

Guide to Publishing Research Data for Secondary Analysis

Heather Gordon, Caroline O’Callahan, Katie Hunter, Hande Inanc, Mary Grider, and Heather Zaveri

Overview

Across the federal government, agencies and their operating divisions collect data through research projects and to support program operations and program improvement efforts. This data collection frequently occurs through contracts. Throughout the proposal and project phases, agencies guide how they manage data, including whether to publish the data for secondary use. That is where archiving of research data comes in. **Data archiving involves the preservation and publishing of data for secondary analysis (reuse) by researchers and others.**

The Office of Planning, Research, and Evaluation (OPRE) within the Administration for Children and Families (ACF) conducts descriptive and impact research studies to build evidence about ACF programs and the populations they serve. Securely publishing data generated from OPRE research projects is important to ACF, the field, and the public. It helps ensure the data are useful beyond their initial collection purposes by contributing to secondary analysis and supporting future research (Borgerud and Borglund 2020).

This brief highlights current practices throughout the contract life cycle for publishing research data gathered and prepared as part of a federal contract. It provides useful information more broadly related to making data available for reuse. The focus of the brief is to identify the stages of data archiving and capture lessons learned from OPRE-contracted projects, but the information likely applies to other research contracts that have a data archiving component.

To inform the brief, Mathematica completed an environmental scan and interviewed key informants at OPRE, other federal agencies, and various federal contractors (see Appendix B) to learn what staff would find useful to improve their work and common challenges encountered while archiving data from research studies. Staff at the Inter-University Consortium for Political and Social Research (ICPSR) and the Harvard Dataverse Repository—two archives that house federal research data—discussed recommended practices for researchers seeking to publish data files. The interviews provided rich information that built upon what we learned in the environmental scan and specified for the context of publishing data from a federal contract. OPRE seeks to understand what OPRE, its contractors, and data repositories have learned about making research data available for secondary use to improve these practices moving forward.

Key terms for this brief

- **Data:** Quantitative or qualitative information collected for research purposes.
- **Data archiving:** Preservation of data for secondary analysis. Publishing archived data involves decisions about the type of access and level of security required for archived data.
- **Metadata:** Information about data. In this brief, metadata are defined as “basic descriptive content for variables, files, source material, and study level information” (Data Documentation Initiative 2015).

Key principles of publishing data for reuse

Transparency, usability, and accessibility are three key principles for data archiving highlighted in the literature. These principles closely relate to the FAIR (findable, accessible, interoperable, and reusable) data principles (Wilkinson et al. 2016). The FAIR principles were developed as best practice guidance for publishing data and serve as guideposts for researchers when considering optimal reuse of their data.

Transparency refers to the disclosure of processes and documentation for the purposes of improving data usability and ensuring ethical data collection. “Transparency is at the heart of accountability” and it is important to “make every effort to inform the public about our work by making our data public and accessible” (Millennium Challenge Corporation 2020). This principle, also included in ACF’s Evaluation Policy, includes two important elements:

- **Publishing detailed documentation** of methods used for data collection, analysis, and archiving
- **Outlining any restrictions** that might exist for accessing and using the data

Usability refers to the degree to which data are reusable for secondary analysis after primary research is completed. This principle includes three elements:

- **Following the data archiving entities’ recommendations** for data format and documentation and metadata (descriptions of your data and its usability) (Kapiszewski and Karcher 2020)
- Using **common data standards**, such as controlled vocabularies for metadata, when possible, to improve the feasibility of integration with other data sets (Columbia University n.d.)
- Using **a logical file structure** written for a nontechnical reader and easy for research staff to replicate (Duke University 2020)

Accessibility refers to the degree to which people of diverse backgrounds can find and understand the data. This principle includes four elements:

- Publishing data **in a location that can be found easily** by interested parties and is available to them (Neild et al. 2022)
- **Creating metadata and a data dictionary** that explains variables and making it highly readable for a nontechnical audience (Kapiszewski and Karcher 2020) and machines (Columbia University n.d.)
- **Supplying well-structured data and codebooks** to ensure others can analyze the data and replicate findings (Neild et al. 2022)¹
- **Sharing user guides** that maximize usability for secondary analysis (ICPSR 2020)

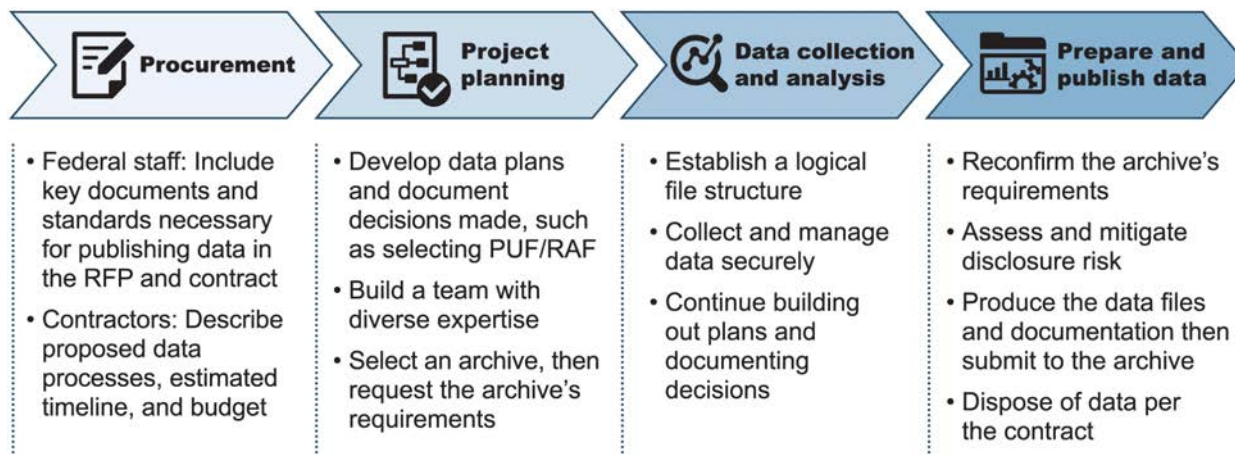
The principles of transparency, usability, and accessibility inform the processes and practices for archiving. They provide high-level guidance for teams to consider when publishing data. The examples of these principles closely reflect the recommendations summarized in this brief.

¹ Neild et al. (2022) recommended that teams share analytic data files at a minimum and potentially include raw data files, which could be a subset of the data used in the analysis. Files can include identified or deidentified data depending on the sensitivity of the data and level of restriction to access the files.

Four stages of publishing data for reuse

This brief is structured around three stages of the project life cycle pertaining to publishing data identified in the literature. Federally contracted work includes an additional step of planning during the procurement stage. Combined, this leads to four stages to consider when publishing data for reuse: (1) procurement, (2) project planning, (3) data collection and analysis, and (4) prepare and publish data. We describe the important considerations for each of these stages below, drawing on findings from interviews and resources reviewed during the environmental scan.

Exhibit 1. Stages of publishing data



PUF = public-use file; RAF = restricted-access file; RFP = request for proposal.

Key informant interviews with OPRE project officers, federal contractor staff, and data repositories underscored that efficient and effective archiving requires early planning, beginning at project conception when developing the project's request for proposal (RFP).

Stage 1: Procurement

In the first stage, federal agencies release an RFP outlining the study and contracting requirements. It is best practice to include the data archiving and data security requirements of the federal agency releasing the opportunity. The RFP requirements and accompanying list of deliverables, when provided, will become the template for the schedule of deliverables in the contract. Contracted research teams (referred to as contractors in this brief) respond to the RFP with a proposal. In the proposal, the contractor will describe a data management process and provide an accompanying budget to conduct the work. OPRE provides its staff with standardized contract language for archiving activities. The Office routinely updates the language to incorporate lessons learned from previous archiving efforts and any updates to ACF's privacy and security requirements.

Documents to consider including in a contract's schedule of deliverables

- A **data management plan** (DMP) that documents how data are collected, stored, analyzed, protected, and disposed; this plan will include procedures, timelines, and key roles and responsibilities (more detail in Stage 2)
- **Consent forms and data use agreements** that include clauses on the rights to publish data (potentially deidentified) in an archive external to the project
- **Data files** that will be published with both metadata and codebooks (more detail in Stage 3)
- **Data security requirements** and documentation, considering both contractual requirements, such as a data security plan, and agency requirements, such as completing a privacy impact assessment and obtaining an Authority to Operate (ATO)

In addition to required deliverables, the RFP might include provisions that address questions about publishing processes and intended future use, such as secondary analysis. The RFP might also specify whether the contractor should propose a plan that addresses these questions. Another question to address as part of the proposal or at the outset of a study is whether the research data should be published in public-use files (PUFs) or restricted-access files (RAFs) (definitions follow). Privacy, audience, and available resources are critical considerations when deciding whether to produce one type of file or both. Privacy protections will dictate the level of restriction for the published data. Various privacy requirements may apply to federal research projects; project staff should consult their agency's privacy team to identify requirements and how to archive the data in a permitted manner.

Two types of published data files

- **PUFs:** Data that are downloadable and available to the public (for example, files housed on Data.gov). Researchers often use PUFs to publish data with low to no risk of identifying a participant. The files are easier for secondary users to access. PUFs often require more resources to produce due to their increased level of data masking, which can limit the ability of secondary users to reproduce the original research exactly.
- **RAFs:** Data that are accessible for an approved reason, such as secondary analysis, with security restrictions. This is better suited for data that might contain personally identifiable information (PII) or other sensitive information that should not be accessible to everyone. Although data in RAFs can contain either identified or deidentified data, any deidentification procedures are typically less extensive than for PUFs, so RAFs tend to be less expensive and quicker to prepare. Researchers can also typically use RAFs to reproduce the original research exactly. Because RAFs can contain identified and deidentified data, access will likely be more restricted and involve special requirements to protect the privacy, confidentiality, and security of these sensitive data, such as secure workspaces.

Stage 2: Project planning by the contractor in consultation with the federal agency

After contract award, OPRE staff and contractors work to further specify plans for publishing data, and contractors take concrete steps preparing to conduct the upcoming work. Key informants noted that publishing data is often most efficient when archiving plans begin alongside planning for data collection, analysis, and dissemination; decisions made early in the project will influence future activities. For example, the terms specified in consent forms, data use agreements, and memorandums of understanding developed at the beginning of data collection will influence how data can be published toward the end of a study.

For OPRE projects, data planning is iterative and developed in consultation with the contracting officer's representative (COR) and federal project staff. Staff will make key project decisions in consultation with CORs. For example, federal staff will review consent forms alongside other materials associated with Paperwork Reduction Act clearance and help select how and where to publish data. It is important that contractors inform CORs of activities throughout the project and discuss how decisions might affect publishing data.

Plans for publishing data are most efficient when they anticipate how to use the collected data to inform the project's deliverables and future research, according to literature and key informants interviewed for this brief. The CORs and contractors will consider who, when, and how people will use the project's published data, as well as questions on equity, described in more detail later. This will inform other decisions, such as creating a DMP, planning for privacy, building a team, and selecting an archive, discussed later in Stage 2.

- **Who:** Deciding who will want or need access to the data will help guide decision making. Considering why the public or researchers might be interested in your data can help define the audience and what data to publish. For example, will data be used for secondary analysis? Will researchers want raw data in addition to analytic files so they can replicate the analysis? Defining who will need the data will help contractors think about how it will be used, which might affect decisions about choosing an archive, data elements to include in the published files, and documentation.
- **When:** Key informants from federal contractors, OPRE CORs, and archives recommended that project schedules include time after the data are analyzed to prepare data files and accompanying documentation. It is important to know when the data will be released and to plan backwards from that date, so the contractors have dedicated time to answer the archive's questions and requests. Archives' timelines to publish the data will vary by facility and level of access restrictions. For federal agencies, it is also helpful for CORs to consider if any planned subsequent studies might require use of the published data when reviewing the contractor's proposed timeline.
- **How:** The amount of sensitive and identifiable information and any restrictions for use will affect how others can access published data. When constructing data files for publication, contractors and CORs consider the need to mitigate disclosure risk by deidentifying data against the potential usability of data sets that include identifiers. Sensitive and identifiable data require more stringent access restrictions. Contractors and CORs will decide whether to publish data in PUFs, RAFs, or both, considering questions such as: Will researchers who want to analyze the data be able to get it? Will data be restricted and accessible only at certain locations? What type of files will be available, and will they stand on their own without experience working on the original research study? Consider future users when creating documentation and files, ensuring they are clear without additional context.
- **Equity considerations:** Will the participating community want and reasonably be able to access their own data? In addition, what are the risks to participants' privacy and what can be done to mitigate those risks?



Creating a data management plan

Project requirements often stipulate that contractors prepare a DMP that includes information on publishing data in an archive.

A DMP documents how the contractor will securely collect, store, analyze, publish and dispose of the data. DMPs are living documents that project teams update as they make decisions and data collection evolves throughout the project. The sidebar on this page describes the essential components of a DMP, which will also serve as the blueprint for archiving activities.

The plan requires documentation to ensure transparency, usability, and accessibility. This might include consent forms, security procedures, metadata, codebooks, data user guides, and any additional documentation required by the RFP. Not all documentation for the DMP will be available at the start of a project. However, as work progresses and decisions are made, contractors can update DMPs to include more of these components.

Data management plan

A DMP often includes several key items:

- Details about the types of data the research team will collect and analyze
- Documentation on data decisions and procedures to track that all applicable data policies are followed
- Consent procedures used in the study
- The responsibilities of individuals on the project team—this holds the team accountable to data integrity standards, which are key in preparing data for the archive
- Disclosure control methodology

Include documentation for data decisions in DMPs and follow standard templates to ensure consistent organization. Project decisions might include inclusionary criteria, variable naming protocols, data cleaning mechanisms, procedures for administering data collection, and additional background on the project (Neild et al. 2022).



Planning for privacy

During the planning stage, contractors ensure they follow contractual policies to protect PII. PII is data that can identify an individual or determine their identity when used alone or in combination with other data (Johnson and Grandison 2007). Contractors think about what PII is necessary to collect, how to store it, and what linkages might be possible. It is best practice for the COR and contractor to discuss whether to store PII in the data files sent back to the federal agency. Retaining PII can help the agency link the project's data to other data records and conduct follow-up studies, but linkages also raise privacy issues and the need for strategies to mitigate risks and enhance data stewardship (Government Accountability Office 2001). Consider the possibility of future linkage to data that might not be available (or exist) at the time of the original study.² These discussions will inform data storage throughout the project and during archiving.



Building a team

A qualified and diverse team is essential to produce and publish data files securely. Throughout the project, data activities will require expertise from project leaders, data collection staff, programmers, and legal and security personnel, among others. Many of these individuals will be critical for producing the data files and documentation to publish research data.

Some projects might also benefit from having community advisory groups, which might weigh in on data collection and publishing activities. An example of engaging the community at this stage is requesting a review of data collection instruments and consent language for clarity. The project could engage with people with lived experience or program participants for this review.



Selecting an archive

Some federal agencies publish and house their own data and others publish data with a third party. Third parties might include contractors, other agencies, and archives (sometimes referred to as repositories). OPRE often publishes research data with archives, many of which support project teams with preparing data files, publicizing the data's availability, and managing users' access. In consultation with CORs, contractors might reach out to multiple archives to decide where to publish their data. Considerations for selecting an archive include cost, timeline, established contracts with a specific archive, and data restriction options. Exhibit 2 contains examples of archives that house federal data, but this list is not exhaustive.

Literature suggests that project teams contact the data archiving entity as soon as possible (ICPSR 2020). Archiving entities might have requirements or recommendations for standards that inform how to complete the work. For example, some archiving entities provide a metadata standard or data documentation specifications that the plan should include or implement as soon as possible during data collection. It is also useful to contact the archive to understand the timeline, collect templates for data documentation, and confirm best practices.

² See this [OPRE guide](#) for more information on assessing the feasibility of linking existing data sets to examine long-term outcomes.

Exhibit 2. Sample of potential third-party data archive options for policy researchers

Archive	Description
Dataverse 	<p>Dataverse is an open-source platform for publishing data developed at Harvard University's Institute for Quantitative Social Science. There are more than 340,000 data sets stored in Dataverses worldwide.</p>
Federal Statistical Research Data Centers (FSRDCs) 	<p>These are partnerships between federal statistical agencies and research institutions. FSRDCs provide secure environments to support researchers using restricted-access data.</p>
Inter-University Consortium for Political and Social Research (ICPSR) 	<p>ICPSR archives research in the social and behavioral sciences and is a frequent option for OPRE projects. The archive offers Open ICPSR, an unrestricted access archive that does not require the same level of curation and cleaning as ICPSR.</p>
National Archive of Criminal Justice Data (NACJD) 	<p>This archive houses data on crime and justice. The Bureau of Justice Statistics, the National Institute of Justice, and the Office of Juvenile Justice and Delinquency Prevention of the Department of Justice sponsor the NACJD.</p>
National Data Archive on Child Abuse and Neglect (NDACAN) 	<p>The ACF Children's Bureau funds this archive, which includes data sets relevant to the study of child abuse and neglect. Data sets include data files, complete documentation, and technical support from the NDACAN staff.</p>
National Institute of Mental Health Data Archive 	<p>This archive facilitates data sharing across mental health and other research communities. It provides curated human subjects research and longitudinal storage of a research participant's information generated by one or more research studies.</p>
Substance Abuse and Mental Health Services Administration (SAMHSA) Data Archive 	<p>The Center for Behavioral Health Statistics and Quality (CBHSQ) maintains this archive. CBHSQ has primary responsibility for collecting, analyzing, and disseminating SAMHSA's mental health data.</p>

Stage 3: Data collection and analysis

OPRE project staff, contractors, and data repositories all noted that considering data publishing standards during data collection helps set up researchers for success. Developing clear documentation, using common file formats, and other activities make preparing to publish data more efficient.

During this stage, key informants recommend contractors continue to think about how their data will be reused. When future users download the data, they will have to search through archived files to locate the relevant information they seek. If data are hard to find or are in one large file, it could pose a burden for future users when extracting the information they need. To avoid this situation, it is important to establish an organizational structure for data files as data are collected. Logical structures will make it easier for future users to navigate the published data. A logical file organization system will consider the number of files, their content, linkages, and type of file (Duke University 2020). The sidebar provides examples. Ensure the structure and labeling make sense for both the project team and future users.

Future users might encounter challenges, such as software compatibility, when attempting to analyze data. Publishing data in formats that are accessible to others, such as American Standard Code for Information Interchange (ASCII) delimited, Extensible Markup Language (XML), or Microsoft Excel can prevent this issue.

Including metadata in the published files will help make data readable for the current research team and future users. It is often helpful for future researchers to have access to the maximum amount of information, including documents such as survey questionnaires, syntax and code, and user guides. Ideally, both raw and analytical data are available to allow for re-creating the analysis.

One resource available for contractors is the [Data Documentation Initiative](#), which provides a set of XML rules specifically for describing social, behavioral, and economic data. These standards ensure consistent data and metadata structures across research projects. Consistency in these structures can help make data easier to use in the future.

Stage 4: Prepare and publish data

Preparing data and documentation. Before the end of the contract, key informants from data archives recommended contractors contact the selected archive to discuss publishing data in greater detail. Even if contractors contacted the archive early in the project, it is possible that requirements for submission changed. It is helpful to ensure the project team has a clear understanding of what archives need and how to successfully submit files with desired documentation the first time. Contractors must develop accessible codebooks, a data dictionary, and metadata for future users to understand the published data. Without clear information, archivists preparing the data for publishing might need to clarify and request additional information from the project team, which can cause publishing delays. When using a third-party archive with curation support, archival experts might be the first external users of the data, so they will need the same context and information as future users.

Examples of organizational structures

Files by entity type

- Person level
- Household level
- Organization level (for example, grantee or school)
- Geographic (for example, county, state, or region)

Files by data source

- Baseline survey
- 12-month follow-up survey
- 36-month follow-up survey
- Service receipt data
- Employment records
- Program staff interviews
- Participant interviews

Accessible data formats. According to ICPSR (2020), there are two common approaches for preparing files that are accessible to a broad range of users:

1. Providing data in raw ASCII format, along with set-up files to read them into standard statistical programs
2. Providing the data as a software-specific system file (such as a Stata data file) or in a portable file produced by a statistical program (such as an SPSS portable file or SAS transfer file) designed to support data transfer between different systems

Privacy assessment, disclosure risk mitigation, and submission. Depending on the archive selected, a data curation team might be available for support to assess disclosure risk or highlight possible concerns. Although all data publishing involves some risk, contractors can take several steps to avoid pitfalls. Neild et al. (2022) outlined a plan for assessing and mitigating disclosure risk. After locating potential direct identifiers (such as name) and indirect identifiers (features that might identify an individual when combined with other information), contractors can take steps to protect individuals' privacy. For example, deidentifying files to reflect greater levels of privacy might include assigning new identifiers, deleting or editing variables, or recoding data. Contractors have to assess trade-offs between privacy and utility of the data. Contractors and CORs will work together to weigh trade-offs and priorities and ensure the strategies the project adopts conform with the agency's privacy and risk assessment requirements.

Neild et al. (2022) recommended that the privacy assessment process determine which files could pose an identification risk if linked to data available outside of the archived project. If disclosure continues to be a potential risk after implementing technical solutions, it might be necessary to protect data through limiting access by selecting a higher level of restriction from among archiving options. When the contractor addresses concerns about identification, key informants from data archives recommend that contractors follow the archive's procedures to submit data files to the archive securely.

Disposing of data. At the end of a project, data management plans will outline the data disposition steps to destroy original data per contract requirements. Contractors might have specific obligations in their contracts for keeping or destroying information. Before executing this step, it is recommended that contractors ensure the data have been submitted to and received by the federal agency, if required by the contract, as well as published in the archive. As with safe storage and secure submission, data disposition is a critical part of protecting the privacy of research subjects and honoring the guiding principles of ethical data use.

Conclusion

Publishing data is complex and requires planning throughout the contract life cycle to ensure contractors meet requirements to release accurate and usable data that follow privacy standards. Ample communication and proper documentation of decisions is vital. Key informants from archives and CORs suggest communication include federal staff, contractors, and data archive representatives. Employing data archiving standards to publish research data will ultimately make these data more accessible to the public and research community.

References

- Borgerud, Charlotte, and Erik Borglund. "Open Research Data, An Archival Challenge?" *Archival Science*, vol. 20, no. 3, 2020, pp. 279–302.
- Columbia University Irving Medical Center Augustus C. Long Health Sciences Library. "What are the FAIR Data Principles?" New York: Columbia University, n.d. <https://library.cumc.columbia.edu/insight/what-are-fair-data-principles>.
- Data Documentation Initiative. "Document, Discover and Interoperate." 2023. <https://ddialliance.org>.
- Data Documentation Initiative. "DDI Codebook 2.5." 2014. <https://ddialliance.org/Specification/DDI-Codebook/2.5/>.
- Duke University Office of Scientific Integrity. "Duke Data Management Plan Guidance Document." Durham, NC: Duke University, January 27, 2020. <https://dosi.duke.edu/advancing-scientific-integrity-services-and-training/accountability-research/data-management-plan>.
- Government Accountability Office. "Record Linkage and Privacy—Issues in Creating New Federal Research and Statistical Information." GAO Publication No. 01-126SP. Washington, DC: U.S. Government Printing Office, 2001. <https://www.gao.gov/assets/gao-01-126sp.pdf>.
- Inter-University Consortium for Political and Social Research. "Guide to Social Science Data Preparation and Archiving Best Practice Throughout the Data Life Cycle." Ann Arbor, MI: University of Michigan, 2020.
- Johnson, C.M., and T.W.A. Grandison. "Compliance with Data Protection Laws Using Hippocratic Database Active Enforcement and Auditing." *IBM Systems Journal*, vol. 46, no. 2, 2007, pp. 255–264.
- Kapiszewski, Diana, and Sebastian Karcher. "Making Research Data Accessible." In *The Production of Knowledge: Enhancing Progress in Social Science*, edited by Colin Elman, James Mahoney, and John Gerring. Cambridge, UK: Cambridge University Press, 2020. <https://doi.org/10.1017/9781108762519.008>.
- Millennium Challenge Corporation. "Open Government." n.d. <https://www.mcc.gov/initiatives/initiative/open>.
- Millennium Challenge Corporation. "MCC Guidelines for Transparent, Reproducible, and Ethical Data and Documentation (TREDD)." March 6, 2020. <https://www.mcc.gov/resources/pub-pdf/guidance-mcc-guidelines-tredd>.
- Neild, R.C., D. Robinson, and J. Agufa. "Sharing Study Data: A Guide for Education Researchers." NCEE 2022-004. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2022. <https://ies.ed.gov/ncee/pubs/2022004/pdf/2022004.pdf>.
- Wilkinson, M., M. Dumontier, I. Aalbersberg, et al. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data*, vol. 3, article no. 160018, 2016. <https://doi.org/10.1038/sdata.2016.18>.

Appendix A: Guiding Questions

Six learning questions guided the data archiving learning agenda for the Office of Planning, Research, and Evaluation (OPRE). This brief is a component of the learning agenda and informs OPRE's efforts to develop a systematic data archiving approach.

1. How does the Administration for Children and Families (ACF) decide to archive research study data?
2. How does ACF decide to archive administrative data?
3. What has ACF learned from the archiving process?
4. What can OPRE learn from others?
5. Does a single conceptual framework apply to archiving research study data and archiving administrative data?
6. What are open questions about archiving and opportunities for future learning?

Appendix B: Organizations that Participated in Data Collection

Exhibit B.1 lists the organizations of the staff we interviewed from federal agencies, archives and repositories, and federal contractors. The staff discussed recommended practices and difficulties encountered while archiving data from research studies.

Exhibit B.1. List of organizations of staff who participated in interviews

Organizations
Group 1. Federal agencies
Administration for Children and Families' Office of Planning, Research, and Evaluation
U.S. Department of Agriculture
U.S. Department of Education
Millennium Challenge Corporation
Group 2. Data archives and repositories
Harvard Dataverse Repository
Inter-University Consortium for Political and Social Research
Group 3. Contractors
Abt Associates
MDRC
Mathematica

Disclaimers: This brief is not official legal or regulatory guidance, and to the extent there is any conflict between this brief and governing contracts, regulations, or laws, those contracts, regulations, or laws take precedence. The views expressed in this publication do not necessarily reflect the views or policies of the Office of Planning, Research, and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services.

August 2023

OPRE brief: OPRE Report #2023-227

Authors: Heather Gordon, Caroline O'Callahan, Katie Hunter, Hande Inanc, Mary Grider, and Heather Zaveri

Contributors: The authors are grateful to the following OPRE staff who contributed time, expertise, and program materials to support the development of this brief: Brett Brown, Valeria Butler, Wendy DeCoursey, Christine Fortunato, Aaron Goldstein, Lauren Griffin, Daniella Lowenberg, Nancy Margie, and Girley Wright. We also would like to thank the OPRE project officers and staff at research firms, data archives and repositories, and other federal agencies who shared their experiences with data archiving.

Project officer: Nicole Deterding, Office of Planning, Research, and Evaluation

Project director: Heather Zaveri
Mathematica
1100 First Street, NE, 12th Floor
Washington, DC 20002-4221

Contract number: HHSP2332015000351/75P00120F37052

Suggested citation: Gordon, Heather, Caroline O'Callahan, Katie Hunter, Hande Inanc, Mary Grider, and Heather Zaveri. "Guide to Publishing Research Data for Secondary Analysis." OPRE Report #2023-227. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2023.

This report and other reports sponsored by the Office of Planning, Research, and Evaluation are available at <https://www.acf.hhs.gov/opre>.

Connect with OPRE

