Guide to Accelerate Public Access to Research Data

Association of American Universities
Inquiry · Innovation · Impact

Association of Public & Land-grant Universities

ACKNOWLEDGEMENTS

The Association of American Universities (AAU) and the Association of Public and Land-grant Universities (APLU) have collaborated and led national discussions to improve public access to data resulting from federally funded research. The current *Guide to Accelerate Public Access to Research Data* builds on many prior efforts and is consistent with national and global open science efforts as well as international declarations, such as the Sorbonne declaration on research data rights.

- In 2016 AAU and APLU formed a working group to examine issues relating to public access to the results from federally funded research. This working group examined how to improve public access to data resulting from federally funded research. In 2017, the group issued a **report** with a series of recommendations to universities on how to increase public access to research data on their own campuses and how they might work together to advance these efforts. The group also made recommendations concerning how federal agencies could help facilitate sharing of research data at universities.
- In 2018, APLU and AAU hosted a National Science Foundation (NSF) funded workshop (NSF #1837847) that convened 30 cross-institutional teams with the goal of developing campus-specific strategies for making data from federally funded research publicly available. The two associations issued a **report** chronicling learning from the workshop.
- As part of the NSF-funded (NSF # 1939279) Accelerating Public Access to Research Data Initiative, AAU and APLU reconvened representatives from the university teams at an Acceleration Conference in 2020, to share progress to date, successes and challenges. The associations also facilitated two national Summits to help universities create robust systems for ensuring effective public access to high-quality research data and develop the current *Guide.* The *Guide* has been informed by 261 campus representatives from 111 institutions, representatives from several federal agencies, and other key stakeholders.

Many collaborators and staff members played key roles in the design and implementation of the Accelerating Public Access to Research Data Initiative as well as the creation of this *Guide*: Tobin L. Smith, AAU Vice President for Science Policy & Global Affairs; Kacy Redd, APLU Associate Vice President, Research and STEM Education; Gregory Madden, Chief Information Officer, University Corporation for Atmospheric Research (UCAR); Emily R. Miller, AAU Deputy Vice President of Institutional Policy; Sarah Nusser, AAU Senior Fellow and Professor of Statistics, Iowa State University; Robert Samors, AAU Senior Scholar; and Katie Steen, former AAU Federal Relations Officer and current Manager of Public Policy & Advocacy at SPARC.

We are also grateful for the guidance provided by the Steering Committee to help advance the Initiative: Sarah Nusser (Chair); Jeff Chasen, University of Kansas; Tom Cramer, Stanford University; Kevin Gardner, University of Louisville; Heidi Imker, University of Illinois at Urbana – Champaign; Jim Luther, Duke University; Greg Madden; Pegah K. Parsi, University of California, San Diego; Lawrence Sutter, Michigan Technological University; Tyler Walters, Virginia Tech; James Wilgenbusch, University of Minnesota.

Recommended Citation:
Association of American Universities and Association of Public and Land-grant Universities (2021). *Guide to Accelerate Public Access to Research Data*. Washington, DC.
DOI: https://doi.org/10.31219/osf.io/tjybn
CC BY-NC-SA

This material is based upon work supported by the National Science Foundation under Grant No. 1939279, with additional support from the National Institute of Health. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the associations and do not necessarily reflect the view of the National Science Foundation or the National Institute of Health.

LETTER FROM THE PRESIDENTS

On behalf of the Association of American Universities (AAU) and the Association of Public and Land-grant Universities (APLU), we are pleased to present this *Guide to Accelerate Public Access to Research Data.*

The Guide is intended to serve as a resource to help university administrators develop robust support systems to accelerate sharing of research data. It provides advice to universities concerning actions they can take, as well as the infrastructure and support that may be required to improve access to research data on their respective campuses. It also offers examples of how institutions are approaching specific challenges to providing public access to research data and results.

Advancing public access to research data is important to improving transparency and reproducibility of scientific results, increasing scientific rigor and public trust in science, and -- most importantly -- accelerating the pace of discovery and innovation through the open sharing of research results. Additionally, it is vital that institutions develop and implement policies now to ensure consistency of data management plans across their campuses to guarantee full compliance with federal research agency data sharing requirements. Beyond the establishment of policies, universities must invest in the infrastructure and support necessary to achieve the desired aspirations and aims of the policies.

The open sharing of the results of scientific research is a value our two associations have long fought to protect and preserve. It is also a value we must continue to uphold at all levels within our universities. This will mean overcoming the various institutional and cultural impediments which have, at times, hampered the open sharing of research data.

AAU and APLU hope that this guide will play a useful role in helping universities tackle the ongoing institutional challenges associated with ensuring public access to research data and will accelerate progress toward making research data widely and freely available to those who can benefit from it.

Sincerely,

Barbara R. Snyder
President
Association of American Universities

M. Peter McPherson
President
Association of Public and Land-grant Universities

Introduction

"… Ensuring that research data are more accessible clearly has tremendous potential to fuel scientific analysis and discovery by making data more open to scrutiny, re-analysis, and extension."

—Report of the AAU-APLU Working Group on Public Access, November 2017

Ensuring broad-based public access to research data is fundamental to advancing the research, education, and service missions of institutions of higher education. Public access to research data is, in fact, a natural continuation of academic institutions' research mission and their function in creating and disseminating new knowledge for societal and economic benefit. As an element of open scholarship, public access to research data can help accelerate the pace of discovery and its application to societal problems, as well as heighten the visibility and reputation of an institution and its scientists and scholars. By ensuring transparency and facilitating the reproducibility of research results, data access is also important to preserving research integrity and maintaining public trust in science. Finally, as stewards of taxpayer dollars and innovators in research, research institutions must meet, and make good-faith efforts to exceed, public expectations and government mandates regarding access to the results of their research and scholarship.

While a consensus is emerging among federal policymakers and many in the university and scientific community regarding the value of making research data publicly accessible, many barriers still exist to achieving this goal. Overcoming these barriers will require: a commitment of resources by both universities and federal research agencies; the development of new institutional data policies; the extension or creation of new data services and infrastructure; and a major cultural shift within universities, scientific disciplines, and individual university departments concerning how faculty members are evaluated, assessed, recognized, and rewarded regarding their data stewardship practices. Greater coordination among campus stakeholders (including the university provost, senior research officer, and chief information officer as well as general counsels, compliance, privacy, and security officers, librarians, faculty members, and students)

will also be vital to ensuring broad-based data accessibility and data protections in those instances where it is required.

This Guide provides universities with a road map to initiate or bolster current efforts to create a robust system for ensuring effective public access to high-quality research data. The Guide aims to assist universities and their senior administrators in crafting consistent and uniform approaches to all aspects of research data management and sharing (i.e., data stewardship) on their campuses. The authors also hope the Guide will facilitate development of standard **research data stewardship** practices at AAU and APLU member campuses that will promote compatibility and interoperability among institutions and ensure that institutions are able to retain academic control over their research results and associated data products. The Guide can also help federal research agencies and other government partners understand how universities are ensuring sponsored research is accessible.

Research data stewardship refers to the activities required to plan, acquire, process, document, and package research data for sharing, as well as the acts of reviewing research data for potential restrictions and making the data publicly accessible.

The Guide is divided into three sections: framing an initiative to accelerate public access to research data, establishing the priority and planning structures for implementing the initiative, and implementation areas to consider in developing the plan. Beginning with the argument for embracing rigorous sharing of research data, the Guide outlines a series of recommendations and initial action steps for building and implementing a robust approach to supporting public access to research data. Additional resources and examples from campuses that have begun to build their systems are provided, along with questions to assess progress for each of the recommendations.

Ultimately, it is the aspiration of APLU and AAU that this Guide will facilitate adoption of new institutional policies, procedures, and approaches that actively support and promote research data sharing, while at the same time ensure rigor in the research process and the veracity of its intellectual outputs.

PART 1

Framing a Campus Initiative to Accelerate Public Access to Research Data

PART I

Framing a Campus Initiative to Accelerate Public Access to Research Data

Why Public Access to Research Data is Important

Public access to research data is a key tenet of open scholarship, a paradigm that modernizes how to achieve transparency and collaboration in research and scholarship. **Research data** sharing is strongly aligned with one of the core missions of research institutions: to create and share new knowledge and address societal problems while ensuring scholarly rigor and compliance with federal, state, and local policies. Data transparency also provides opportunities for the research, researchers, and institution to have greater visibility.

Research data as defined by the Code of Federal Regulations (2 CFR 215.36) is *"the recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not any of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, [or] communications with colleagues. This 'recorded' material excludes physical objects (e.g., laboratory specimens)."* Institutions may more broadly define research data. For examples, see Rice University's definition and the University of North Georgia's definition,

Accelerating Discovery and Innovation

Funders and the public seek returns on their investments through timely release of research findings, data, and other outputs. By making the best science widely accessible as quickly as possible, we can increase the speed at which science advances and is translated in ways that address scientific and societal challenges. This confers benefits to researchers and institutions through increased visibility, citations, and impact associated with publicly accessible research findings[1]. Further, the rapid availability of research outputs opens the potential for new collaborations that extend a researcher's work or introduce new threads of inquiry, increasing the opportunity landscape for both scholars and institutions.

Increasing Rigor and Public Trust

Research institutions, funders, disciplinary communities, and the public expect scholarship to be rigorous and defensible. Concerns about research rigor and integrity have

1 Colavizza, G., Hrynaszkiewicz, I., Staden, I., Whitaker, K., & McGillivray, B. (2020). The citation advantage of linking publications to research data. PLOS One, 15(4), e0230416. Available at: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0230416

led to waning public trust and highlighted the need to increase transparency in scholarship. Federal agencies now see transparent sharing of well-documented data as central to addressing issues of research integrity. Transparency enables others to understand the context (goals), process (methods), and products (article, data, code, etc.), and to evaluate the quality, relevance, and limitations of research for the specific question being investigated. In addition, by embedding the intent to share data and other research outputs in the study planning and design phase, actions to meet transparency expectations create and reinforce a research process that increases the rigor and quality of the work. Instituting and sustaining public access policies and practices at both the institutional level and within disciplinary units will create more opportunities for other scientists to examine, test, evaluate, and validate the research methods, data, and scientific findings of research performed by their colleagues.

Meeting Compliance and Other Sponsor Requirements

An important obligation of sponsored funding is to meet privacy, confidentiality, and cyber and national security requirements. Federal agencies and other research sponsors have increasingly developed specific guidance about required behaviors for data sharing during proposal, sponsored project, and post-award phases, and compliance monitoring is expected to follow. Research institutions that fail to systematically comply with common contractual requirements will damage the capacity of their researchers to successfully seek funds and potentially create onerous administrative burdens for the institution.

Accelerating Public Access to Research Data will Require Cultural Change

Scientific research occurs in complex organizational systems. Creating sustainable reforms to accelerate public access to research data requires rethinking institutional structures *and* culture. Drawing on successful systems approaches used to transform undergraduate teaching and learning for science, technology, engineering, and mathematics (STEM)[2], the Guide offers strategies for public access to research data by outlining what academic institutions can do to change campus culture and expand supporting infrastructure to promote research data sharing.

Implementing sustainable change requires individuals finding and using the correct levers for change[3] that will counterbalance forces that reinforce ineffective practices

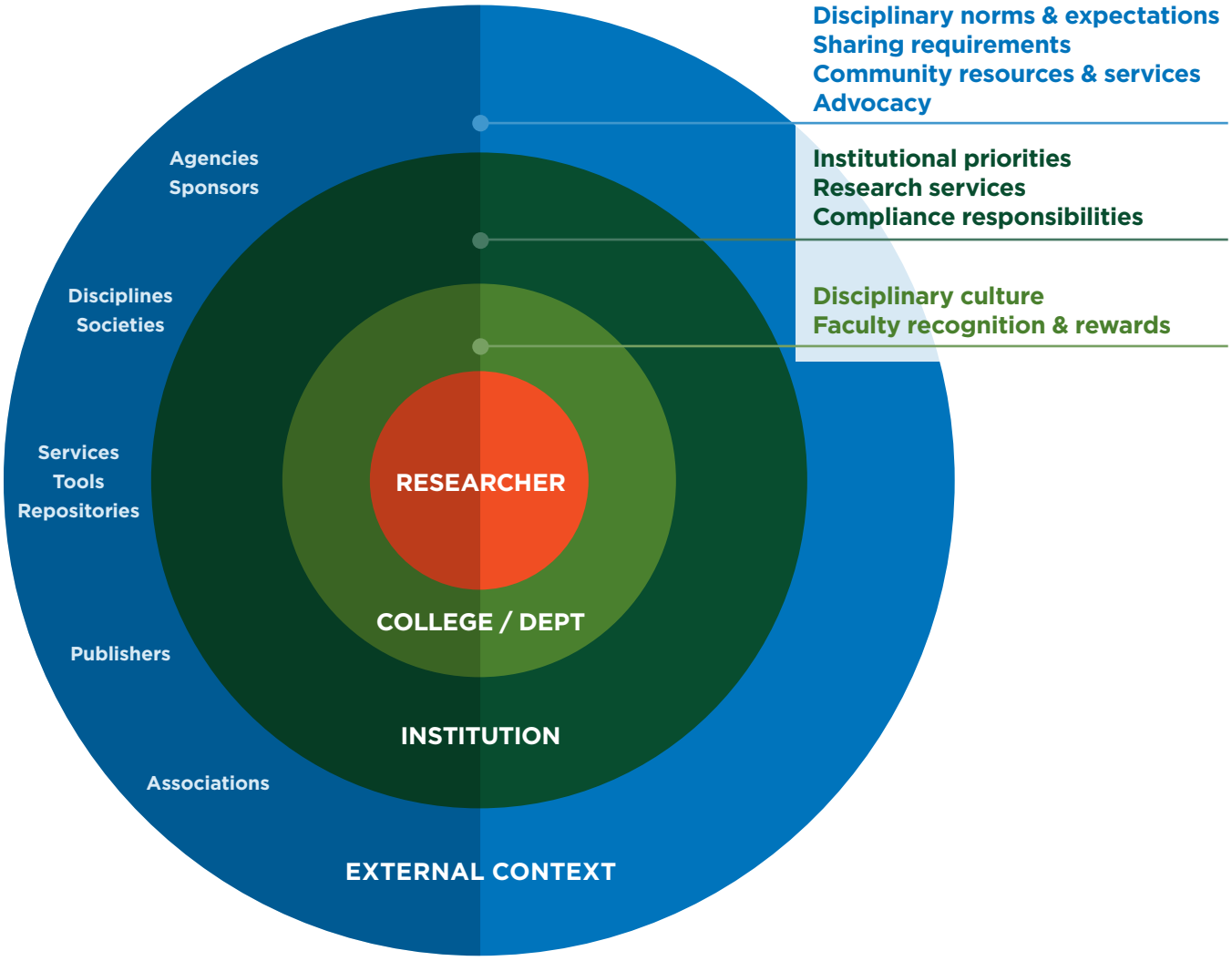
2 Austin, A.E. (2011). Promoting Evidence--Based Change in Undergraduate Science Education. Paper commissioned by the Board on Science Education of the National Academies National Research Council. Washington, D.C.: The National Academies.

3 Austin, A.E. (2014). Barriers to Change in Higher Education: Taking a Systems Approach to Transforming Undergraduate STEM Education. White paper commissioned for Coalition for Reform of Undergraduate STEM Education. Washington, D.C.: Association of American Colleges and Universities. www.aacu.org/CRUSE

and address the obstacles inherent in the system where research practice innovations need to take place.

One important set of levers arises from the external ecosystem promoting public access to research data. For example, disciplinary practice norms, federal research agency policies and publisher requirements, the availability of community-based tools and data repositories, and the views of university associations and disciplinary societies all influence behavior and actions within the university and represent levers or barriers to catalyzing institutional change.

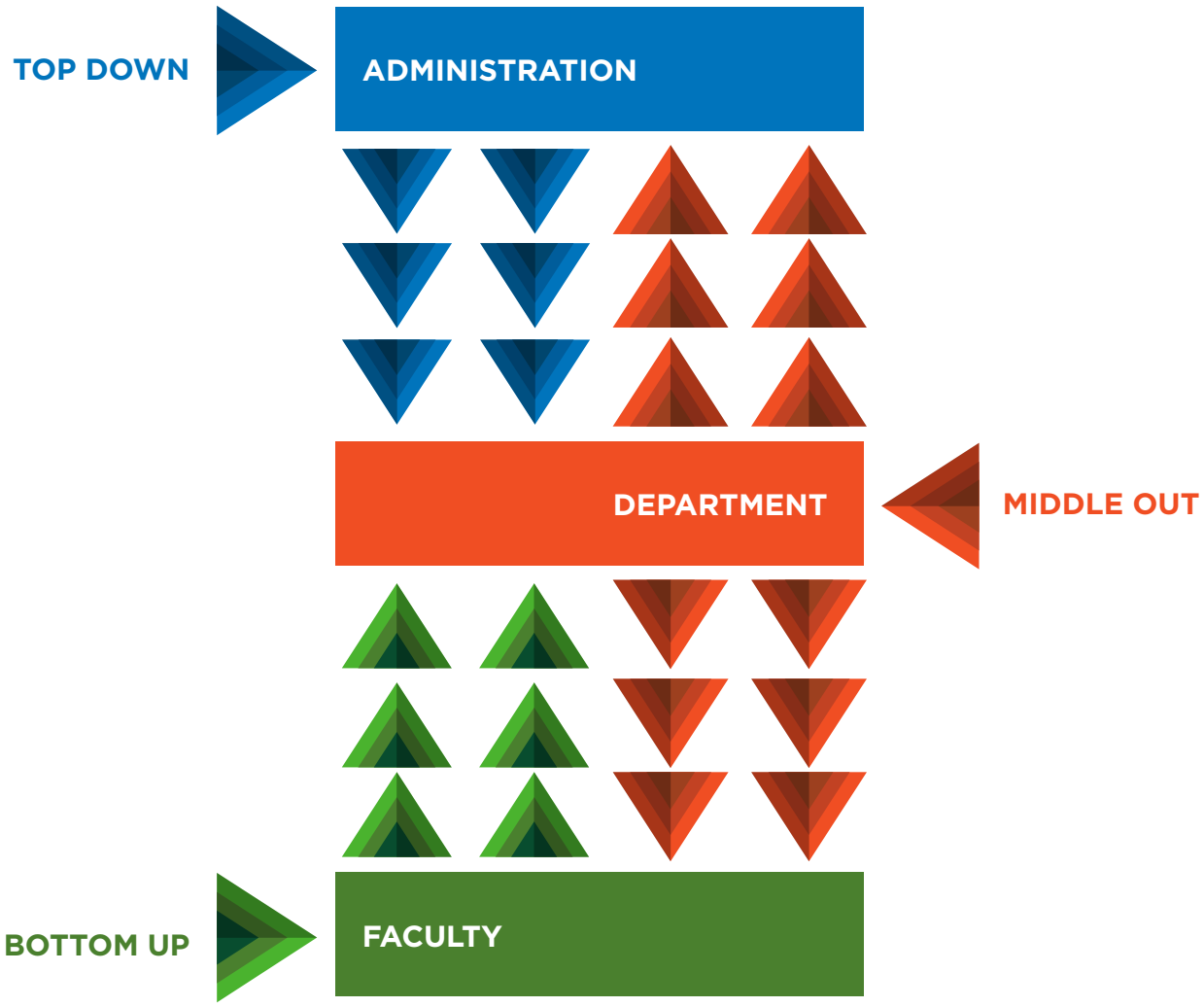
Ecosystem influencing researcher actions



Adapted from Austin, 2011, *Promoting evidence-based change in undergraduate education,* National Research Council.

Essential campus levers include not only top-down (senior campus leaders) and bottom-up (faculty) reforms, but also "middle-out" reforms – facilitated by department chairs, college deans, and other mid-level administrators. These are especially important to fostering culture change[4]. As demonstrated by transformations in STEM education initiatives, central efforts aimed at increased institutional support, adopting new policies and procedures, and providing for additional infrastructure to support data sharing are critical but insufficient to facilitating institutional change. Real and long-lasting change will only occur if *all* levels within the institution take an interest in, and feel that they hold a stake in, seeing such change occur.

Campus Levers for Change[4]



4 Corbo, J. C., et al. (January 26, 2015). "Sustainable Change: A Model for Transforming Departmental Culture to Support STEM Education Innovation," Physics Education Research. http://arxiv.org/abs /1412.3034

Ultimately, institutional change to accelerate public access to research data can only occur if research faculty inherently value and buy into the notion that making their research data accessible is important to advancing their work as a scholar and has value both to themselves and to advancing knowledge within their particular research field. The ease with which faculty are able to make their data accessible can also play a significant role in the ability of faculty to ensure their research data is accessible, vetted for quality, and well-documented so that it can be understood and reused by others.

While not the only factor in achieving sustained cultural change, senior university administrators are a critical ingredient and can play a key catalytic role. It is important for the university administrator to signal from the top that the institution values making research data accessible. Institutional administrators can help to send such signals through steps including:

- Making public pronouncements about the value and importance of public access to research data and articulating this goal as part of university priorities.
- Convening a cross-campus working group on public access to data consisting of key university officials from the research and sponsored programs office, library, office of the chief information officer, faculty, and others as deemed appropriate to each specific institutional context.
- Developing a clear and consistent campus-wide **research data policy** with clearly stated institutional expectations for data management and sharing.
- Providing existing and new resources to build the institutional support and infrastructure required to enable research faculty to easily make their data publicly accessible.
- Encouraging colleges, departments, and/or units to discuss how to account for and recognize public access to data in evaluation and assessment of faculty performance and how to make sure that data accessibility is recognized in the university's overall faculty rewards structure, including annual review, promotion, and tenure.

Research data policy sets the roles, rights, responsibilities, and expectations for researchers and the institution in their research data stewardship activities. Policies are formally adopted by the institution.

The remainder of this Guide builds on these and other recommendations that aim to influence researcher actions in ways that will help change the culture to promote and recognize the importance of public access to research data.

PART 2

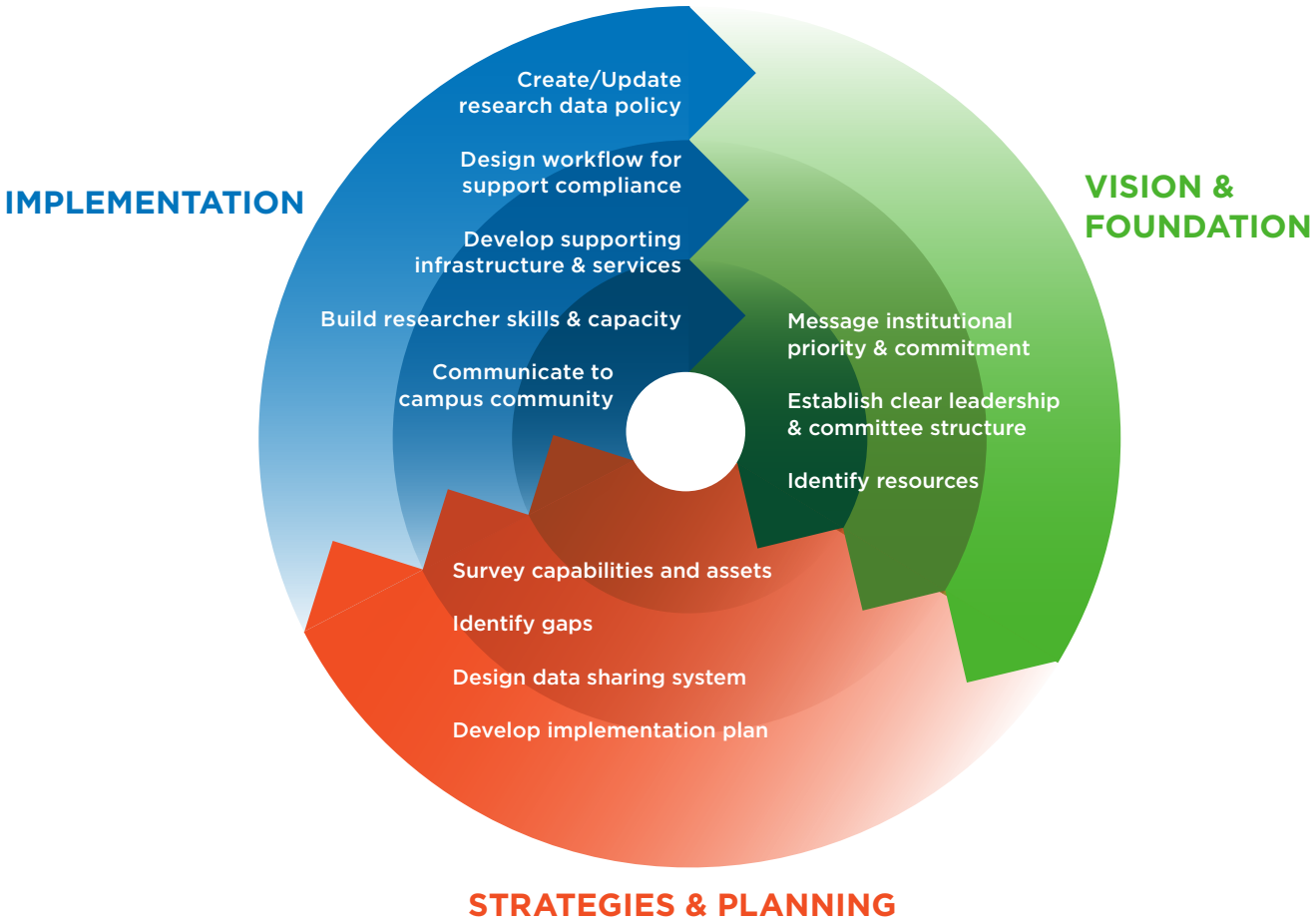
Making Priorities Visible and Establishing a Plan

PART 2

Making Priorities Visible and Establishing a Plan

Accelerating public access to research data at research universities will be a continuous cycle. Each campus has a unique context and starting place in supporting public access to research data. Institutions should consider recommendations in light of their specific challenges, evaluating which are most important, which can be scheduled for later in the process, and which are not applicable. The process itself is not always linear, and as institutions make progress, they may find it beneficial to reflect and evaluate after gaining experience from a pilot or implementation phase.

Cycle to Accelerate Public Access to Research Data



Refine priorities and implementation plan based on an assessment of institutional progress and changes in the external ecosystem.

Establish Data Sharing as a Campus Priority

Senior leaders play an essential role in ensuring the campus culture embraces data sharing. To ensure the success of an initiative to accelerate public access to research data, campus administrators at the highest level must clearly articulate the importance to the institution of sharing research data and continually reinforce the institution's priority and commitment to supporting data sharing. Visible and consistent top-level messaging will help ensure the success of researchers in sharing their data and campus efforts to establish a robust data sharing support system.

Initial Action

- Develop a message for the president and provost to articulate and regularly reaffirm that the institution prioritizes and supports the open sharing of research results, including ensuring public access to research data.

Create a Structure to Guide a Campus Initiative and Provide Support to Ensure its Success

A campus initiative should be a priority project to develop research culture and build a robust infrastructure to support researchers in their data stewardship activities, ultimately leading to publicly accessible research data. To guide the campus initiative in accelerating public access to research data, senior leaders should establish and support an approach for coordinating planning, implementing, and communicating the activities of the initiative. This includes several elements: establishing who is responsible for managing the initiative, a cross-campus structure or **coordinating committee** to develop and manage the initiative, a clear charge for the structure, and resources to support the initiative *(see pg. 15)*.

Coordinating committee refers to a cross-campus structure or a body of representatives from administrative and academic units that are charged with developing and overseeing implementation of a plan to build or enhance the infrastructure needed for a researcher to successfully comply with providing public access to their research data. Institutions have called such structures a task force, steering group, advisory body, or committee.

Just as campus efforts can falter if members of senior leadership fail to consistently send the message that data sharing is a priority, it is also essential that the roles and responsibilities of the participating campus stakeholders are well-defined. Especially critical is clearly establishing who will manage ("own") the planning, development, and implementation process. Another important area to consider is the membership of the cross-campus coordinating structure to integrate perspectives in developing and implementing the initiative, which may take the form of one or more committees (e.g., advisory group and committee on public access to research data, task force with senior leader sponsors). Senior leaders should draw from administrative units that are key to understanding, developing, and/or expanding needed services and infrastructure (e.g., representatives from the research office, research data and library services, information technology, sponsored funding and compliance units, policy and privacy

Data Sharing Task Force and Charge at Iowa State University

In October 2017, Iowa State University established a Data Sharing Task Force (DSTF) to develop a framework for implementing services and policies necessary to support research data sharing. The sponsors of the Task Force were the Vice President for Research (VPR), University Library Dean, and Chief Information Officer (CIO). It was co-chaired by the Associate VPR, Associate Dean of the University Library, and Associate Director of Information Technology Services. Its membership included faculty from diverse fields; directors and/or associate directors from the Office of the CIO, Office of Intellectual Property and Technology Transfer, Office of Research Ethics, Office of Sponsored Programs Administration, Office of the VPR; associate dean of the University Library; and research counsel from the University Counsel.

The DSTF was charged with considering the actions and guidance needed to support researchers and the institution in providing public access to research data. These considerations included support for quality research practices; polices related to data sharing and credit; awareness of and compliance with sponsor requirements, processes, and workflows; and the information systems to support these processes and data sharing. The members were also asked to consider researchers' needs for proposal development, designing and conducting a study, evaluating the data products for sharing, and data dissemination and documentation.

The charge had expected deliverables, which included the need to pilot elements of a data repository system. It also included goals and responsibilities of Task Force members regarding key aspects of this pilot.

DOMAIN	PILOT GOALS	TASK FORCE MEMBERS
Policy	Research data, repository	Research counsel, IT systems lead, IP lead, AVPR for research ethics
Compliance	Awareness and prevention, monitoring, noncompliance	Research ethics and sponsored programs staff
Research practice	Understanding requirements, rigor and reproducibility, workflow and documentation	Faculty
Systems and services	Systems (repository, master record, Kuali, workflow)	Library, IT security lead, AVPR for research ethics

Over the following two years, the DSTF guided development of the Iowa State University DataShare repository, research data policies, workflows, training and other resources. In 2020, the DSTF outlined the next phase of development, and Task Force's work was considered completed.

Full charge is here: https://www.research.iastate.edu/wp-content/uploads/2021/02/Iowa-State-University-Data-Sharing-Task-Force.pdf

offices), as well as researchers and academic units affected by research data stewardship requirements (e.g., faculty, department chairs, associate deans for research). It is also important to set forth a well-delineated charge with the initiative's goals, how the coordinating body will interact with senior leadership and campus stakeholders, and what they should strive to achieve. Senior leaders must also identify and align institutional resources to support the envisioning, planning, and implementation stages required to implement and sustain meaningful actions to support researchers in meeting their requirements to responsibly share their research data.

Initial Actions

- Identify existing campus organizations and structures central to the planning, development, implementation, and governance of research data stewardship policies and practices.
- Identify the lead unit and individual to coordinate the planning and development process.
- Establish an inclusive, institution-wide structure with a clear charge, defined roles and responsibilities, and accountability for measures monitoring the progress and success of the initiative.
- Identify adequate resources to support the planning, development, and implementation process.

Develop a Plan for an Effective Research Data Stewardship System

A well-conceived strategic goal and implementation plan is the foundation for ensuring a successful initiative to accelerate public access to research data. Once a cross-campus coordinating committee has been established, it can begin gathering information and developing a plan to provide or expand campus support for research data stewardship. A common approach in scoping the planning phase is to focus primarily on research data underlying scholarly publications. Early in the process, the structure should engage in information gathering activities such as: developing a broad understanding of sponsor requirements for making data publicly accessible; assessing current campus policies, services, and infrastructure and what may be missing to support research data sharing; and investigating how comparable campuses have approached building their data sharing infrastructure. Ultimately, the cross-campus structure will benefit from drafting a process **workflow** document outlining necessary decisions and actions by researchers and administrative units to help contextualize information that has been accumulated in this phase and to guide the implementation planning. The workflow document should consider what the researcher and institution need throughout the entire research data life cycle associated with data stewardship – e.g., submitting

Workflow represents the steps that need to be taken in research data stewardship as part of data sharing and compliance with institutional and sponsor requirements.

proposals, receiving awards, planning and designing the study, acquiring and processing data, documenting study design and methods as well as data, evaluating data for sensitive information and compliance, and curating and managing the data as it is made publicly accessible. In any of these endeavors, campuses should keep in mind that a workflow will need to be applicable across the institution, regardless of discipline or field. With this information in hand, a strategic plan can be developed that guides implementation and assessment of the campus data-sharing initiative. The actions below provide more detail on some of the elements that need to be considered in developing a plan and should be reviewed as part of the planning process.

Initial Actions

- Define the scope of the research data system that needs to be considered in information gathering and planning.
- Conduct an initial scan of the data sharing landscape, including: an assessment of the institutional, funder, and other organizations' requirements relating to open data/data sharing; an inventory of existing research data policies, procedures, services, systems, and resources in place to meet those requirements; and a gap analysis of expertise, technical services and technology, and financial resources required to meet those requirements.
- Develop a set of goals and a high-level workflow for research data stewardship to serve as a guide for outlining an approach to developing or enhancing campus infrastructure for supporting public access to research data.
- Develop a strategic plan with implementation steps, an assessment approach, and an indication of human, technical, and financial resource needs for the project.

PART 3
Key Implementation Areas to Consider

PART 3

Key Implementation Areas to Consider

Establish or Update Institutional Research Data Policy and Related Guidelines and Procedures

A campus research data policy establishes expected roles, rights, and responsibilities regarding data management and sharing for researchers and the institution. The purpose of the policy often refers to a commitment to data sharing and research integrity as a standard part of the research process, mandates from federal agencies and other organizations, and sometimes a philosophical or principle statement (such as a commitment to transparency and open science). The body of the policy often includes statements relating to ownership, management, and sharing of research data; retention, archiving, and transfer of data; protecting sensitive information; and roles and responsibilities of campus community members (e.g., principal investigators, students, administrative offices). As with any policy, the narrative should be concise and easy to understand and should clearly define terminology. The narrative should be mindful of the variations in research data stewardship practices across all disciplines and funder mandates and strive to minimize administrative burden on researchers. The development of the policy should also involve consultation with relevant stakeholders, including researchers and administrative units relevant to implementing and supporting the policy.

Guidelines include guidance and other resources to help researchers or other campus members execute their research data stewardship activities in compliance with the Research Data Policy.

Procedures are the processes set by the institution to support the researcher in making data publicly accessible and complying with institutional and sponsor requirements.

Practices are the steps researchers take in their research data stewardship activities.

With a policy in hand, associated **guidelines** and **procedures** can be established for researchers to follow in adhering to the policy and updating their research **practices**. These associated guidelines and procedures can provide processes and touch points for complying with the policy and working with offices that support research data stewardship and research compliance. In creating these processes and touch points, coordinating committees should consider the workflow diagram, especially in designing an approach to becoming aware of data management and sharing requirements for sponsored funding awards and tracking researcher compliance with funder requirements for public access to research data. Guidelines are also an opportunity to encourage a life-cycle perspective for data stewardship, e.g., considering data-sharing requirements during the planning phase of a research study; starting early on documenting goals, methods, processes, and data; and addressing how to ensure quality in data stewardship.

Initial Actions

- Review the institution's policies related to research data and explore research data policies of other institutions.
- Discuss and draft content of the policy, including roles, responsibilities, and expected actions of individuals and organizations across the campus.
- Working with the research data workflow developed by the coordinating committee, identify guidelines and procedures necessary to implement the data policy (including how compliance requirements will be monitored).
- Seek feedback from campus stakeholders on drafts of policy, guidelines, and procedures.
- Define a comprehensive strategy around awareness, compliance, and monitoring regarding the campus research data policy.

Model Policy Documents

- **Florida State University**
  https://regulations.fsu.edu/sites/g/files/upcbnu486/files/policies/research/FSU%20Policy%207A-26.pdf
- **Rice University**
  https://policy.rice.edu/308
- **University of Minnesota**
  https://policy.umn.edu/research/researchdata
- **University of New Hampshire**
  https://www.usnh.edu/policy/unh/viii-research-policies/c-unh-policy-ownership-management-and-sharing-research-data
- **University of North Georgia**
  https://policy.ung.edu/policy/research-data-management

Identify Research Data Services and Expertise for Supporting Public Access to Research Data

The success of researchers in navigating the opportunities and risks of data sharing rests, in part, on providing services that reduce researcher burden, improve the quality of shared research outputs, and promote compliance with data stewardship practices. The coordinating committee should consider what research data services are needed to help researchers effectively execute the data stewardship workflow, comply with the research policy, and follow guidelines and procedures for data sharing. Services that would be helpful to researchers include: upfront planning in relation to how data will be developed for sharing; identifying a trusted repository for data; guidance on how to

document data with sufficient context for new users; protecting sensitive information; reviewing of research data outputs in relation to compliance requirements from sponsors and other entities; and curation, documentation, and management of shared data.

Along with identifying these services, it is essential to identify expertise needed to provide knowledgeable and responsive support to faculty, staff, and students as they follow the established policy and procedures. It is critical for campuses to recognize that no single person or department understands or is able to implement all aspects of research data stewardship, and that a broad array of individuals from across the campus will need to be tapped. The needed expertise is likely to come from a range of perspectives and disciplines – e.g., data and information science; research policy and administration; software and information systems; and data-specific realms such as privacy, disclosure limitation, and licensing. In addition, supporting services may need to be staffed with individuals who are experts in specific scholarly areas or specific research approaches (e.g., human subjects) to support the unique expectations for a given discipline or field.

Data Curation Network

The Data Curation Network (https://datacurationnetwork.org/) is a cross-institutional collaboration that supports public research data access. Modeled after successful library consortia and open-source coding communities, the Data Curation Network operates as a central platform for partner institutions to pool data curation experts. Data curators are the "human layer" in an institution's cyberinfrastructure; they collaborate with researchers to ensure that data are shared without ethical, legal, and translational barriers to reuse.

The Data Curation Network expands data curation services beyond what any single institution might offer alone. By building a bridge between institutions grappling with similar research data challenges, the network is a catalyst for sharing best practices, workflows, and resources with the broader data stewardship community. Working groups tackle a wide range of topics, including:

- campus advocacy and outreach on data services
- data curation education
- special interest groups on big data and human subjects data
- diversity, equity, inclusion, accessibility (DEIA) in data curation work
- writing "data curation primers," resources on how to curate specific file formats (e.g., geodatabases, neuroimaging files, and confocal microscopy) and specialized topics (e.g., human subjects, oral histories).

Some research data services will require technical infrastructure to be established or updated. Some examples of the types of infrastructure likely to be needed include: systems to create, update, and manage research data management and sharing plans for proposal submissions and compliance monitoring once an award has been made; an institutional repository or storage approach for making data and associated research outputs publicly accessible; the capacity to obtain persistent identifiers (e.g., ORCID,

ROR, DOIs) for people, organizations, data, and other entities to enable digital tagging of objects and promote FAIR data systems (FAIR data are findable, accessible, interoperable, and reusable). Campuses may opt to develop at least some of these systems internally, or may consider leveraging external resources, e.g., in developing partnerships with other institutions and organizations to share resources and capabilities.

Five Recommended Persistent Identifiers

The Association of Research Libraries, the California Digital Library, APLU, and AAU released a report, ***Implementing Effective Data Practices: Stakeholder Recommendations for Collaborative Research Support***, in 2020 with recommendations for data practices supporting an open research ecosystem. The report identified five core Persistent Identifiers (PIDs) that are fundamental and foundational to an open data ecosystem. Using these PIDs will ensure that basic metadata about research is standardized, networked, and discoverable in scholarly infrastructure:

1. Digital object identifiers (DOIs) to identify research data, as well as publications and other outputs
2. Open Researcher and Contributor (ORCID) IDs to identify researchers
3. Research Organization Registry (ROR) IDs to identify research organization affiliations
4. Crossref Funder Registry IDs to identifier research funders
5. Crossref Grant IDs to identify grants and other types of research awards

Initial Actions

- Utilize the high-level research data stewardship workflow to identify services, technical infrastructure, and expertise needed to support researchers in data sharing.
- Evaluate how other institutions have organized their services and infrastructure and what expertise they provide in promoting and implementing effective data stewardship practices.
- Develop a multiyear plan to establish or update research data services.
- Create a centralized web-based portal for data sharing that will link users to the various campus services and systems, as well as useful external resources.
- Evaluate institutional memberships with services that support persistent identifiers, data citation, and other elements of the public access data system (e.g., ORCID, DataCite, Crossref).
- Explore the potential to work with external partners to share technical resources and associated costs.

Identify Costs and Resources to Address Them

Identifying, examining, and securing the financial resources necessary to support the services, staffing, and technology necessary for open scholarship is essential to effective implementation of research data stewardship practices. A thorough examination of the costs involved in campus-wide research data stewardship is critical to successful implementation. While it can be difficult to broach the topic of funding, each campus will need to determine the initial and annual, sustaining investments required to successfully implement research data stewardship systems and services outlined in the previous section. A number of resources and templates exist that could provide guidance to campuses developing research data stewardship cost models. Part of the solution to financial support lies in reducing costs. For example, institutions may develop fee-for-service structures for campus research data services that enable researchers to appropriately budget and charge their sponsored awards for data management and sharing costs. Alternatively, many scholarly fields have established repositories that could be leveraged to reduce campus storage costs. Institutions may explore other creative models to fund these initiatives, such as agency and nonprofit grant programs that support pilot tests or infrastructure development, or further reducing costs through shared hosting of services with other organizations.

Guidance on Costs

The National Institute of Standards and Technology (NIST) has released a Research Data Framework (RDaF) that has both a research data ecosystem and data lifecycle approach. The RDaF core in Appendix E can provide insight to campuses trying to estimate where costs might be accrued.

Kaiser, D. L., & Hanisch, R. J. (2021). Research Data Framework (RDaF): Motivation, Development, and a Preliminary Framework Core. Available at https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-18.pdf.

Life Cycle Decisions for Biomedical Data: The Challenge of Forecasting Costs provides a framework for cost-effective decision making for biomedical research data preservation, discoverability, and use. The appendices include salary ranges for relevant jobs for the data life cycle (Appendix C), soft costs for digital preservation (Appendix D), and a template to map cost drivers (Appendix E).

National Academies of Sciences, Engineering, and Medicine. (2020). Life-Cycle Decisions for Biomedical Data: The Challenge of Forecasting Costs. Available at https://www.nationalacademies.org/our-work/forecasting-costs-for-preserving-archiving-and-promoting-access-to-biomedical-data

Initial Actions

- Develop a multiyear budget for the staff and infrastructure necessary to implement effective research data stewardship practices across the institution.
- Encourage researchers to include research data curation and storage costs in their research budgets whenever allowed by research sponsors.
- Develop pilot efforts to better estimate the component costs of effective research data stewardship programs.
- Encourage researchers to use domain repositories associated with fields of study or organizations to reduce institutional costs.
- Explore potential consortia approaches to research data storage and management as ways to control costs.

Build Capacity and Skills for Good Research Data Stewardship Practice

Building the capacity and skills among faculty, graduate students, and other researchers is a critical element to successfully implementing research data stewardship practices and procedures. The work put into developing services and infrastructure can be leveraged by experts creating a robust professional development program. Approaches may range from workshops that discuss research data stewardship for the full data life cycle to short courses on key software tools (e.g., GitHub, Jupyter notebooks) to written or video content that explains how to execute a specific task (e.g., obtain a persistent identifier). Services should also have some capacity for consulting directly with departments, research groups, or individuals, ideally with the ability to discuss research practices in a specific field of scholarship. Another option is to include data stewardship as a topic in responsible conduct of research curriculum or research methods coursework for a specific discipline. In addition, awareness of support and training can be built through a communications program and training outreach that targets researchers of all types, including immediate touch points with new faculty, students, and staff, as well as other regular email or newsletter updates to campus researchers. In addition, external organizations frequently host workshops that could be made known to campus researchers.

Cornell University–
Research Data Management Service Group
Comprehensive Data Management
Planning & Consultation Services

The Research Data Management Service Group (RDMSG) is a collaborative, campus-wide organization that links Cornell University faculty, staff, and students with data management services to meet their research needs. The RDMSG's broad range of science, policy, data, and information technology experts provide timely and professional assistance for the creation and implementation of data management plans (DMPs), and help researchers find specialized data management services they require at any stage of the research process.

As a virtual organization, the RDMSG is composed of a management group, a consultant group, and implementation teams. The management group includes a body of decision-making administrators from campus service providers such as the Center for Advanced Computing, the Cornell University Library, Cornell Center for Social Sciences, and central IT, other stakeholders (e.g., chief information officers from the Ithaca and Weill campuses), and a staff coordinator. The RDMSG consultant group consists of science, GIS, social science, digital humanities, metadata, and medical librarians, scholarly communication experts, a senior policy advisor, and other campus service provider staff data experts. Implementation teams conduct assessments, provide outreach and training, and initiate new projects. RDMSG is sponsored by the Cornell Office of the Vice President for Research and Innovation, the University Librarian, and the Cornell IT Chief Information Officer, and is guided by a faculty advisory board.

Consultants
RDMSG consultants are available to meet with researchers upon request to assist with a wide range of data needs, including DMP proposal preparation, finding data, evaluating and selecting data storage tools, analyzing data, creating metadata, sharing and publishing, and more. The RDMSG encourages best practices in data management, including those that promote sharing, reuse, and preservation of research data following FAIR principles, while respecting the concerns and practical constraints researchers face. The consultants bring diverse backgrounds and expertise, share information with each other to provide the best possible service, and collectively treat information in grant proposals as confidential.

Outreach and Training
Complementing their strong web presence, RDMSG members offer guidance and instruction on various aspects of research data management, such as preparing data management plans, privacy, data publication, metadata and documentation, data cleaning and analysis, and best practices for data management. Training occurs throughout the year as workshops, information sessions, customized small group sessions, and as course-related instruction or one-on-one interactions.

RDMSG
RESEARCH DATA MANAGEMENT SERVICE GROUP
Comprehensive Data Management Planning & Services

Initial Actions

- Develop consensus of research data stewardship training priorities and goals, and outline a plan with roles and responsibilities to support research data stewardship education and training.
- Identify specific opportunities to educate and inform faculty/researchers – especially Early Career scholars – of the research data policy and the importance of effective research data stewardship practices. Such opportunities could include the hiring/onboarding process, department-level seminars, workshops, information sessions, and other professional training opportunities.
- Identify research data stewardship education and training opportunities for graduate and undergraduate students, such as Responsible Conduct of Research modules.
- Develop case studies, using campus-based examples, highlighting exemplary research data stewardship practices, including descriptions of the outcomes/benefits of the practices employed.

Build Support for Embracing the Research Data Policy and Research Data Sharing

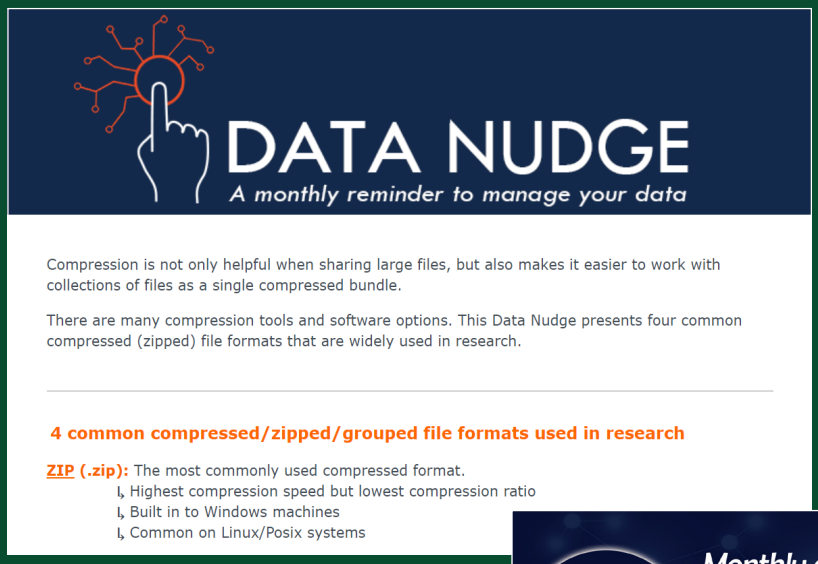
A significant challenge to successful implementation of a campus initiative to support public access to research data is building awareness and support among faculty and other researchers for the necessity and benefits of practicing effective research data stewardship. Besides the importance of messaging from the president and Provost, a campus-wide communications strategy should be considered. Communications should emphasize how the systems, services, training, and support being established across the institution are designed to enable researcher success in sharing data while minimizing burdens and otherwise making the adoption of new research data stewardship practices as seamless as possible. Messaging should also emphasize how effective research data stewardship practices and procedures can contribute to research integrity and scholarly excellence, increase the value and visibility of one's research, make scientific collaboration easier, and provide a return on the public's investment in research funding. Institutions should seek opportunities to engage with faculty, students, and other researchers within individual research groups, centers, or institutes, as well as within departments, schools, and disciplines. In addition, as part of resetting the culture, administrators may consider local discussions within departments to gather ideas on how incentives, recognition, and rewards can be aligned to support public access to research data.

University of Illinois Urbana-Champaign Data Nudge: A Monthly Reminder to Manage Your Data

The Data Nudge is a succinct, topic-focused monthly email to help "nudge" researchers towards better data practices. It is created by the Research Data Service (RDS) in the University Library but taps experts across campus to ensure accurate and high-quality content. For example, while creating Data Nudge content, the RDS staff will contact IT, IRB, or other professionals on campus to vet materials to ensure accuracy and consistency with university policies and best practices.

Begun in 2017, the open rates average 52% per month and has never dipped below 40%, which is more than 2 times the industry standard for higher education. The current 400+ subscribers include individuals from all career stages and domains--faculty, staff, postdocs, graduate students, and undergrads representing 13 of the University's 16 colleges--as well as others from other U of I research institutions and administrative units. The Data Nudge often receives spontaneous positive feedback about its content, clarity, and relevance, even from outside the U of I, from both other universities and even government organizations.

Some topics are U of I-specific (e.g., local storage options) but others are universal, such as data analysis, data cleaning, data destruction, data loss, data sharing, and data visualization. Although the target audience is researchers at the U of I, anyone can sign up, and the content is freely open to anyone to use directly or adapt for their campuses under a CC-BY license. View past Data Nudges and subscribe to the Data Nudge: http://go.illinois.edu/nudge



Initial Actions

- Create a research data stewardship communications and marketing plan with materials to support that plan.
- Develop a narrative to explain how data stewardship services and systems can be leveraged to enhance researcher success and reduce researcher burdens.
- Identify research data stewardship "champions" to share "success stories" that support institutional data sharing goals and would have a positive impact on faculty/researchers. Champions could be on-campus or off-campus leaders in specific disciplines, or meet other criteria.
- Identify or create opportunities to recognize and celebrate exemplary research data stewardship practices newly adopted by faculty and other researchers, such as awards.
- Encourage colleges, departments, and/or units to discuss how to account for and recognize public access to data in evaluation and assessment of faculty performance and ensure data accessibility is recognized in the university's overall faculty rewards structure, including annual review, promotion, and tenure.

Conclusion

Ensuring public access to research data is fundamental to the role of research universities and their function in creating and disseminating new knowledge for societal and economic impact. The Guide provides universities with strategies and approaches to address the cultural and structural barriers that often hinder efforts to create robust systems for ensuring effective public access to high quality research data. In conclusion, we encourage universities to consider the following essential questions. By reflecting on these essential questions, institutions can engage in thoughtful discussions and assess progress toward establishing data sharing as an institutional priority and improve public access to research data.

PART 4
Essential Questions

PART 4

Essential Questions

The following set of questions aligns with the recommendations in the Guide and draws primarily from the experience of the Association of American Universities and the Association of Public and Land-grant Universities in engaging research universities in a series of grant-funded efforts aimed at understanding the critical levers to build support for research data policies, sharing, and stewardship.

1. What messaging would be most helpful for the president and provost to articulate about the priority and rationale for the public access to research data? How does this messaging align with the institutional mission and priorities? To whom and at what venues would these messages have the most impact? In what ways does this message need to be targeted at specific key constituencies? Which other institutional leaders might be helpful in sharing this message with key constituents?

2. Does the coordinating committee include the various campus communities critical to successful data stewardship that can help drive the planning, development, implementation, and assessment and iteration of research data stewardship policies and practices? Who or what expertise might be missing (e.g., different disciplinary perspectives)?

3. What elements are needed in the coordinating committee's charge to help ensure success? Does it need to address actionable milestones, expected deliverables, indicators of success, responsible parties, a plan for reflection and iteration, and a plan for sharing the work of the committee with campus stakeholders?

4. Does the charge to the committee need to encompass the whole research life cycle, or begin at the point where data is ready to be deposited, or at some intermediate point? Has the institution developed a "research data workflow" that could serve as the basis for developing its research data policies, supporting systems, and practices?

5. Is the committee adequately resourced to do this work? Does it have dedicated staff time and capabilities to help project-manage the work of the committee? What internal and external sources of funding can be leveraged to implement components of the project?

6. In what ways might the coordinating committee design a collaborative and inclusive process to gather input for developing or reviewing a draft data policy? How are other institutions approaching their research data policy? Which approaches are most applicable, and how can they be adapted to the institution's context to develop an initial policy draft?

7. Has the institution identified, marshalled and, if necessary, built the appropriate technological resources to make it as easy as possible for faculty, staff, and students to follow the data management practices and procedures being established? If not, what is needed? Which constituents are best served by the currently available resources? Which constituents are not well served? What is your evidence?

8. In what ways might the institution leverage external resources (e.g., discipline-based or other external repositories), as well as partnerships with other institutions, government agencies, non-profit and for-profit providers, and consortia to share resources and capabilities?

9. How might the institution gain a clearer understanding of the costs of the staff and infrastructure necessary to implement effective research data stewardship policies and practices across the institution? For example, are there any campus pilots that track the component costs of an effective research data stewardship program, including post-grant storage and curation costs? How will the institution cover ongoing costs?

10. What guidance do researchers need to estimate costs to be covered in their sponsored funding proposals? Who else must be involved in helping set these estimates?

11. What are the data stewardship training priorities and goals of the institution and/or the departments? Have departments been consulted for discipline-specific training opportunities?

12. What can departments and/or the institution do to help educate faculty, staff, and students about its data stewardship policies and practices and how to meet those expectations? Have targets been set (e.g., required of all active grantees)? Is there consensus on what a quality curriculum would entail?

13. Has the institution developed a data stewardship communications and marketing plan that addresses all campus communities affected by the institution's research data stewardship policies and practices? How might the coordinating committee identify and engage allies or champions both to advise and carry the message within the communication plan?

14. In what ways does the institution recognize and/or reward students, staff, and faculty who use excellent data stewardship practices? How might these be expanded?

15. Does the institution track metrics related to data sharing (e.g., number of data sets shared, reuse of data sets) that might inform a reward/recognition system? If so, are they used and how?

16. How will the institution assess whether data stewardship practices are being adopted by researchers, staff, and faculty? Are there opportunities to collect other relevant data (e.g., actual costs for research projects, feedback from stakeholders) that can help the institution make data-informed decisions and inform a continuous improvement process?

17. How is the institution addressing the need for culture change? Has the institution developed an approach to convening departmental conversations around sharing data and other research outputs, and how these can be included in performance assessments? Have well-respected faculty who engage in sharing data and open science/scholarship been identified for sharing success stories? Have awards that recognize data sharing and other open research practices been developed to signal the value the institution places on making research outputs publicly accessible?

Association
of American
Universities
Inquiry · Innovation · Impact

Association of
Public &
Land-grant
Universities