

TEACHING RESEARCH DATA MANAGEMENT

Edited by
JULIA BAUDER

JULIA BAUDER is the social studies and data services librarian at the Grinnell College Libraries. She is the author of the book *The Reference Guide to Data Sources* and the editor of *Data Literacy in Academic Libraries: Teaching Critical Thinking with Numbers*. She has also published and presented about information literacy, data literacy, and data visualization in venues including *Information Technology and Libraries*, *College & Undergraduate Libraries*, and the LITA National Forum.

© 2022 by the American Library Association

Extensive effort has gone into ensuring the reliability of the information in this book; however, the publisher makes no warranty, express or implied, with respect to the material contained herein.

ISBN: 978-0-8389-3797-6 (paper)

Library of Congress Cataloging-in-Publication Data

Names: Bauder, Julia, editor.

Title: Teaching research data management / edited by Julia Bauder.

Description: Chicago : ALA Editions, 2022. | Includes bibliographical references and index. |

Summary: "This collection gathers practitioners from a broad range of academic libraries to describe their services and instruction around research data. You will learn about such topics as integrating research data management into information literacy instruction; threshold concepts for novice learners of data management; designing a data management workshop series; and key competencies that are entry points for library-faculty collaboration in data instruction" —Provided by publisher.

Identifiers: LCCN 2021043154 | ISBN 9780838937976 (paperback)

Subjects: LCSH: Electronic information resource literacy—Study and teaching (Higher) | Electronic information resource literacy—Study and teaching (Higher)—Case studies. | Information literacy—Study and teaching (Higher) | Information literacy—Study and teaching (Higher)—Case studies.

Classification: LCC ZA4065 .T43 2022 | DDC 028.7071/1—dc23/eng/20211020

LC record available at <https://lccn.loc.gov/2021043154>

Composition design by Alejandra Diaz in the Utopia Std, Bebas Neue Pro, and Gotham typefaces. Cover design by Kimberly Hudgins. Cover image © max_776/Adobe Stock.

☺ This paper meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).

Printed in the United States of America

26 25 24 23 22 5 4 3 2 1



CONTENTS

INTRODUCTION by Julia Bauder vii

PART I Teaching Research Data Management in the Classroom

- 1 Pieces of the Whole** 3
Using the Research Process to Integrate Data Management
and Information Literacy Skills
Patricia B. Condon, Megan Bresnahan, and Eugenia Opuda
- 2 Research Data Management Needs of Music
Graduate Students** 27
Stacy Winchester and Ana Dubnjakovic
- 3 Research Data Management Instruction
in Undergraduate Research Programs** 39
Case Studies
Laura Palumbo

PART II Teaching Research Data Management as Data Information Literacy

- 4 Disciplinary Data Literacy** 57
Expanding on What Instructors Already Teach
Kay K. Bjornen, Megan Carlton, and Jacinta Mutambuki
- 5 Data Literacy on Demand** 73
Creating a Set of Data Literacy Modules for Remote Instruction
Ginny Boehme, Kristen Adams, and Matt Benzing

6 Leveraging the Utility of RStudio to Teach Data Management, Data Visualization, and Research Reproducibility..... 91
Briana E. Wham

PART III Teaching Research Data Management as an Aspect of Data Science

7 Building Data Science and Data Literacy Communities of Practice at Academic Libraries..... 115
Daniel Chen, Jonathan S. Briganti, and Anne M. Brown

8 Complementary Scales for Learning..... 133
The Feedback Loop between Short-Form Technical Workshops and Long-Form Carpentries Workshops
Karl Benedict and Jonathan Wheeler

PART IV Outreach and Capacity-Building for Research Data Management Instruction

9 Creating Capacity for Research Data Services at Regional Universities..... 155
A Case Study
Omer Farooq, Jason A. Heppler, and Kate M. Ehrig-Page

10 Expanding Research Data Management to UC Berkeley Researchers..... 177
A Targeted Approach to Outreach and Instruction
Anna Sackmann, Amy Neeser, Samantha Teplitzky, Ann Glusker, and Elliott Smith

11 Teaching Data Management Concepts to Research Administrators..... 193
Nick Ruhs, Jesse Klein, and Renaine Julian

ABOUT THE CONTRIBUTORS 205

INDEX 213



INTRODUCTION

by Julia Bauder

At first glance, teaching research data management (RDM) may seem to be an easier job for librarians than teaching information literacy. Historically, the typically envisioned audience for RDM instruction has been faculty, postdoctoral researchers, or advanced graduate students in science, technology, engineering, and medicine (STEM) fields: people who generally are very comfortable with the concepts and mechanics of data and scientific research and who, as a rule, just need some help figuring out the best procedures and technological platforms for organizing, preserving, and sharing the data that they generate. This is not to say that RDM is ever easy—creating a plan to manage gigabytes of heterogeneous data across an array of formats, confidentiality levels, and collaborators can present quite a challenge—but it does limit the scope of the instruction that librarians are asked to provide.

However, as the contributors to this volume show, teaching RDM has expanded far beyond this initial vision. As access to data has increased, the tools for working with data have become more user-friendly, and data literacy has become a sought-after skill in the workforce. Data—and the need to manage it—has spread to all levels and disciplines within academia. As audiences for RDM instruction have grown to include doctoral students in music, the undergraduate staff of a data consulting service, and grants compliance officers, to take three examples that appear in this volume, librarians who are called upon to teach RDM have found themselves working with people who are novices in some of the concepts that fall under research and/or data as well as with the management of research data. Many of the chapters in this volume, but particularly those in the section “Teaching Research Data Management in the Classroom,” describe techniques for teaching RDM to an audience of students who are simultaneously learning how to do research and how to work with data.

THE BROAD VIEW OF RESEARCH DATA MANAGEMENT: DATA INFORMATION LITERACY

RDM, as a set of skills, falls under the larger umbrella of data information literacy. This phrase, popularized by Jake Carlson, Michael Fosmire, Chris Miller, and Megan R. Sapp Nelson in an influential report that was later incorporated into the 2015 book *Data Information Literacy: Librarians, Data, and the Education of a New Generation of Researchers*, encompasses a range of data-related competencies needed by researchers in a variety of technology-heavy fields. These fields include not only the sciences, technology, and medicine, but also digital humanities, computational social science, and other areas where cutting-edge, computationally intensive methods of data collection, analysis, and visualization come into play.

According to Carlson and co-authors, the following are the core competencies that fall under data information literacy:

- Introduction to databases and data formats
- Discovery and acquisition of data
- Data management and organization
- Data conversion and interoperability
- Quality assurance
- Metadata
- Data curation and reuse
- Cultures of practice
- Data preservation
- Data analysis
- Data visualization
- Ethics, including citation of data¹

Although data information literacy is broader than RDM, some of these competencies, such as data management and organization, data preservation, and metadata, are clearly core competencies for managing research data as well. Others—for example, data analysis and data visualization—while being important skills for data-driven research, are not directly related to managing data. However, especially when working with novice researchers, it can be productive to teach RDM by integrating RDM instruction into data information literacy instruction. In an ideal scientific workflow, RDM is an integral

part of the research process, not, as so often happens in reality, something tacked on at the end after the research is largely complete. Teaching RDM in an integrated fashion to new researchers who are just forming their research habits can help to naturalize the idea that managing data is part of the scientific process. The chapters in the section “Teaching Research Data Management as Data Information Literacy” explore teaching RDM in this way as part of the broader suite of data information literacy competencies.

DATA SCIENCE AND RESEARCH DATA MANAGEMENT

From the perspective of those outside of librarianship, research data management skills are often seen not as a part of data information literacy but as an aspect of data science. This is particularly evident in the success and influence of The Carpentries, a volunteer-led effort to teach computational skills to researchers at all levels. Originally launched in 1998 as Software Carpentry, a course in which programming experts taught scientists how to write better code for their research, The Carpentries now also incorporates workshops on Data Carpentry and Library Carpentry.² Although the original Software Carpentry and Data Carpentry efforts originated in the scientific community rather than in libraries, The Carpentries’ philosophy, which focuses on empowering novices, sharing freely, valuing all contributions, and building an inclusive interdisciplinary community around best practices for using computing in scientific research, aligns well with the philosophy of librarianship.³ The Carpentries doesn’t emphasize the language of research data management in its lesson plans, but skills and concepts that fall under RDM are covered. Many Carpentries lessons focus on tools and workflows that help to ensure the integrity, organization, and preservation of research data, such as using Git for version control and scripting the data cleaning and analysis process to ensure that the research is reproducible. Because of these synergies, many libraries are now involved with hosting The Carpentries’ coding workshops and otherwise building The Carpentries’ community. The chapters in the section “Teaching Research Data Management as an Aspect of Data Science” discuss the experiences of two institutions that have incorporated The Carpentries and a data-science-focused perspective into their efforts to teach RDM on their campuses.

BUILDING INSTITUTIONAL CAPACITY FOR TEACHING RESEARCH DATA MANAGEMENT

Teaching RDM can be a labor-intensive service for libraries to provide. This is especially true in a research-intensive environment where many faculty and graduate students work on a broad spectrum of projects, each of which may require specialized, individualized instruction for each researcher or research group. Identifying the aspects of RDM about which researchers have the least knowledge, designing instructional programs and support services that are scalable, and creating knowledgeable advocates for good RDM practices in other administrative offices can all help to ease the burden on library staff who teach RDM. The chapters in the final section of this volume, “Outreach and Capacity-Building for Research Data Management Instruction,” consider strategies for making RDM instruction a feasible and scalable service for librarians to provide.

Wherever your library is with respect to teaching RDM, and whether you’re thinking about providing it as a new service and looking for ideas or your long-standing RDM instruction program has grown so much that you need solutions for making it more efficient and sustainable, I hope that this book will provide you with inspiration.

Notes

1. Jake Carlson, Michael Fosmire, C. C. Miller, and Megan Sapp Nelson, “Determining Data Information Literacy Needs: A Study of Students and Research Faculty,” in Jake Carlson and Lisa R. Johnston, eds., *Data Information Literacy: Librarians, Data, and the Education of a New Generation of Researchers* (Lafayette, IN: Purdue University Press, 2015), 29–30.
2. Greg Wilson, “Software Carpentry: Lessons Learned,” *F1000Research* 3 (2014), www.ncbi.nlm.nih.gov/pmc/articles/PMC3976103; Tracy Teal and Chris Erdmann, “Library Carpentry Is Now Officially a Lesson Program!” The Carpentries Blog, November 2, 2018, <https://carpentries.org/blog/2018/11/welcoming-library-carpentry>.
3. “About Us,” The Carpentries, <https://carpentries.org/about>; “The Nine Core Values of The Carpentries,” The Carpentries, <https://carpentries.org/values>.

PART I

TEACHING RESEARCH DATA MANAGEMENT IN THE CLASSROOM

R - 35

alastore.ala.org

PIECES OF THE WHOLE

Using the Research Process to Integrate Data Management and Information Literacy Skills

Patricia B. Condon, Megan Bresnahan, and Eugenia Opuda

University of New Hampshire

The research process is naturally embraced as part of the academic curriculum in higher education. Graduate students write theses and dissertations based on their original scholarships, undergraduate students produce papers for courses and work in labs or in the field, and both participate in faculty-led research projects. The research process is tackled holistically through coursework, yet when library instructors are invited to teach students about information literacy and research data management topics, these may be presented as tangential to or mistimed with other course content and learning activities.¹ In this chapter, the authors present a comprehensive, student-centered model for teaching research data management and information literacy as components embedded in the research process. The approach presented in this chapter is comprehensive in several ways:

- It merges research data management and information literacy instruction into a single session.
- It positions these library instruction components within the context of the research process.
- It aligns instruction outcomes with course objectives.

The authors describe two courses in which they have adopted this approach—an undergraduate science class and a doctoral-level nursing seminar—to highlight how it can be repurposed for different learning environments.

BACKGROUND AND LESSON PLAN OVERVIEW

Over the past several years, course instructors in many disciplines at the University of New Hampshire (UNH) have invited subject and data services librarians into their classrooms to teach often separate, one-shot sessions about research skills. While course instructors have typically been interested in making space for librarians to work with their students, these invitations have almost always been made in isolation from one another. As a result, students were exposed to siloed conversations about information literacy and research data management and were not presented a cohesive picture of how information is produced and used.

The research data services librarian and two subject librarians—specifically, the life sciences and agriculture librarian and the health and human services librarian—decided to design a more integrated approach to their work in these classrooms. In this new model, a subject librarian and the research data services librarian worked together to teach collaboratively and to integrate their respective instruction components. Instead of having separate sessions for the data services and subject librarians, this approach combined two sessions for a more cohesive and holistic library instruction session. The lesson plans for the combined sessions focused on situating library instruction within the ongoing work of the course or program and reframed the research skills presented and practiced in the sessions as part of a research process in which the students were active participants. Students were invited to shape their own research process through activities and apply newly acquired data and information literacy skills in their coursework.

There were three key components to the approach:

1. A visual representation of the research process that emphasized components of research data management and information literacy
2. Active learning activities
3. A case study

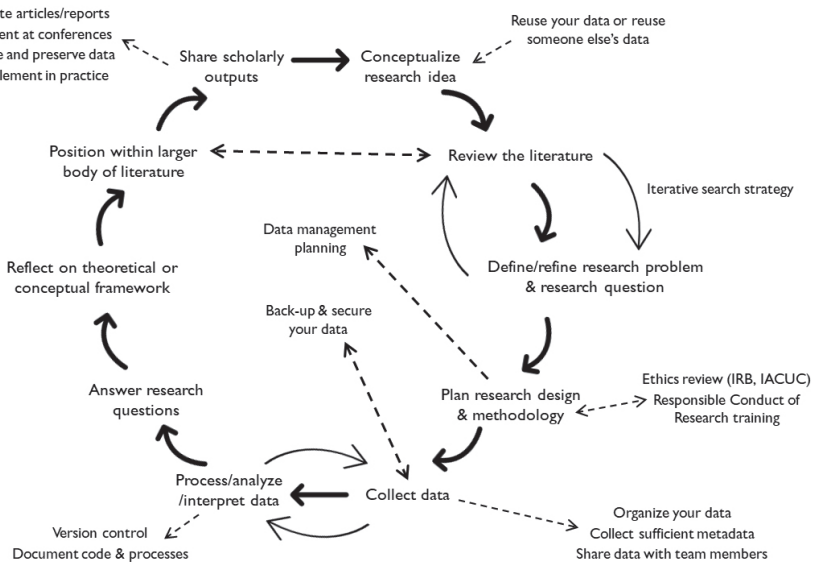
The librarians used a visual representation of the research process that depicted a detailed, iterative lifecycle model spanning from ideation to dissemination

of information (see figure 1.1). It was intended to be adaptable and applicable across multiple disciplines, and it combined elements of library research, the scientific method, scholarly inquiry, and data management. The visual representation illustrated a non-linear approach to the research process, which underscored the iterative process of research. The purpose of the visual representation was to show the many ways in which literature searching and research data management are key pieces that make up parts of the whole research process. The librarians aligned their discussion of the research process to course or curricular learning objectives. Students were encouraged to identify how their work prior to the session aligned with the model.

Active learning shapes the approach to all library instruction sessions at UNH. There is a growing body of research that suggests active learning has the potential to improve learning outcomes, critical thinking, and information retention.² In an active learning approach, the librarians “engage students to

FIGURE 1.1

Visual representation of the research process, a lifecycle model



Copyright and related rights for this image waived via CC0.

be fully involved and to participate in the learning process.”³ In the approach described here, the librarian pairs used course objectives to design case study discussions, interactive online quizzes, hands-on activities, and small group interactions to “[encourage students] to use their higher-order thinking skills (e.g., analysis, synthesis, reflection, evaluation), while [they engage] in activities that help them think critically and explore their own attitudes and values.”⁴ This strategy helped students connect with the content and engage with the session’s learning objectives. Students were also able to immediately and practically apply session concepts to their own research needs.

Case-based learning is one method of active learning in which “problems are embedded in cases to allow students to make real-world connections and apply knowledge in authentic problem situations.”⁵ This approach to instruction has a long history, with experience and research suggesting that it can be effective across disciplines for improving student engagement with and understanding of course content.⁶ The case studies that were selected for the sessions were tailored to the course. The librarians selected an appropriate narrative and developed discussion questions based on the discipline, the course assignment, and both the course learning objectives and their own session learning objectives. In both examples presented in this chapter, the librarians adapted research cases published by the New England Collaborative Data Management Curriculum.⁷ The librarian pairs used the case studies to encourage student discussion and exploration, to examine the assignment that the session supported, and to design the session theme.

COLLABORATION

When scheduling library instruction, librarians are often put in a position of negotiation. Course instructors who request one-shot library instruction sessions may have a limited understanding of the diverse educational opportunities that librarians can bring to their classrooms. To develop effective research skills in their students, course instructors may default to focusing on scholarly information access through the library, either via physical or virtual services points. Yet the contributions that librarians make to student learning extend well beyond using “library resources.” They engage students in actively finding, using, and contributing to scientific information and research. Librarians, as educators, can support student research by encouraging students to

explore and think critically about the information choices they make in each part of the research process.

Collaboration is a critical element to the success of the integrated instruction model presented here. Collaboration between librarian pairs and between librarians and course instructors strengthens the creation and delivery of course content. Engagement with a course instructor can help to integrate library content more seamlessly into course objectives and outcomes and to enhance the impact to student learning by identifying the needs of students within the course.⁸ In this case, the librarians were fortunate to have established relationships with highly collaborative and engaged course instructors who saw librarians as teaching partners; however, the authors recognize that it can be difficult to develop these types of collaborations and that “[m]any faculty do not perceive librarians as partners in teaching and learning, especially to the degree of involvement at the assignment design level.”⁹ Yet, identifying and building these collaborative relationships is the first step to integrating information literacy and research data management concepts meaningfully into a course.

In the two cases presented here, the subject librarian was the main contact with the course instructor and initially established an understanding of what the instructor needed for their course. The librarian pairs worked together to design the lesson plan, identify and create activities, and set an agenda for the session. Each librarian was then responsible for their area of expertise. The research data services librarian introduced the research process, provided the lecture on data management, and led the case study discussions. The subject librarian was responsible for literature searching, citation management, and subject-specific content. The librarian who was not facilitating or presenting provided support, helping students during literature searches, fielding questions, and posing guiding questions to the class to introduce a concept or to initiate student engagement.

IMPLEMENTATION

To implement this approach, the librarians selected courses or programs with the following characteristics: the course or program included a significant research component as part of its requirements, the course instructor was willing to explore a new approach, and the length of the session was substantial

enough to allow the learning outcomes for research data management and information literacy to be integrated. The first case took place in an undergraduate science course and included a topic mapping exercise, keyword development worksheet, and discussions on planning research projects, preparing for data collection, and managing field research data. The second case took place in a doctoral-level nursing seminar and included hands-on literature searching, interactive online quizzes, and discussions on ethics, data management challenges for clinical research, and health disparities in patient populations. In both examples, the model is now established, and the librarian pairs teach the session annually.

Undergraduate Science Course

Each fall, the life sciences and agriculture librarian and research data services librarian teach an instruction session in a research-based, undergraduate wildlife management course with an enrollment of about thirty students. The students in this course are assigned a semester-long group project to design a field research study and write a mock grant proposal to the US Fish & Wildlife Service to fund the study. The librarians teach a five-hour lab session when the student groups are beginning to plan their research projects. The session's activities use their existing project groups and provide an opportunity for groups to make progress on their research project's planning. The learning objectives include the following:

Students will:

- recognize key data management concepts (data organization, description, and storage) as applied to their field research projects
- design a research plan for a wildlife field study
- develop a research question that is right-sized and researchable
- select keywords and practice search strategies
- organize and share sources in research groups

The session begins with an overview of the research process, which emphasizes where in the process the learning objectives for the lab are positioned. This content portion also includes a high-level overview of research data management foundational practices and their positions in the research lifecycle. Students are

then introduced to the lab's first activity: a case study discussion that includes guiding questions that focus on research design and data management planning (see appendix 1.1). Students are asked to do a close reading of the case study while annotating areas of the case where they identify components of the research process or data issues. After reading the case, the research data services librarian leads a discussion. This first exercise is meant to serve as a warm-up for students and allow them to begin to explore some of the data management topics presented. The case study also provides context for the main lab activity, where groups are asked to develop a research plan and to apply some of the same concepts within a more complex and novel framework.

Next, each group is asked to convene to discuss their initial research topic idea. The life sciences and agriculture librarian provides an introduction to shaping and resizing a research question so that the question will inspire an appropriately scoped review of the literature and then asks groups to work on a topic mapping exercise. As groups create their topic maps, they are asked to brainstorm different lines of inquiry related to their topics, think about what background information and data they will need to seek out, and suggest different ways to refer to the concepts outlined in their maps. After completing the topic mapping exercise, each group shares their refined research question with the instructors and their peers for discussion. Students are asked to consider if the question is too specific or too broad or to identify different ways to frame the question as they begin to think about their study's design. Research questions tend to transform from topics such as "the movement of gray wolves" to "How do food choice behaviors impact the migratory patterns of the gray wolf in northern New England forests?"

This second exercise sets up a librarian-led discussion about preparing for a literature search, gathering keywords, and developing search strategies. Next, the librarians and students discuss strategies for literature searching and explore the process of building a search strategy using Web of Science. The life sciences and agriculture librarian asks students to help her design a search strategy for an example topic by brainstorming keywords and creating a search structure. Using the terms that the students generate, the librarian demonstrates the search in the Web of Science, discusses the results, and asks students to help revise the search strategy to retrieve a more relevant set of results. Students are encouraged to recognize their own agency in designing search strategies and that the choices they make will shape their research

outcomes. Students, as scholars, make decisions about what is examined and emphasized in the review of the literature. They are also encouraged to learn from the literature that they find and to explore new angles of their research topic, potentially using the process to refine their research questions further.

The main activity of the lab is the research plan worksheet. The groups spend time together discussing their topic map, conducting literature searches, and beginning to outline key questions that need to be addressed in the final project deliverables. The worksheet is modeled after a Kanban board, a popular project management tool, which has been customized to suit the criteria required for the course assignment, such as what research questions are being asked, what areas of the literature need to be explored, what kinds of data will need to be collected, and how will they share the data and reports. Similar boards can be found at Academic Toolkit.¹⁰ The librarians, along with the instructor, facilitate this activity, providing one-on-one support to groups as they work independently through the initial planning phase.

Finally, in the last part of the session, groups return to discuss the outcomes of their research plan with the full class. Groups share their process and expected next steps. The librarians also review the importance of citation management and sharing citations in their groups and briefly demonstrate the citation management software Zotero.

Doctoral-Level Nursing Seminar

Each semester, the health and human services librarian and research data services librarian conduct an instruction session in a doctoral-level nursing seminar that takes place during a two-day intensive residency. The residency lays the foundation for students who are beginning their doctoral program, which includes a capstone project related to quality improvement in a clinical practice environment. The librarians provide an hour-and-a-half-long orientation session to a class of ten to fifteen students. The session uses a narrative presented in the case study to provide overall context for the learning objectives. Students in the nursing program are intimately familiar with the case study approach to learning within healthcare and often use this approach in other education settings. The librarians leverage this and provide a case study ahead of the session for students to review. Learning objectives include the following:

Students will:

- understand the value of digital information management to the success of a project
- describe an approach for planning data storage, backup, and security for a project
- articulate data management concepts as a component of the responsible conduct of research and scholarly activity
- strategically search for relevant literature to answer clinical questions in biomedical resources
- recognize the importance of feasibility and availability of interventions for specific patient populations
- use Zotero to gather, organize, and cite sources

The session begins with an overview of the research process emphasizing where in the process the learning objectives for the orientation are positioned. The research data services librarian presents two versions of the research process. The first version is a linear, simplified model of the research process that describes key activities, including reviewing the literature, refining the research question, data collection, and reporting results as distinct actions. The other version (see figure 1.1) is a lifecycle model that depicts research as an iterative, interconnected process and emphasizes the often obscured underlying actions of data management and literature analysis. Then, the librarian has students describe their unique approaches to research and compare them with the presented models. This introduction helps students visually frame data management and information literacy as pieces of the research process.

Following the introduction, the health and human services librarian provides a short lecture on evidence-based practice with an emphasis on health equity and the unique needs of patients using the narrative from the case study. Students share with one another their experiences working with a diverse set of patients and meeting patient needs, particularly patients from marginalized communities. This exercise helps prepare students to critically reflect on their current practices, center the results of research and clinical practice around evidence-based practice and patient outcomes, and critically interrogate patient populations in published research. The exercise underscores the values of the health and human services librarian as an educator and the mission of the nursing program, which embraces reflective practice, healthcare as a partnership, and evidence-based nursing knowledge.¹¹

INDEX

A

Abdallah, Samer, 30
Academic Toolkit, 10
accessibility, 160
ACRL Research Data Management
Roadshow, 40
active learning approach, 5–6
administrators, teaching data management
concepts to, 193–203
Adobe products, 142–143
advocacy, 16
Akers, Katherine, 30
All European Academies, 161
Aresty Program, 44–48, 49, 52–53
Arts and Humanities Research Council (UK),
30
assessment, four levels of, 144
Assessment in Action program, 134, 144
asynchronous instruction model, 83, 88

B

Baumer, Ben, 135
Berkeley Research Computing (BRC),
180–181
Biomedical Big Data Program, 181
Bloom's taxonomy, 125, 125*fig*
buy-in, building, 15

C

Calzada Prado, Javier, 60
Canvas, 73, 82, 85
Canvas Commons, 86, 87
Cao, Longbing, 135
capacity
analyzing, 161–162
creating, for RDM services, 155–176
Carlson, Jacob (Jake), viii, 60, 77, 78, 81, 135
Carpentries, The, ix, 115, 126–127, 133–134,
135–151

case studies

on creating capacity for RDM services,
155–176
in doctoral-level nursing seminar, 10, 12
in undergraduate research programs,
39–53
in undergraduate science course, 8–9
at University of California, Berkeley,
180–191, 189*t*
case-based learning, 6
Chicago elections, 124
citation management, 10, 12
Clark, Joe C., 28, 29
Code for Her program, 99
coding and software instruction, 68
Coffee & Code workshop series, 137–145,
147–148
collaboration, 6–7
communities of practice, building, 115–131
competencies
of data literacy, 118, 119*fig*
list of core, viii
complementary scales for learning, 133–151
computational reproducibility, 93
Computational Social Science Training
Program, 181, 184
Consortium for the Advancement of
Research Methods and Analysis (CARMA),
164
copyright, 34
core competencies, list of, viii
Cornell, 76
COVID-19 pandemic, 12–14, 16, 49–50, 84,
124, 142–143, 144–145, 149, 183
Creative Commons, 34, 198
culture of practice for data, 69

D

data analysis, 32, 69, 81–82, 82*fig*, 87–88

- Data Carpentry, ix, 68, 135
- data collection
- coursework on, 32
 - preparing for, 8
 - in research model, 11
- data containers, 108
- data conversion, 68
- data curation and reuse, 81–82, *82fig*
- data discovery, 67–68
- data documentation, 34, 37
- data information literacy
- broad view of RDM and, viii–ix
 - description of, 60
- Data Initiatives Plan, 179
- data interoperability, 161–162
- data lifecycle, 45, 116–117*fig*
- data literacy
- communities of practice for, 115–131
 - computational, 60
 - core competencies of, 78, *78fig*, 81
 - on demand, 73–90
 - disciplinary, 57–72
 - statistical, 60
 - in undergraduate chemistry courses, 57–72, 62–66*t*
- data management, 68, 81–82, *82fig*, 83–84, *84fig*, 200–201
- data management plans (DMPs), 27, 34–35, 41, 42–44, 46–47, 49, 77, 91, 159–160, 165
- data management skills in higher education, 75–76
- data organization, 81–82, *82fig*, 83–84, *84fig*, 198, 200
- data preservation, 37
- data production, 76
- data quality assurance, 68
- data repositories, 46–47, 181, 191, 197–198
- data science, ix, 115–131
- “Data Sharing and Management SNAFU in 3 Short Acts,” 185
- Data Stewardship Wizard, 160
- data storage practices, 34
- data visualization, 66–67, 69, 81–82, *82fig*, 87–88, 91–111, 163
- data wrangling, 36, 94, 97, 99, 103–104
- DataBridge program, 115, 123–127
- data-driven concepts, inclusion of, 59–60
- data-intensive science, 177
- data-loss scenarios, 185
- DataONE’s Education EVALuation (EEVA) tool, 144
- datasets, 35–36
- data.table package, 103–104
- DC Capital City Bikeshare anonymized trip dataset, 100, 103
- Department of Education’s Federal TRIO Programs, 40
- Department of Energy (DOE), 194, 197
- digital badges, 79
- Digital Music Lab (DML), 30
- disciplinary data literacy, 57–72
- Division of Computing, Data Science, and Society, 181
- DMPOnline, 41, 47, 160
- DMPTool, 41, 46, 47, 160, 181
- Docker, 108, 141–142
- doctoral-level nursing seminar
- case study from, 22–25
 - integrated approach in, 10–12, 13–14, 15–16
- documenting course content, 14–15
- Doty, Jennifer, 30
- Dougan, Kirstin, 29
- Downie, J. Stephen, 30–31
- Dryad repository, 136, 181
- Duke University, 94
- E**
- Eastern University Library, 79
- end-of-life care, improving, 22–25
- Engineering & Physical Sciences Division, UC Berkeley Library, 186
- Environmental Data Initiative project, 135
- e-research, 60
- e-science, 120
- eSPEAR (Electronic Sponsored Projects Education and Resources) Certificate Series, 196
- ethical issues, 12, 58, 66–67, 69, 183–185
- evidence synthesis course, 16
- evidence-based practice, 11–12
- Excel, 36, 165
- F**
- Faculty Leadership Forum (FLF), 164
- FAIR (findable, accessible, interoperable, and reusable) practices, 118
- field research data, managing, 8

file format selection, 31, 34
 file naming conventions, 34, 198, 200–201
 findable, accessible, interoperable, and reusable (FAIR) research, 158
 Florida State University (FSU), 193–203
 focus groups, 161–162, 176
 Forbes Technology Council, 58
 formative assessment, 47–48
 Fosmire, Michael, viii
Framework for Information Literacy for Higher Education, 40, 41–42, 45, 46, 77–78, 156, 166
 Fund for Undergraduate Scholarly Experiences (FUSE), 157

G

General Refine Expression Language (GREL), 36
 ggplot2 package, 96
 Git, 94, 96, 104, 138
 GitHub, 94, 96, 104, 124, 138, 141–142
 Graduate Research and Creative Activity (GRACA), 157
 “grammar of graphics” strategy, 96
 grants compliance process, 197
 Guo, Philip J., 135

H

habitat selection, 19–21
 health disparities, 8, 12
 Holdren, J. P., 200
 holistic thinking, 16

I

identifiers, unique, 76
 IMLS-funded Assessment in Action Program, 144
 information literacy, 3–18, 29, 76–79
Information Literacy Competency Standards for Higher Education (ACRL), 77
Information Literacy Competency Standards for Science and Engineering/Technology (ACRL), 78
 information sharing, 129
 institutional capacity, building, x
 integrated instruction, model for, 3–18
 interdisciplinarity skills, 120
 iterative lifecycle model of research process, 4–5, 5*fig*, 11

J

Jackson, Amy, 31
 job preparedness, 58
 Jupyter notebooks, 138, 142

K

Kanban board, 10
 keyword development, 8, 9
 Kirkpatrick, Donald L., 144
 knit, 96, 104
 Kross, Sean, 135

L

LabArchives, 94
 Lawrence Berkeley National Laboratory, 178
 less is more, 15–16
 lesson planning, 14–15
 Library Carpentry, ix
 library subject specialists, 165
 listening sessions, 162–163
 literate programming, 96, 104
 literature searching/review, 8, 9–10, 11
 “live-coding” model, 143
 LoboCloud, 136
 LoboGit, 136
 LoboVault, 136
 Long Term Ecological Research project, 135
 Lorente, Clara Llebot, 35
 Love Data Week event, 181

M

Marwick, Ben, 93
 Marzal, Miguel Ángel, 60
 Mayer, Jennifer, 28–29
 McNair Postbaccalaureate Achievement Program, Ronald E., 40–42, 49–50
 measurement coursework, 37
 Mentimeter, 12, 13–14
 metadata
 descriptive, 31, 34, 36, 68
 digital badges and, 79
 lack of understanding of, 160–161
 Mette, Beth, 80–81
 Miami Ecological Big Data Initiative (MiEBDI), 73, 80–81, 88–89
 Miami University, 73–90
 micro-credential programs, 74, 79–81, 88
 Microsoft Excel, 36, 165
 Miller, Chris, viii

Monash University, 94
 multimodal modules, 83
 music graduate students, 27–38
 music information retrieval (MIR), 30

N

narrative approach, 158
 National Academy of Sciences, Engineering,
 and Medicine (NASEM), 58
 National Endowment for the Humanities
 (NEH), 41
 National Institute of Food and Agriculture
 (NIFA), 183
 National Institutes of Health (NIH), 41, 45,
 158, 180, 181, 183, 194, 197
 National Oceanographic and Atmospheric
 Administration (NOAA), 197
 National Science Foundation (NSF), 41, 45,
 47, 76–77, 135, 145, 183, 194, 197
 needs assessment, 158–159
 New England Collaborative Data
 Management Curriculum, 6
 New Mexico Computer Science Alliance, 146
 New Mexico Cyberinfrastructure Training
 (NMCIT) program, 145–147
 New Mexico Established Program to
 Stimulate Competitive Research (NM
 EPSCoR), 136, 137, 145–146
 NSF Center for Excellence in Science and
 Technology (NSF CREST), 136
 NYU Health Sciences Library, 185

O

OneDrive, 136
 online curriculum/instruction
 for Aresty Program, 44–48, 49, 52–53
 assessment of, 47–48
 Coffee & Code workshop series and,
 144–145, 149
 data literacy and, 73–90
 for music graduate students, 32
 perceptions of, 49–50
 “pre-flight” sessions, 145
 recommendations for, 49–50
 in undergraduate research programs, 39
 at University of New Hampshire, 6, 8,
 12–14
 open access, 158, 197–198
 open formats, 34

open research, 158
 Open Science Communities, 188, 190
 Open Science workshop series, 180
 OpenDOAR, 41
 OpenRefine, 34, 36
 Organizational Change for Gender Equity in
 STEM Academic Professions (ADVANCE)
 program, 136
 OSE, 94
 outreach, targeted approach to, 177–192

P

Pangeo project, 141–142
 Pascuzzi, Pete, 135
 Pennsylvania State University Libraries, 79,
 91–111
 pilot testers, 84–87, 85*t*, 88
 planning research projects, 8, 10, 15
 Portland State University, 79
 Pratt, Katie, 190
 “pre-flight” online sessions, 145
 preservation, 160, 166
 privacy issues, 58
 products of research, definition of, 30
 Project Dragonfly (PD), 87
*Proposal and Award Policies and Procedures
 Guide* (National Science Foundation),
 76–77
 Public Health, School of (UCB), 182–183, 184
 PubMed, 12
 Purdue University, 75, 76
 Python, 20, 137, 138, 140–143, 148, 165

Q

Qualtrics, 159, 163, 168–175
 quantitative analysis course, 27–38
 Quinn, Todd, 31

R

R Markdown, 94, 96, 104
 R programming language, 87–88, 91–94, 95–97,
 99–100, 103–104, 108, 138, 142–143, 165
 re3data.org, 36, 41, 46
 reflection, 16
 Registry of Research Data Repositories, 36
 remote instruction, data literacy and, 73–90
 Rempel, Hannah, 35
 reproducibility, 91–111, 98*t*, 101–102*t*,
 105–107*t*, 179–180, 200

- Research and Creative Activity Fair, 157
 research data lifecycle, 198
 research data management
 broad view of, viii–ix
 case studies, 19–25
 creating capacity and, 155–167
 expanding to UC Berkeley researchers,
 177–192
 information literacy and, 3–18
 introduction to, vii–x
 music graduate students and, 27–38
 teaching to research administrators,
 193–203
 in undergraduate research programs,
 39–51
 Research Data Management Policy (FSU),
 195
 Research Data Services Committee (RDSC),
 158–159
 Research Development Program, 157
 Research Experience for Undergraduates
 (REU) programs, 80–81, 84, 88–89
 research guides, development of, 164–165
 Research IT, 180–183
 research lifecycle, 4–5, *5fig*, 11
 research process, 3–18
 research projects, planning, 8, 10, 15
 research reproducibility, 91–111, *98t*,
 101–102*t*, 105–107*t*
 research workflow cycles, 186, 187*fig*
 Responsible Conduct of Research
 Certification programs, 136, 148
 Responsible Conduct of Research Program,
 180
 Responsible Conduct of Research (RCR)
 ethics requirements, 183–185
 RStudio, 91–111
 Rutgers University, 39–53
- S**
- sampling procedures, coursework on, 32
 Sapp Nelson, Megan R., viii, 135
 Saucedo, Jonathan, 29
 scientific method, 45
 search strategies, 9
 sharing requirements, 91
 SMART Grid Center, 145–146
 Social Science Methodology & Methods
 (SSM&M) Faculty Leadership Forum, 164
- Social Science Research Commons (SSRC),
 163–164, 166
 Software Carpentry, ix, 68, 135, 137, 138, 139
 Sponsored Research Administration (SRA) at
 FSU, 195–202
 SPSS, 165
 SRA Certificate Series, 196
 statistical methods, coursework on, 32–33
 STEM Ambassador Program (STEMAP), 146
 storage, 34, 159, 166, 198
 Stormes, Sheridan, 29
 student consultants, 121
 student work, involvement with, 16
 summative assessment, 47
 Summer Science Research Program, 44
 Suskie, Linda, 144
 sustainable energy technologies, 145
- T**
- teaching philosophy, articulating, 14
 team-teaching approach, 148–149
 threshold concepts, 140, 166
 topic mapping, 8, 9
 Tropy, 94
- U**
- undergraduate chemistry courses, 57–72
 Undergraduate Research Forum, Miami
 University, 74
 undergraduate science course
 case study from, 19–21
 integrated approach in, 8–10, 12–13, 15–16
 University of California, Berkeley, 177–192
 University of Illinois Urbana-Champaign, 75
 University of Minnesota, 76
 University of Nebraska at Omaha (UNO),
 155–176
 University of Nebraska Consortium of
 Libraries (UNCL), 157, 162–163
 University of New Hampshire (UNH), 3–18
 University of New Mexico (UNM), 133–151
 University of Oregon, 76
 University of Oslo, 75
 University of Sheffield, 94
 University of South Carolina, 27–38
- V**
- version control, 96, 104
 Virginia Tech, 115–131

W

Web of Science, 9
Wheeler, Jonathan, 31
White House, 77
Wilson, Greg, 135
Woodley, Lou, 190
workflow cycles, research, 186, 187*fig*

Y

Yeager, K., 29
YouTube, 29

Z

Zotero, 10, 12