your department/unit, and can include collaborations with colleagues at other institutions (public or private)
- ☐ Software and/or instrumentation used for research.
- ☐ High performance computing: supercomputers and computer clusters used to solve advanced research problems
- ☐ Learning and workforce development: training resources and capabilities available to you and/or your research staff related to research technology, both for current and future needs

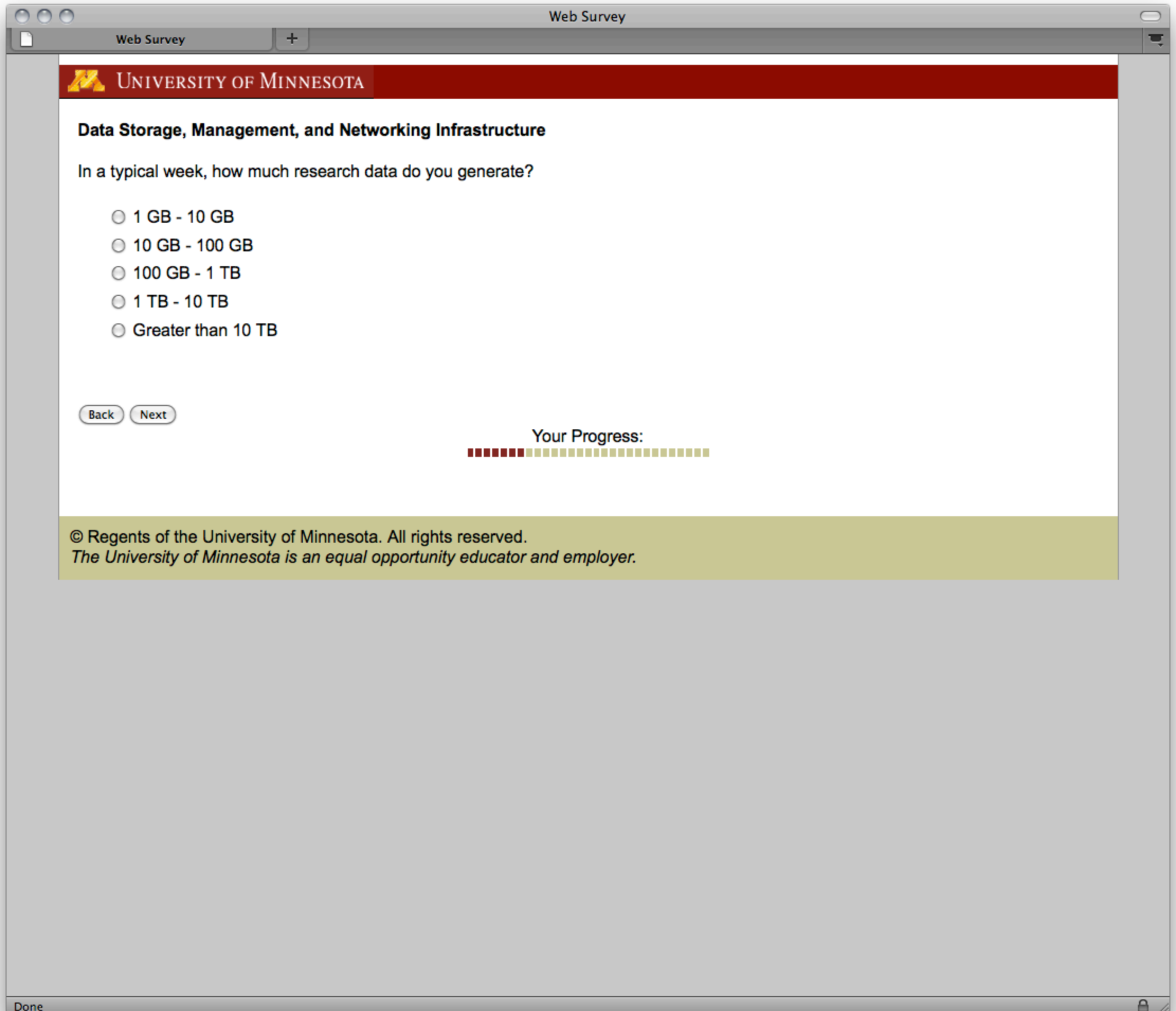
Back Next

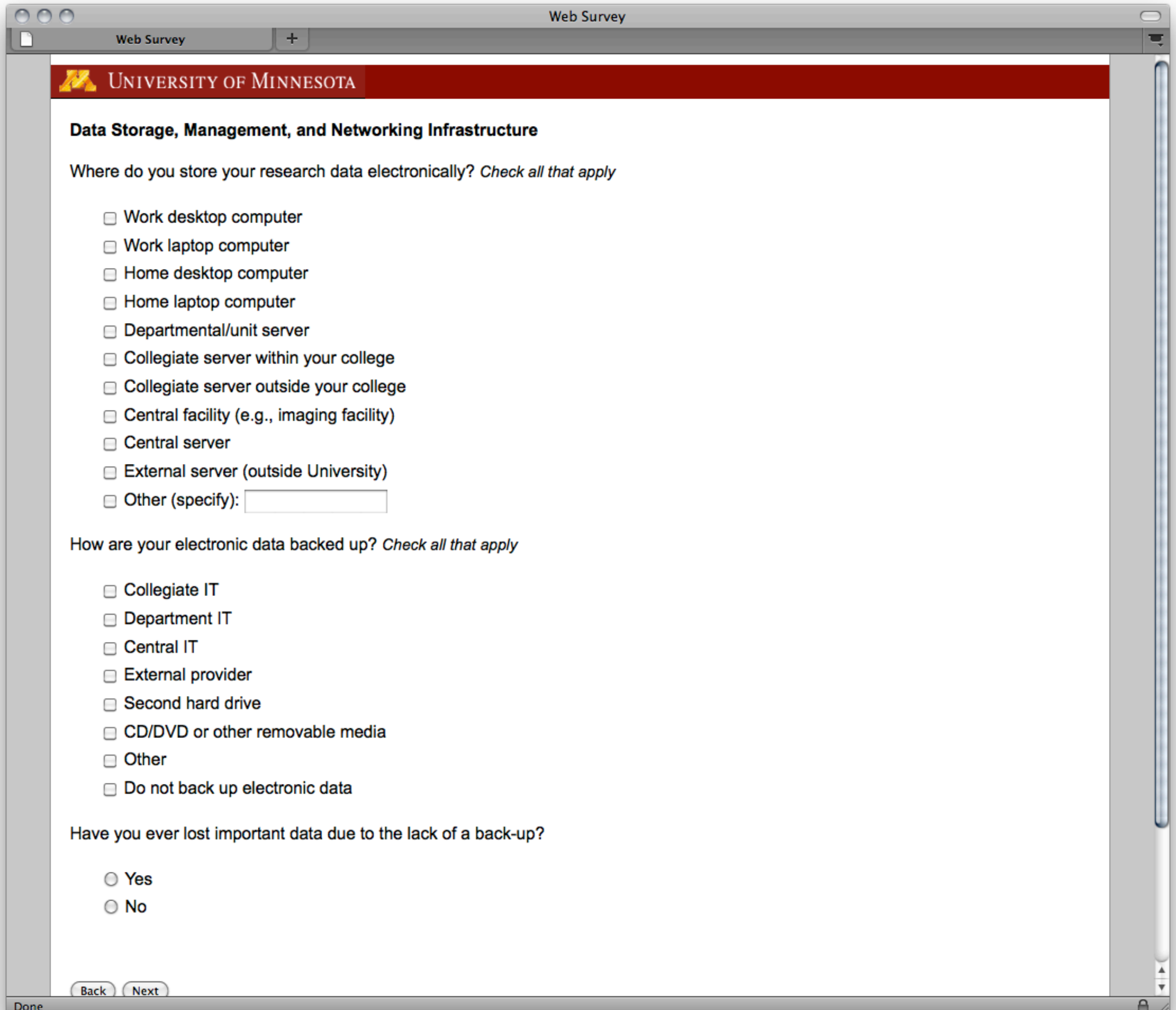
Your Progress:

© Regents of the University of Minnesota. All rights reserved.
The University of Minnesota is an equal opportunity educator and employer.

Done


Web Survey
Web Survey
University of Minnesota
Data Storage, Management, and Networking Infrastructure
This section deals with research data storage space and any resources used to manage, backup, and archive these data.
In a typical week, do you personally generate more than one GB of research data?
Yes
No
Back
Next
Your Progress:
© Regents of the University of Minnesota. All rights reserved.
The University of Minnesota is an equal opportunity educator and employer.
Done


Web Survey
Web Survey
University of Minnesota
Data Storage, Management, and Networking Infrastructure
In a typical week, how much research data do you generate?
1 GB - 10 GB
10 GB - 100 GB
100 GB - 1 TB
1 TB - 10 TB
Greater than 10 TB
Back
Next
Your Progress:
© Regents of the University of Minnesota. All rights reserved.
The University of Minnesota is an equal opportunity educator and employer.
Done



Web Survey

University of Minnesota

Data Storage, Management, and Networking Infrastructure

Where do you store your research data electronically? *Check all that apply*

- ☐ Work desktop computer
- ☐ Work laptop computer
- ☐ Home desktop computer
- ☐ Home laptop computer
- ☐ Departmental/unit server
- ☐ Collegiate server within your college
- ☐ Collegiate server outside your college
- ☐ Central facility (e.g., imaging facility)
- ☐ Central server
- ☐ External server (outside University)
- ☐ Other (specify):

How are your electronic data backed up? *Check all that apply*

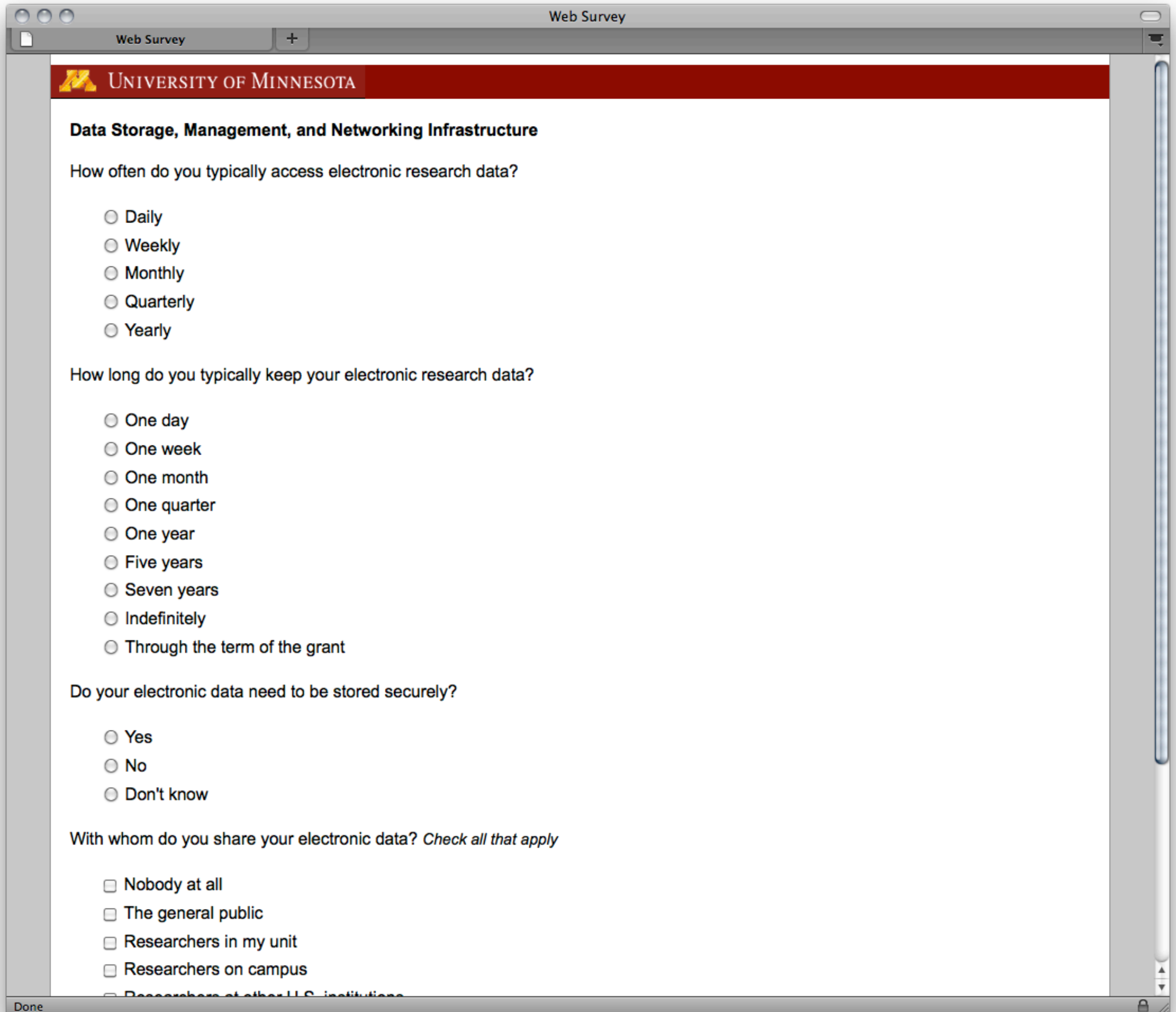
- ☐ Collegiate IT
- ☐ Department IT
- ☐ Central IT
- ☐ External provider
- ☐ Second hard drive
- ☐ CD/DVD or other removable media
- ☐ Other
- ☐ Do not back up electronic data

Have you ever lost important data due to the lack of a back-up?

- ○ Yes
- ○ No

Back Next

Done



UNIVERSITY OF MINNESOTA

Data Storage, Management, and Networking Infrastructure

How often do you typically access electronic research data?

- Daily
- Weekly
- Monthly
- Quarterly
- Yearly

How long do you typically keep your electronic research data?

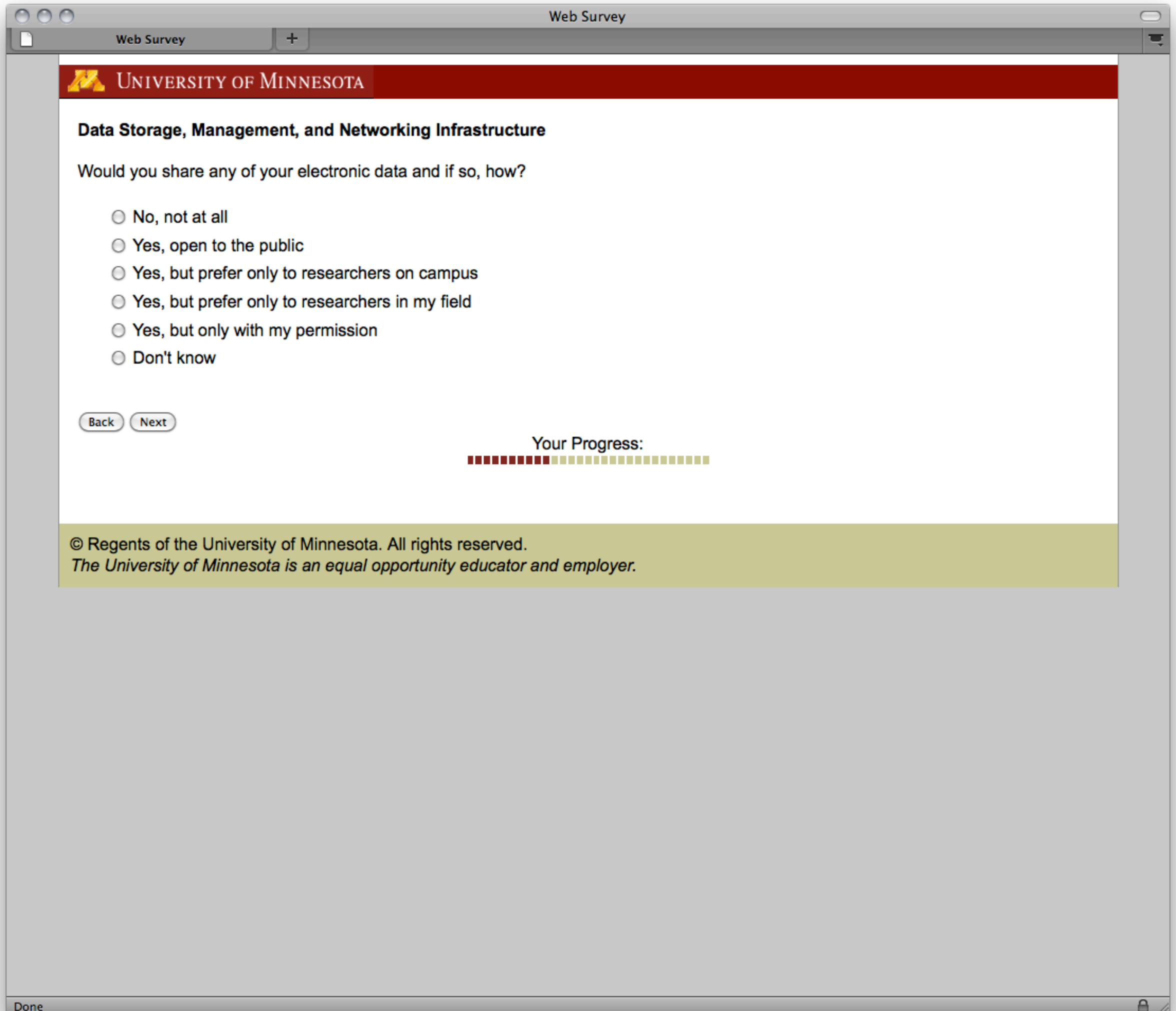
- One day
- One week
- One month
- One quarter
- One year
- Five years
- Seven years
- Indefinitely
- Through the term of the grant

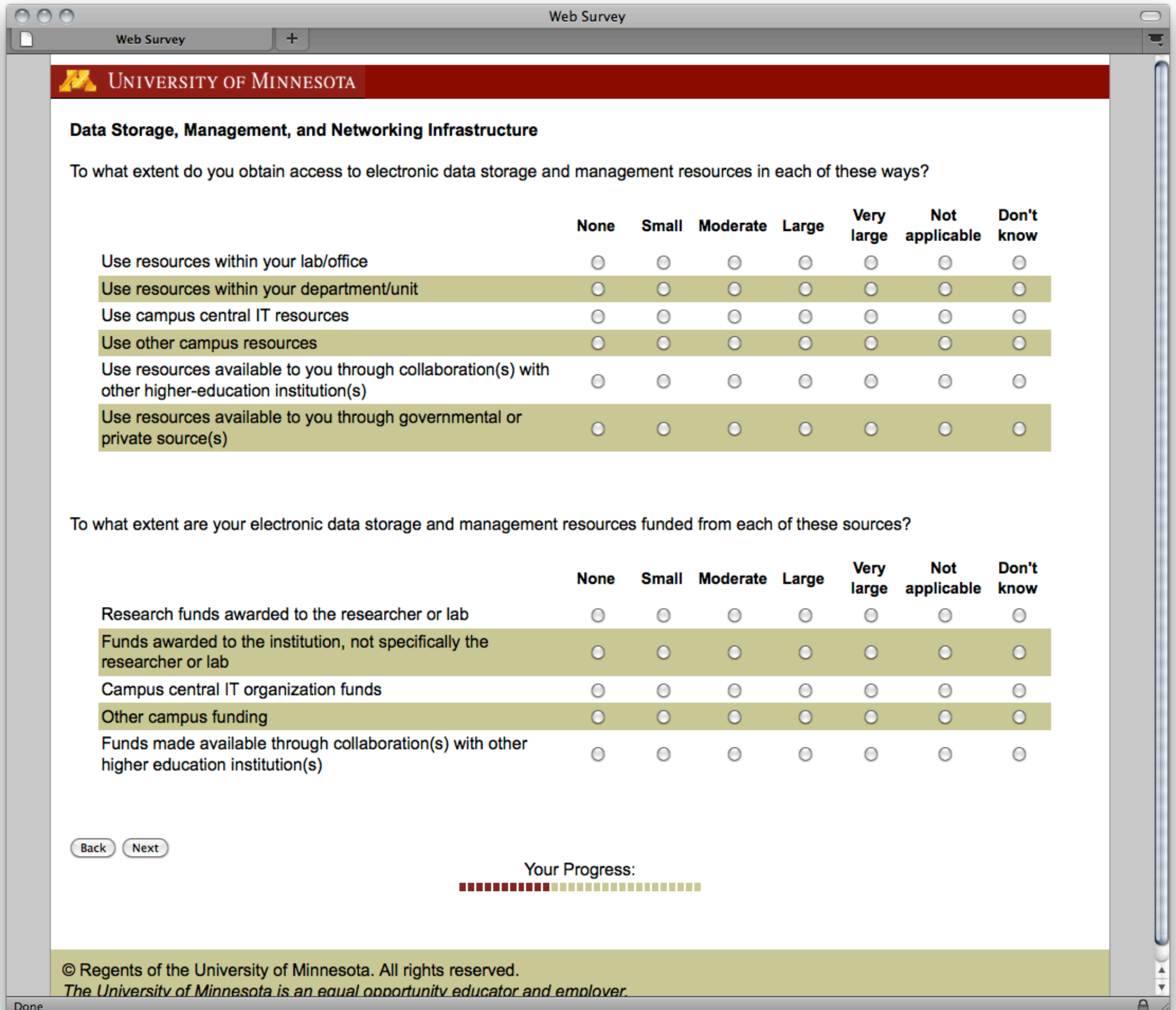
Do your electronic data need to be stored securely?

- Yes
- No
- Don't know

With whom do you share your electronic data? *Check all that apply*

- Nobody at all
- The general public
- Researchers in my unit
- Researchers on campus


Web Survey
Web Survey
University of Minnesota
Data Storage, Management, and Networking Infrastructure
Would you share any of your electronic data and if so, how?
No, not at all
Yes, open to the public
Yes, but prefer only to researchers on campus
Yes, but prefer only to researchers in my field
Yes, but only with my permission
Don't know
Back
Next
Your Progress:
© Regents of the University of Minnesota. All rights reserved.
The University of Minnesota is an equal opportunity educator and employer.
Done



Web Survey

UNIVERSITY OF MINNESOTA

Data Storage, Management, and Networking Infrastructure

To what extent do you obtain access to electronic data storage and management resources in each of these ways?

	None	Small	Moderate	Large	Very large	Not applicable	Don't know
Use resources within your lab/office							
Use resources within your department/unit							
Use campus central IT resources							
Use other campus resources							
Use resources available to you through collaboration(s) with other higher-education institution(s)							
Use resources available to you through governmental or private source(s)							

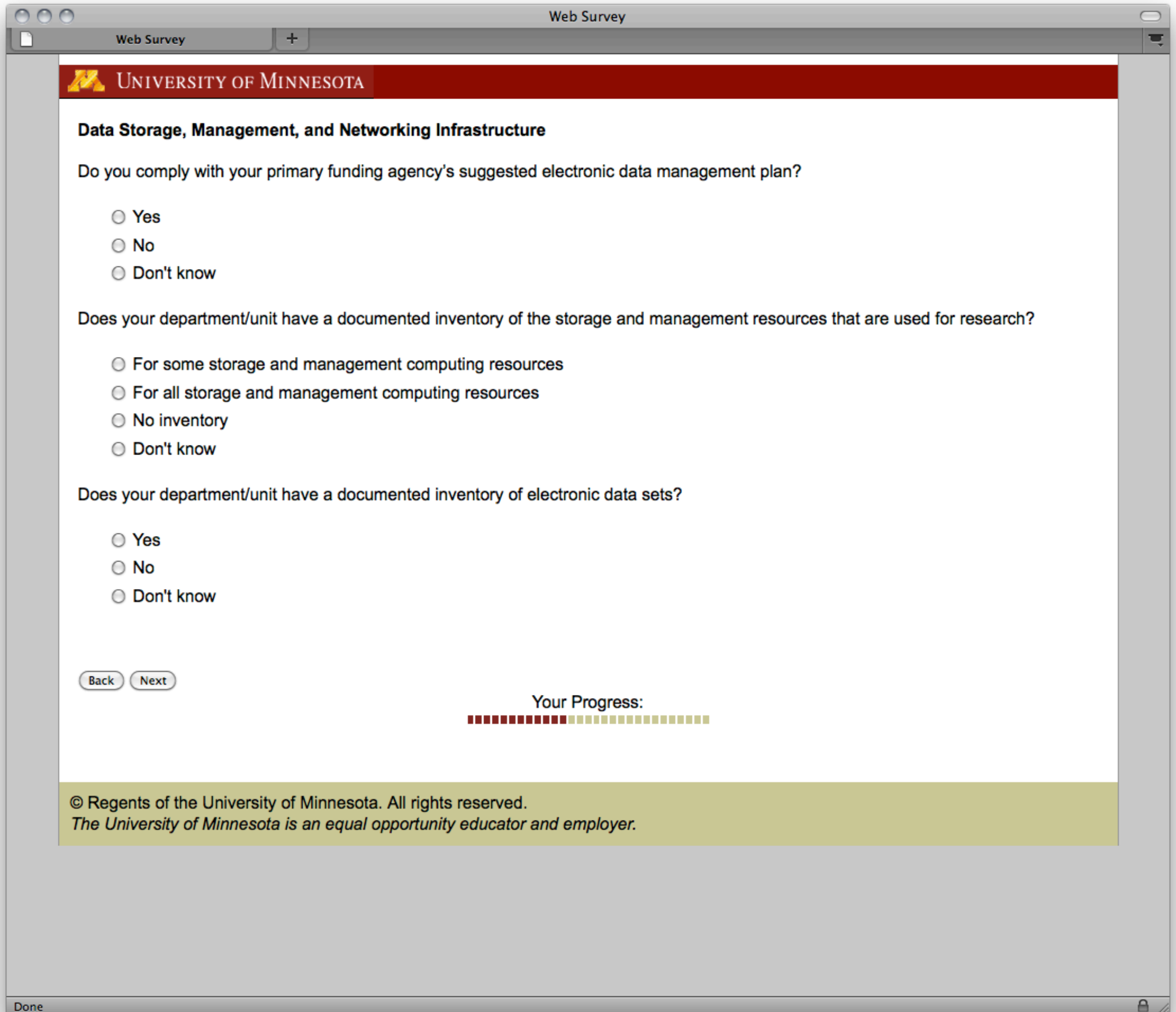
To what extent are your electronic data storage and management resources funded from each of these sources?

	None	Small	Moderate	Large	Very large	Not applicable	Don't know
Research funds awarded to the researcher or lab							
Funds awarded to the institution, not specifically the researcher or lab							
Campus central IT organization funds							
Other campus funding							
Funds made available through collaboration(s) with other higher education institution(s)							

Back Next

Your Progress:

Done



Web Survey

University of Minnesota

Data Storage, Management, and Networking Infrastructure

Do you comply with your primary funding agency's suggested electronic data management plan?

- Yes
- No
- Don't know

Does your department/unit have a documented inventory of the storage and management resources that are used for research?

- For some storage and management computing resources
- For all storage and management computing resources
- No inventory
- Don't know

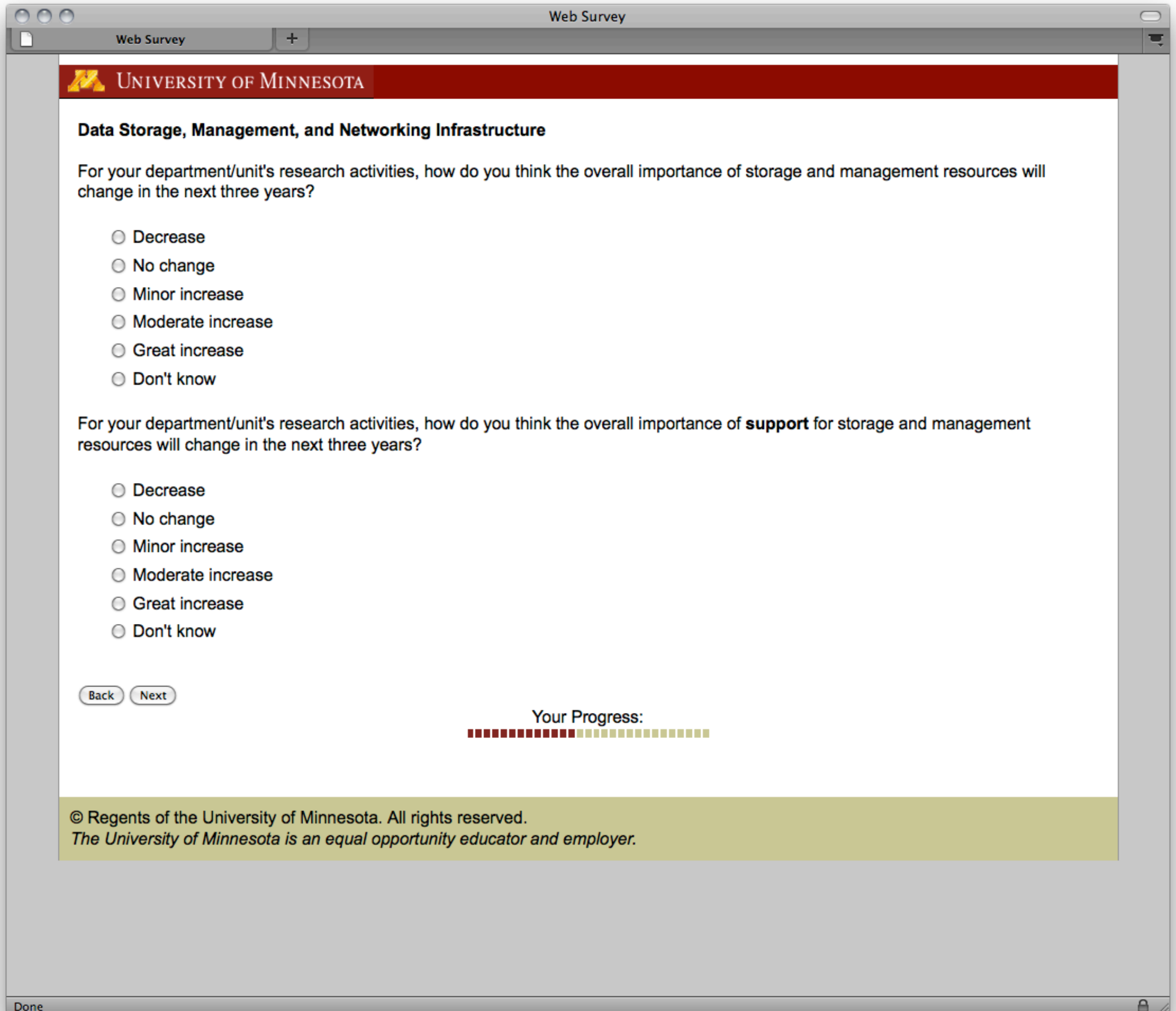
Does your department/unit have a documented inventory of electronic data sets?

- Yes
- No
- Don't know

Back Next

Your Progress:

Done



Web Survey

University of Minnesota

Data Storage, Management, and Networking Infrastructure

For your department/unit's research activities, how do you think the overall importance of storage and management resources will change in the next three years?

- Decrease
- No change
- Minor increase
- Moderate increase
- Great increase
- Don't know

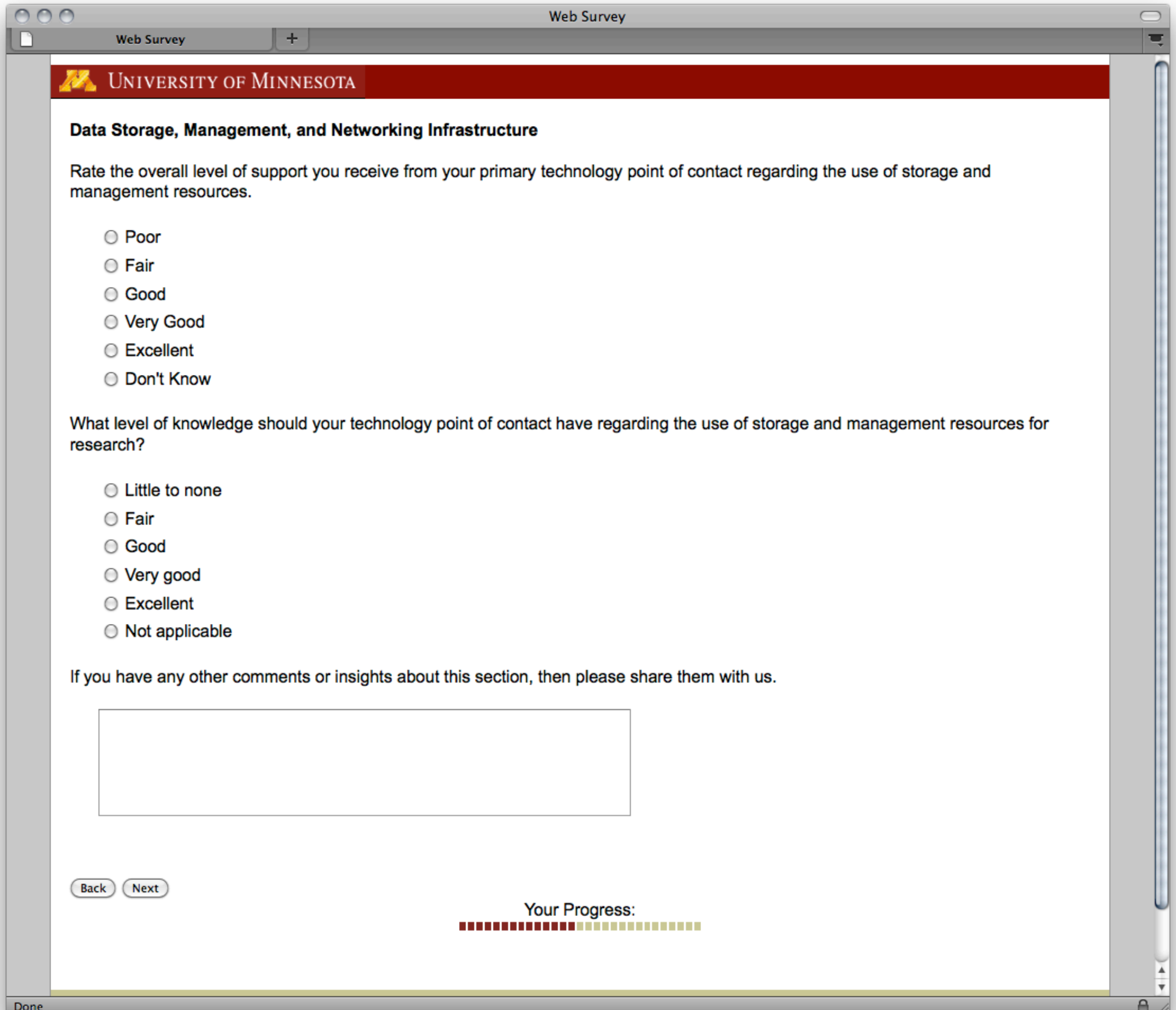
For your department/unit's research activities, how do you think the overall importance of **support** for storage and management resources will change in the next three years?

- Decrease
- No change
- Minor increase
- Moderate increase
- Great increase
- Don't know

Back Next

Your Progress:

Done



Web Survey

University of Minnesota

Data Storage, Management, and Networking Infrastructure

Rate the overall level of support you receive from your primary technology point of contact regarding the use of storage and management resources.

- Poor
- Fair
- Good
- Very Good
- Excellent
- Don't Know

What level of knowledge should your technology point of contact have regarding the use of storage and management resources for research?

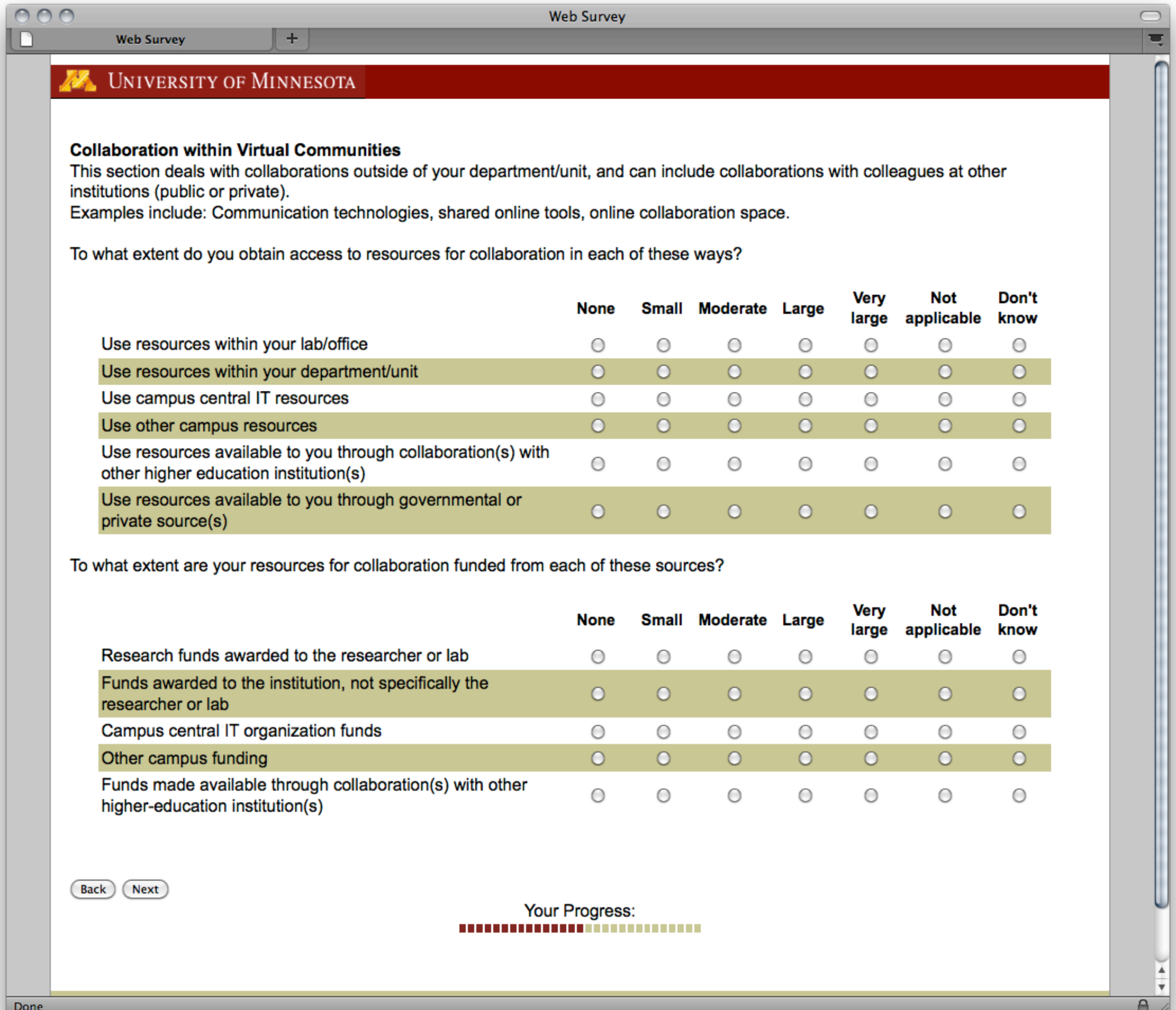
- Little to none
- Fair
- Good
- Very good
- Excellent
- Not applicable

If you have any other comments or insights about this section, then please share them with us.

Back Next

Your Progress:

Done



Web Survey

University of Minnesota

Collaboration within Virtual Communities

This section deals with collaborations outside of your department/unit, and can include collaborations with colleagues at other institutions (public or private).

Examples include: Communication technologies, shared online tools, online collaboration space.

To what extent do you obtain access to resources for collaboration in each of these ways?

	None	Small	Moderate	Large	Very large	Not applicable	Don't know
Use resources within your lab/office	○	○	○	○	○	○	○
Use resources within your department/unit	○	○	○	○	○	○	○
Use campus central IT resources	○	○	○	○	○	○	○
Use other campus resources	○	○	○	○	○	○	○
Use resources available to you through collaboration(s) with other higher education institution(s)	○	○	○	○	○	○	○
Use resources available to you through governmental or private source(s)	○	○	○	○	○	○	○

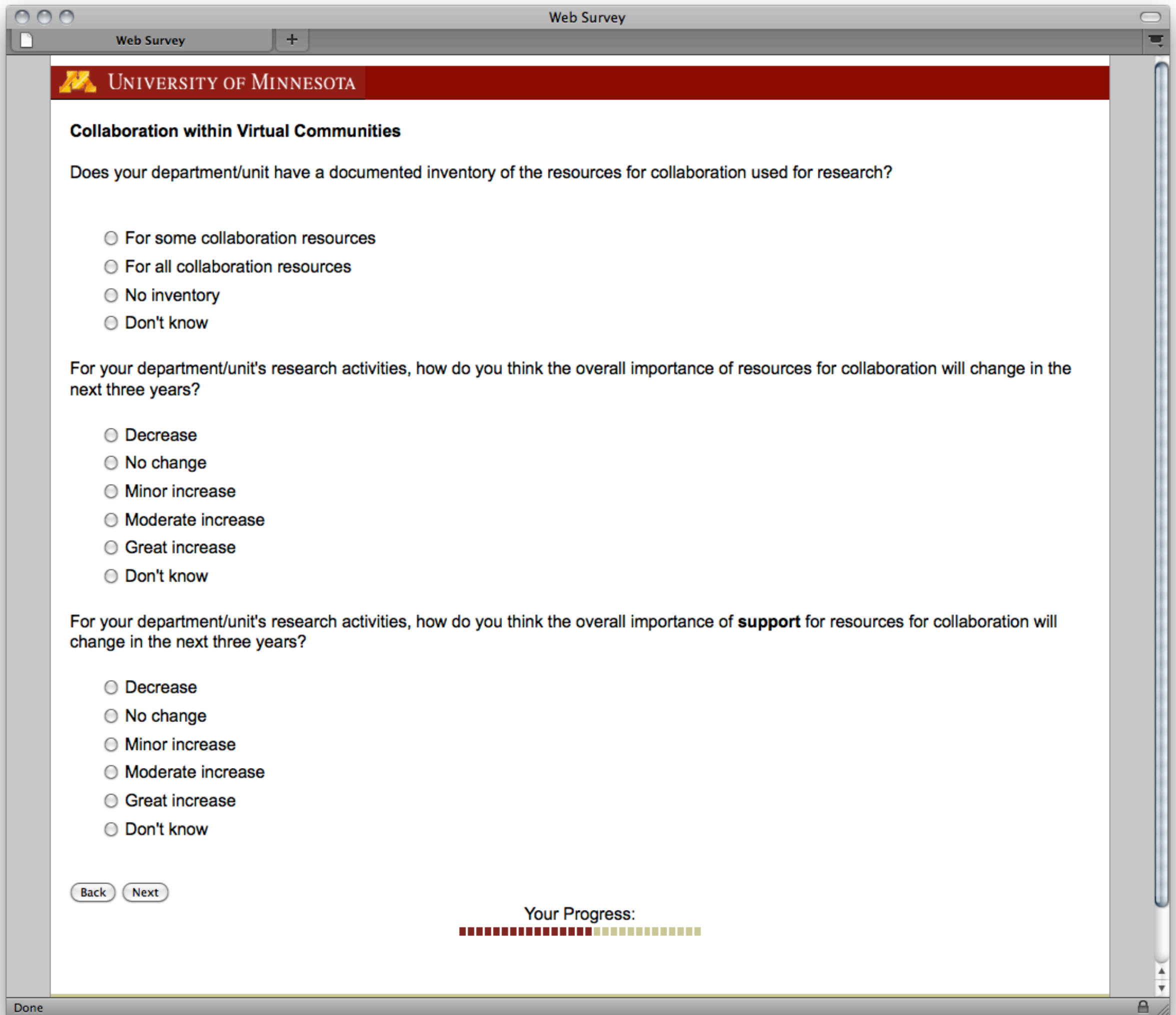
To what extent are your resources for collaboration funded from each of these sources?

	None	Small	Moderate	Large	Very large	Not applicable	Don't know
Research funds awarded to the researcher or lab	○	○	○	○	○	○	○
Funds awarded to the institution, not specifically the researcher or lab	○	○	○	○	○	○	○
Campus central IT organization funds	○	○	○	○	○	○	○
Other campus funding	○	○	○	○	○	○	○
Funds made available through collaboration(s) with other higher-education institution(s)	○	○	○	○	○	○	○

Back Next

Your Progress:

Done



Web Survey

University of Minnesota

Collaboration within Virtual Communities

Does your department/unit have a documented inventory of the resources for collaboration used for research?

- ○ For some collaboration resources
- ○ For all collaboration resources
- ○ No inventory
- ○ Don't know

For your department/unit's research activities, how do you think the overall importance of resources for collaboration will change in the next three years?

- ○ Decrease
- ○ No change
- ○ Minor increase
- ○ Moderate increase
- ○ Great increase
- ○ Don't know

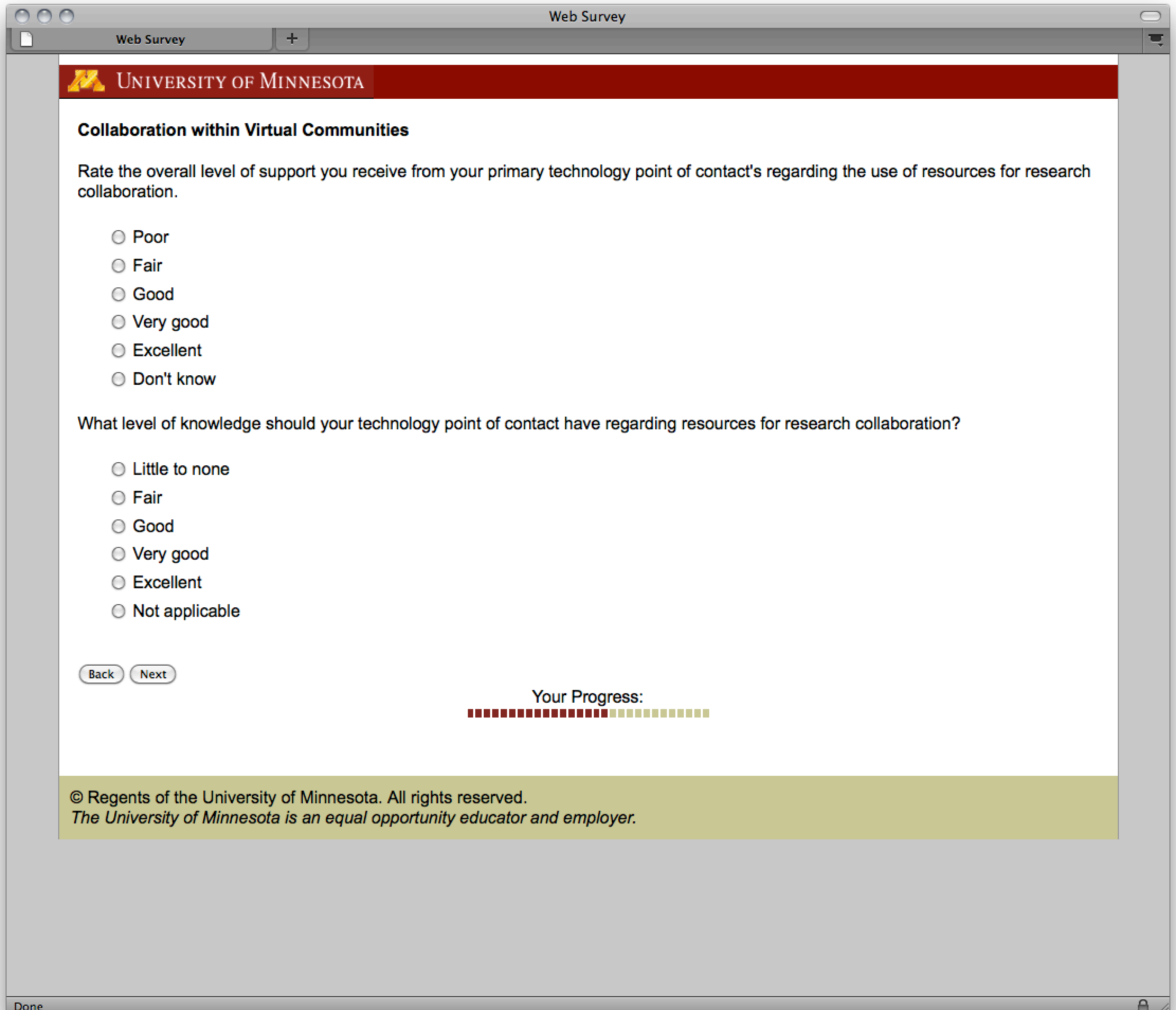
For your department/unit's research activities, how do you think the overall importance of **support** for resources for collaboration will change in the next three years?

- ○ Decrease
- ○ No change
- ○ Minor increase
- ○ Moderate increase
- ○ Great increase
- ○ Don't know

Back Next

Your Progress:

Done


Web Survey

Web Survey

UNIVERSITY OF MINNESOTA

Collaboration within Virtual Communities

Rate the overall level of support you receive from your primary technology point of contact's regarding the use of resources for research collaboration.

- ○ Poor
- ○ Fair
- ○ Good
- ○ Very good
- ○ Excellent
- ○ Don't know

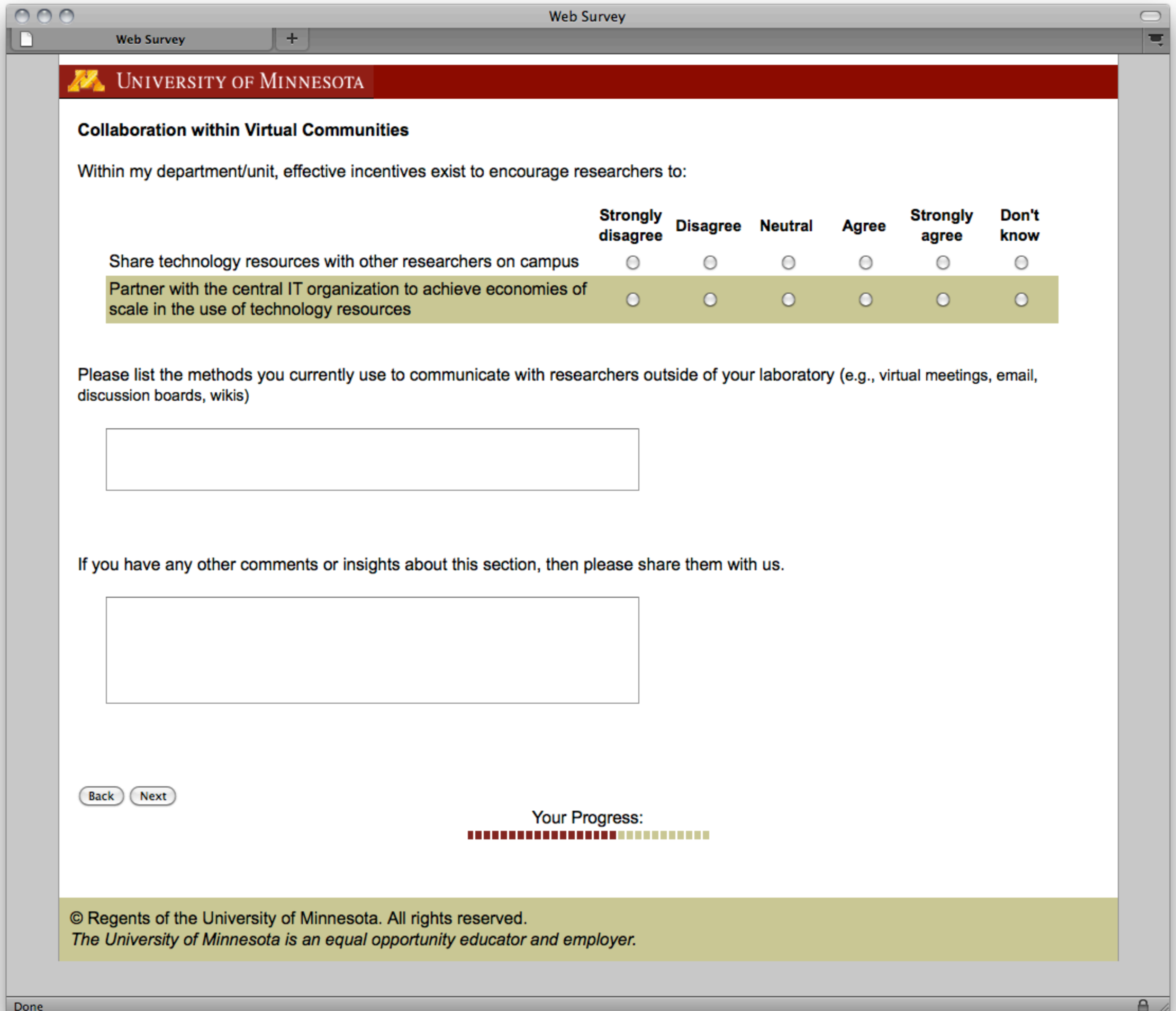
What level of knowledge should your technology point of contact have regarding resources for research collaboration?

- ○ Little to none
- ○ Fair
- ○ Good
- ○ Very good
- ○ Excellent
- ○ Not applicable

Back Next

Your Progress:

Done



Web Survey

UNIVERSITY OF MINNESOTA

Collaboration within Virtual Communities

Within my department/unit, effective incentives exist to encourage researchers to:

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Don't know
Share technology resources with other researchers on campus	○	○	○	○	○	○
Partner with the central IT organization to achieve economies of scale in the use of technology resources	○	○	○	○	○	○

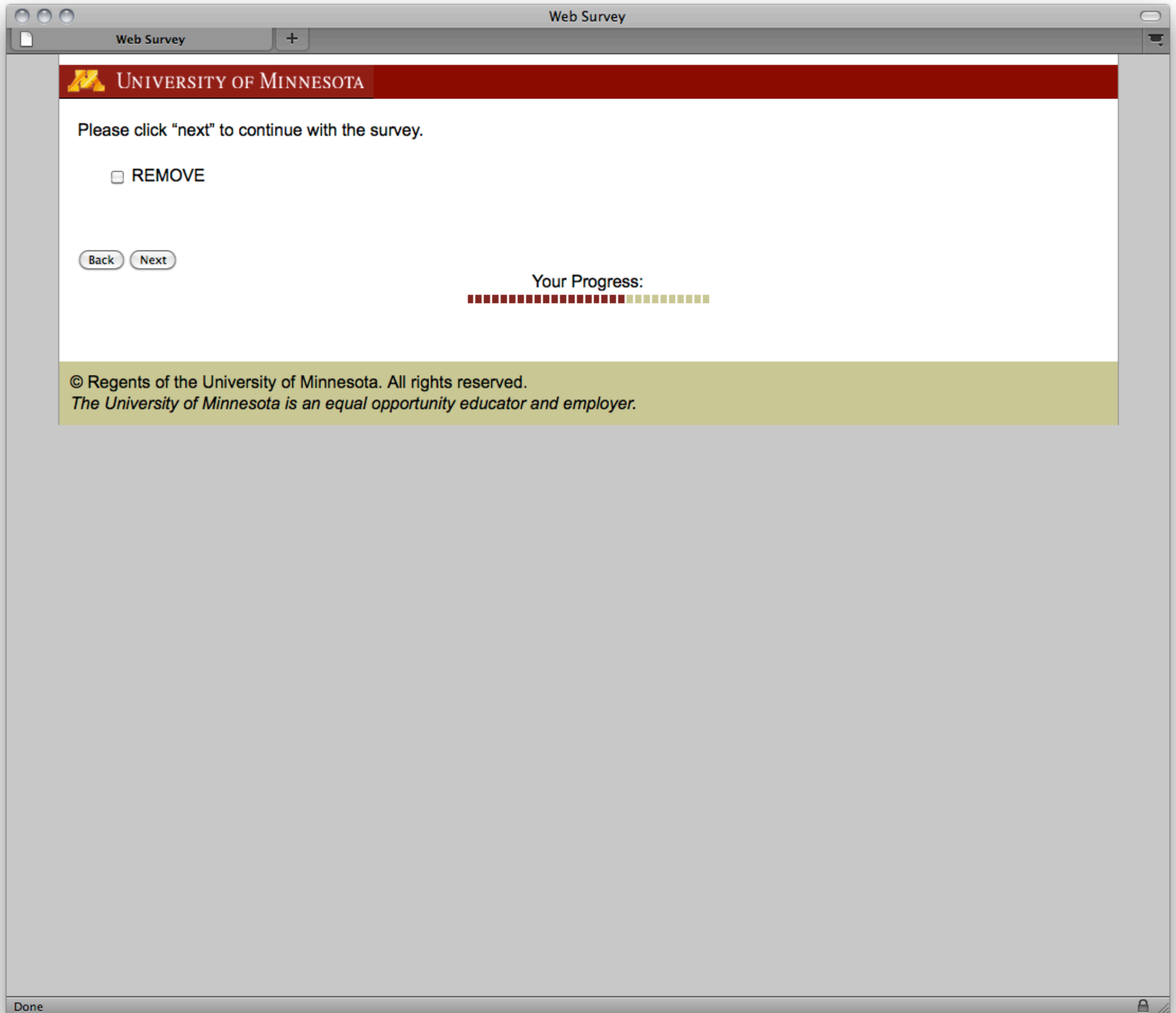
Please list the methods you currently use to communicate with researchers outside of your laboratory (e.g., virtual meetings, email, discussion boards, wikis)

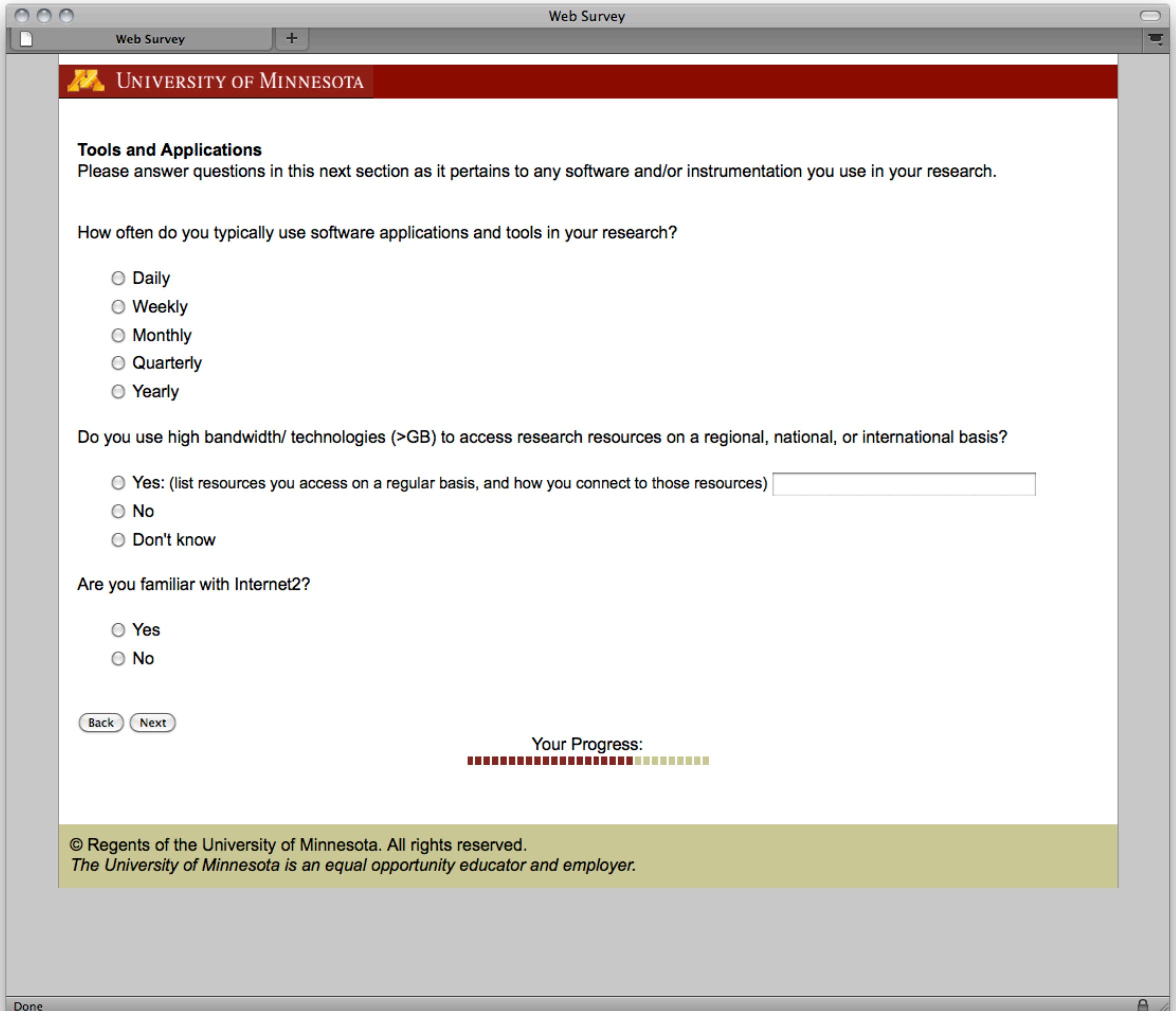
If you have any other comments or insights about this section, then please share them with us.

Back Next

Your Progress:

Done


Web Survey
Web Survey
University of Minnesota
Please click "next" to continue with the survey.
REMOVE
Back
Next
Your Progress:
© Regents of the University of Minnesota. All rights reserved.
The University of Minnesota is an equal opportunity educator and employer.
Done



Web Survey

UNIVERSITY OF MINNESOTA

Tools and Applications
Please answer questions in this next section as it pertains to any software and/or instrumentation you use in your research.

How often do you typically use software applications and tools in your research?

- ○ Daily
- ○ Weekly
- ○ Monthly
- ○ Quarterly
- ○ Yearly

Do you use high bandwidth/ technologies (>GB) to access research resources on a regional, national, or international basis?

- ○ Yes: (list resources you access on a regular basis, and how you connect to those resources)
- ○ No
- ○ Don't know

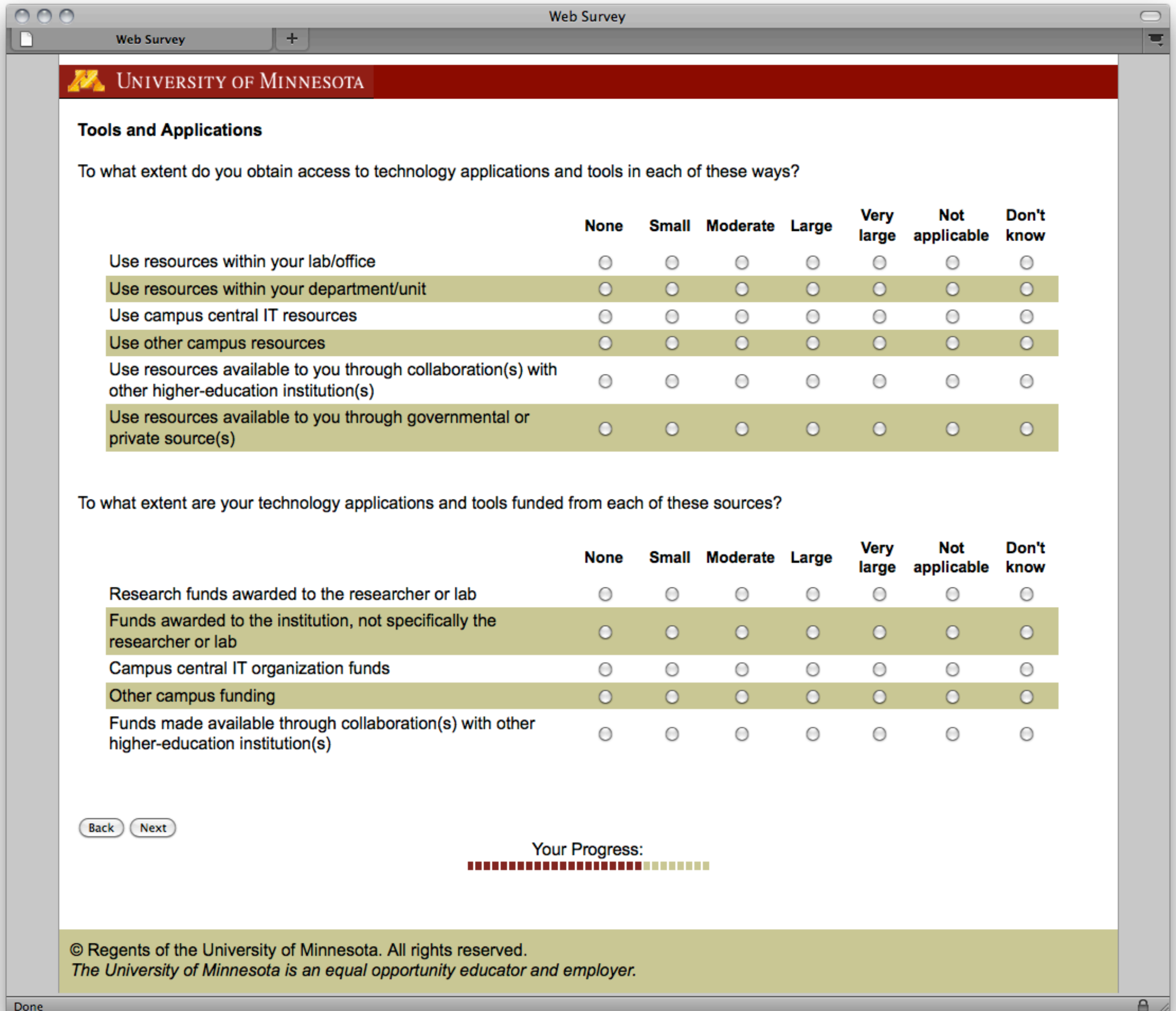
Are you familiar with Internet2?

- ○ Yes
- ○ No

Back Next

Your Progress:

Done



Web Survey

UNIVERSITY OF MINNESOTA

Tools and Applications

To what extent do you obtain access to technology applications and tools in each of these ways?

	None	Small	Moderate	Large	Very large	Not applicable	Don't know
Use resources within your lab/office							
Use resources within your department/unit							
Use campus central IT resources							
Use other campus resources							
Use resources available to you through collaboration(s) with other higher-education institution(s)							
Use resources available to you through governmental or private source(s)							

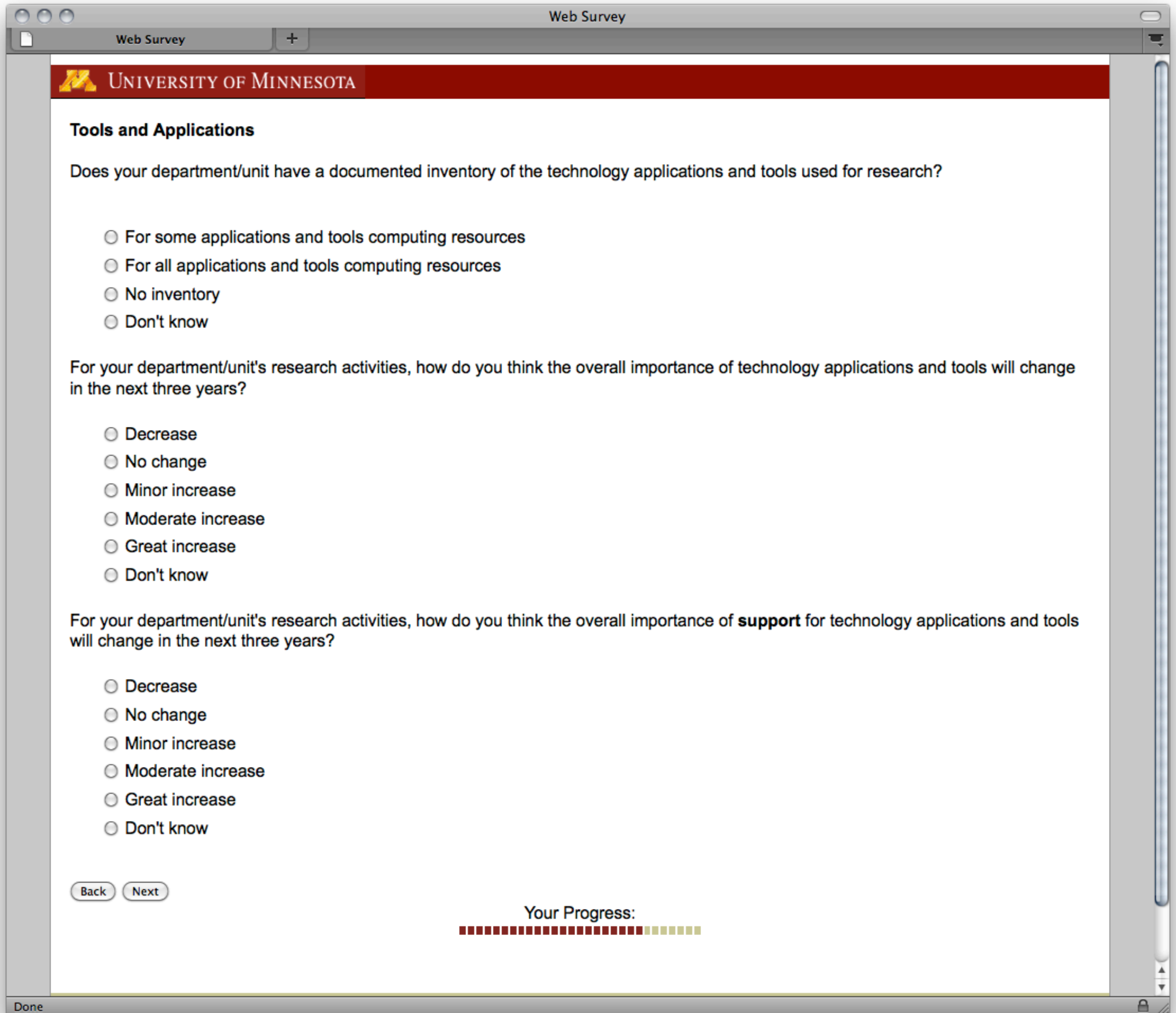
To what extent are your technology applications and tools funded from each of these sources?

	None	Small	Moderate	Large	Very large	Not applicable	Don't know
Research funds awarded to the researcher or lab							
Funds awarded to the institution, not specifically the researcher or lab							
Campus central IT organization funds							
Other campus funding							
Funds made available through collaboration(s) with other higher-education institution(s)							

Back Next

Your Progress:

Done


Web Survey

Web Survey

UNIVERSITY OF MINNESOTA

Tools and Applications

Does your department/unit have a documented inventory of the technology applications and tools used for research?

- ○ For some applications and tools computing resources
- ○ For all applications and tools computing resources
- ○ No inventory
- ○ Don't know

For your department/unit's research activities, how do you think the overall importance of technology applications and tools will change in the next three years?

- ○ Decrease
- ○ No change
- ○ Minor increase
- ○ Moderate increase
- ○ Great increase
- ○ Don't know

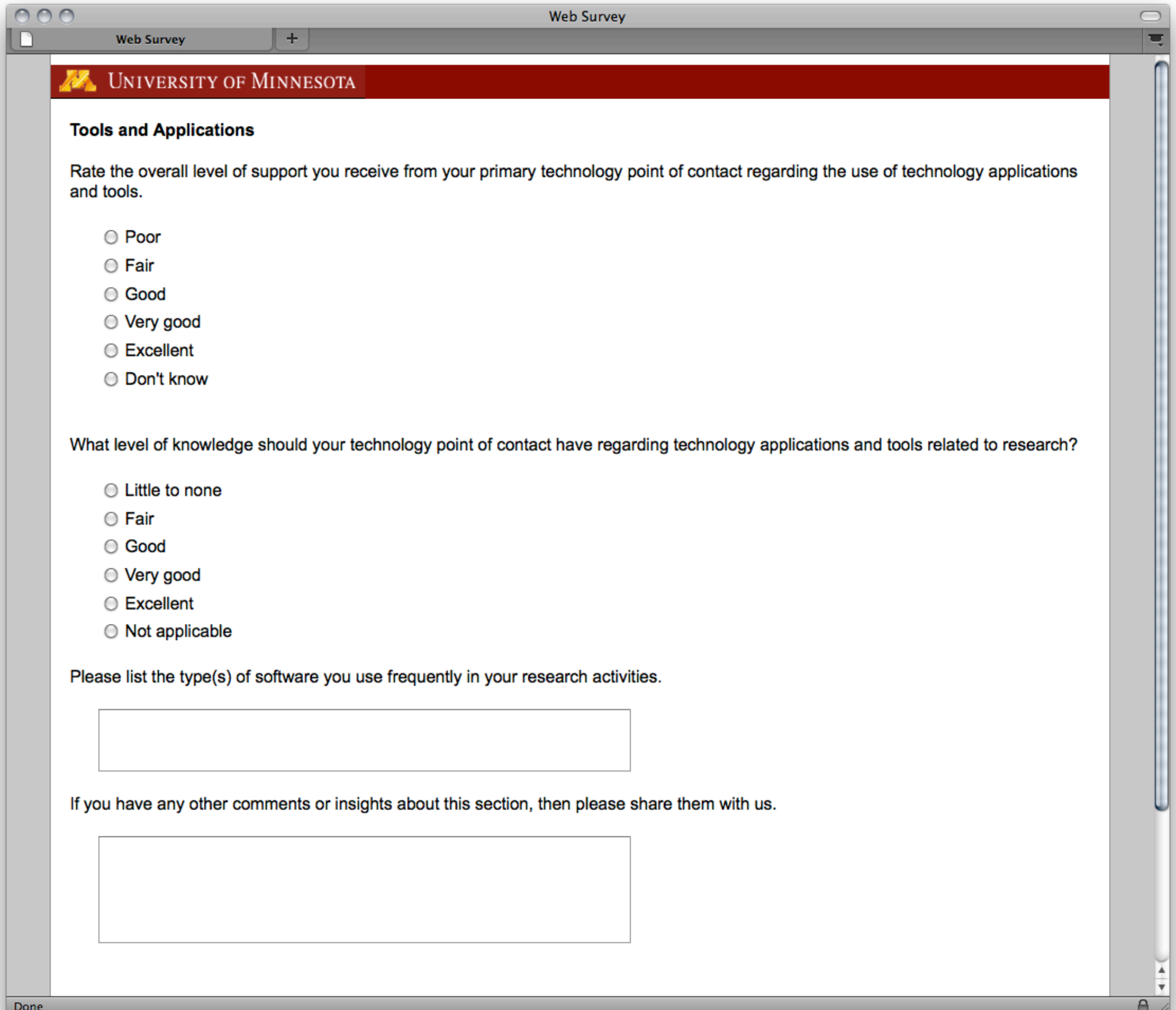
For your department/unit's research activities, how do you think the overall importance of **support** for technology applications and tools will change in the next three years?

- ○ Decrease
- ○ No change
- ○ Minor increase
- ○ Moderate increase
- ○ Great increase
- ○ Don't know

Back Next

Your Progress:

Done


Web Survey

Web Survey

UNIVERSITY OF MINNESOTA

Tools and Applications

Rate the overall level of support you receive from your primary technology point of contact regarding the use of technology applications and tools.

- Poor
- Fair
- Good
- Very good
- Excellent
- Don't know

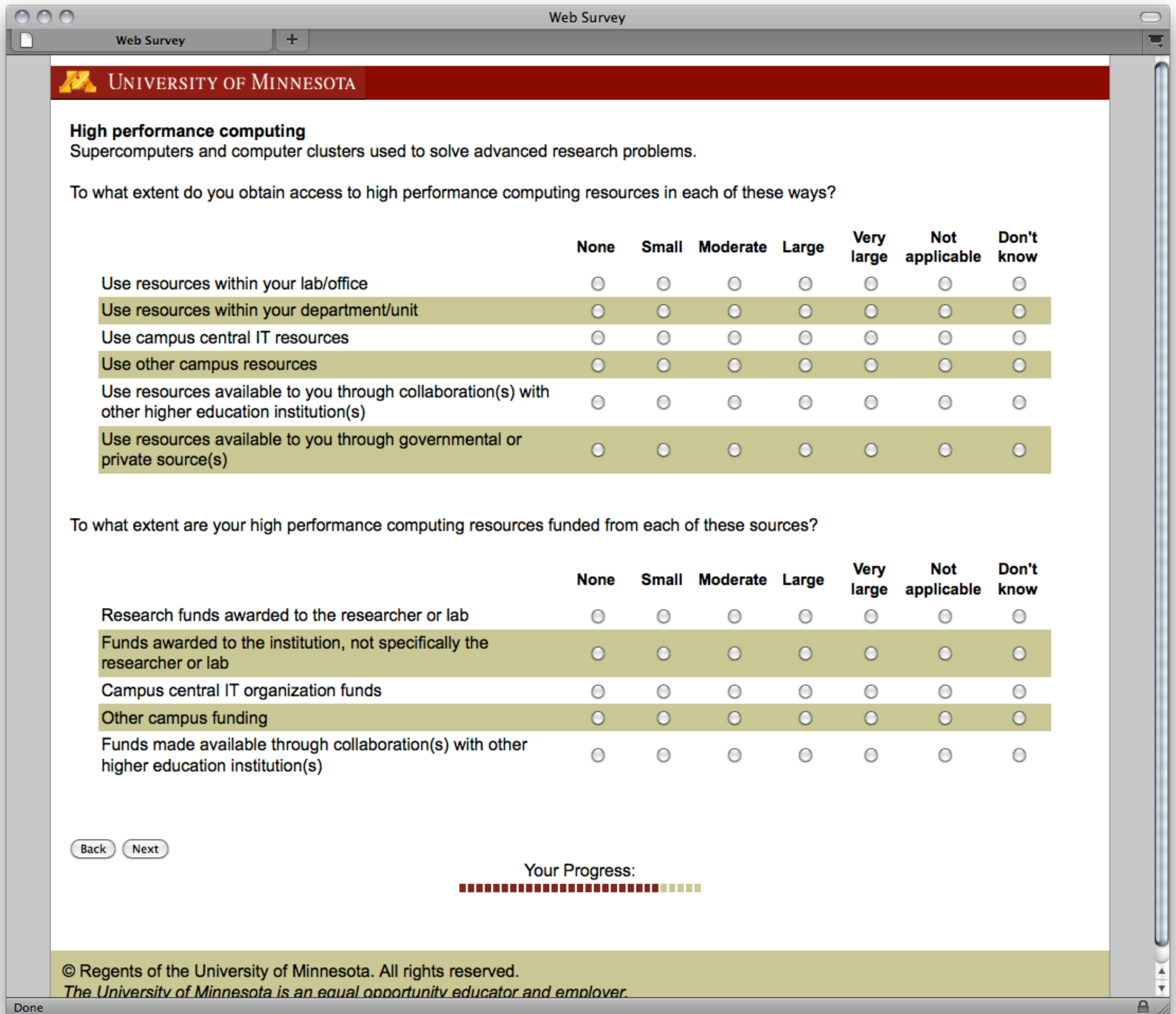
What level of knowledge should your technology point of contact have regarding technology applications and tools related to research?

- Little to none
- Fair
- Good
- Very good
- Excellent
- Not applicable

Please list the type(s) of software you use frequently in your research activities.

If you have any other comments or insights about this section, then please share them with us.

Done



Web Survey

UNIVERSITY OF MINNESOTA

High performance computing
Supercomputers and computer clusters used to solve advanced research problems.

To what extent do you obtain access to high performance computing resources in each of these ways?

	None	Small	Moderate	Large	Very large	Not applicable	Don't know
Use resources within your lab/office	○	○	○	○	○	○	○
Use resources within your department/unit	○	○	○	○	○	○	○
Use campus central IT resources	○	○	○	○	○	○	○
Use other campus resources	○	○	○	○	○	○	○
Use resources available to you through collaboration(s) with other higher education institution(s)	○	○	○	○	○	○	○
Use resources available to you through governmental or private source(s)	○	○	○	○	○	○	○

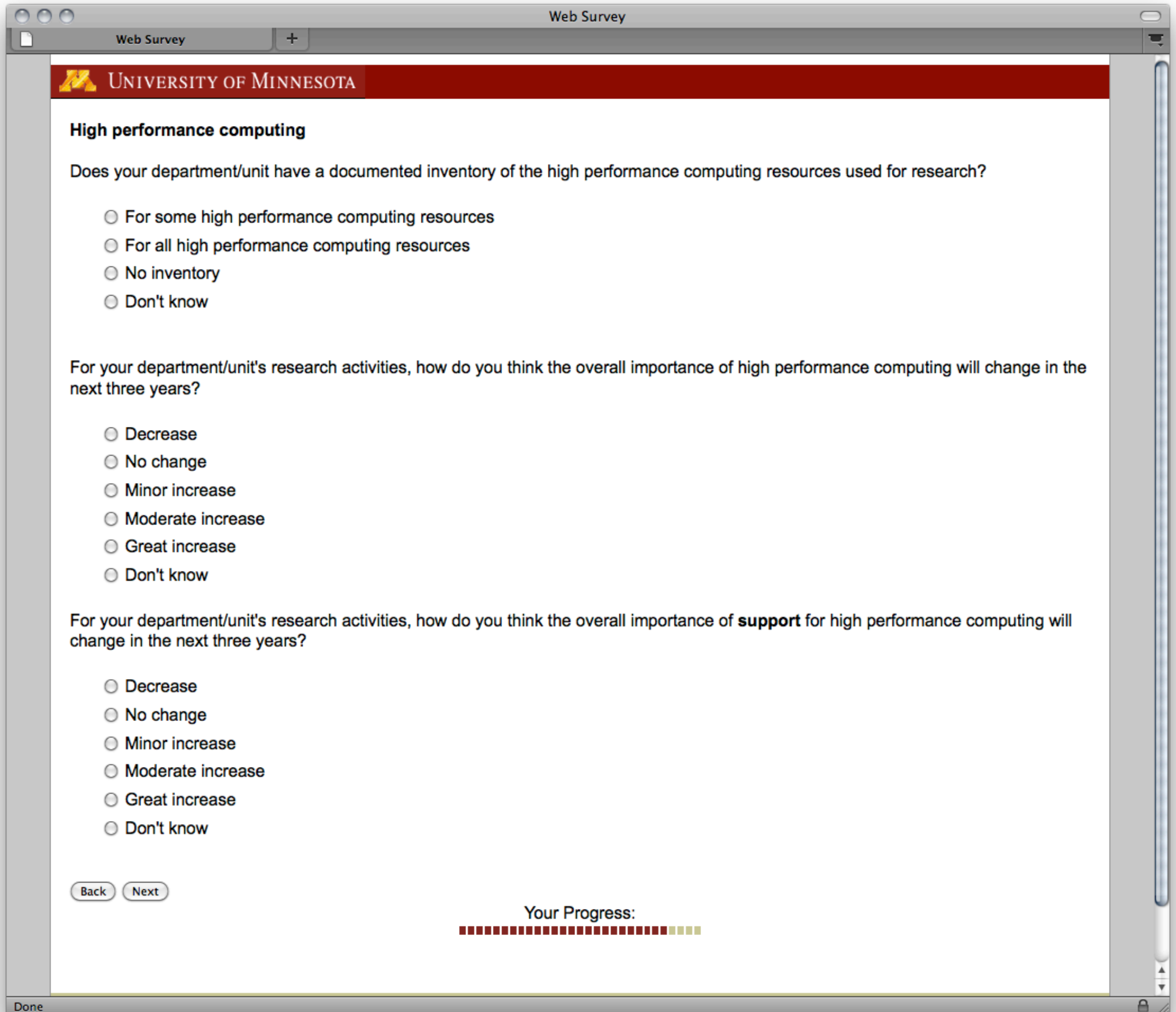
To what extent are your high performance computing resources funded from each of these sources?

	None	Small	Moderate	Large	Very large	Not applicable	Don't know
Research funds awarded to the researcher or lab	○	○	○	○	○	○	○
Funds awarded to the institution, not specifically the researcher or lab	○	○	○	○	○	○	○
Campus central IT organization funds	○	○	○	○	○	○	○
Other campus funding	○	○	○	○	○	○	○
Funds made available through collaboration(s) with other higher education institution(s)	○	○	○	○	○	○	○

Back Next

Your Progress:

Done



Web Survey

UNIVERSITY OF MINNESOTA

High performance computing

Does your department/unit have a documented inventory of the high performance computing resources used for research?

- ○ For some high performance computing resources
- ○ For all high performance computing resources
- ○ No inventory
- ○ Don't know

For your department/unit's research activities, how do you think the overall importance of high performance computing will change in the next three years?

- ○ Decrease
- ○ No change
- ○ Minor increase
- ○ Moderate increase
- ○ Great increase
- ○ Don't know

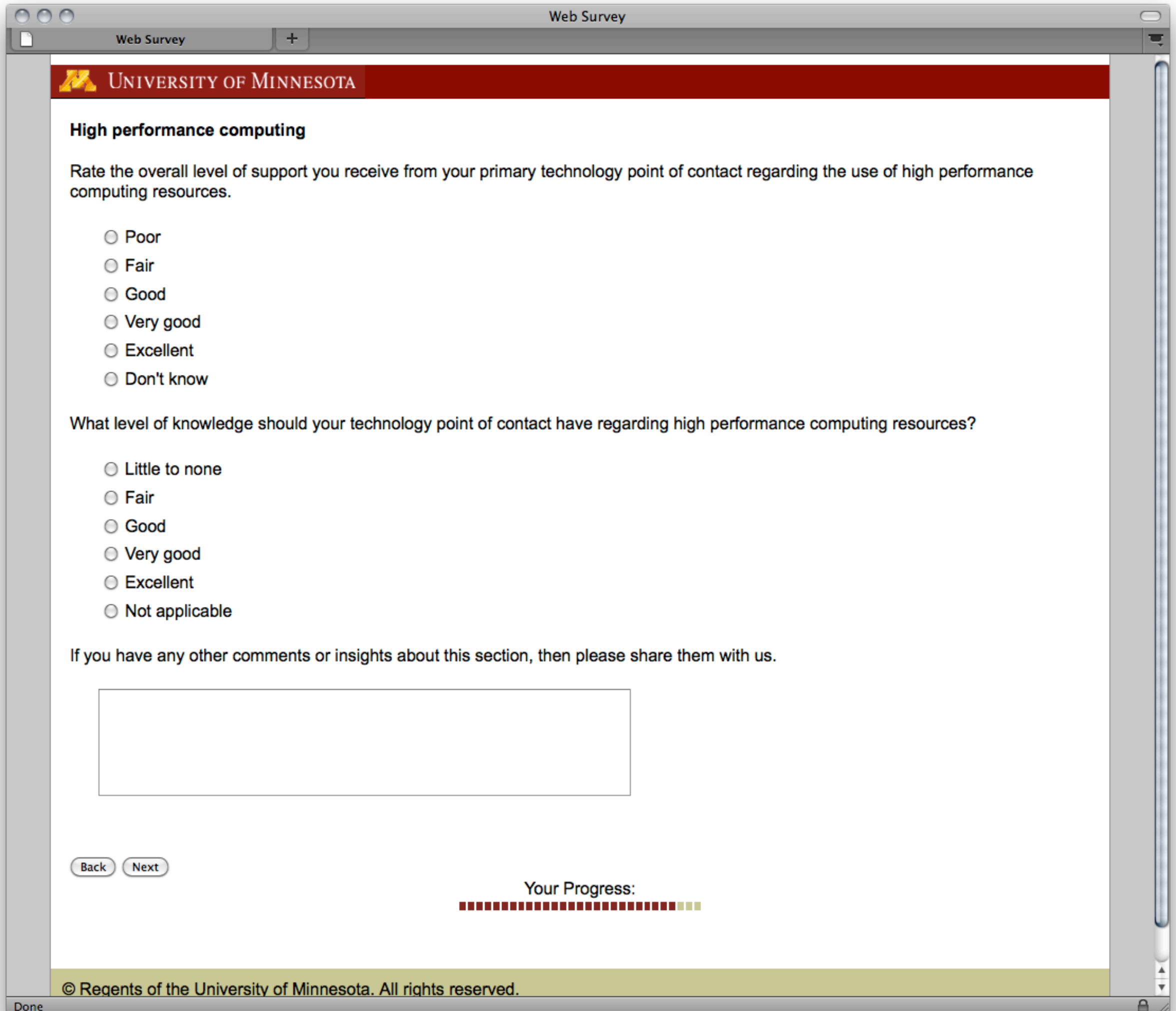
For your department/unit's research activities, how do you think the overall importance of **support** for high performance computing will change in the next three years?

- ○ Decrease
- ○ No change
- ○ Minor increase
- ○ Moderate increase
- ○ Great increase
- ○ Don't know

Back Next

Your Progress:

Done


Web Survey

Web Survey

UNIVERSITY OF MINNESOTA

High performance computing

Rate the overall level of support you receive from your primary technology point of contact regarding the use of high performance computing resources.

- ○ Poor
- ○ Fair
- ○ Good
- ○ Very good
- ○ Excellent
- ○ Don't know

What level of knowledge should your technology point of contact have regarding high performance computing resources?

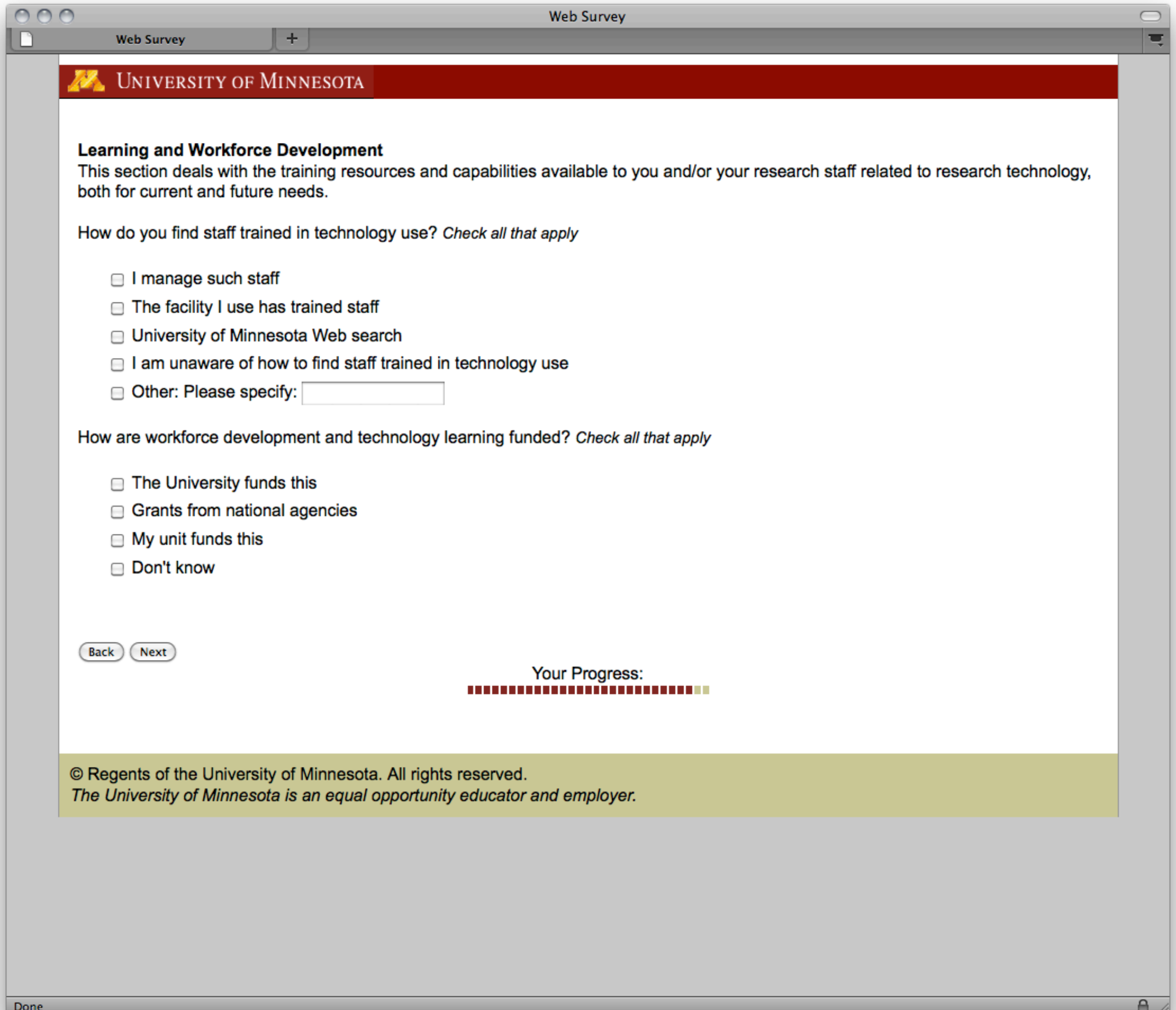
- ○ Little to none
- ○ Fair
- ○ Good
- ○ Very good
- ○ Excellent
- ○ Not applicable

If you have any other comments or insights about this section, then please share them with us.

Back Next

Your Progress:

Done



Web Survey

UNIVERSITY OF MINNESOTA

Learning and Workforce Development
This section deals with the training resources and capabilities available to you and/or your research staff related to research technology, both for current and future needs.

How do you find staff trained in technology use? *Check all that apply*

- ☐ I manage such staff
- ☐ The facility I use has trained staff
- ☐ University of Minnesota Web search
- ☐ I am unaware of how to find staff trained in technology use
- ☐ Other: Please specify:

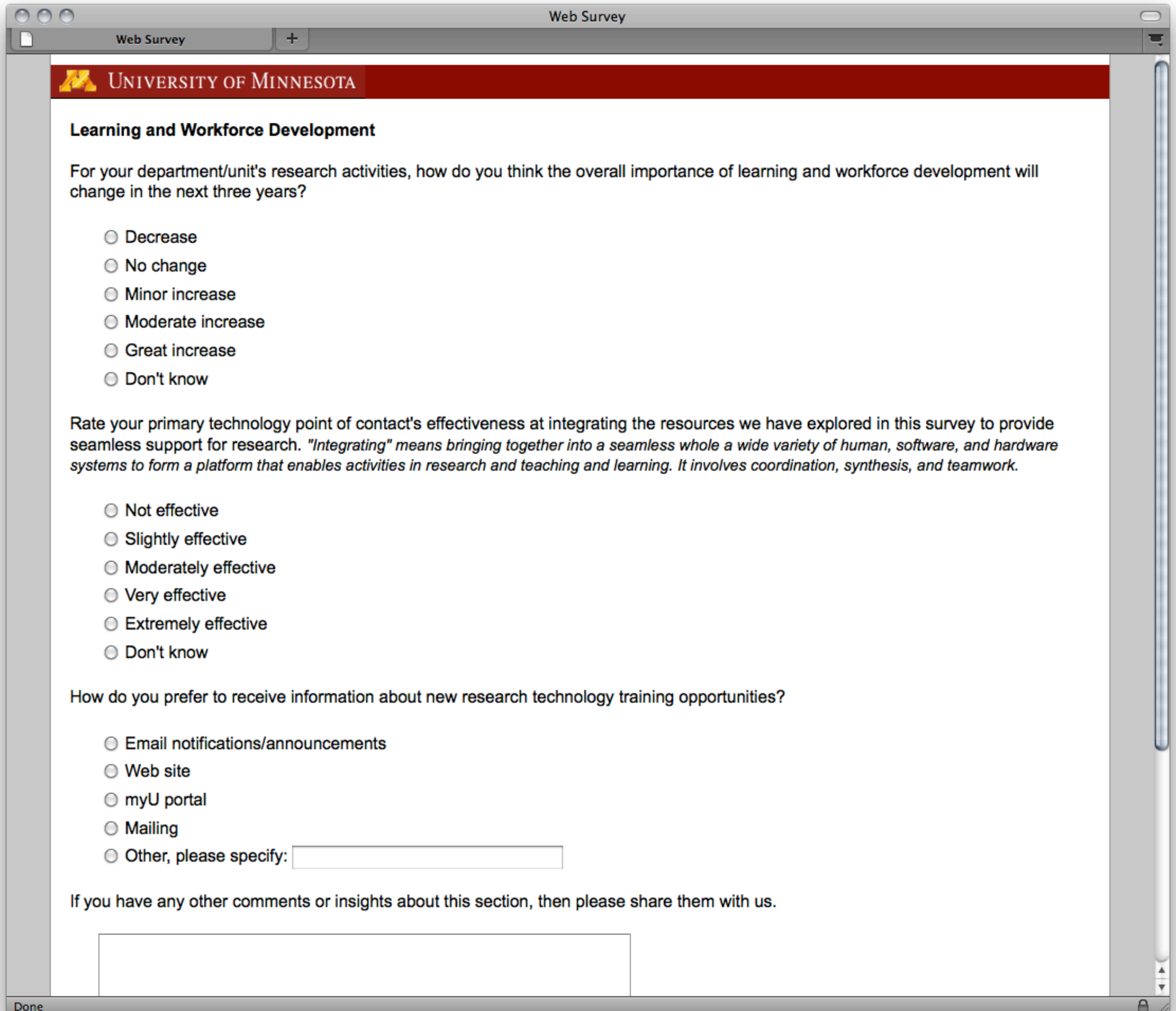
How are workforce development and technology learning funded? *Check all that apply*

- ☐ The University funds this
- ☐ Grants from national agencies
- ☐ My unit funds this
- ☐ Don't know

Back Next

Your Progress:

Done



Web Survey

UNIVERSITY OF MINNESOTA

Learning and Workforce Development

For your department/unit's research activities, how do you think the overall importance of learning and workforce development will change in the next three years?

- ○ Decrease
- ○ No change
- ○ Minor increase
- ○ Moderate increase
- ○ Great increase
- ○ Don't know

Rate your primary technology point of contact's effectiveness at integrating the resources we have explored in this survey to provide seamless support for research. *"Integrating" means bringing together into a seamless whole a wide variety of human, software, and hardware systems to form a platform that enables activities in research and teaching and learning. It involves coordination, synthesis, and teamwork.*

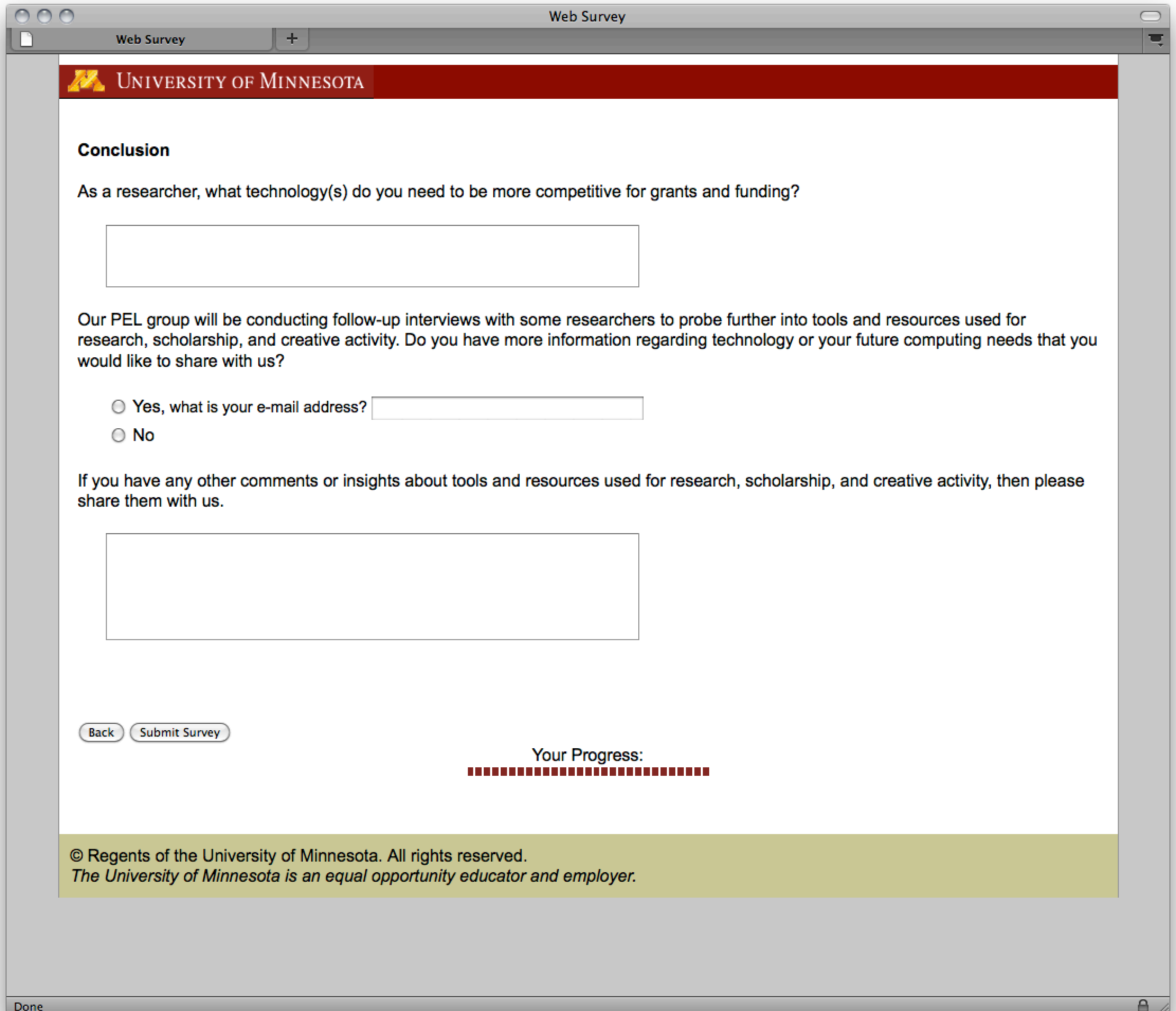
- ○ Not effective
- ○ Slightly effective
- ○ Moderately effective
- ○ Very effective
- ○ Extremely effective
- ○ Don't know

How do you prefer to receive information about new research technology training opportunities?

- ○ Email notifications/announcements
- ○ Web site
- ○ myU portal
- ○ Mailing
- ○ Other, please specify: ______

If you have any other comments or insights about this section, then please share them with us.

Done



Web Survey

Web Survey

UNIVERSITY OF MINNESOTA

Conclusion

As a researcher, what technology(s) do you need to be more competitive for grants and funding?

Our PEL group will be conducting follow-up interviews with some researchers to probe further into tools and resources used for research, scholarship, and creative activity. Do you have more information regarding technology or your future computing needs that you would like to share with us?

- Yes, what is your e-mail address?
- No

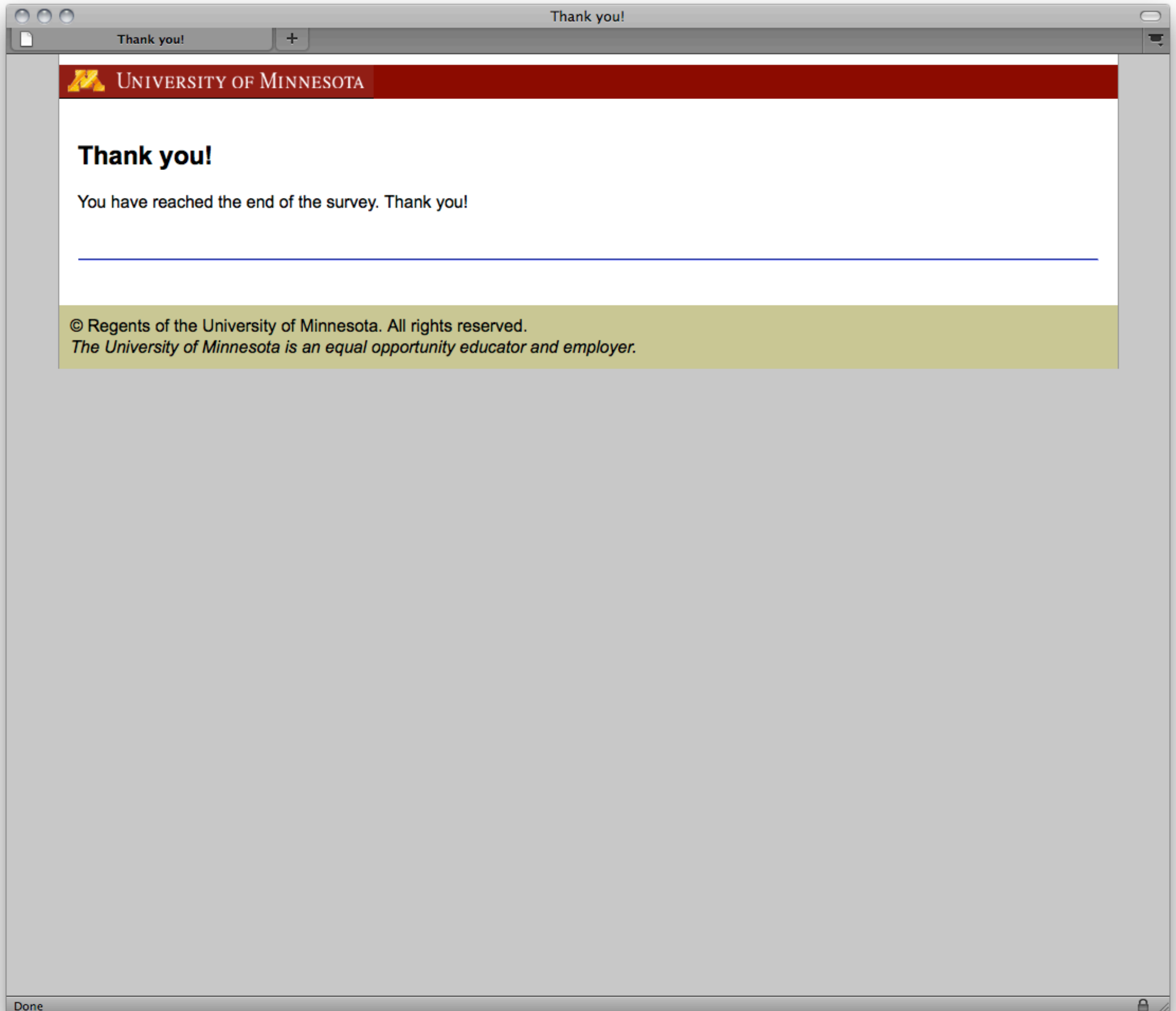
If you have any other comments or insights about tools and resources used for research, scholarship, and creative activity, then please share them with us.

Back | Submit Survey

Your Progress:

The University of Minnesota is an equal opportunity educator and employer.

Done


Thank you!
Thank you!
University of Minnesota
Thank you!
You have reached the end of the survey. Thank you!
© Regents of the University of Minnesota. All rights reserved.
The University of Minnesota is an equal opportunity educator and employer.
Done

Appendix G. PEL Survey Result Data, Graphs, Handouts, and Poster (Starts on next page.)

What are ways to forge a common cyberinfrastructure with central and distributed support? ▶
How can the University compete at the national level for cyberinfrastructure grants and resources? ▶

Observed Trend:

Local data storage access and back-up support is more commonly used over University- or cloud-level storage. However comments reflect a need/desire for more central support.

Frequency of data access is daily by most researchers and nearly all in the physical sciences. Most researchers want to keep their data forever.

What they are doing:

Only 14% of respondents use central data storage solutions. More prevalent were work computer desktops and laptops (63% and 39% or users) and departmental servers (40%). External hard drives and other media were also noted by 15% of respondents.

Back-up solutions show a preference for local and "low-tech" storage. Nearly half of respondents (43%) reported using a secondary hard drive to back-up their data, and 29% use CD/DVD or other removable media. Interestingly, 72% of researcher have never lost data due to the lack of a back-up (27% have).

Compared by Discipline:

Storage solutions and backup were consistent across all disciplines but vary by user type.

Grad students more likely to store data on personal computers. Faculty are more likely to store data on Central.

What they say:

"In my research group, research data is mainly managed by the respective owner. We evaluated the file sharing service from central IT but found it too cumbersome to use. It works best if research groups can set up their own servers to have control over their data storage platform."

-Post Doc in the Institute of Technology

"There needs to be a mechanism for centralized storage of data with high security for individual research groups. This at least partially removes the responsibility for secure backups from the research group alone."

-PhD Research Assistant, Institute of Technology

PEL Recommendation:

1

Develop enterprise-wide, integrated cyber-infrastructure to align with national efforts.

Accountability at the senior leader level is needed to ensure the University's competitiveness in the future (i.e., obtain grants, become a top three public research university, recruit, evolve research and establish new programs). At the same time, the University has the opportunity to leverage existing 'cottage industries' and to align internal services.

Tracy Anderson, College of Biological Sciences <> Craig Gjerdingen, Carlson School of Management <>Bryan Herrmann, Office of Admissions, Morris Campus <> Katherine Himes, Office of the Senior Vice President for Academic Affairs and Provost <> Lisa Johnston, University Libraries

Are colleges and research centers meeting the needs of their researchers? ▶
How can units across the University best work in partnership? ▶

Observed Trend:

Local IT department/unit contact is primary end-user resource for research needs including storage, electronic security, space & environment, and system administration. Central University level services are contacted for compliance & privacy issues, and bandwidth needs. Outside data services are not needed for many.

Compared by Discipline:

Health Sciences are more likely to contact university-level service for support (15% vs~5%). Could be an artifact of how AHC views "university level". Outside data services are not needed particularly in the Health Sciences.

What they are doing:

A majority of respondents access the following resources and support through their "local IT point of contact" based in their department/college/institute: Data storage (68%), electronic security (65%), space/environment (56%), systems administration (64%). However local connections are much more split with University level service centers when it came to Compliance & privacy (local 41% vs. univ 28%), bandwidth (local 46% vs. univ 34%), and obtaining outside data (26% vs. 16% univ).

The extent to obtain access for the following resources also indicated a preference (very large to moderate) for local support in the lab and departments for data storage (83% and 55% respectfully), virtual collaboration (64% and 41% respectfully), software and tools (85% and 55% respectfully), and, to a lesser extent, high performance computing (47% and 23%).

What they say:

"It's hard to know who to contact in the tech area for some research tech support. The basic tech person assigned to the dept. is fine for routine computer and computer related software things, but beyond the basics it's not clear who to contact for what."

-Faculty Member in the College of Liberal Arts

"As far as I know, there is no inventory for data storage and we don't have a contact person for data storage. We just save our data in multiple places and on an external hard drive. There is no oversight for data storage."

- Post-Doctoral Fellow, School of Dentistry

PEL Recommendation:

2

Better position data storage, management, and networking services around the end-user.

Provide IT contacts with the tools to share & leverage existing resources across University service providers. Researchers do not care who offers the service; they just want the opportunity to take advantage of the offering.

Tracy Anderson, College of Biological Sciences <> Craig Gjerdingen, Carlson School of Management <>Bryan Herrmann, Office of Admissions, Morris Campus <> Katherine Himes, Office of the Senior Vice President for Academic Affairs and Provost <> Lisa Johnston, University Libraries

Implementing Cyberinfrastructure for 21st Century Research

2008-2009 President's Emerging Leaders Program

What issues are researchers dealing with in research computing?

Observed Trend:

The amount of data storage may be driven by a high demand from a small number of users from all disciplines.

Frequency of data access is daily by most researchers and nearly all in the physical sciences. Most researchers want to keep their data forever.

What they are doing:

The majority of researchers reported generating less than 1 GB of data per week (80%). 19% of respondents do generate greater than 1GB of data per week.

Data are accessed daily by 64% of researchers, weekly by 25%, and monthly, quarterly, yearly at 5%, 3%, and 1% respectfully.

Data are typically stored indefinitely by 70% of the population and are shared with departmental colleagues.

Compared by Discipline:

Of the 19% who do generate over 1 gigabyte of data per week, interestingly the rate of responses by discipline did not suggest a significant difference between Biological (21%), Physical (15%), Engineering (14%), Health (19%), Interdisciplinary (11%) and Social Sciences (15%), as each ranged between 11-21%. Humanities and Arts represented 2% respectively.

Humanities researchers access their data as frequently (75% daily) as Biological and Engineering Researchers.

Daily access is most consistent in the physical sciences (89%) and by many in Arts (56%), Biological (75%), Engineering (71%), Health (59%), Humanities (75%), Interdisciplinary (62%). Social Sciences access their data daily (48%), weekly (33%) or monthly (10%)

What they say:

"Once or twice a year I have to negotiate for storage space; I always feel as if my research is at odds with IT policy. I always feel as if I'm living on borrowed time, no confidence in having access to adequate data storage for research in the future."

-Faculty Member in the College of Education & Human Development

"Instead of looking at individual requirements for storage, management, software and security take a look at how CLA now manages its research services. There is one single point of contact , they provide space, security, management, software -- all you do is apply and you get an account WITH phone or email support!"

-Faculty Member in the College of Liberal Arts

PEL Recommendation:

3

Improve and expand the University's ability to handle data.

In order to meet the growing needs of researchers, IT service providers should provide short-term data storage as needed (ie. a bank model), utilize external cloud storage for long-term archiving, and provide high-speed data connectivity and automatic back-up with the appropriate compliance and privacy mechanisms.

Tracy Anderson, College of Biological Sciences <> Craig Gjerdingen, Carlson School of Management <>Bryan Herrmann, Office of Admissions, Morris Campus <> Katherine Himes, Office of the Senior Vice President for Academic Affairs and Provost <> Lisa Johnston, University Libraries

How do we engage key stakeholders across campus to spread the vision of cyberinfrastructure? ▶

Observed Trend:

Humanities feel that this survey did not 'speak' to them; however, humanities has self-identified cyberinfrastructure needs. These trends were evident in the comment section of the survey.

What they say:

"This instrument seems to have been constructed using the sciences as its template. I find it hard to translate the terms into my research in the Humanities."

- Faculty member in the College of Liberal Arts

PEL Recommendation:

In order to bridge the jargon gap among the research community, the service providers and University leaders should re-brand cyberinfrastructure to ensure efficient workforce training and awareness of technology needs.

4

Re-brand Research Cyberinfrastructure at the University to improve accessibility for all disciplines.

What are the strengths of the University in research computing? ▶
What are the best practices for articulating central and distributed support? ▶

Observed Trend:

Data is shared by nearly all researchers, most share locally on campus.

Department data sets and resources are not well documented.

The majority of all disciplines did not have or know of any documented inventory of data sets.

What they are doing:

Data is shared by 92% of all researchers, but primarily with researchers in their own unit or on campus (51%, 18% respectively).

Only 8% of researchers do not share their data at some level. Of those, 52% would share their data, 30% would not and 19% did not know.

What they say:

"International collaboration will increase; how do we encourage & assure compliance with security standards from our collaborators?"

-Faculty in the College of Education and Human Development

5

Create incentives for virtual collaborative relationships.

PEL Recommendation:

Give researchers tools, such as an informative web site, access to data inventories, and awareness of virtual conferencing, to help develop collaborative relationships that will better align the University as top research institution.

Tracy Anderson, College of Biological Sciences <> Craig Gjerdingen, Carlson School of Management <>Bryan Herrmann, Office of Admissions, Morris Campus <> Katherine Himes, Office of the Senior Vice President for Academic Affairs and Provost <> Lisa Johnston, University Libraries

President's Emerging Leaders (PEL) Research Cyberinfrastructure Project

2008-2009 President's Emerging Leaders Program

Tracy Anderson, College of Biological Sciences
Craig Gjerdingen, Carlson School of Management
Bryan Herrmann, Office of Admissions, Morris Campus
Katherine Himes, Office of the Senior Vice President for Academic Affairs and Provost
Lisa Johnston, University Libraries

The 2008-2009 PEL Project Survey Overview:

- Released March 24th – April 8th 2009.
- Delivered to 8603 research faculty, postdocs, student research assistants, and center and institute directors on all UMN campuses.
- Received 780 responses overall.

Goals of the PEL Survey:

- To assess the current and future needs of research cyberinfrastructure (technology + support) at the University of Minnesota
- To help our sponsors and members of RCA shape the future direction of research technology support at the University

PEL RCA Project: Survey Sections

1. **Quantitative Results**
   - Demographics*
   - Data Management and Storage
   - Collaboration and Virtual Communities
   - Software Applications and/or Instrumentation
   - High Performance Computing
   - Learning and Workforce Development
   - Combined Sections
2. **Qualitative Results**
   - Comment Trends Analysis
   - Fill-in Answer Trends
3. **Cross-Tabulation by Subject Discipline**
4. **Retrospective: Survey Response and Related Issues**
5. **Final Thoughts**

**Our survey results are broken down into 5 areas that 780 respondents self-selected to answer.*

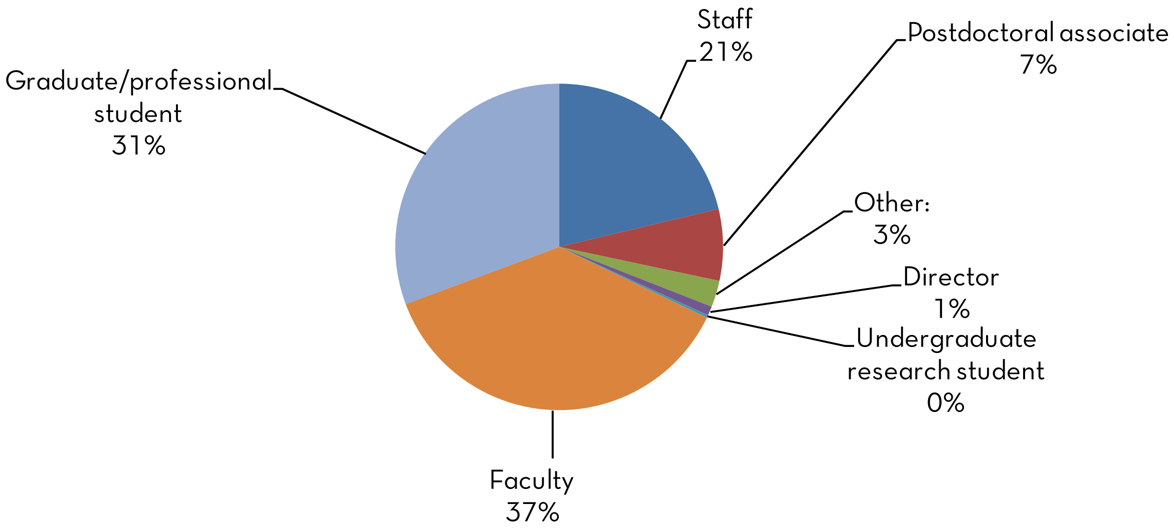
Quantitative Results

Demographics

780 responses

Demographics

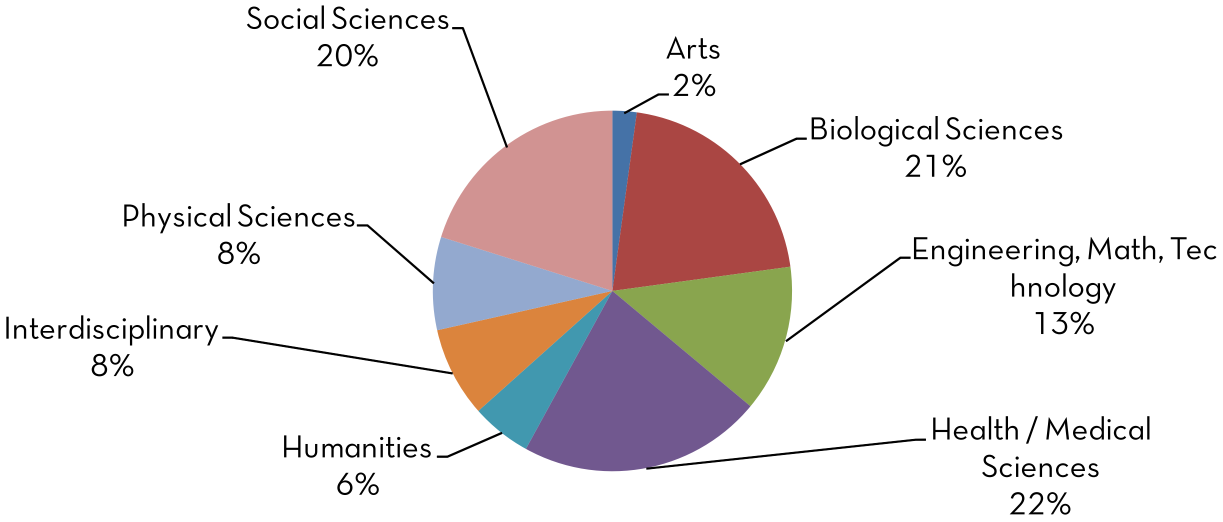
What is your primary role here at the University?



Cross-section of researcher roles.

Demographics

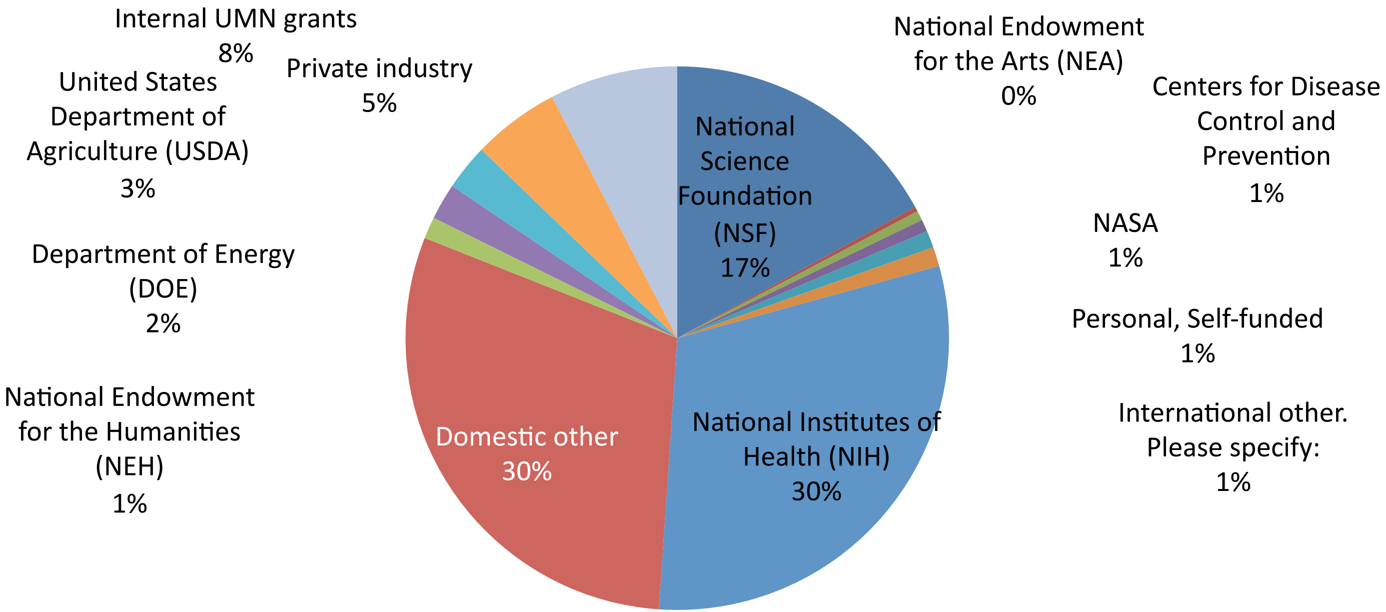
What best describes your research area?



Cross-section of academic area representation. Fields were chosen based on major areas on the five campuses.

Demographics

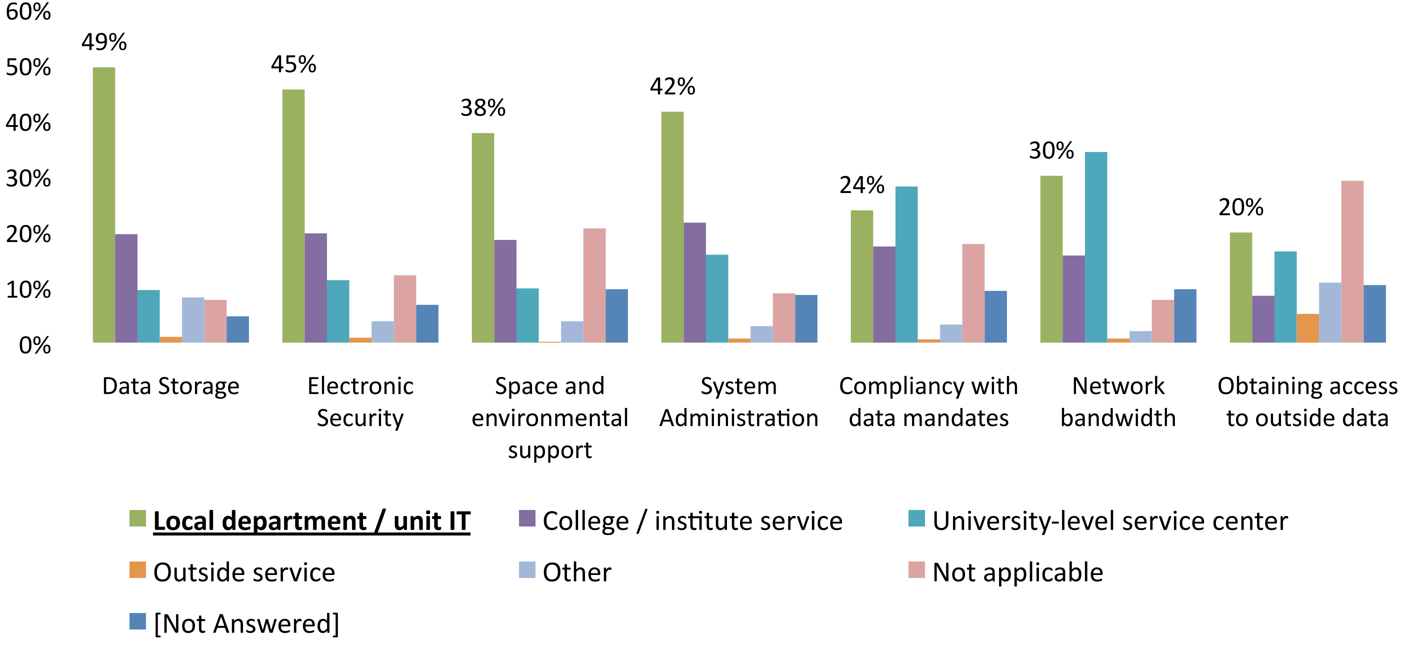
What is the primary source of your research funding?



"Domestic Other" includes specific internal, private, and government grant funding bodies.

Demographics

What department/facility at the University do you contact for the following support?



Respondents were instructed to refer to these people as their "technology points of contact" for the remainder of survey.

Quantitative Results

Data Management and Storage

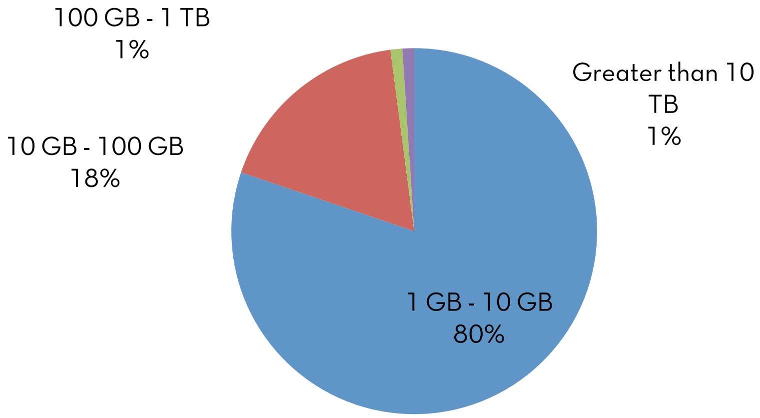
494 responses

Data Management and Storage

In a typical week do you generate more than one GB of research data?

- **No 80%**
- **Yes 19%**

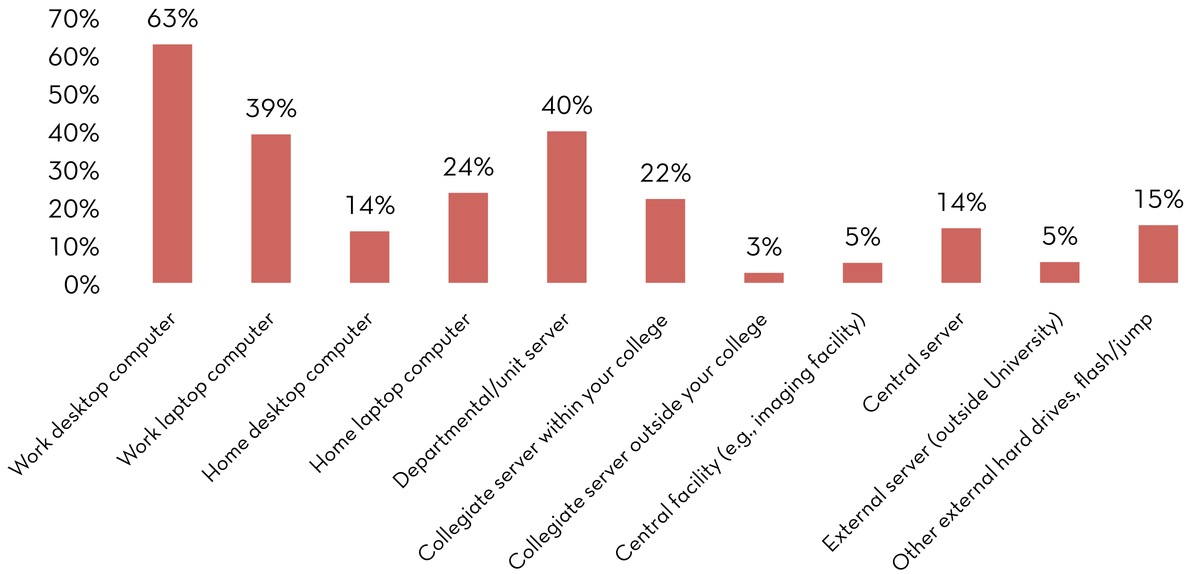
If so, how much research data do you generate?



The majority of respondents generate less than 1 GB of data per week.

Data Management and Storage

Where do you store your research data electronically?



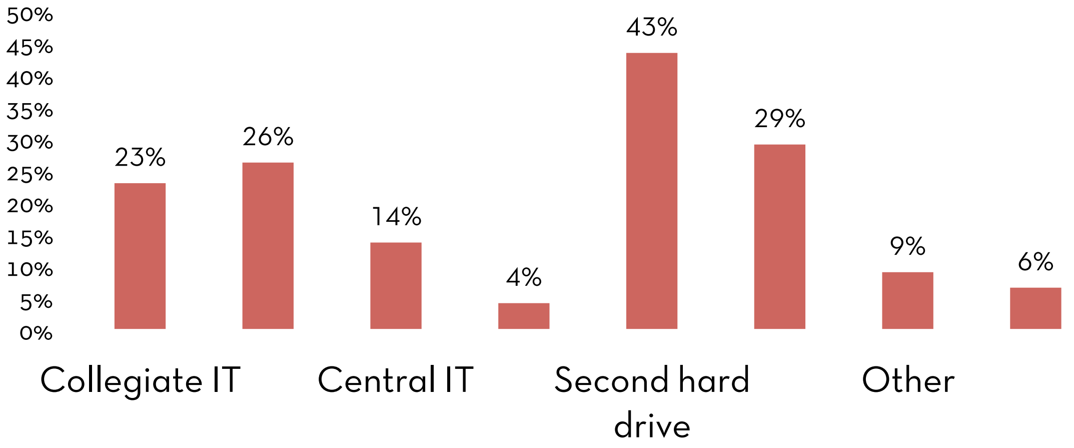
Less than 10% of respondents use central data storage solutions (collegiate, OIT, or University-level servers).

Data Management and Storage

Have you ever lost important data due to the lack of a back up?

- **Yes 27%**
- **No 72%**

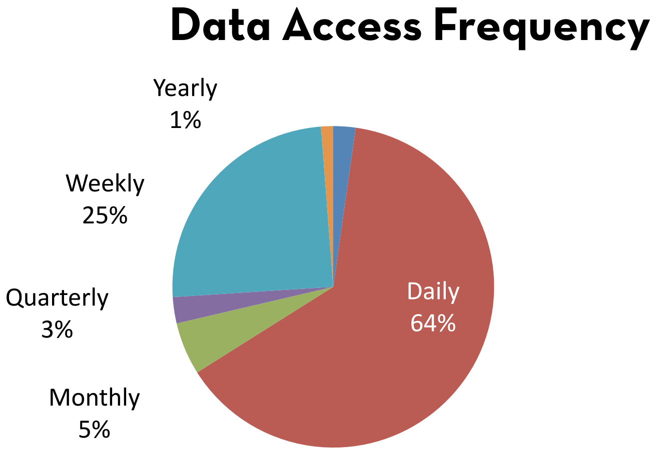
How are your electronic data backed up?

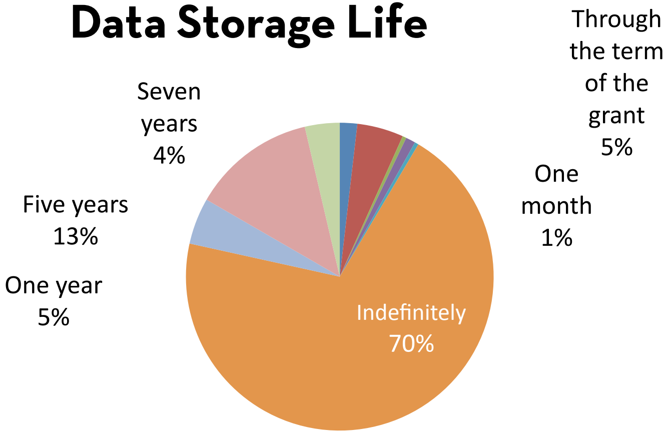


A second, or external, hard drive is a primary storage solution among respondents.

Data Management and Storage

How often do you access electronic data? How long do you typically keep your electronic research data?

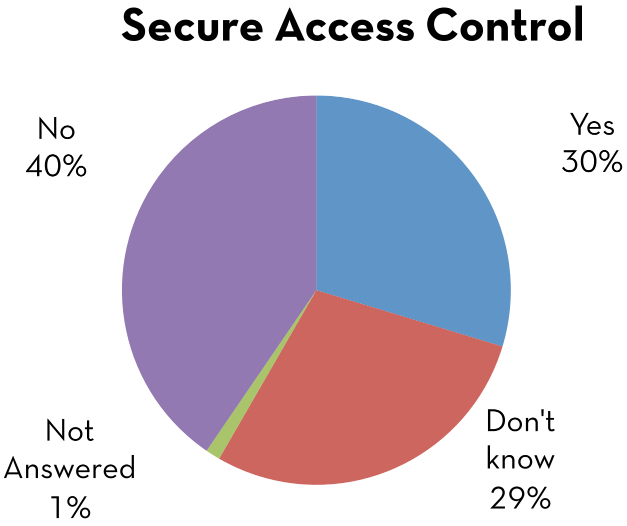


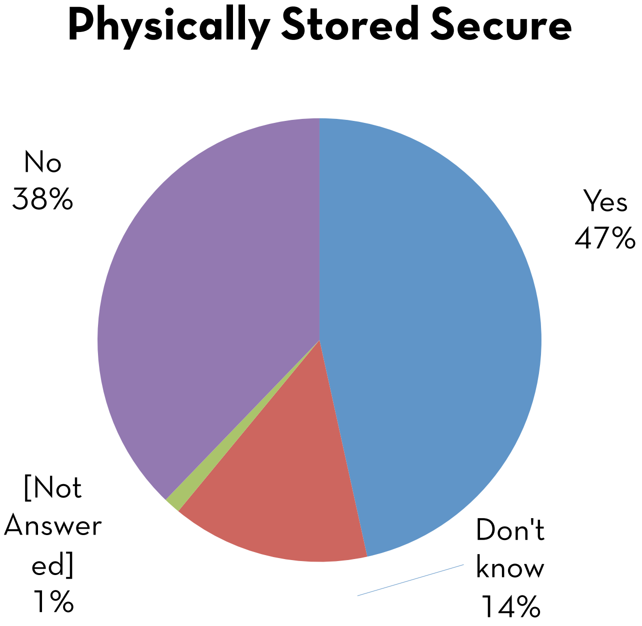


Data storage must provide easy access and last forever.

Data Management and Storage

Do your data need to be stored securely?

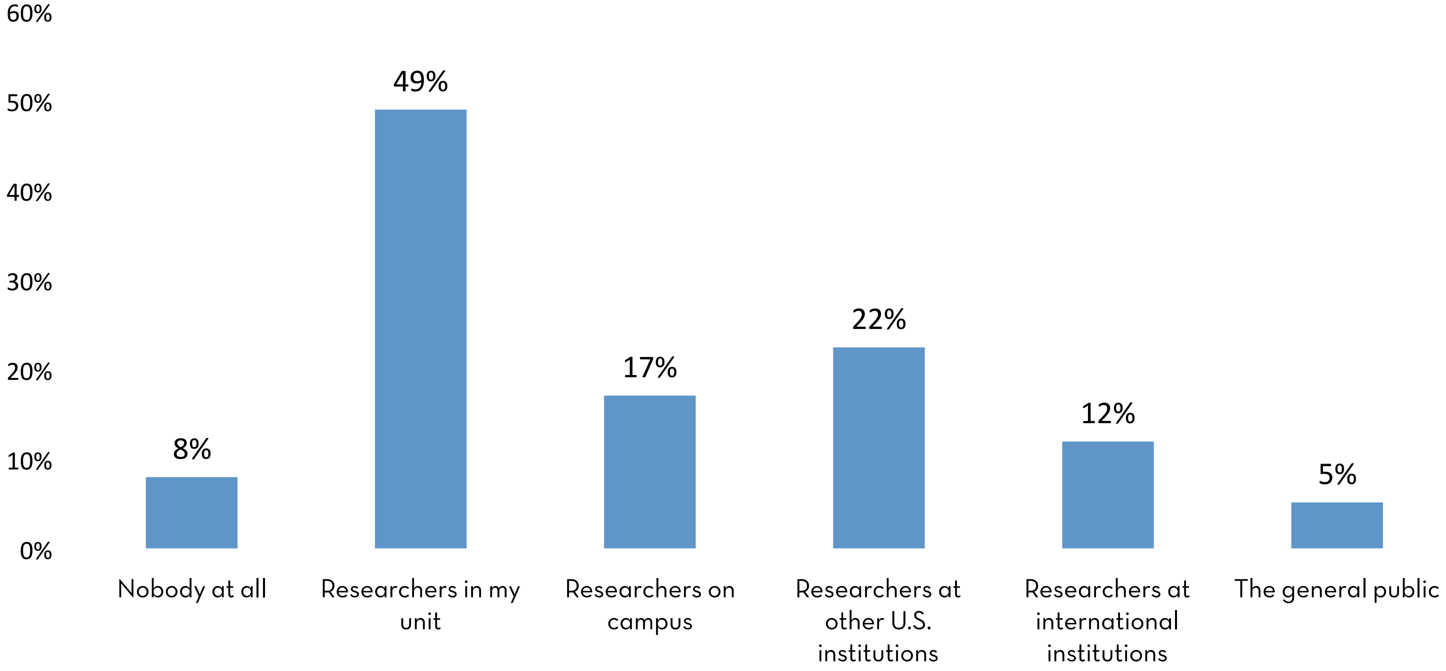




Many researchers do not require secure data storage.

Data Management and Storage

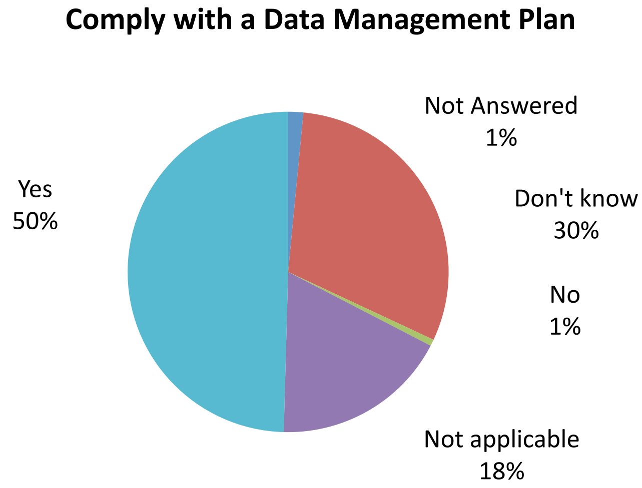
Would you share any of your data? If so, with whom?

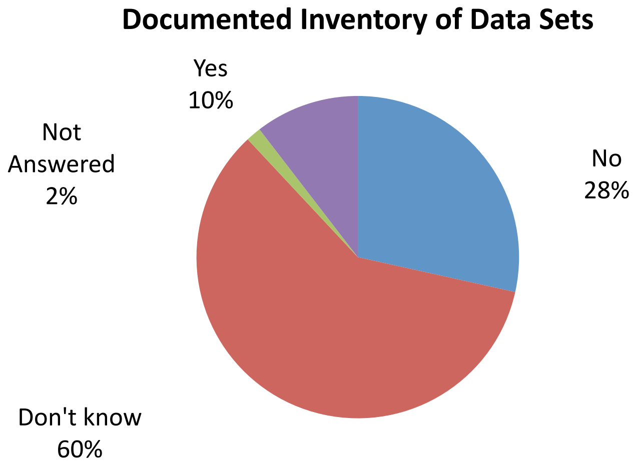


Nearly 92% of researchers are currently sharing their data.

Data Management and Storage

Do you comply with your primary funding agency's suggested data management plan? Does your department/unit have a documented inventory of electronic data sets?





Half of respondents are adhering to funding agency requests to successfully store and manage data.

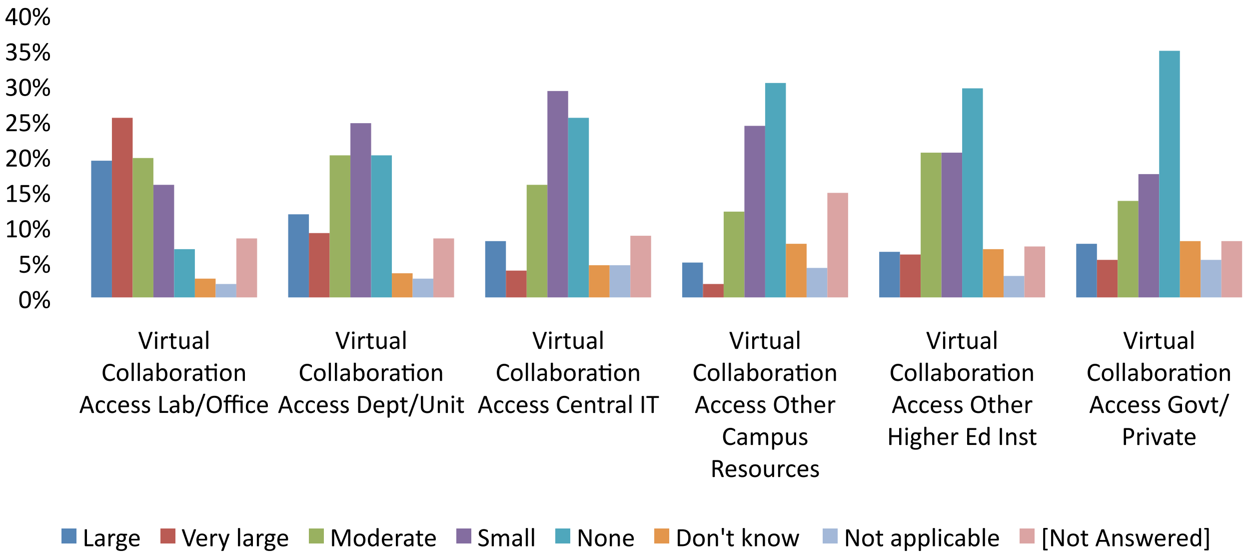
Quantitative Results

Collaboration and Virtual Communities

289 Responses

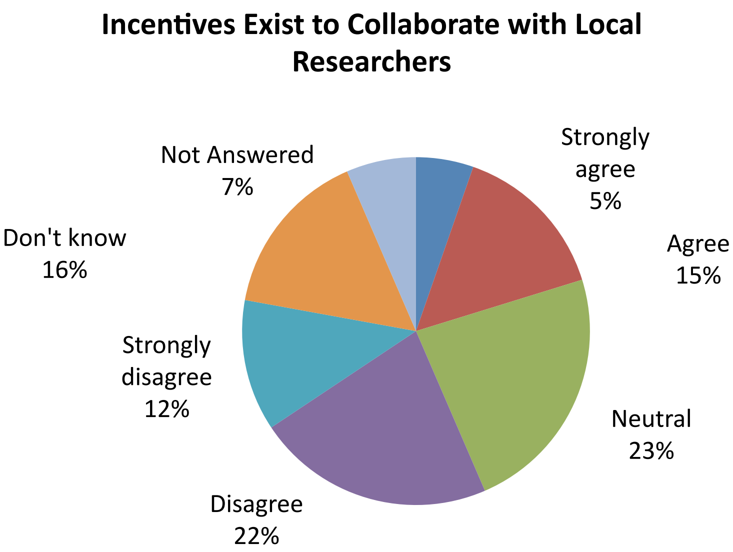
Collaboration and Virtual Communities

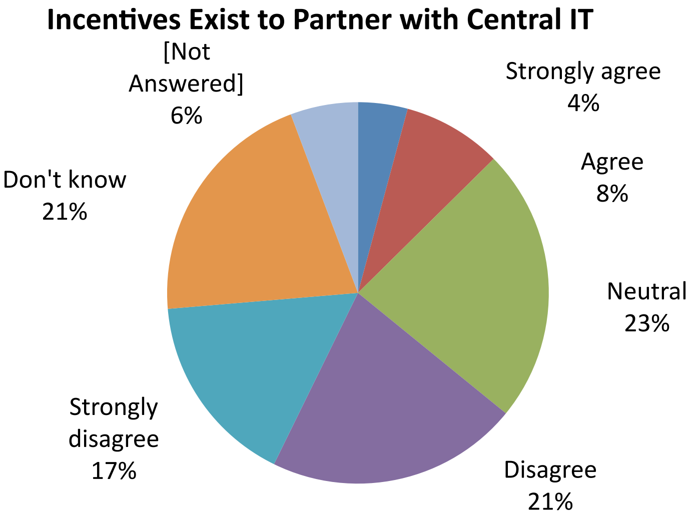
To what extent do you obtain access to resources for collaboration in each of these ways?



Collaboration and Virtual Communities

Within my department/unit, effective incentives exist to encourage researchers to share technology with other researchers on campus and partner with central IT to achieve economies of scale.





There is no clear consensus on incentives to collaborate with others on campus.

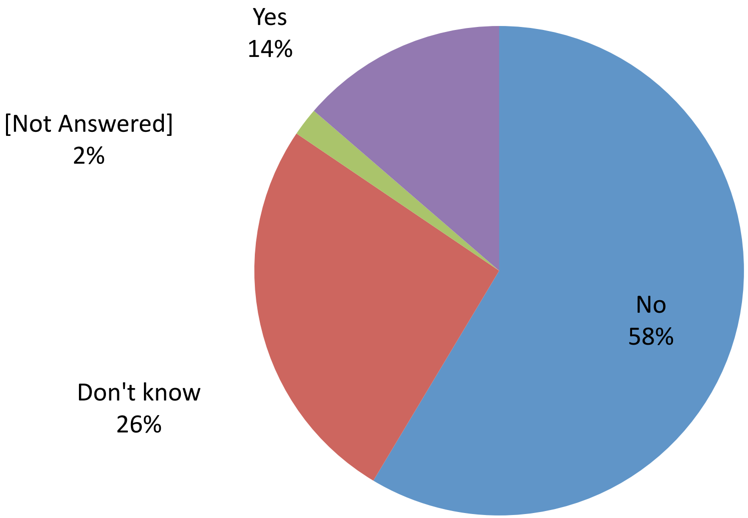
Quantitative Results

Software Applications and/or Instrumentation

521 Responses

Software Applications and/or Instrumentation

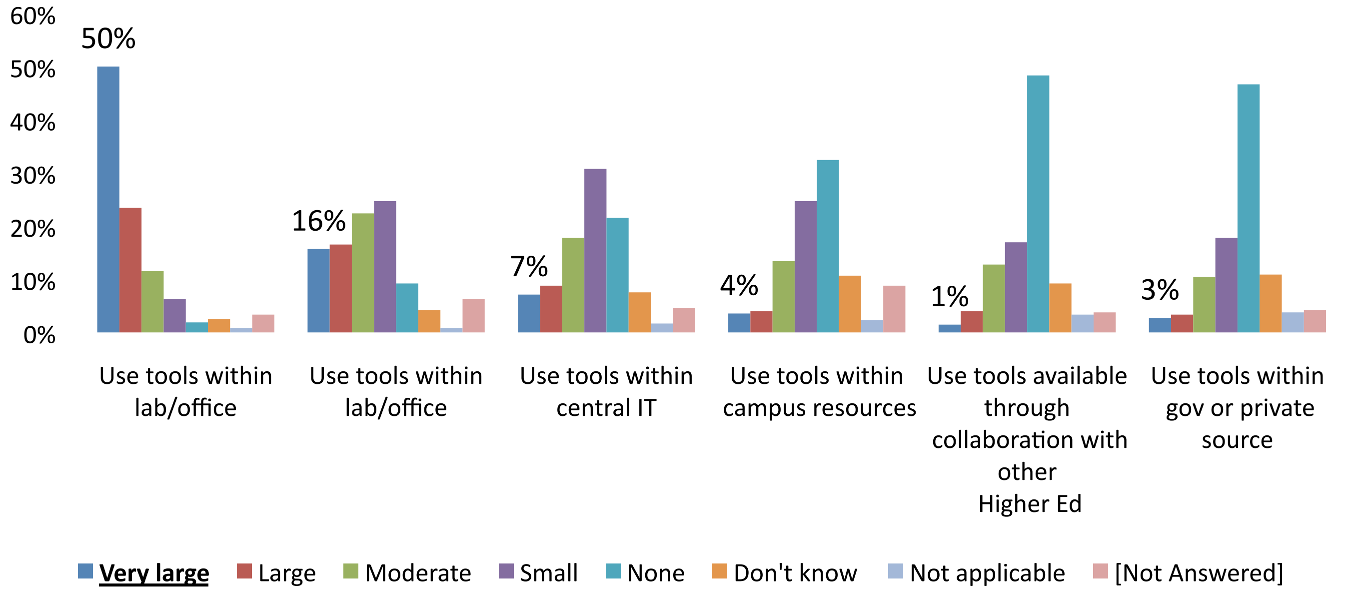
Do you use high bandwidth technologies (greater than GB) to access research resources on a regional, national, or international basis?



Researchers are not in need of or unaware of the high bandwidth resources on campus.

Software Applications and/or Instrumentation

To what extent do you obtain access to technology applications and tools in each of these ways?



Outside of their unit, researchers are less likely to obtain access to software and instrumentation.

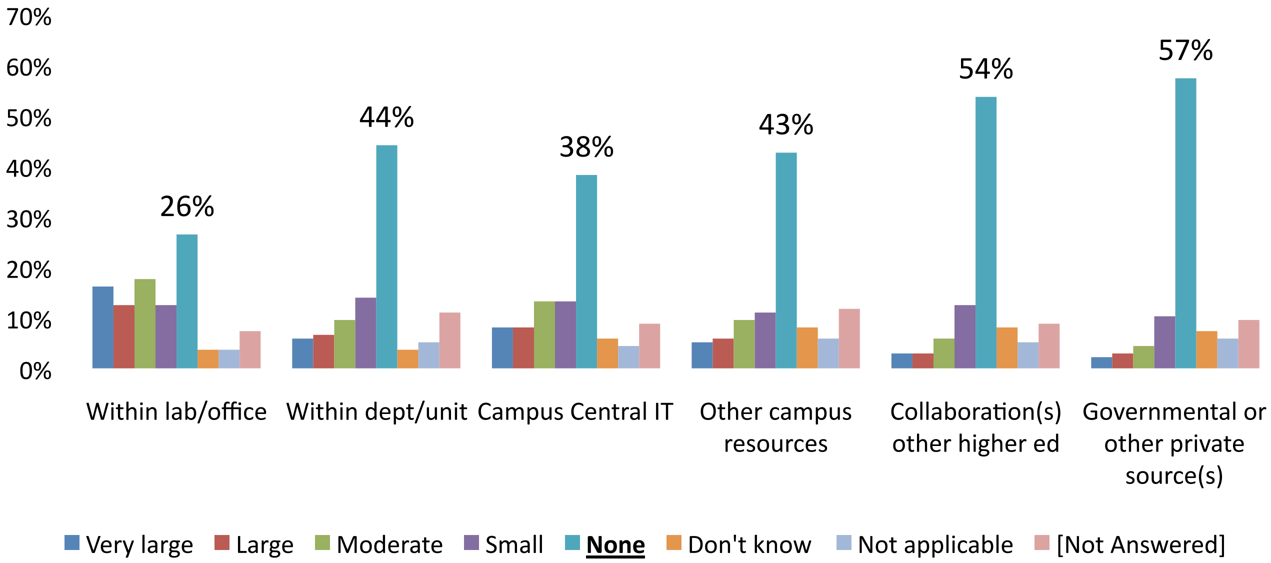
Quantitative Results

High Performance Computing

147 Responses

High Performance Computing

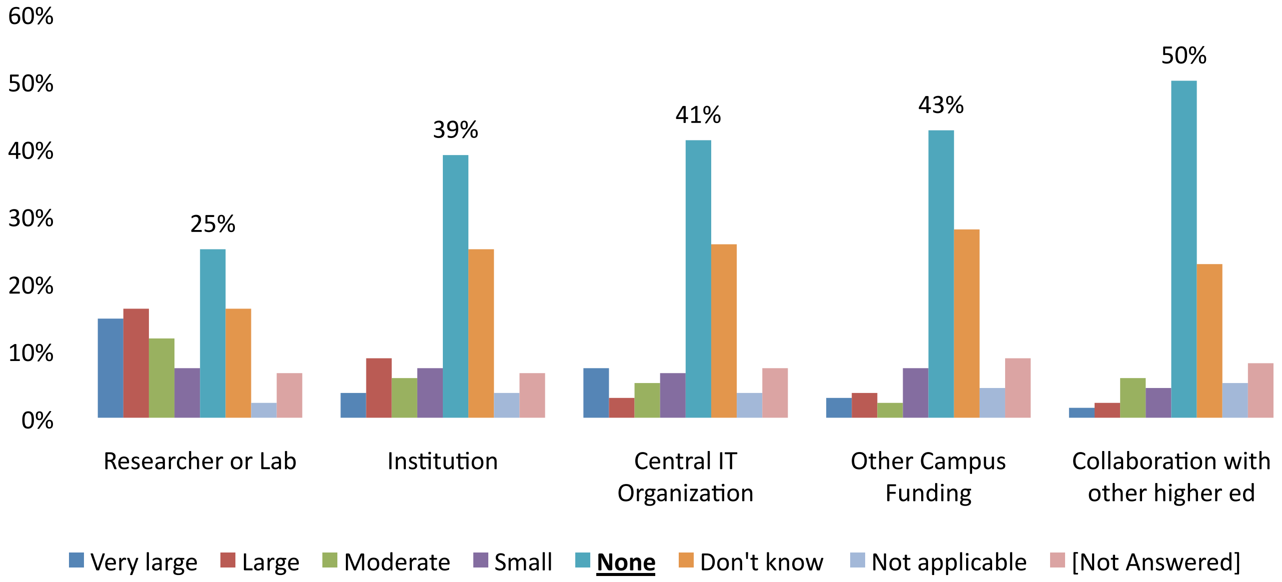
To what extent do you obtain access to high performance computing resources in each of these ways?



Many respondents who selected this area as part of their job do not obtain access to high performance computing.

High Performance Computing

To what extent are your high performance computing resources funded from each of these sources?



Many researchers receive no funds for supercomputing. Need to explore further these responses.

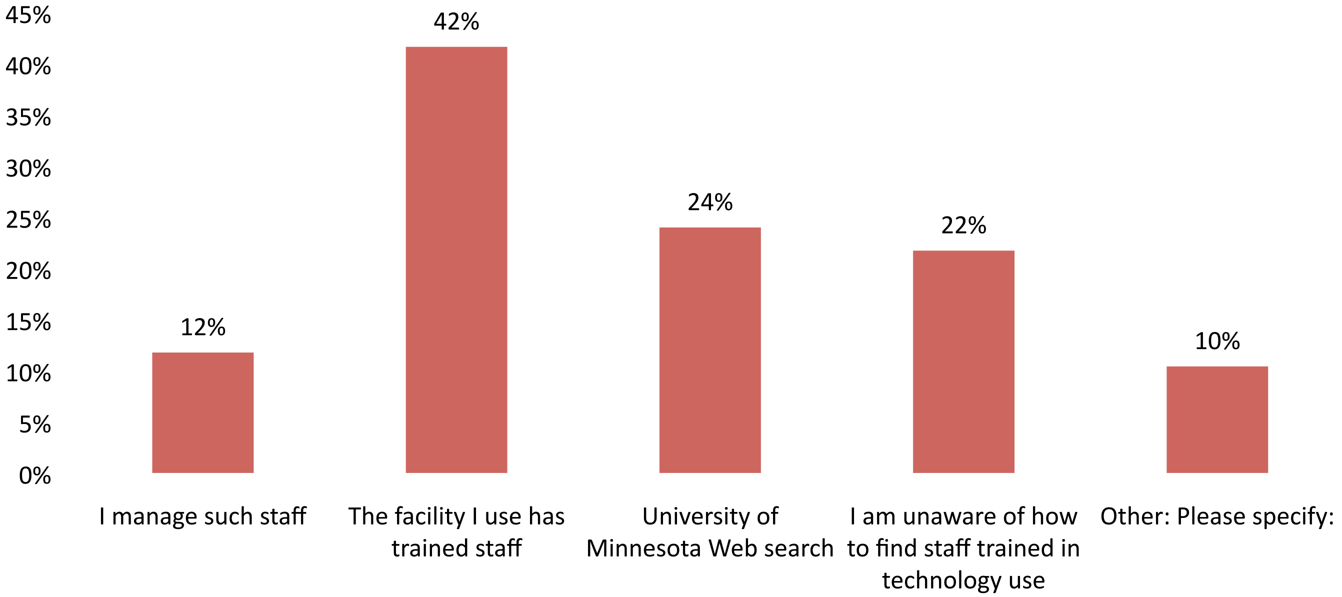
Quantitative Results

Learning and Workforce Development

237 Responses

Learning and Workforce Development

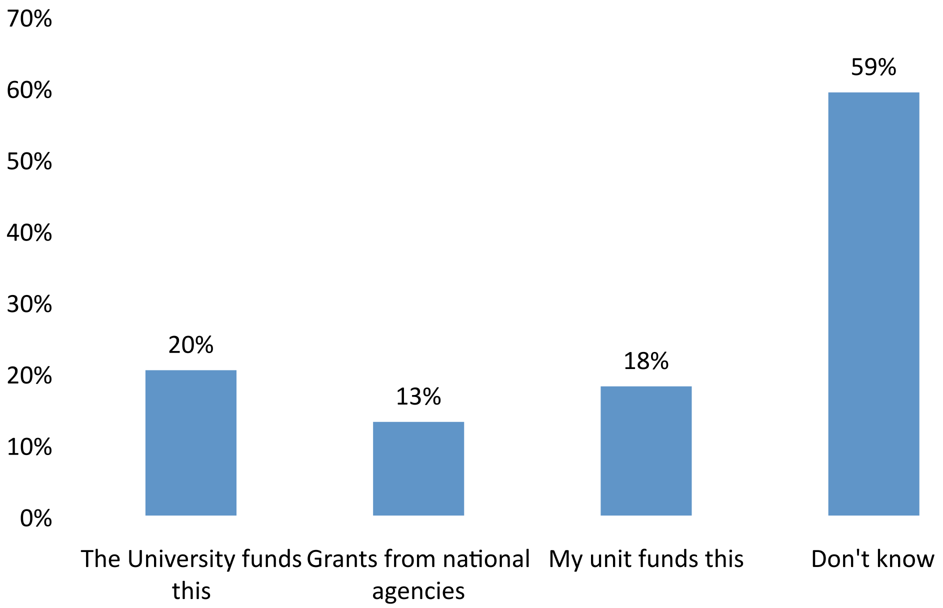
How do you find staff trained in technology use? Check all that apply.



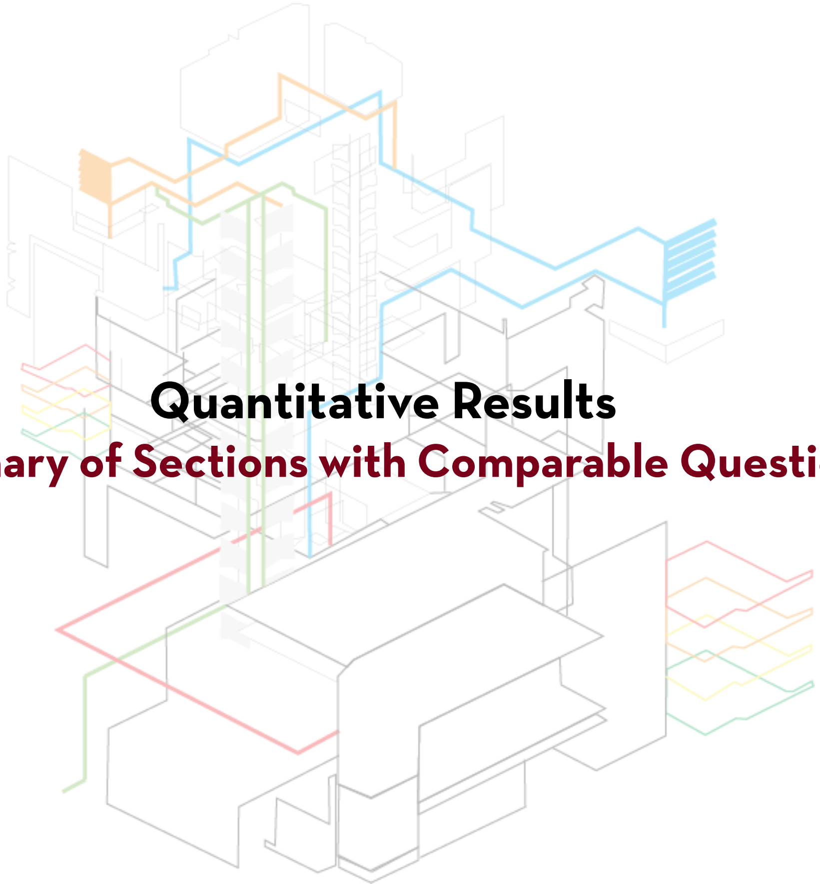
Nearly a quarter of respondents are unaware of how to locate technology staff on campus.

Learning and Workforce Development

How are workforce development and technology learning funded?



Researchers are unaware of staff technology training funds.

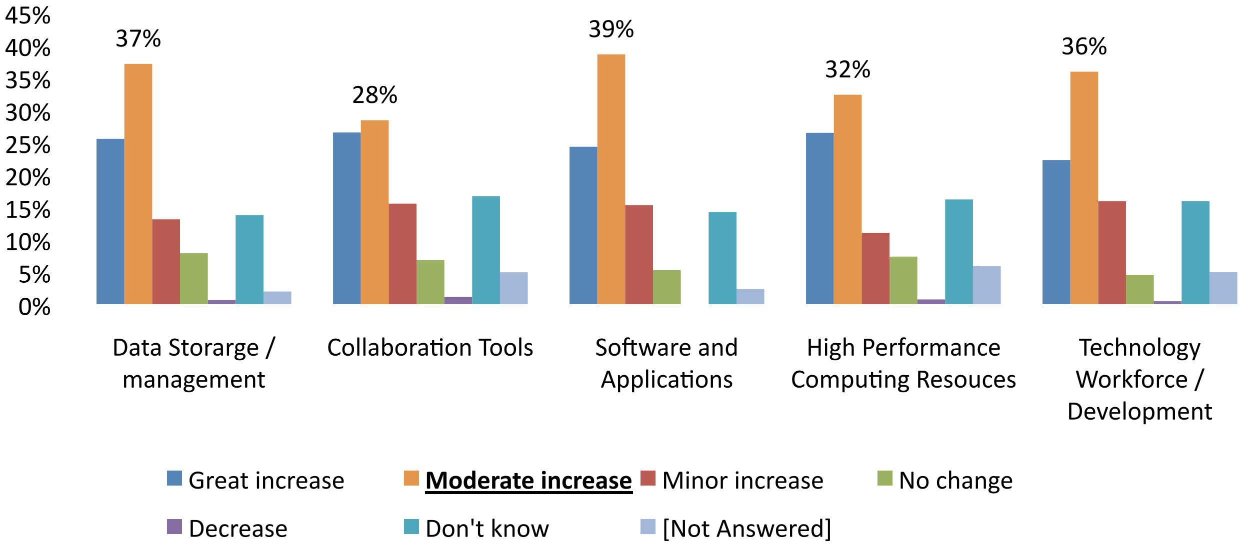


Quantitative Results

Summary of Sections with Comparable Questions

All Sections

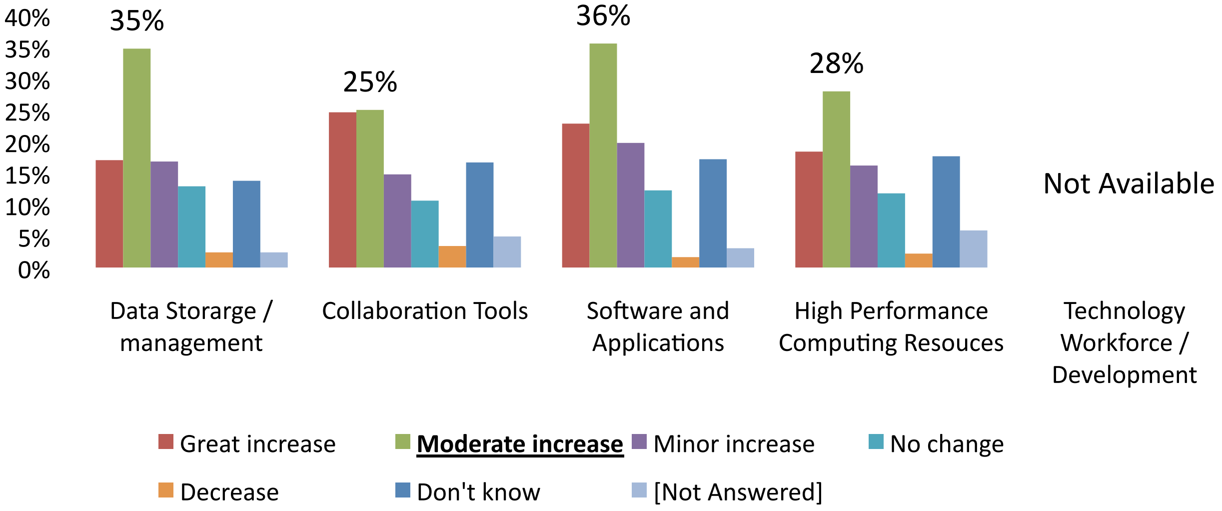
For your department/unit's research activities, how do you think the overall importance of the following areas will change in the next three years?



Over half of researchers foresee a great to moderate increase of each area's importance in the next three years.

All Sections

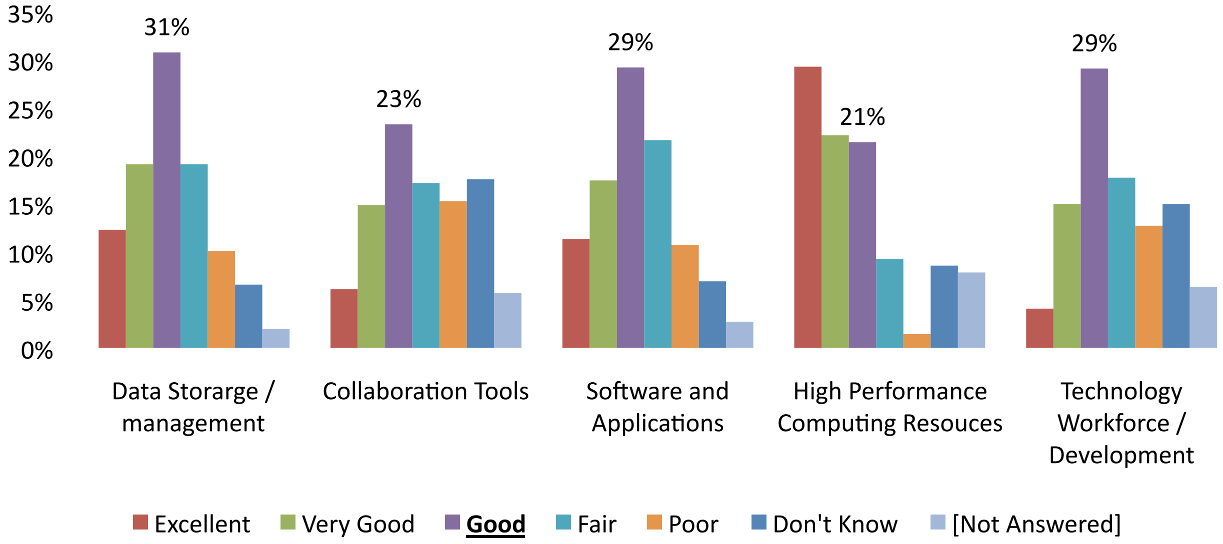
For your department/unit's research activities, how do you think the overall importance of *support* for the following areas will change in the next three years?



The importance of support in each area is expected to moderately to greatly increase by half of researchers.

All Sections

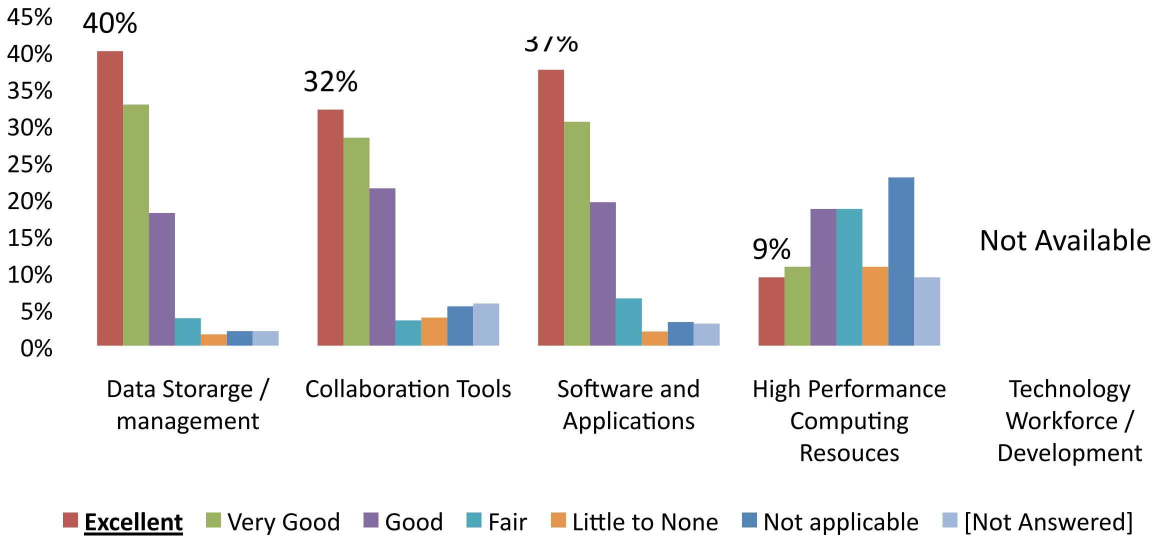
Rate the overall level of support you receive from your primary technology point of contact in the following areas.



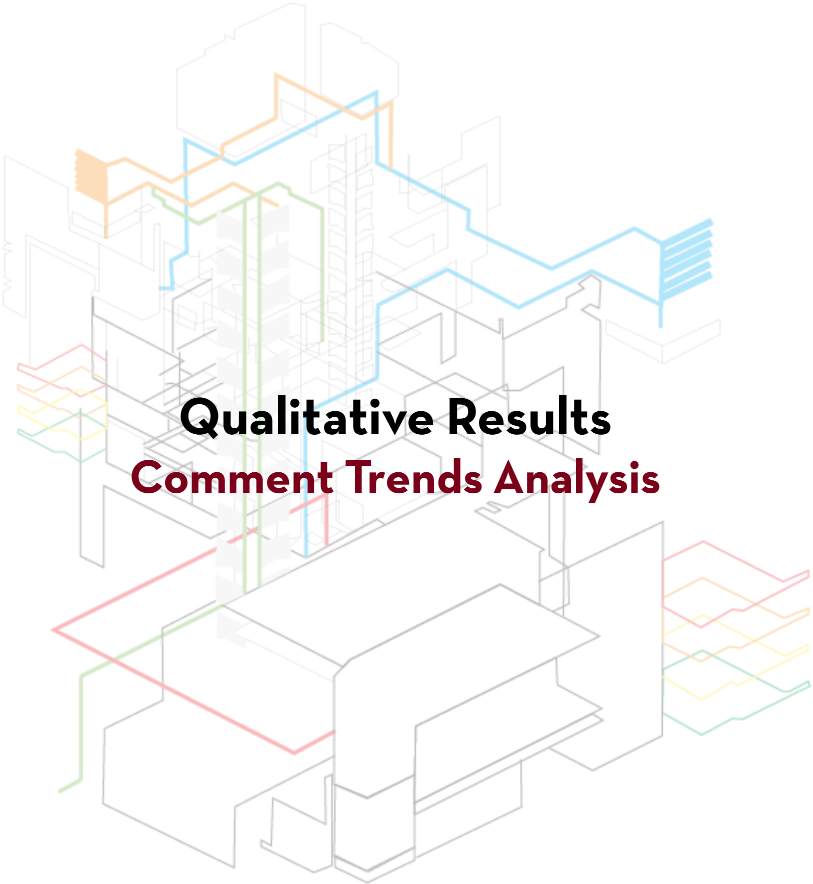
High Performance Computing was the only area where a majority of researchers felt their level of support excellent.

All Sections

What level of knowledge should your technology point of contact have regarding in the following areas.



A technology point of contact is expected to have excellent or very good knowledge for a majority of researchers.



Qualitative Results

Comment Trends Analysis

PEL RCA Charge Areas of Interest: Comment Trends

What are the University of Minnesota's…

- Infrastructure Needs
- Opportunities to Align Existing Resources
- Central vs. Local Support Needs
- Core/Specialized Resources Requirements
- Broad and Strategic Essentials: Systems/Service/Capacity
- Needs to Enable intra/inter-University collaboration
- Strategic/existing capacity ("Low hanging fruit")
- Budget Models for Tech Support
- Policy for sustaining Campus involvement

Comment Trend: Infrastructure Needs

A lack of backup at the department level was mentioned in a variety of ways.

- *When I started working at this lab (2.5 years ago) there was no tech support or data backup system in place. I asked my department administrators (at ALRT) how I can backup my lab data and they did not know. From searching the U website I found out about the existence of AHC-IS and was finally able to get funding approved to pay for an account with AHC-IS…*

- *The big problem we have is the difficulty to back up files. Some Universities have automatic backup. In our department there is no automatic back-up. It has to be done manually and the places we have to save data are too small to hold all of the files that need to be backed up.*

- *There is little to know support or services for IT, including/especially data backup within Medical School/ Dept. of Medicine. There is no automated backup of office or lab computers to central University server provided, as I have had at other Big 10 universities.*

Comment Trend: Alignment of Existing Resources

Support. Many departments are doing a good job of meeting researcher needs, some are not. How to model good behavior?

- *The DMC does the best job of providing incentives and excellent people for those of us interested in creative collaborative research projects. And the DMC is fabulous. But I get no credit for the 'extra' work I do with the DMC.*

- *It's hard to know who to contact in the tech area for some research tech support. The basic tech person assigned to the dept. is fine for routine computer and computer related software things, but beyond the basics it's not clear who to contact for what.*

- *Research computing at CLA-OIT, headed by Kemal Badur, has been very helpful with both data storage as well as access to servers with licensed statistical packages (SPSS, SAS, etc.). This help has been wonderful. The 3D modeling and animation software is not really in the purview of the CLA-OIT staff and so the problems encountered in my lab with these software have required us to seek help from outside of the University.*

Comment Trend: Central vs. Department level support

Researchers want control over their data at the unit level

- *In my research group, research data is mainly managed by the respective owner. we evaluated the file sharing service from central IT but found it too cumbersome to use. it works best if research groups can set up their own servers to have control over their data storage platform.*

- *The department satisfies all my needs at a cost lower than what is currently charged by similar central services. Having local services also increases greatly the responsiveness and flexibility. I have much more control on the services I get from the Department than what I would get from central.*

Comment Trend: Central vs. Department level support

There is a clear role for centralized support of technology workforce development and training.

- *Please look at how CLA has incorporated* [Tech learning/development] *into their faculty and departmental needs. They work with many different departments , colleges and schools to bring a 'most used' variety to their research facilities. A centralized research need is evident, but not how Central services currently handles its disparate groups. A single point of contact with automatic needs for software, disk space , hardware and network needs are handled as a system rather than separate departments that don't or cant work together.*

- *It seems that there are training opportunities out there, but no centralized way of viewing University-wide programs. Meaning, I might have to go to my unit to see what's offered there, then to OIT to look there, then to other units. It would be nice to have one database that pulls in all training opportunities to one website.*

Comment Trend: Core/Specialized Resources Requirements

Large data storage, privacy, and specialized software are common needs among researchers.

•Once or twice a year I have to negotiate for storage space; I always feel as if my research is at odds with IT policy. I always feel as if I'm living on borrowed time, no confidence in having access to adequate data storage for research in the future.

•We have remote access file servers and a wiki deployed for internal group use, but virtually no support from our department admin/IT staff in implementing systems which would be accessible to external collaborators.

•Challenges with the tight AHC administrative controls on research computers. This makes it difficult/ impossible to install software updates especially evenings and weekends when their help line is not manned. Controls of research computers should be handled differently from those in more public areas.

•University-wide software agreements are nice for general computing applications, but not for specialized software applications. Further, support for these specific applications in general-use labs, even those used specifically by faculty and students in my area, is very weak and not directed at the special needs (installation, maintenance, use) of these applications.

Comment Trend: Broad and Strategic Essentials for Systems/Service/Capacity

Service Liaison/Consultation

- *I don't think the Tech. Point of Contact should actually have to know how to use these tools [software/applications for research], but they should know about the tools that exist on the University.*

- *I have found it necessary to hire my own bioinformaticist to handle data storage and analysis*

- *I wish I had a consultant re SPSS*

Comment Trend: Enabling University Collaboration

Video Conferencing

•*Research communication compromised by the security/firewalls required by academic health system storing private health information. We need two systems, one for PHI, and one for research.*

•*What's really missing is serious support for high-quality teleconferencing, both within the University (different sites) and outside. It would be quite useful to have teleconference-ready conference rooms all over campus.*

•*I am currently working with a group at another institution on research projects based on the data we have here as well as other projects' data. They hold meetings at the other University for people to get together, it would be helpful if we had a way to hold 'virtual' meetings with cameras that go to each location so we could see the individuals we are meeting with rather than just having a phone conversation.*

•*There are really no good collaboration tools for conducting research other than assistance with communication. For wet labs, on-line electronic lab books would be helpful especially when working with other labs or at decentralized locations. It would be nice it equipment relying on visualization (i.e. microscopy) was adapted to delivery live images over the internet for remote viewing. Especially useful with distant collaborators.*

Comment Trend: Strategic/existing capacity ("Low hanging fruit")

EDUCATION

- *In general, most faculty in my area/department have very little knowledge of computing systems and storage opportunities and methods.*

- *In my experience, IT staff are woefully ignorant (and incurious) of resource needs and University/ College/Dept. resources made available are paltry.*

- *We need more transparency and organization in these matters. Is there a single informative website about this?*

- *Bringing people into the department to tell us about options and opportunities would be very helpful. Or offering tours and opportunities to meet with technology staff before help is needed would be great.*

- *If infrastructure exists for sharing data, the knowledge has not been imparted on me.*

Comment Trend: Strategic/existing capacity (“Low hanging fruit”)

PROMOTION

- *Our department and college offer exceptional support and opportunities to learn and apply a range of technology available. The support persons make these opportunities as available as possible. I have not yet taken the best advantage of what is offered, but I am working on it.*

- *BSCL/MSI has been invaluable in my obtaining research grant support*

Comment Trend: Budget Models for Tech Support

Balancing Local vs Central support in a fiscally responsible way

•*Central IT organization is extremely important in streamlining infrastructure and establishing standard protocols, however, it does not necessary always translate to cost savings. As central IT gets larger and complex, the cost for high end hardware equipment and technical staff to manage all the systems can become a heavy financial burden. There should always be room for research autonomy. Not every IT task needs to be high tech. Nothing wrong with backing up on a 1 Tb external hard drive myself that cost $300 verses on a 1 TB central IT file server that costs $30K and a half FTE that manages the system.*

•*There is a lot of high-quality open software available that the University seems to ignore in favor of proprietary and more expensive solutions.*

•*Funding for support and hardware has been and continues to be cut so severely that soon I will be carving my own notes on stone tablets.*

Comment Trend: Policy for sustaining Campus involvement

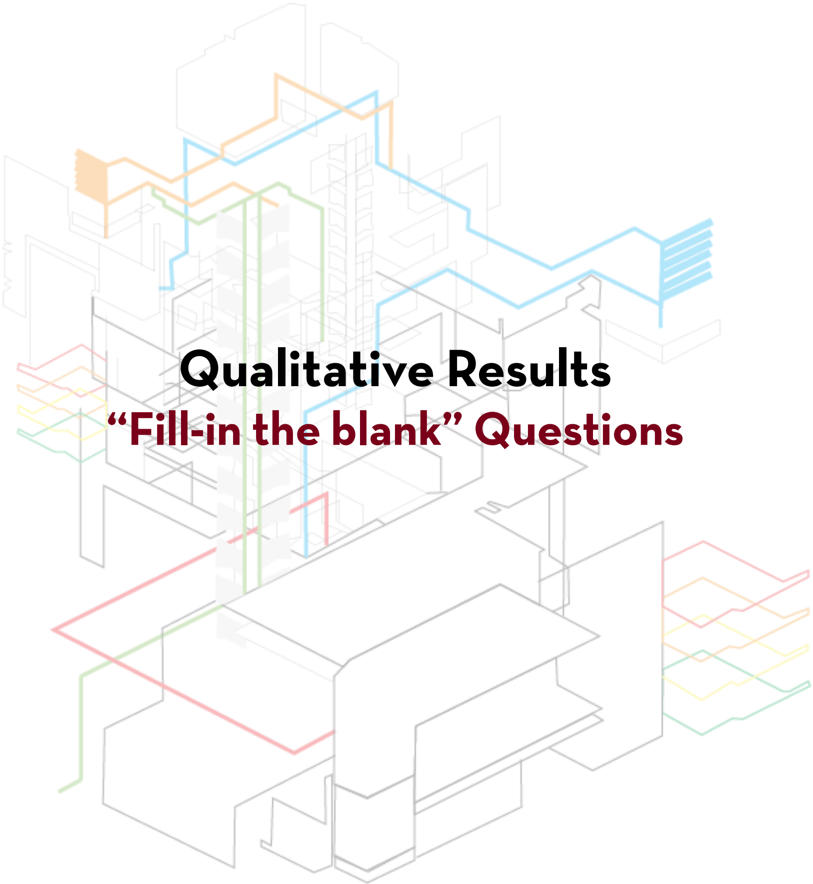
Mandate management, security issues, and supporting software licenses are concerns.

- *If the U mandates more from the departments without providing full support, research will become more difficult.*

- *International collaboration will increase; how do we encourage/assure compliance with security standards from our collaborators?*

- *We get very good support on the tech side (e.g., servers, security) but minimal on data tracking or compliance issues.*

- *For a small lab---it is very difficult to maintain expensive licenses---difficult to write this type of 'maintenance' into a grant---our funders just don't understand it.*



Qualitative Results

"Fill-in the blank" Questions

Fill-in: Technology needed to be competitive

303 Answers to this question

access, bandwidth, better, collaboration tools, computer, data, data management, data storage, database, digital media, GIS/spatial, grants, help, High-Performance Computing, informatics, library/biblio tools, Proteomics/spectroscopy, software, space, statistical support, support, text mining, training/awareness, upgrades, video, visualization, web development, Studios/performance,

Count
1-9
10-14
15-22
22-28
28-35
35-50
50-80

Fill-in: Software used in research

322 Answers to this question

ArcGIS , endnote, Image analysis and visualization, Internet, matlab ,

Microsoft Office,

Photoshop, R , SAS, SPSS,

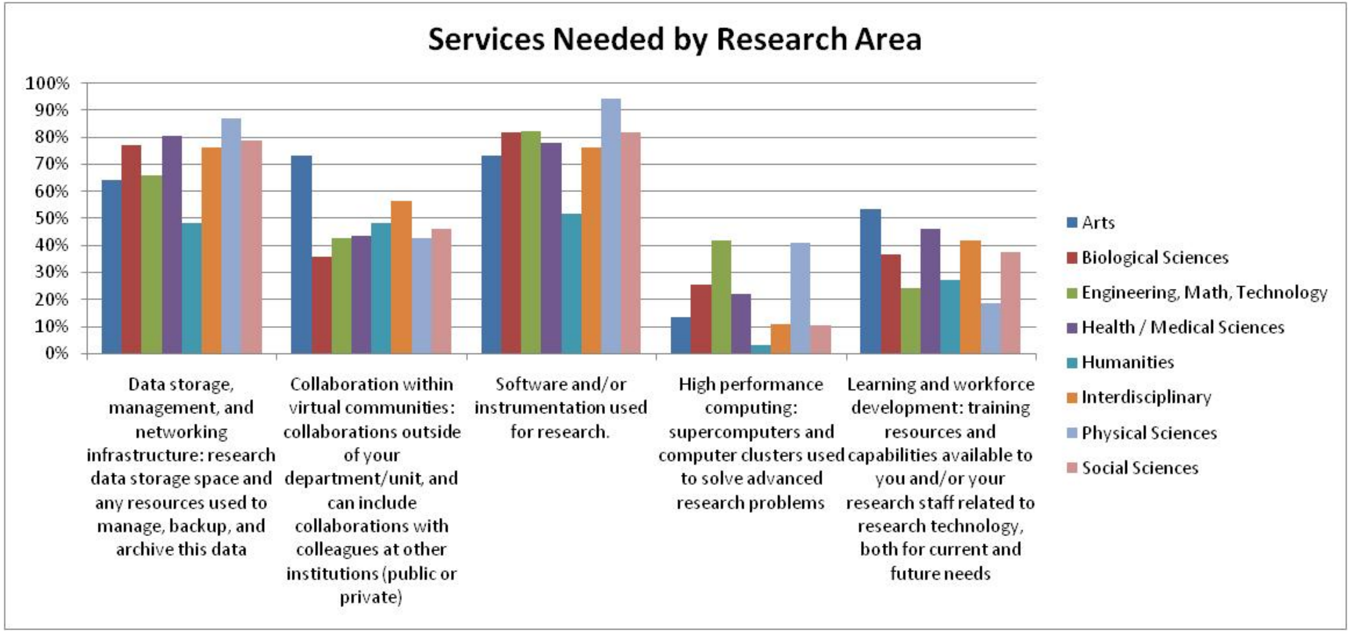
Count
1-9
10-14
15-22
22-28
28-35
35-50
50-80

Cross-Tabulation

Survey Response by Subject Discipline

Cross-Tabulation: Demographics

What support do you currently use or need in your research activities?

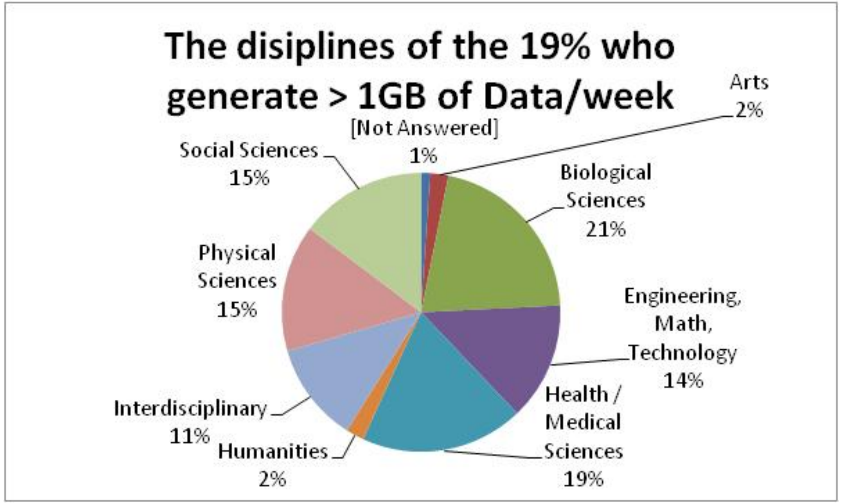


Services needed did not vary much by discipline.

Cross-Tabulation: Data Management and Storage

In a typical week do you generate more than one GB of research data?

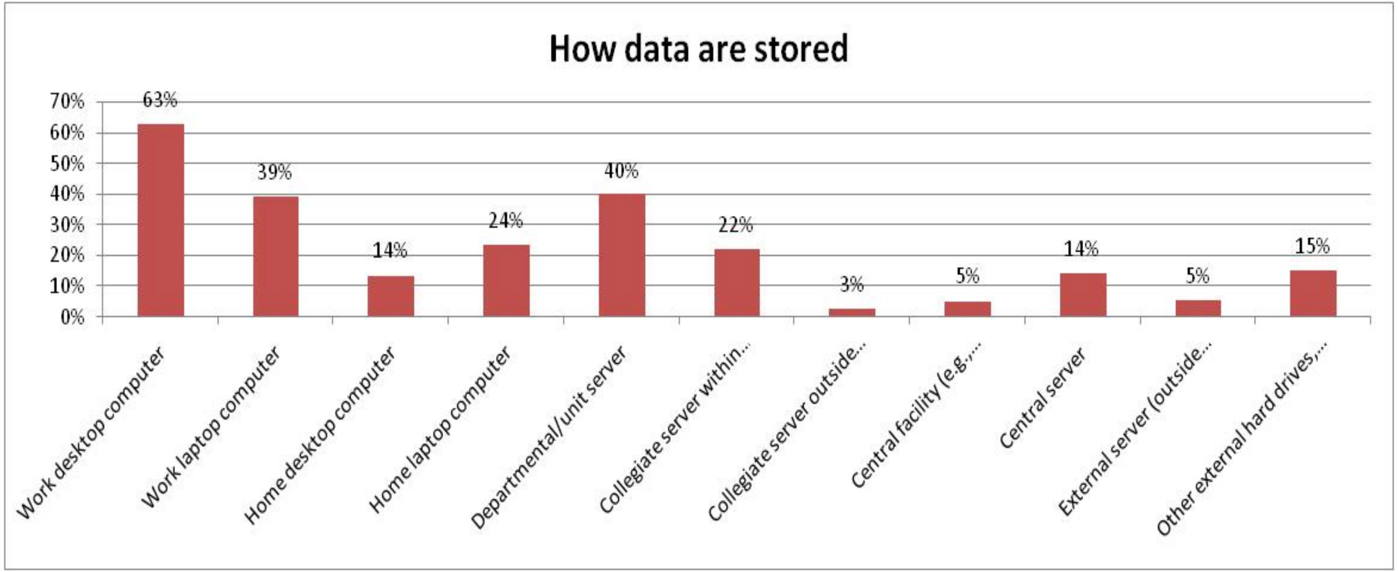
- **No** **80%**
- **Yes** **19%**



Researchers from all disciplines generate over 1 GB of data per week.

Cross-Tabulation: Data Management and Storage

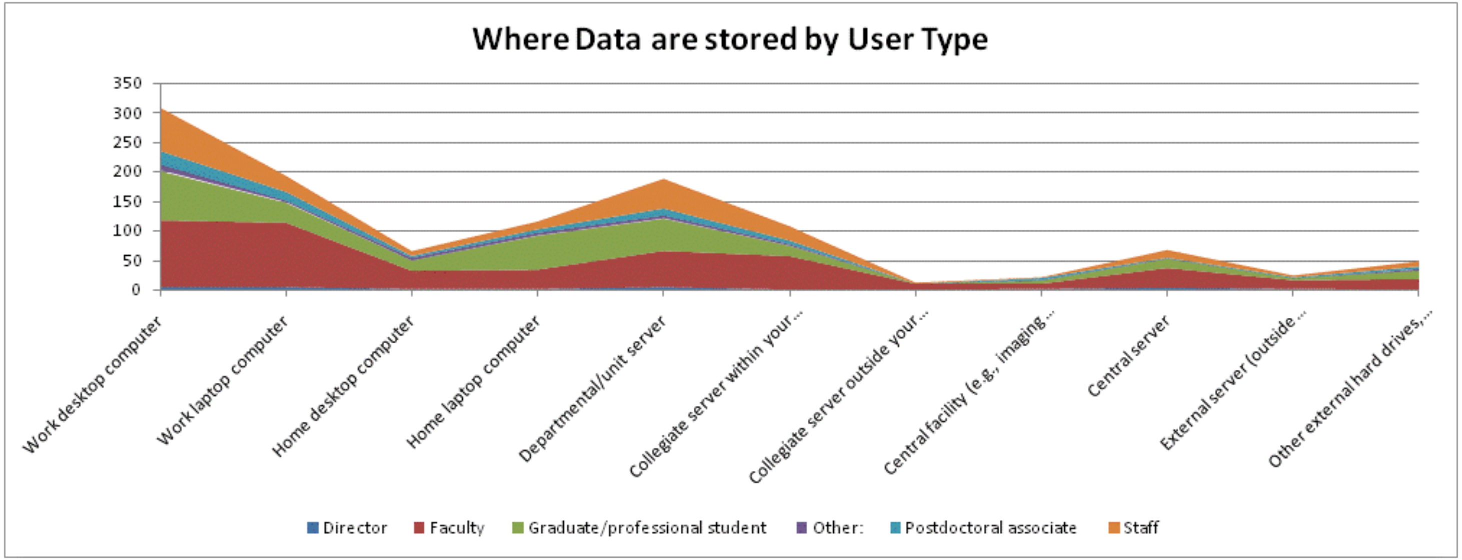
Where do you store your research data electronically?



Less than 14% of respondents use central data storage solutions (collegiate, OIT, or University-level servers).

Cross-Tabulation: Data Management and Storage

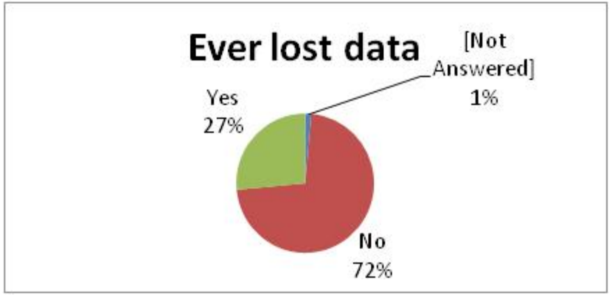
Where do you store your research data electronically?

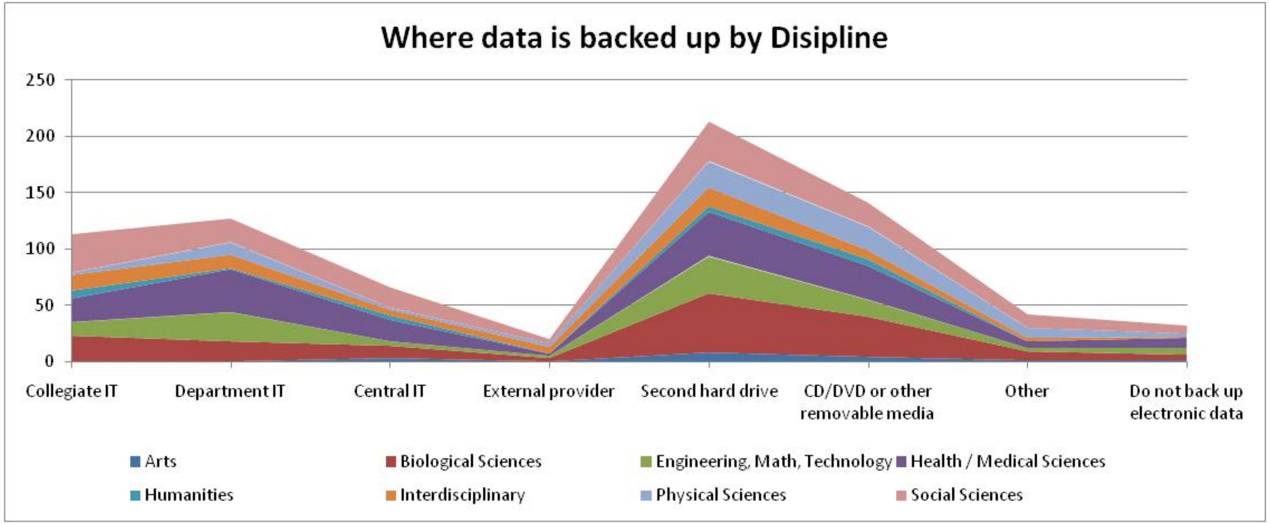


Grad students more likely to store data on personal computers. Faculty are more likely to store data on Central.

Cross-Tabulation: Data Management and Storage

Have you ever lost important data due to the lack of a back up?

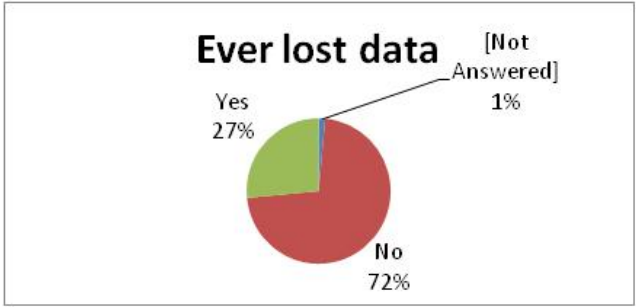


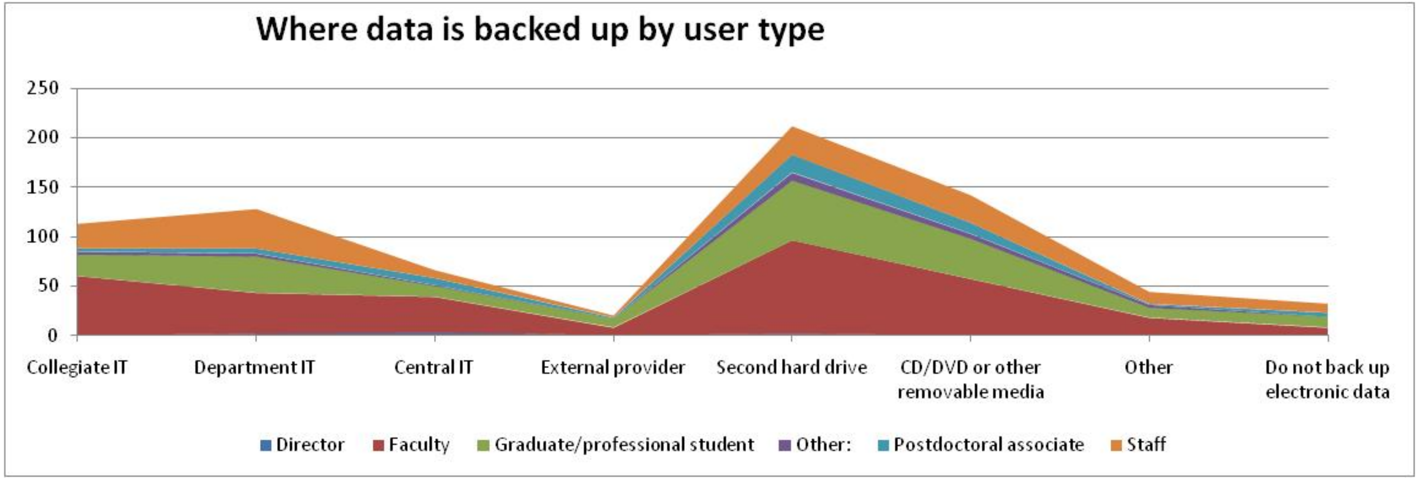


A second, or external, hard drive is a primary storage solution among respondents.

Cross-Tabulation: Data Management and Storage

Have you ever lost important data due to the lack of a back up?

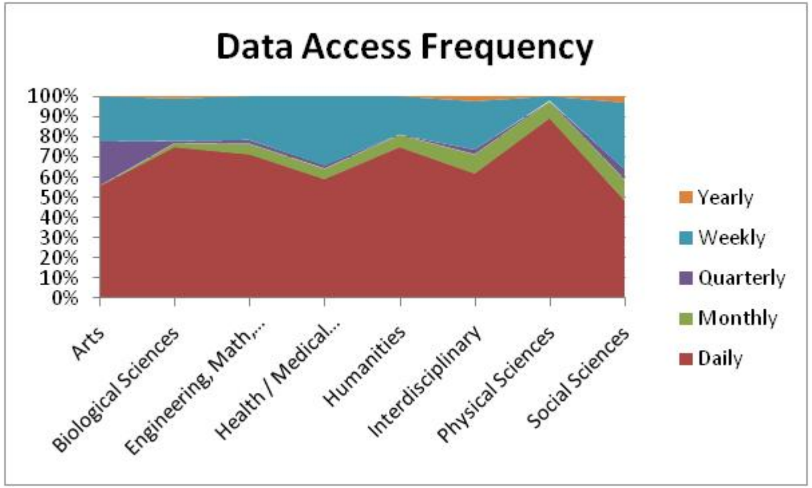


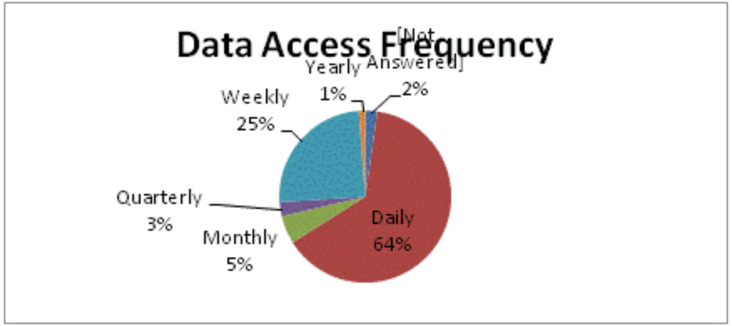


A second, or external, hard drive is a primary storage solution among respondents.

Cross-Tabulation: Data Management and Storage

How often do you access electronic data?

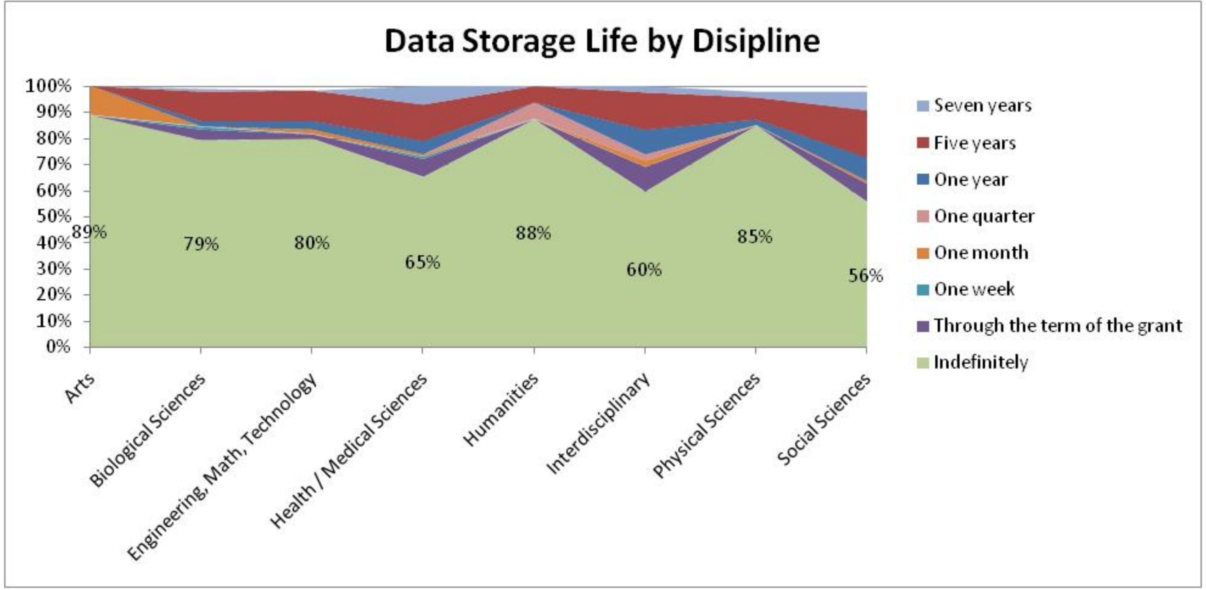


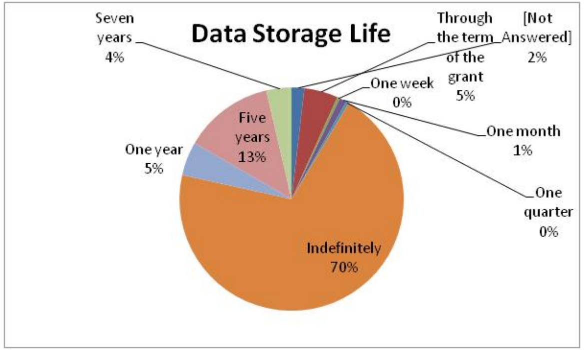


Humanities researchers access their data as frequently (75% daily) as Biological and Engineering Researchers

Cross-Tabulation: Data Management and Storage

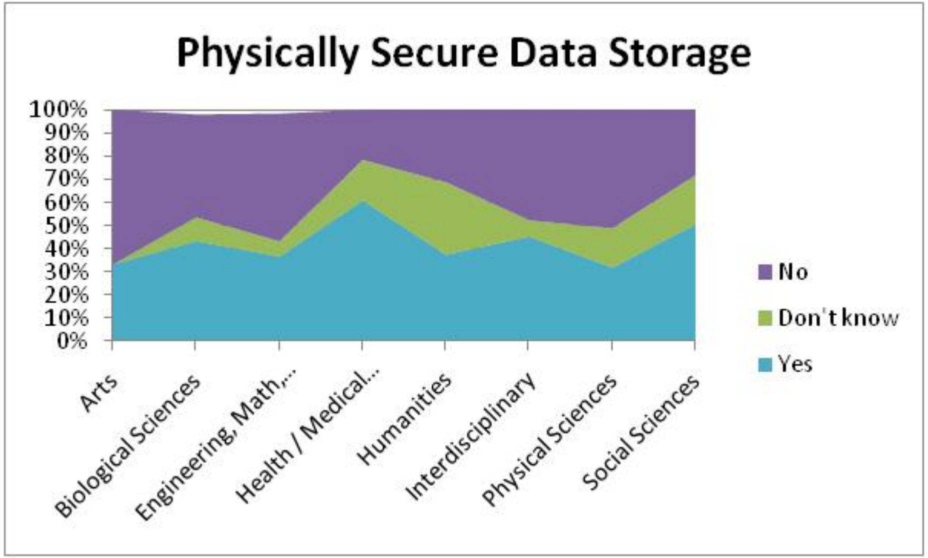
How long do you typically keep your electronic research data?

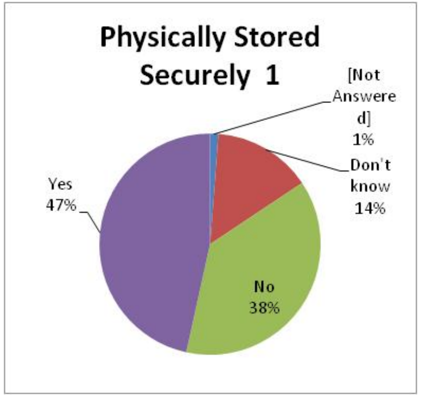




Cross-Tabulation: Data Management and Storage

Do your data need to be stored securely?

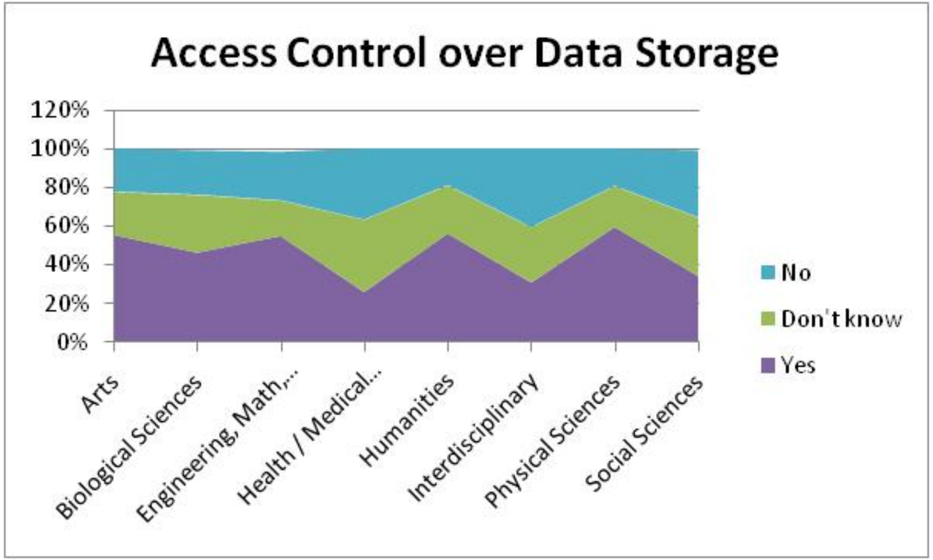


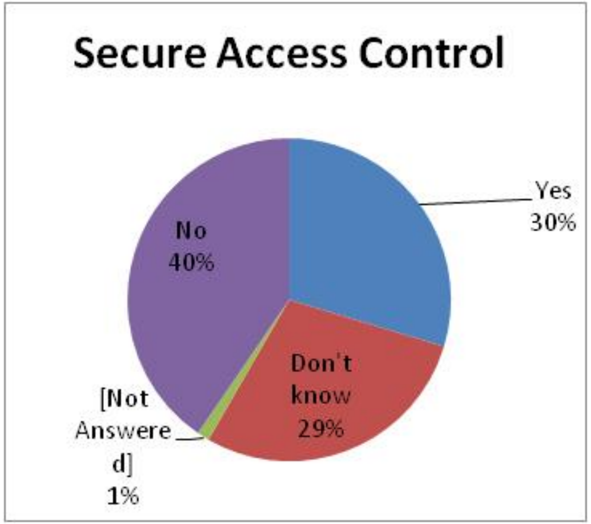


Many researchers do not require secure data storage.

Cross-Tabulation: Data Management and Storage

Do your data need to be stored securely?

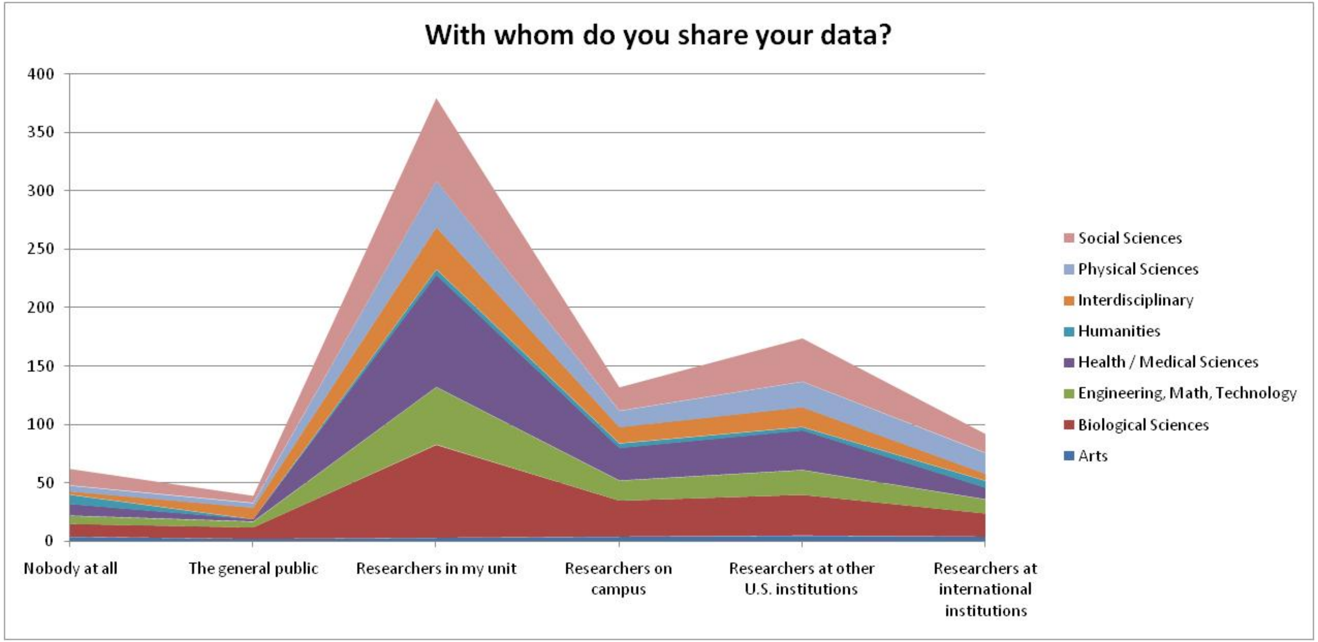




Many researchers do not require secure data storage.

Cross-Tabulation: Data Management and Storage

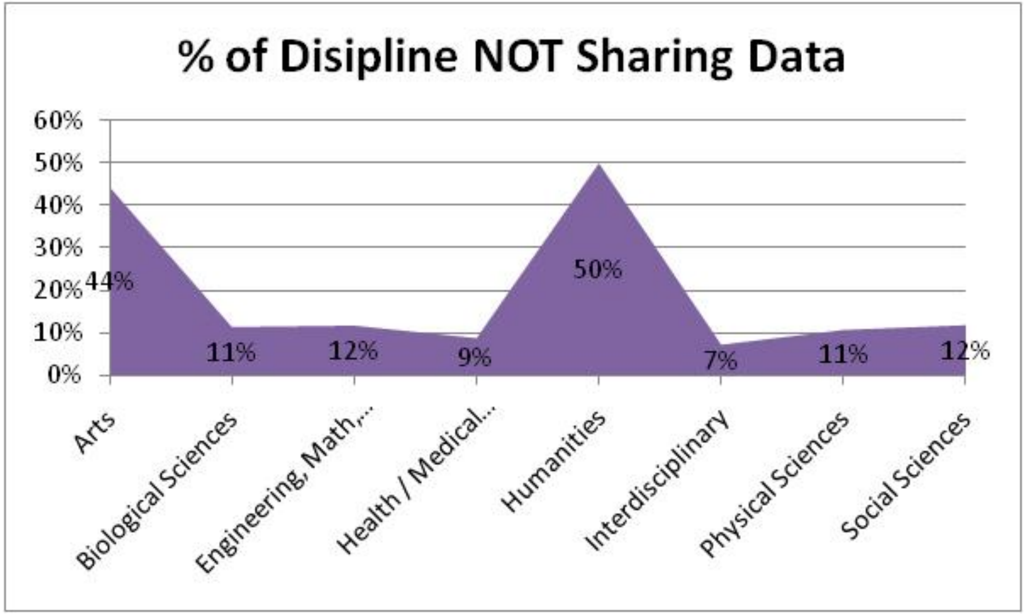
With whom do you share data?



Nearly 92% of researchers are sharing their data.

Cross-Tabulation: Data Management and Storage

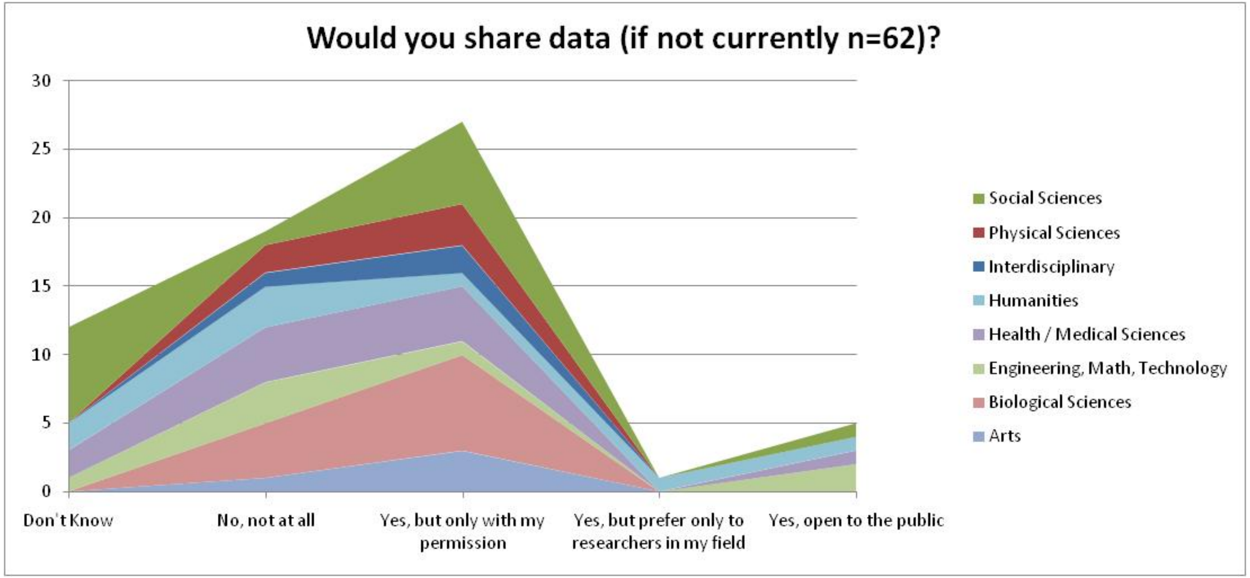
With Whom do you share data?



Only 8% of researchers are not currently sharing data.

Cross-Tabulation: Data Management and Storage

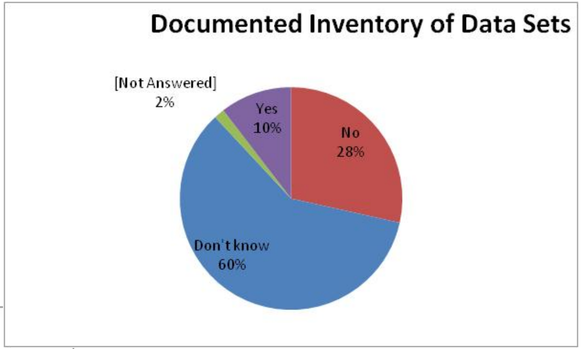
If you don’t share data, would you with the following caveats?

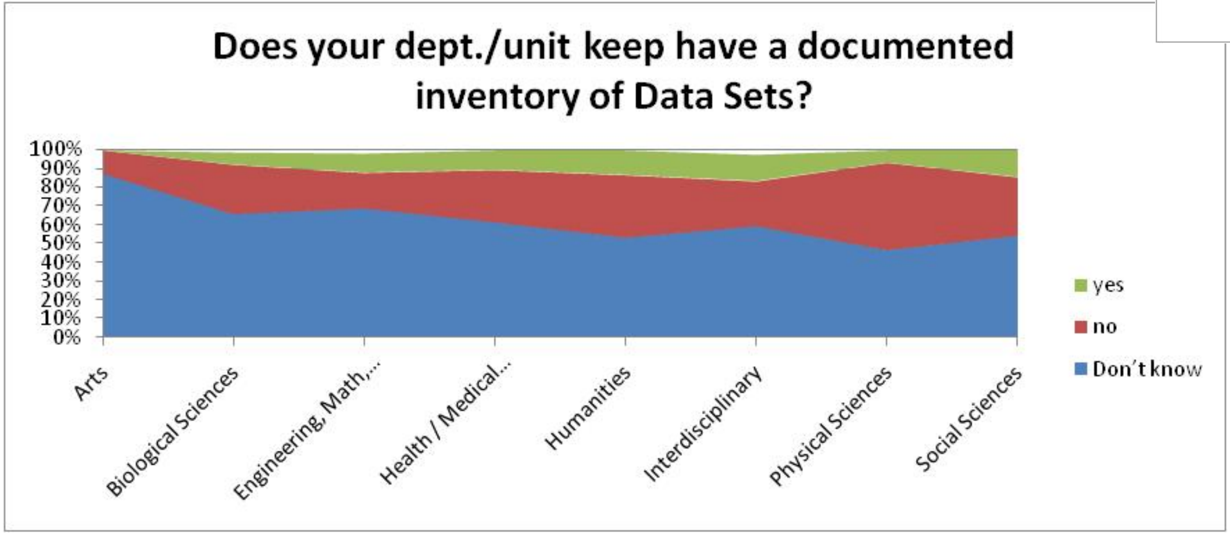


Of those 8% who don’t , 52% would share their data (with limitations), 30% would not and 19% did not know.

Cross-Tabulation: Data Management and Storage

Does your department/unit have a documented inventory of electronic data sets?

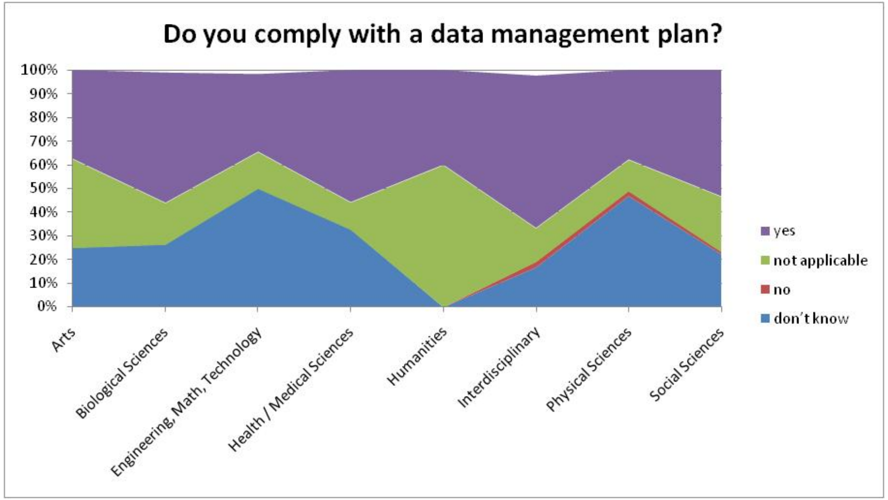


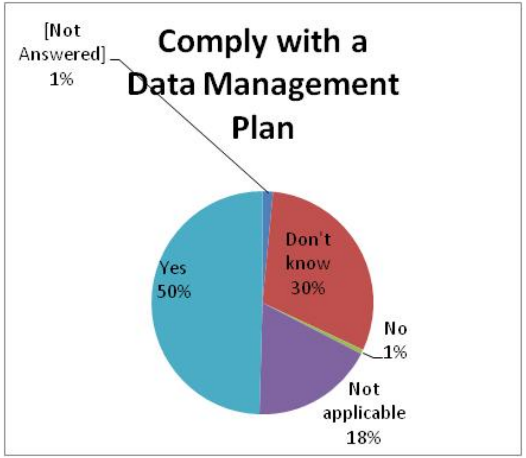


Half of respondents are adhering to funding agency requests to successfully store and manage data.

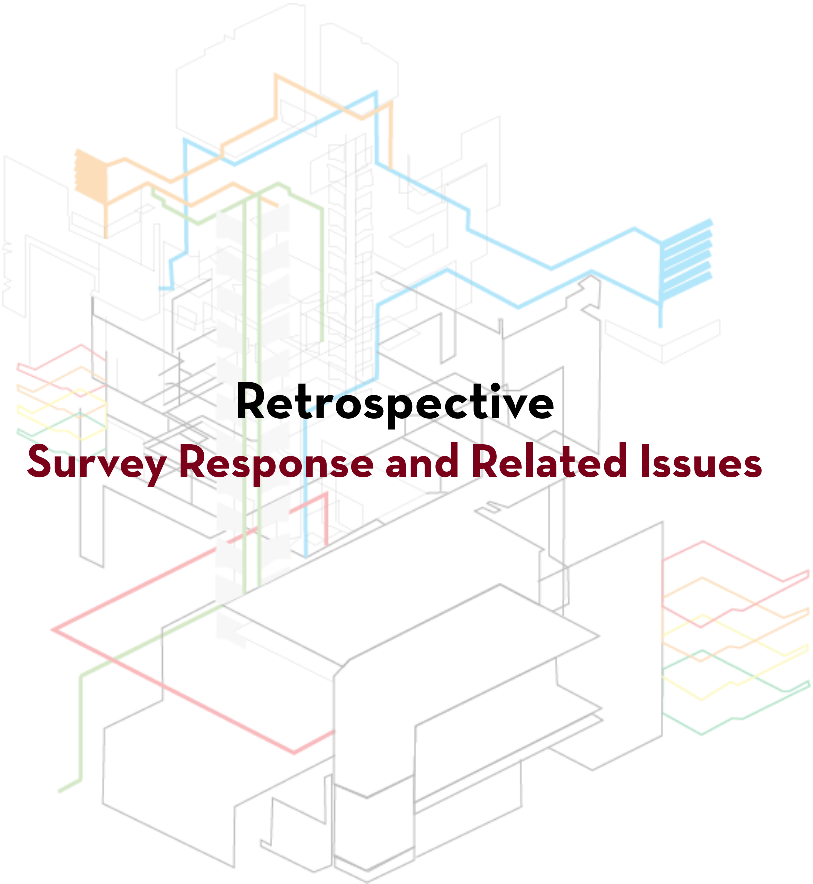
Cross-Tabulation: Data Management and Storage

Do you comply with your primary funding agency's suggested data management plan?





Half of respondents are adhering to funding agency requests to successfully store and manage data.

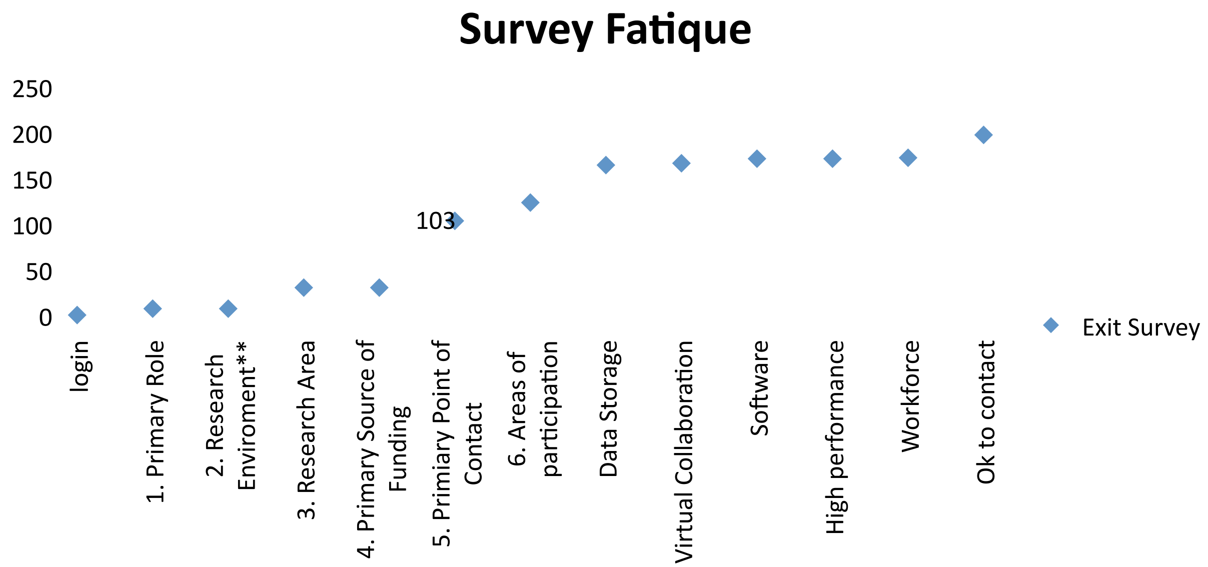


Retrospective

Survey Response and Related Issues

Retrospective: Survey Fatigue

Of the 780 responses, 197 participants exited prematurely.
583 successful responses.



**Note: 14 users were exited from the survey due to lack of research environment (Q. 2)

Retrospective: Survey Terminology

Of the 302 comments, 29 comments were complaints on the survey itself. 4 comments spoke positively.

- *This is a very confusing questionnaire; I am very tempted to close the window and give up on it.*
- *I wanted to let the author of this survey know that the anchors are confusing and that the items are also confusing. This is not straightforward, and I am concerned that your survey is biased.*
- *This is not a helpful survey.*
- *Too long and too 'who cares.' I have work to do.*
- *I found this survey difficult, because too often you were looking for one answer when the right answer is 'it depends on the specific project, data set, etc.'*
- *This survey was very poorly worded, difficult to understand, and difficult to follow.*

Retrospective: Humanities Biases

Humanities accounted for 5% of our responses.

- *The assumption here seems to be that there is grant support for all research on campus, but most Humanities scholars do the majority of their research without such support.*

- *This instrument seems to have been constructed using the sciences as its template. I find it hard to translate the terms into my research in the Humanities.*

- *Your language is impenetrable. I couldn't answer most of the questions because I didn't know what they mean. How about revising the survey for humanities faculty & ask what programs they have and use on their PCs; what programs, instruction, and services they would like; and what kinds of research they use them for? Some of us speak English, not techno-babble!*

- *Technology use may be more relevant to teaching for humanities scholars.*

- *This survey is largely geared to researchers outside the Humanities. Consequently, many if not most of the questions are stretch -- in terms of relevance. Please find a way to incorporate a genuine commitment to research in the humanities (without which no University with which we aspire to compete, either public or private, is ranked highly).*

Final Thought

Next Steps

The 2008-2009 PEL Project Survey Recap:

- What we learned overall…
- Over 50 respondents agreed to meet with us to discuss these questions in further detail.

Next Steps of the PEL Survey:

- Identify trends and requirements for computational infrastructure.
- Analyze the qualitative data; each section received individuals comments and suggestions. Further explore the funding-related questions that have high “I don’t know” responses.
- Formulate Recommendations and present data findings to several campus groups.
  - The 2008-9 PEL Poster Session took place on June 16th 2009 where we presented our final recommendations.
- Submit a report on our findings that compares these data to information from other recent University surveys.

And ….

Thank You!

Implementing Cyberinfrastructure for 21st Century Research

2008-2009 President's Emerging Leaders Program

Tracy Anderson, College of Biological Sciences
Craig Gjerdingen, Carlson School of Management
Bryan Herrmann, Office of Admissions, Morris Campus
Katherine Himes, Office of the Senior Vice President for Academic Affairs and Provost
Lisa Johnston, University Libraries

Driven to Discover℠

Project Charge

To develop recommendations for a strategic approach to Research Cyberinfrastructure that addresses the growing need on campus and leverages partnerships across the University. Currently, research technology and emerging global partnerships are generating data with computational and storage needs that outpace the information infrastructure available.

What is Research Cyberinfrastructure?

Research Cyberinfrastructure includes the instruments, sensors, high performance computational systems, massive storage systems, data resources, and visualization facilities, tied together by high speed networks and made to work together by advanced software to accomplish goals that would not be possible by any single information technology system. It also includes the people, processes, training, security, policies, and capabilities to sustain the systems and networks over time. Implementing Research Cyberinfrastructure requires a high level of coordination and collaboration between researchers and an information technology workforce with expertise in scientific computing.

Methodology

We conducted an online user-needs survey to 8,424 University faculty, research staff, and students asking them to report the current state of Research Cyberinfrastructure support at the University as well as to assess their future needs. The survey was comprised of 130 questions on the following Research Cyberinfrastructure trends: data storage, data management, and networking infrastructure; collaboration with other researchers; tools and applications; high performance computing; and learning and workforce development, as well as future trends within each of these areas. We had 780 successful responses (9.2% response rate). These data formed the basis for our trends and recommendations. In addition, we performed a meta-analysis of recent University reports and surveys on this topic to enhance our trends and recommendations.

Recommendations

1. Re-brand Research Cyberinfrastructure at the University to improve accessibility for all disciplines.
2. Develop enterprise-wide, integrated cyberinfrastructure to align with national cyberinfrastructure.
3. Improve and expand the University's ability to handle data.
4. Better position services around the end-user.
5. Create incentives for virtual collaborative relationships.

What are the benefits?

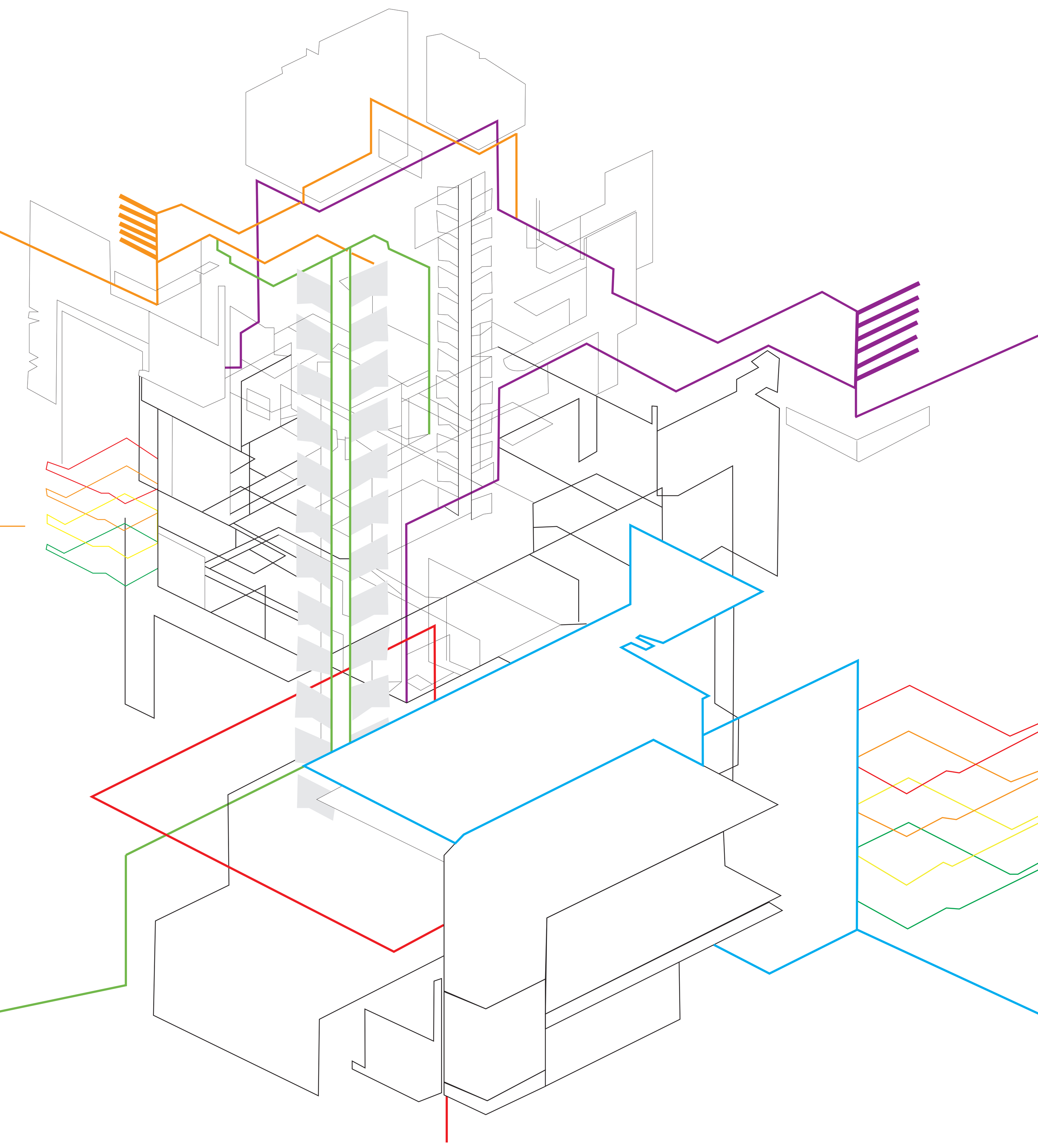
Continued focus on Research Cyberinfrastructure will help the University become a top three public research university through accelerating research-enabled discovery; increasing partnerships, research collaborations, and communication within the University; increasing the University's competitiveness for funding from federal agencies; and, finally, sparking new types of research and Research Cyberinfrastructure collaborations.

Implementing Cyberinfrastructure for 21st Century Research

2008-2009 President's Emerging Leaders Program

Tracy Anderson, College of Biological Sciences
Craig Gjerdingen, Carlson School of Management
Bryan Herrmann, Office of Admissions, Morris Campus
Katherine Himes, Office of the Senior Vice President for Academic Affairs and Provost
Lisa Johnston, University Libraries

Driven to Discover℠



"I use a USB drive purchased with personal funds to backup data."

"I use a USB drive purchased with personal funds to backup data. Note that this is data about mice and is not related to any personal data that may need to be encrypted. We have an imaging system for viewing gels that is connected to a computer that can only store data on the hard drive and floppy disk. A floppy disk can hold only four images. Other storage or a network connection would be nice." -Scientist in the Medical School

"There needs to be a mechanism for centralized storage of data with high security for individual research groups." -PhD Research Assistant in the Institute of Technology

"In my research group, research data is mainly managed by the respective owner. We evaluated the file sharing service from central IT but found it too cumbersome to use. It works best if research groups can set up their own servers to have control over their data storage platform." -Post Doc in the Institute of Technology

1

Develop enterprise-wise integrated cyberinfrastructure to align with national efforts

Accountability at the senior leader level is needed to ensure the University's competitiveness in the future (i.e., obtain grants, become a top three public research university, recruit, evolve research and programs). At the same time, the University has the opportunity to leverage existing 'cottage industries' and to align internal services.

"...it's not clear who to contact for what."

"The basic tech person assigned to the deptartment is fine for routine computer and computer related software things, but beyond the basics it's not clear who to contact for what." -Faculty Member in the College of Liberal Arts

"Attention needs to be given to the organization of the multiple electronic records we store for multiple researchers at the University. There needs to be a management plan that we can follow and somebody needs to be doing the inventory of these records." -Senior Research Fellow at the Academic Health Center.

2

Better position services around the end-user

Provide IT contacts with the tools to share/leverage existing resources across University service providers. Researchers do not care who offers the service; they just want the opportunity to take advantage of the offering.

"...I have storage requirements on the order of terabytes."

"5 GB space is paltry for backing up systems on active directory. I have storage requirements on the order of terabytes." -Faculty Member in the College of Food, Agriculture & Natural Resource Sciences

"Once or twice a year I have to negotiate for storage space; I always feel as if my research is at odds with IT policy. I always feel as if I'm living on borrowed time, no confidence in having access to adequate data storage for research in the future." -Faculty Member in the College of Education and Human Development

3

Improve and expand the University's ability to handle data

In order to meet the growing needs of researchers, IT service providers should provide short-term data storage as needed (i.e. a bank model), utilize external cloud storage for long-term archiving, provide high speed data connectivity, and automatic back-up, with the appropriate compliance and privacy mechanisms.

"I find it hard to translate the terms..."

"This instrument seems to have been constructed using the sciences as its template. I find it hard to translate the terms into my research in the Humanities.." -Faculty member in the College of Liberal Arts

"This survey needs to take into account that some of us are not at all familar with technology or the language used to describe it." -Faculty member in Classical and Near Eastern Studies

"This survey is largely geared to researchers outside the Humanities. Consequently, many if not most of the questions are stretch -- in terms of relevance. Please find a way to incorporate a genuine commitment to research in the humanities (without which no university with which we aspire to compete, either public or private, is ranked highly)." -Faculty member in Cultural Studies and Comparative Literature

4

Re-brand cyberinfastructure at the University to improve accessibility for all disciplines

In order to bridge the jargon gap among the research community, the service providers and University leaders should re-brand cyberinfrastructure to ensure efficent workforce training and awareness of technology needs.

"International collaboration will increase..."

"International collaboration will increase; how do we encourage/assure compliance with security standards from our collaborators?" -Faculty Member in the College of Education and Human Development

"As far as I know, there is no inventory for data storage and we don't have a contact person for data storage. We just save our data in multiple places and on an external hard drive. There is no oversight for data storage." - Post-Doc in the School of Dentistry

5

Create incentives for virtual collaborative relationships

Give researchers the necessary tools to help develop collaborative relationships and facilitate local and external data sharing in order to establish the University as a top research institution.

What is Research Cyberinfrastructure?

Research Cyberinfrastructure (CI) is a key ingredient in fostering interdisciplinary research, garnering national funding, and transforming the University of Minnesota into a top three public research university. Increasingly, research technology and emerging global partnerships are generating data that have computational and storage needs that outpace the information infrastructure currently available. Our project addressed the growing needs for CI at the University and determined strategic ways that the University's Research Cyberinfrastruture Alliance (RCA) and other University partners might respond.

Methodology

During March 24st-April 8th, 2009, the PEL project on Research Cyberinfrastructure at the University of Minnesota emailed our online user-needs survey to 8,424 faculty, research staff, and students asking them to report the current state of cyberinfrastructure support at the University as well as to assess their future needs.

The survey was comprised of 130 questions on the following cyberinfrastructure trends: data storage, data management, and networking infrastructure; collaboration with other researchers; tools and applications; high performance computing; and learning and workforce development, as well as trends within each of these areas.

After two weeks, our survey generated 780 successful responses (a 9.2% response rate) from a broad cross-section of affiliates representing all research disciplines, environments, and campuses.

Acknowledgments:

Thanks to our sponsors Ann Hill Duin and Bernard Gulachek, our project leads Kemal Badur and John Sonnack, and our advisors from the University of Minnesota Research Cyberinfrastructure Alliance. Thanks also to CLA Survey Services and OIT Communications for survey assistance. Thanks to Alex Cramer for help with graphic design.