

2021

Data Literacy in Undergraduate Education: Faculty Perspectives and Pedagogical Approaches

Theresa G. Burress

University of South Florida St. Petersburg, tburress@usf.edu

Emily Zoe Mann

University of South Florida St. Petersburg, ezmann@usf.edu

Susan Montgomery

Rollins College, smontgomery@rollins.edu

Rachel Walton

Rollins College, rwalton@rollins.edu

Follow this and additional works at: https://digitalcommons.usf.edu/fac_publications

 Part of the [Information Literacy Commons](#)

Recommended Citation

Burress, T., Mann, E.Z., Montgomery, S., & Walton, R. (2021). Data Literacy in Undergraduate Education: Faculty Perspectives and Pedagogical Approaches. In J. Bauder (Ed.), *Teaching Critical Thinking with Numbers: Data Literacy and the Framework for Information Literacy for Higher Education* (pp. 1-22). Chicago: Association of College & Research Libraries.

This Book Chapter is brought to you for free and open access by the USF Faculty Publications at Digital Commons @ University of South Florida. It has been accepted for inclusion in USF St. Petersburg campus Faculty Publications by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact scholarcommons@usf.edu.

DATA LITERACY IN ACADEMIC LIBRARIES

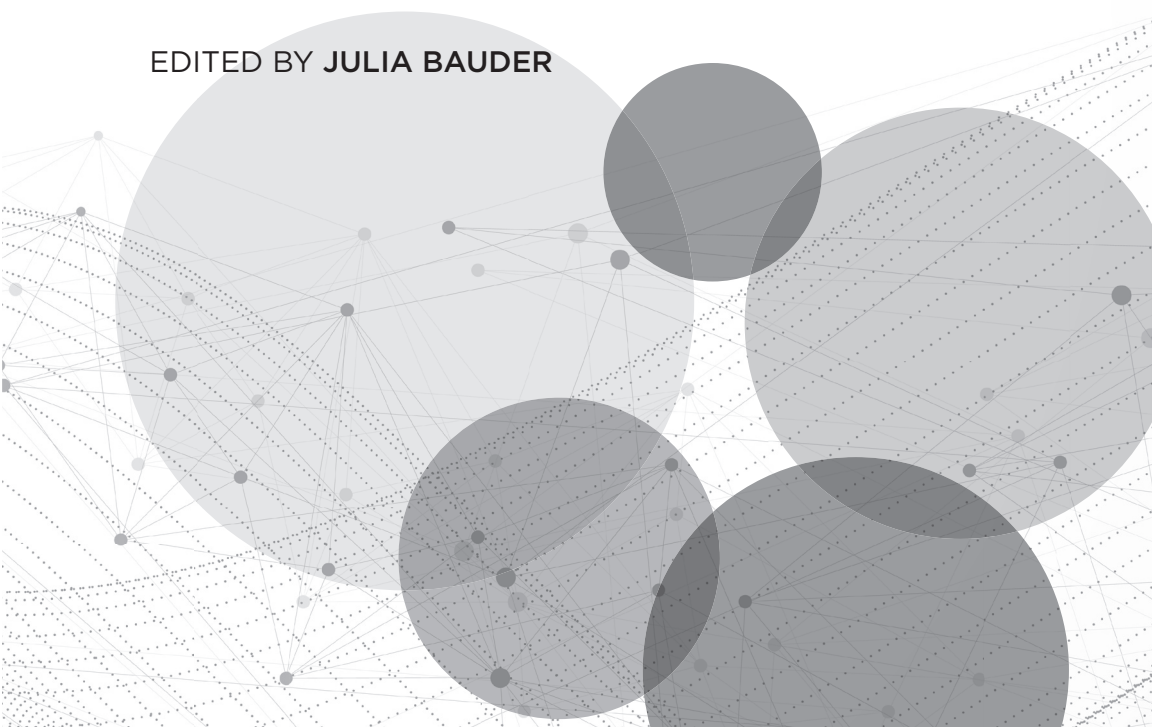
ALA Editions purchases fund advocacy, awareness, and accreditation programs for library professionals worldwide.

DATA LITERACY

IN ACADEMIC LIBRARIES

Teaching Critical Thinking
with Numbers

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 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.

© 2021 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.

ISBNs

978-0-8389-4883-5 (paper)

978-0-8389-3750-1 (PDF)

978-0-8389-3749-5 (ePub)

Library of Congress Cataloging-in-Publication Data [to come]

-
-
-
-
-
-

Composition design by Alejandra Diaz in the Utopia Std, Bebas Neue Pro, and Gotham typefaces.
Cover design by Kimberly Hudgins.

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

Printed in the United States of America

25 24 23 22 21 5 4 3 2 1



CONTENTS

INTRODUCTION vii

1	Data Literacy in Undergraduate Education 1
	Faculty Perspectives and Pedagogical Approaches Theresa Burress, Emily Z. Mann, Susan E. Montgomery, and Rachel Walton
2	Building Undergraduate Data Literacy through Faculty Development 23
	Clarence Maybee, Guang Lin, Wei Zakharov, Chao Cai, Jason FitzSimmons, and Yixuan Sun
3	Beyond the Numbers 39
	Building a Data Information Literacy Program for Undergraduate Instruction Veronica Arellano Douglas, Wenli Gao, Emma Fontenot, and Andrea Malone
4	The Barnard College Empirical Reasoning Center 59
	A Case Study in Developing a Service Model for Teaching Data Literacy across Disciplines Alisa B. Rod, David F. Weiman, Fatima Koli, Anna Carlson, and Jennifer A. Green
5	Making the Most of a Data One-Shot with Mapping 77
	Multidisciplinary Data and Spatial Literacy Kathryn M. Wissel
6	A Data Discovery Project 91
	Seeking Truth in a Post-Truth World Allison Brungard and Langdon Smith

7	More Data, Less Dewey	105
	Modernizing a For-Credit Library Course to Support Data and Algorithmic Literacy for Undergraduates Elizabeth Blackwood and Debra Hoffmann	
8	Data Literacy for Entrepreneurs	123
	Exploring the Integration of Pedagogy, Practice, and Research at MIT Nicholas Albaugh, Elizabeth Soergel, and Micah Altman	
9	Data Literacy and the Framework for Information Literacy	139
	Julia Bauder	

ABOUT THE CONTRIBUTORS 147

INDEX 153

DATA LITERACY IN UNDERGRADUATE EDUCATION

Faculty Perspectives and Pedagogical Approaches

Theresa Burress, University of South Florida

Emily Z. Mann, University of South Florida

Susan E. Montgomery, Rollins College

Rachel Walton, Rollins College

Data is becoming increasingly ingrained in the everyday information landscape of our society. Navigating this reality makes data literacy an essential skill for all educated citizens. Data can take various forms and be presented in a multitude of ways. For example, quantitative data is typically numeric, while qualitative data is often text based. Data-literate students understand the different types of data and how to find reliable data, how to consider the source of the data as well as its purpose, how to determine if the data is appropriate for a given context, and how to use data effectively for research and writing.

Responding to the growing data demands of their research communities, many academic libraries have data specialists who provide a suite of “data services,” including data repositories, curation tools, metadata creation, grant writing assistance, and data citation. However, these services focus primarily on the needs of graduate students and faculty and mainly involve aspects of data management. Fewer libraries have created data literacy education and programming specifically targeting undergraduate students.

As data literacy becomes an essential part of higher education, it *must* be taught to undergraduate students; faculty are already incorporating data-driven assignments into their curriculum and need support. Librarians are practiced in teaching information literacy skills in general education and discipline specific contexts. The Association of College and Research Library's *Framework for Information Literacy* provides librarians with a set of learning competencies for information literacy instruction.¹ The *Framework* allows librarians to design their instruction in a way that enforces certain key competencies. However, there is little direction on how librarians can support data literacy for undergraduate students and what key competencies these students need to master in order to be data literate. In fact, the definition of data literacy itself is still evolving.

This research study examines this issue further and asks the following central questions:

1. What are the main data literacy competencies that faculty expect of undergraduate students?
2. What are the issues faculty have in teaching data literacy?
3. Considering the above, how can librarians partner with faculty to facilitate the development of data literacy competencies for undergraduate students?

To begin answering these questions, instructors from a variety of disciplines were interviewed and asked about data literacy in the undergraduate classroom setting. Next, a content analysis of interview transcripts was conducted. Significant themes and instructors' direct comments point to some critical findings. Data literacy is linked directly to information literacy, regardless of discipline. There are practical, meaningful ways that librarians, as information literacy experts and teaching partners, can contribute to data literacy development for undergraduate students. Further, many faculty members welcome the opportunity to collaborate with librarians in this arena.

LITERATURE REVIEW

The majority of the professional literature discussing the academic library and the librarian's role with regard to data focus on research data management (RDM) activities and services.² Librarians at small and large institutions have

argued that they have expertise in all stages of the research process including data management activities.³

After an analysis and review of more than fifty needs assessment case studies focusing on RDM services in libraries, Goben and Griffen found that libraries were overwhelmingly prioritizing faculty and their data-related requirements for large scale, grant-driven, or other intensive research projects.⁴ No case study considered the needs of undergraduate students. Further, it was determined that best teaching practices surrounding the “finding, sharing, and reusing of data” and its “organization, documentation, and metadata” represented a major lacuna in library services and research.⁵

Some librarians have been working to address this gap and determine best practices for meeting the data literacy needs of undergraduate students, well before Goben and Griffen’s work. For example, Stephenson and Caravello designed and taught instruction sessions that incorporated both information literacy and data literacy concepts to help sociology students develop the essential skills needed for research in the field.⁶ Carlson and colleagues’ 2015 work on data information literacy recognized that new graduate students may not be prepared for e-research that requires engagement with digital datasets and the associated responsibilities.⁷ Carlson also affirmed that in addition to RDM-driven work, data-related instruction led by librarians would address an unmet need on college campuses.⁸

Meryl Brodsky made the same argument in the context of liaison librarianship for business disciplines.⁹ Likewise, a 2019 article in *Fortune* made the claim that “librarians have a big role to play in improved data literacy”—specifically, in developing the ability of consumers to understand and analyze data in our now big-data business world.¹⁰ A collaborative white paper titled *Future Themes and Forecasts for Research Libraries and Emerging Technologies* asserts that “a continuum of data analysis skills will be required in every discipline,” not only for students but also for instructors and librarians.¹¹ The report goes on to say that research libraries will be central partners to this endeavor and emphasizes that data literacy activities related to analysis and ethics “will move beyond the classroom,” helping students prepare for jobs and providing a model for lifelong learning in a changing society.¹²

There is some disagreement about whether librarians are suited to teach data literacy concepts to undergraduates. Shorish argues that data-related learning is best led by faculty subject specialists, while Burrell, Mann, and Neville facilitated a recent campus-wide collaboration to begin embedding

data literacy across disciplines, acknowledging the distinct instruction roles for both teaching faculty and librarians.¹³

Though many librarians do data-related work and many large research institutions employ data services librarians, these librarians are often seen as functional specialists. There is not yet consensus around the question of whether data literacy should be in the instruction portfolios of all liaison and/or instruction librarians. After surveying faculty who taught courses that included numeric or spatial data, Hogenboom, Phillips, and Hensley concluded that it is not enough to simply cultivate strong teaching partnerships with faculty members; libraries need to provide the proper training and support for librarians to take on leadership roles and develop a deeper understanding of how campus stakeholders are using and working with data.¹⁴

Prado and Marzal have defined data literacy as “the component of information literacy that enables individuals to access, interpret, critically assess, manage, handle and ethically use data. From that perspective, information literacy and data literacy form part of a continuum.”¹⁵ They used this definition to develop a broad framework for data literacy competencies that can be adapted to any type of library: K-12, public, or academic. More recently, Burrell, Mann, and Neville used these competencies as a starting point, working with teaching faculty across disciplines at a midsize university campus to adapt them for a higher education curriculum.¹⁶ The selected competencies were grouped into three categories: general undergraduate education, undergraduate research, and graduate education. The present study aims to test the relevance of these data literacy competencies across the spectrum of undergraduate curricula at two institutions. The authors also explore faculty needs and opportunities for librarians to support data-related instruction and student learning at the undergraduate level.

METHODS

This study takes place at two institutions of higher education in Florida: University of South Florida (USF) St. Petersburg campus and Rollins College. The USF St. Petersburg is a branch campus of the University of South Florida, serving approximately 4,500 students, with a focus on undergraduate and some master’s programs in three colleges: Arts and Sciences, Business, and Education. Information and data literacy is a core component of the USF Enhanced General Education curriculum.¹⁷ Rollins College is an independently accredited private liberal arts college with an FTE of approximately three thousand.

Rollins comprises three schools: the College of Liberal Arts; the Hamilton Holt School, which offers evening and weekend undergraduate and graduate programs; and the Crummer Graduate School of Business, which includes an MBA program and a small doctoral program. At the time of the data collection for this study in spring 2020, both institutions were independently accredited by the Southern Association of Colleges and Schools Commission on Colleges (SACSCOC), and held the Carnegie Classification as Master's Colleges and Universities: Larger programs.¹⁸

This investigation is informed primarily by interviews with college instructors discussing undergraduate teaching practices specifically relating to data literacy competencies and expectations. A small ($n = 22$) but diverse sample of teaching faculty were recruited for interviews, eleven from Rollins College and eleven from the USF St. Petersburg campus. Disciplines included social sciences, natural sciences, and applied disciplines such as business and education. The authors identified interviewees by contacting individual faculty across subject areas who taught some kind of research methods course or another aspect of data literacy as their disciplinary area of expertise. The goal was to identify faculty in as many different disciplinary areas as possible. The population of faculty interviewees comprised various genders and ranks. Interview subjects were a mix of tenure-track faculty with research requirements and full-time instructors, some of whom teach exclusively and others who teach and conduct research.

Various techniques for collecting relevant data were considered, including individual interviews, focus groups, surveys, and syllabi analysis. Stamatoplos, Neville, and Henry have raised a concern about survey methods limiting discovery in studies “where a topic is relatively new, or thinking and language about it may not be standard.”¹⁹ Because data literacy pedagogy and related competencies may be unfamiliar to some faculty members, a qualitative method was chosen in order to build campus conversations around this topic.

The authors' goals for this research were to strengthen liaison librarian-faculty relationships, raise faculty awareness of librarian interest in and knowledge of data literacy, and determine whether faculty see data literacy as within the scope of library instruction. Semistructured interviews and an inductive approach allowed the authors the flexibility to “establish clear links between the research objectives and the summary findings derived from the raw data.”²⁰ The institutional review boards at Rollins College and University of South Florida approved this study, including the interview script, prior to implementation.

The final interview script and protocol were modeled after the work by Stamatoplos, Neville, and Henry in 2016 and the earlier work of Carlson and Johnston in 2015.²¹ The interview script (see the appendix at the end of the chapter) began with a discussion of current and past courses taught by each faculty member, followed by a series of questions about their understanding of data literacy and whether or how they integrate data literacy into their teaching. The interviewees were also provided a list of data literacy competencies previously developed for USF St. Petersburg campus, and asked about the applicability and importance of each competency in the context of their courses.²²

Semistructured virtual interviews took place in April, May, and June 2020. A standard protocol was followed for each interview. Each author met virtually with a single participant, and all sessions were recorded and transcribed (with one exception made at a faculty member's request). Webex and Microsoft Teams were used for recording each interview, and then the interviews were transcribed using Microsoft Stream transcription. Each interview lasted approximately 30 to 40 minutes, and all interview transcripts were anonymized by an assigned sequential ID number.

The interview transcripts, video recordings, and associated written notes make up the dataset. Each researcher examined the transcripts of the interviews they conducted, and, based on initial analysis, the research team collaboratively identified attributes of interest. These attributes were developed into a set of six thematic codes and compiled into a common codebook using First Cycle Descriptive Codes as described by Johnny Saldaña.²³ These themes are defined in the results according to the respective pedagogical discourses of instruction librarians and teaching faculty.

The next stage of the data analysis included systematic coding, again as outlined in Saldaña with a process to ensure the validity of assigning the codes.²⁴ As stressed by Connaway and Radford, "It is important to remember that it is inadequate to have only one coder."²⁵ The interrater reliability check (ICR) contributed to the agreement of assigned codes and gave reviewers the opportunity to discuss codes, obtain feedback, and build confidence in the coding process. Researchers completed independent close readings of at least eleven interviews they did not conduct. The researchers then compared and combined codes from each transcript in norming meetings to create consensus and reliability around the final coding choices.

Finally, the researchers drew conclusions and developed recommendations based on the significant themes that emerged from the descriptive codes. These

conclusions, explained herein, are supplemented by basic descriptive statistics relating to interviewee responses to qualitative interview questions as well as by research participant quotes, which have been lightly edited for readability.

RESULTS

Conversations with twenty-two faculty from various disciplines and teaching backgrounds revealed acute interest and enthusiasm for the topic of data literacy, with high levels of agreement about the value of this skill set for undergraduate students. Data literacy was also seen as critical for students' postgraduate entry into today's competitive job market and success in adult life.

The researchers presented faculty with a list of competencies without scaffolding and asked them to select the competencies most applicable to their undergraduate courses (see figure 1.1 and the appendix at the end of the chapter). While data analysis is a crucial component of data literacy, there are entire courses devoted to this subject. This list of data literacy competencies focuses on the data-related skills that complement data analysis and therefore may be embedded throughout the curriculum.

FIGURE 1.1

Faculty agreement on data literacy competencies

Select data literacy competencies with the percentage of interviewed faculty who agree these are relevant for undergraduate education.

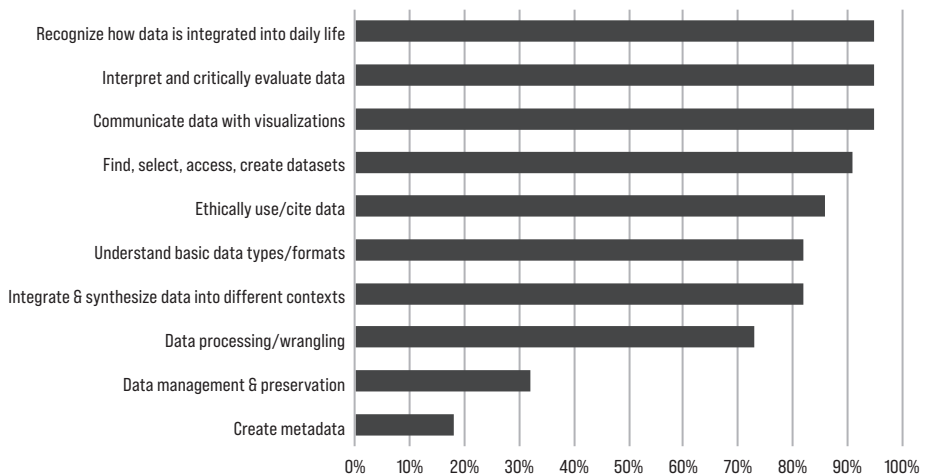


Figure 1.1 reflects interviewees' agreement that many data literacy competencies are applicable to undergraduate education across disciplines. Almost unanimously (95 percent) faculty found the following three competencies relevant for undergraduate learners: recognizing how data is integrated into daily life; interpreting and critically evaluating data and their sources; and communicating data effectively to different audiences in part by using visualizations. Close readings of the interview transcripts affirm this agreement, and indicate that faculty from a range of disciplines consider many of these competencies to be not only applicable to but essential for undergraduate student education. For example, one faculty member explained the importance of recognizing how data is integrated into daily life, saying, "I think that once [students] become comfortable with data meaning more than just numbers, they can better see data in all its forms in their life." Regarding interpreting and critically evaluating data and their sources, a faculty member stated, "[Students] need to be critical of things we know from data; they need to assess data rather than just take it as something that we know . . . some monolithic body of knowledge." Another interviewee spoke about the importance of communicating data as follows: "Being able to write and communicate the data to other people is as important as doing the experiment itself in lab. That's the first real objective."

Further, a large majority of faculty found eight of the ten competencies appropriate for the undergraduate curriculum. Eighty percent or more of faculty interviewees confirmed that competencies originally designated as undergraduate-level skills aligned with the capacities they were trying to instill in their students, whether in general education courses or upper-level, major coursework. Data management and preservation as well as publication metadata activities were, for the most part, seen as better suited to graduate study. However, one competency originally marked as useful for graduate students—data processing and wrangling—was deemed as useful for undergraduates by 73 percent of faculty. It is important to note that there was also variation in how faculty described the scaffolding between general education and upper level or capstone experiences for undergraduate students. (This will be discussed further later in the chapter.)

The authors identified six main concepts and used these to code all transcripts. The concepts are defined below:

1. *Value of Data Literacy*: The participant shares their perception of the overarching value of data literacy to students.

2. *Data Ethics*: The participant shares their concerns relating to ethical use, collection, and citation of data or datasets.
3. *Information Literacy and Research Process*: The participant uses language that also describes major/typical tenets or components of information literacy.
4. *Datasets*: The participant discusses their strategies for how they provide students opportunities to utilize datasets in their course (create new data versus find/compile data versus use “approved” datasets).
5. *Math Anxiety*: The participant discusses obstacles to student readiness for learning data-related competencies with regard to numerical or math literacy.
6. *Pedagogy*: The participant discusses different techniques/methods used to teach students data literacy concepts.

During the coding process, the authors quickly realized that Pedagogy was infused in almost every aspect of our conversations, and thus, while a critical point of analysis, was difficult to code effectively. Information Literacy and Research Process, the Value of Data Literacy, and Datasets were well represented in the interview transcripts; Data Ethics was represented in the transcripts to a lesser extent. Initially, the authors thought Math Anxiety was thought to be a major theme; however, the coding process revealed that it was mentioned less frequently.

The authors examined the most significant themes to better understand their relevance in data literacy instruction. The expectation was to learn ways librarians could engage with faculty colleagues to enhance student learning.

DISCUSSION

Faculty interview responses about data literacy emphasized students’ ability to make informed decisions or conclusions based on data, both inside and outside of academic pursuits. The majority of interviewees, regardless of discipline, saw data literacy as the foundational understanding that data is part of our everyday lives (see the top-rated competency listed in figure 1.1 and table 1.1). As one faculty member commented, “The entire world is data. How could we not make sure that students are a part of that?” This also speaks to the larger value proposition of data literacy as a critical life skill.

Based on the results of the coding and analysis of faculty interviews, these authors propose a shift in the scaffolding previously outlined by Burress, Mann, and Neville.²⁶ The updated scaffolding proposed in table 1.1 reflects a number of insights that arose from analysis of the transcripts and accompanying metadata.

The data reflected in figure 1.1 indicate that nearly three quarters of faculty considered data processing and wrangling to be a relevant undergraduate competency. A number of faculty asked for clarification of the term *data wrangling*, borrowed from data science, because they were unfamiliar with it. Therefore, these authors suggest that this language be adapted and customized for different disciplinary contexts to better reflect the activities of students,

TABLE 1.1
Revised data literacy competencies

Data literacy competency	Undergraduate scaffolding	
Recognize how data is integrated into daily life	Introductory level* All disciplines	
Interpret and critically evaluate data and their sources Analyze data Communicate data effectively to different audiences, in part by using visualizations	All levels, with increasing complexity All disciplines	
Read/understand data types and formats Clean / process / convert data Find, select, access, or create datasets in order to test a hypothesis or answer a research question	Introductory level* Natural sciences	Upper-level** or undergraduate research Social sciences
Ethically collect / use / cite data Integrate and synthesize data into different contexts with other sources and prior knowledge	Upper-level** or undergraduate Research Natural sciences	All levels, increasing complexity Social sciences

* "Introductory level" refers to first-/second-year undergraduates.

** "Upper level" refers to third-/fourth-year undergraduates.

Revised data literacy competencies with flexible scaffolding for the undergraduate curriculum, acknowledging potential disciplinary differences in approaches to data. (Note: Disciplinary groups are organized alphabetically from left to right, not by value.)

who may be required to clean, process, or convert data into various formats in order to use specific tools to practice data analysis (see table 1.1).

The second adjustment the authors suggest making to the competencies is to explicitly mention data analysis (displayed in table 1.1). While data analysis is implicitly referenced in a number of the competencies in the original scaffolding, it should be listed explicitly because analysis is a central skill regardless of the type and complexity of the data. One faculty member stated, “We do data first thing . . . the first thing in the first week, because if you want them to be able to analyze data, they have to believe that they can.” Another faculty member spoke about how data analysis is embedded within multiple competencies, stating, “You can’t answer the research question without the analysis . . . you often have to do data analysis, even if it’s just providing a summary, a graph, a histogram, a bar graph.”

Finally, while interviews with faculty in natural sciences and social sciences reflected agreement about the importance of data literacy in undergraduate learning, close readings of the individual transcripts revealed nuanced disciplinary differences in the way competencies were scaffolded from general education to upper-level/undergraduate research and beyond. As previously mentioned, the concept of Datasets was coded to glean insights about how faculty from various disciplines approach the use of datasets, both original and external, in their undergraduate courses. It was determined that the way faculty use datasets in the classroom may differ depending on disciplinary context. The authors suspect these different approaches to data could influence the scaffolding of data literacy competencies.

For example, faculty in natural science disciplines prioritized the collection of original data and testing hypotheses as a component of introductory courses and labs, while faculty in social science disciplines tended to reserve the collection of original data to senior or capstone-level courses. One faculty member stated, “We have them gather data in the beginning, at the 100 level in our major. What that looks like is they’re given cookbooks of instructions of ‘this is how you put things together.’” In contrast, a social science faculty member stated, “In the lower level classes, I’ll usually give them datasets to work with so that they can practice analysis without having to go through the time-consuming work of data collection.” While these findings are qualitative in nature, this contrasting scaffolding pattern appeared in multiple transcripts and provide new guidance for the evolving data literacy competency recommendations mentioned in table 1.1. Research methods and approaches to

data vary widely within and across disciplines, and the authors acknowledge that the use of disciplinary groups is a simplistic way to begin exploration of a complex topic. Future research to explore pedagogical approaches to data-related competencies based on research methods or data types would be beneficial.

The researchers also noted an interesting and unexpected trend during this investigation. Many faculty interviewees specifically noted the value of students understanding *both* qualitative and quantitative data types. However, some faculty comments suggested that students do not necessarily see both types of data as equally useful (at least, at first), and may not have been exposed to qualitative data as a concept in previous learning contexts. One elaborated, “Data can mean a lot of different things. Students, especially in the lower-level classes, their first thought is ‘Quantitative data is the only thing that counts.’ In my research methods class, we spend quite a lot of time talking about both qualitative and quantitative approaches to gathering and analyzing data.” Another professor reinforced this concept: “Textual data is very important and it’s critical for students to understand. In my methods class I explain how text is data . . . and I find that lots of students have not actually been exposed to this idea before. But it’s very important because so much of what we learn about the world and so much of what we want to understand is not quantitative.” Further investigation is needed to determine the full range of implications of this reality for librarians and other academic learning partners.

In addition to examining faculty approaches to different types of data and their differential use of datasets in the classroom, the research team was interested in faculty pedagogical practices and challenges in teaching data literacy. Specifically, they considered whether faculty taught data literacy alongside traditional information literacy concepts and how they approached data ethics in their courses more broadly. Findings in these areas reinforce the competencies listed above and reveal a number of ways for librarians to develop or add value to existing data literacy curricula in college classrooms.

Information and Data Literacy Pedagogy

Librarians and instructors alike understand that information literacy is a skill set that needs to be embedded throughout an undergraduate student’s ongoing academic journey. Students must develop the skills necessary to be savvy

information consumers and creators regardless of their major, and faculty interviews in this study make it clear that becoming data literate demands the same pedagogical attention and recursive practice.

Similar to information literacy practices, faculty stressed the point that students need to be able to identify and evaluate types of data sources much as they would any other information source. As stated earlier, ensuring that students recognize both qualitative and quantitative data formats are an essential piece of their pedagogy. Alongside this knowledge, students must learn to view data with a critical eye.

Faculty reported the importance of students evaluating data sources for quality and authority. One professor indicated, “It’s so much more important today to be able to figure out whether your sources are reputable.” One professor noted that they teach students, in addition to evaluating data sources, the “importance of knowing the big picture, knowing the data, and going beyond what you’re given to the source.” Another professor stated that when reviewing a research article with students, faculty need to question the data presented and ask, “Did the answer come from a valid process and methodology?” Librarians often emphasize the need for students to be critical of their sources and to use them appropriately in answering a given research question; datasets require a similar approach for undergraduate learners, and faculty seem to be readily recognizing these pedagogical connections.

Data literacy instruction builds on the students’ existing knowledge base, scaffolding from foundational concepts to advanced approaches and analysis, sometimes all within a single semester. A physics professor explained that, following lab experiments, they remind students: “You wrote down some numbers, and then you made a graph, and then what does your graph tell you?”—which reveals the progression from data collection to visualization and then to interpretation. Furthermore, professors shared their extensive efforts to reinforce and revisit data literacy concepts as outlined in table 1.1. One professor shared that data literacy instruction in their course requires reviewing the same concepts “multiple times” within the same semester. Another stated that their stepped approach helps “to intentionally revisit concepts in a more and more complex way.”

Professors also noted that many of their students struggled to synthesize data. One professor said, “We call it the pancake method, where they’re stacking it on top of each other as opposed to interweaving themes.” Another professor said students are often “trying to take something and fit into some

other context.” Students need to contextualize data order to support an argument. Synthesizing data and information is a necessary skill for students to develop, and because many librarians already teach this type of skill as part of information literacy instruction, it would be very feasible to address this data-related competency.

Data Ethics

The findings indicate that 86 percent of all faculty interviewed identified ethical use and citation of data as essential for all undergraduate students. The researchers’ analysis of interview transcripts allowed for insights regarding the complexity and meaning of data ethics.

Many faculty members focused on the difficulty of teaching proper data citation. One interviewee admitted their concern about the reliability of data used by students saying, “Where’d [they] get that data?” Another interviewee stated, “I think that they understand why it’s important to ethically use data. But when it comes to the actual practice of it, I find that semester after semester they aren’t doing it correctly.” These concerns result in faculty sidestepping the issue by preselecting datasets for classroom use. The desire for students to use and cite data ethically and find reliable and relevant data directly intersects with librarian expertise in scholarly communication concepts. Thus, this is an area where librarians can play a strong role in data literacy instruction.

Data ethics encompasses more than citation; it is an ongoing consideration throughout the research process, including data collection, selection, and the process of participant consent. One professor stated, “When [students] think about ethics, they typically just think about the consent process, but they don’t think about their own subjective perspectives and how that might change some of the ways in which data is constructed, collected, interpreted, and presented.” Again, librarians can contribute to students’ data literacy learning by engaging with them throughout the research process.

Professors also talked about social media and their students’ understanding of the data they themselves are creating and sharing. One faculty member said, “Recognizing how data is integrated into daily life and workplaces is really important, especially in light of social media and issues of privacy. . . . People on the user end think that they are just using the social media because it’s fun . . . and you realize these are corporations who are making money off of your data, right?” As Head, Fister, and MacMillan point out, students know

that data is being collected and used to build algorithms and are paradoxically “both resigned to and indignant about algorithmic-driven ads, news, and information.”²⁷

Librarian Roles in Undergraduate Data Literacy

In addition to the themes discussed above, faculty shared many ideas about promising collaborations with librarians related to data literacy. These included facilitating conversations about data across campus, creating a database of local data experts to connect faculty to one another, providing boot camps for students on basics of data literacy, teaching faculty and students how to find data, hosting workshops on data ethics and citation, and creating a warehouse of data literacy tools and resources for campus.

One professor noted that learning how librarians and other faculty colleagues are teaching data literacy would “help us create better assignments or brainstorm ideas that were more intentional.” Another faculty member said, “I would love it if I knew for sure that there were other people who were doing similar things so that it wouldn’t feel so out of the ordinary. Perhaps there could be a cohort amongst faculty that really thought about this collectively or did something programmatic or just encourage[d] communication about it.” At the USF St. Petersburg campus, librarians designed a Data Literacy Teaching Toolkit, which provides access to resources to help faculty in teaching data literacy concepts.²⁸ From this research, the authors plan to include additional activities for faculty to incorporate in their teaching.

Other faculty members expressed a desire to learn more about what their colleagues are doing with respect to data. A professor noted that she would like to learn “more about how we can share data [with colleagues] so that different perspectives and eyes get a look at it and we can learn more with each other.” Librarians could assist faculty members by creating a professional learning group focusing on data literacy at their institution, or could become conduits for data literacy information on campus. Librarians could also host workshops where faculty members present about their data literacy approaches in their classes and the data they use in their research.

Cocurricular library programming was also mentioned as a way to address the time constraints that make it challenging to cover foundational data skills in a content-focused course. One faculty member said, “There should be some type of boot camp that provides a certificate so that the students could

go through that mini boot camp and then show me the certificate so that I know that they've gotten all the things that they need to start the project faster.”

Some areas with potential for further research include an examination of student data practices at different levels of undergraduate education, a syllabi analysis focusing on data-related course objectives and learning outcomes, and focus groups and surveys to test the revised, scaffolded competencies across the full range of academic disciplines. During the interviews, multiple faculty members expressed interest in collaborating with the librarians on future research, building on the present work with longitudinal investigations into student development of data-savvy research practices. These potential librarian-faculty research collaborations are especially valuable in that they represent a model of integrated research to practice. Long-term collaborative partnerships such as these have the potential to impact higher education in a meaningful way.

Research Limitations

Because participants self-selected to take part in the study, they may have had a particular affinity for the topic that could have resulted in selection bias. Although the authors interviewed faculty members from a wide variety of disciplines, humanities faculty were not represented in the sample. In future research that may include focus groups or larger samples of faculty for purposes of surveys, it will be important to ensure that the full range of academic disciplines is adequately covered.

The differing academic portfolios of each institution must also be considered. There is not universal consensus in librarianship or academia more broadly as to a uniform or consistent way of aggregating social sciences and natural sciences disciplines. However, the authors are librarians who have all worked as liaisons for a variety of academic departments and therefore bring to this discussion a general understanding of typical disciplinary differences between college and departmental structures.

CONCLUSION

Information and data literacy are closely related skill sets that are essential for undergraduate students to master, and librarians are well positioned to foster student learning and faculty development in both areas. This research

establishes eight data literacy competencies that are relevant to undergraduate students across an array of social science, natural science, and applied disciplines. Further, these authors found that the variation in reported scaffolding approaches may have disciplinary nuances, and this should be considered by librarians who collaborate with faculty to ensure the appropriate design of data-inflected library instruction.

Primary competencies where librarian expertise can add value to data literacy instruction for students include:

- evaluating data and data sources
- synthesizing data in different contexts
- data ethics and citation

Furthermore, librarians can provide value for faculty directly by:

- facilitating professional development workshops on topics that are relevant to faculty teaching and research, such as data ethics and citation
- fostering faculty communication and exchange of best teaching practices across departments
- partnering on research related to scholarship of teaching and learning (SOTL)

Like information literacy, data literacy instruction requires that students learn through recursive practice. Both data literacy and information literacy are ubiquitous and interdisciplinary. Faculty interviewees showed a real enthusiasm to hone their data literacy pedagogy and practices to better inform their teaching and academic research. Librarians are well positioned to assist students and faculty in the growth and improvement of data literacy learning and initiatives on college campuses.

Notes

1. Association of College and Research Libraries, *Framework for Information Literacy for Higher Education* (Chicago: American Library Association, 2015), <http://www.ala.org/acrl/files/issues/infolit/framework.pdf>.
2. See Robin Rice and John Southall, *The Data Librarian's Handbook* (London: Facet Publishing, 2016).

3. See Tibor Koltay, "Data Literacy for Researchers and Data Librarians," *Journal of Librarianship and Information Science* 49, no. 1 (2017): 3–14; Rachel Walton and Patti McCall-Wright, "What about the Little Guys? How to Approach Supporting Research Data Management at a Small Liberal Arts College," *ResearchDataQ*, November 6, 2019, <https://researchdataq.org/editorials/what-about-the-little-guys-how-to-approach-supporting-research-data-management-at-a-small-liberal-arts-college/>.
4. Abigail Goben and Tina Griffin, "In Aggregate: Trends, Needs, and Opportunities from Research Data Management Surveys," *College & Research Libraries* 80, no. 7 (2019): 903–24.
5. Goben and Griffin, "In Aggregate," 914.
6. Elizabeth Stephenson and Patti Schifter Caravello, "Incorporating Data Literacy into Undergraduate Information Literacy Programs in the Social Sciences: A Pilot Project," *Reference Services Review* 35, no. 4 (2007): 525–40.
- 7.. Jake Carlson et al., "Determining Data Information Literacy Needs: A Study of Students and Research Faculty," in *Data Information Literacy: Librarians, Data, and the Education of a New Generation of Researchers*, eds. Jake Carlson and Lisa R. Johnston (West Lafayette, IN: Purdue University Press), 11–33.
8. Jake Carlson, "Future Directions for Data Information Literacy: Growing Programs and Communities of Practice," in *Data Information Literacy: Librarians, Data, and the Education of a New Generation of Researchers*, eds. Jake Carlson and Lisa R. Johnston (West Lafayette, IN: Purdue University Press), 248–59.
9. Meryl Brodsky, "The Role of Business Librarians in Teaching Data Literacy," *Ticker: The Academic Business Librarian Review* 1, no. 1 (2016): 1–6.
10. Gwen Moran, "We're in a Data Literacy Crisis. Could Librarians Be the Superheroes We Need?," *Fortune*, August 31, 2019.
11. Scout Calvert, *Future Themes and Forecasts for Research Libraries and Emerging Technologies*, ed. Mary Lee Kennedy, Clifford Lynch, and John O'Brien (Association of Research Libraries, Coalition for Networked Information, and Educause, 2020), 7.
12. Calvert, *Future Themes and Forecasts*, 12.
13. Yasmeeen Shorish, "Data Information Literacy and Undergraduates: A Critical Competency," *College and Undergraduate Libraries* 22, no. 1 (2015): 97–106; Theresa Burress, Emily Mann, and Tina Neville, "Exploring Data Literacy via a Librarian-Faculty Learning Community: A Case Study," *Journal of Academic Librarianship* 46, no. 1 (2020): 102076.
14. Karen Hogenboom, Carissa Phillips, and Merinda Hensley, "Show Me the Data! Partnering with Instructors to Teach Data Literacy," in *ACRL Conference Proceedings* (Philadelphia, 2011), 410–17.

15. Javier Calzada Prado and Miguel Ángel Marzal, "Incorporating Data Literacy into Information Literacy Programs: Core Competencies and Contents," *Libri* 63, no. 2 (2013): 126.
16. Burress, Mann, and Neville, "Exploring Data Literacy," 102076.
17. University of South Florida, "Enhanced General Education: Enhanced Curriculum," <https://www.usf.edu/undergrad/general-education-council/enhanced-gened/enhanced-curriculum.aspx>, accessed March 17, 2021.
18. Carnegie Classification of Institutions of Higher Education, A Definition for Master's Colleges and Universities, n.d., accessed June 2020, https://carnegieclassifications.iu.edu/classification_descriptions/basic.php. (Note: In the summer of 2020, USF St. Petersburg accreditation was consolidated with the USF Tampa and Sarasota-Manatee campuses and the Carnegie Classification changed.)
19. Anthony Stamatoplos, Tina Neville, and Deborah Henry, "Analyzing the Data Management Environment in a Master's-Level Institution," *Journal of Academic Librarianship* 42, no. 2 (2016): 156.
20. David R. Thomas, "A General Inductive Approach for Analyzing Qualitative Evaluation Data," *American Journal of Evaluation* 27, no. 2 (2006): 238.
21. See Stamatoplos, Neville, and Henry, "Analyzing the Data Management Environment"; and Jake Carlson and Lisa R. Johnston, eds., *Data Information Literacy: Librarians, Data, and the Education of a New Generation of Researchers* (West Lafayette, IN: Purdue University Press, 2015).
22. Burress, Mann, and Neville, "Exploring Data Literacy," 102076.
23. Johnny Saldaña, *The Coding Manual for Qualitative Researchers* (Thousand Oaks, CA: SAGE Publications, 2009), 70.
24. Saldaña, *Coding Manual for Qualitative Researchers*.
25. Lynn Silipigni Connaway and Marie L. Radford, *Research Methods in Library and Information Science*, 6th ed. (Santa Barbara, CA: Libraries Unlimited, 2017), 316.
26. Burress, Mann, and Neville, "Exploring Data Literacy," 102076.
27. Alison J. Head, Barbara Fister, and Margy MacMillan, *Information Literacy in the Age of Algorithms: Student Experiences with News and Information, and the Need for Change, aka "The Algo Study Report," Project Information Literacy* (Knight Foundation, Harvard Graduate School of Education, and the University of South Carolina School of Information Science, January 15, 2020), 27.
28. See Nelson Poynter Memorial Library, "Data Literacy Teaching Toolkit," USF St. Petersburg Library General Guides, n.d., <https://lib.stpetersburg.usf.edu/dataliteracy>.

APPENDIX

INTERVIEW SCRIPT

NOTES

Date:

Place:

Name of interviewer / data collector / institution:

Name of Participant:

Discipline/Department:

PROCEDURAL NOTE: [Consent form already provided via Qualtrics link before the date of the interview.]

SESSION INTRODUCTION

Read this statement:

Thank you for participating in our research. Before beginning, we would like to tell you about the project. We are conducting interviews to assess faculty perceptions of undergraduate students' data literacy needs. We're curious about which data literacy competencies professors feel are important for their discipline, and we are interested in and how librarians can help meet the needs of these students. As a part of this project, we are asking faculty, such as yourself, who work with undergraduate students and either have implemented or wish to implement data literacy components in their course(s)—classwork, assignments, papers, tests, and so on. During the interview, we will ask you questions about what data literacy means to you, and we'll record your responses in writing and by digital recorder. The purpose of this is to ensure accurate reporting; your responses will not be shared outside of this project, and your name will not be associated with any of these recordings or notes from this session today. The interview will last between 15 and 30 minutes. Responses may be reviewed by any of the investigators of this project [*mention the names here if you wish—e.g., Emily Mann, Theresa Burress, Rachel Walton, or Susan Montgomery*] and aggregated anonymously for the purposes of data analysis, interpretation, scholarly publication, and the development of library services or programs both on and outside of our campus.

RECORDING

Say: Is it all right with you if I record our discussion today?

Inform the interviewee that you are starting recording.

ICEBREAKER/WARM-UP

Ask the interviewee to share a little bit about what classes they regularly teach and what classes they are teaching now.

INTERVIEW QUESTIONS

- What does data literacy mean to you in the context of your teaching?
- These are data literacy competencies identified by librarians in a previous librarian-faculty collaboration. Which of these competencies do you feel are most important for your undergraduate students?
 - to recognize how data is integrated into daily life and workplaces
 - to read and understand basic data types and formats
 - to interpret and critically evaluate data and their sources
 - to integrate and synthesize data into different contexts with other sources and prior knowledge
 - to find, select, access, or create datasets in order to test a hypothesis
 - to find, select, access, or create datasets in order to answer a research question to ethically use/cite data
 - to communicate data effectively to different audiences in part by using visualizations
 - to data processing and wrangling
 - to data management and preservation
 - to metadata to meet data publication requirements
- Do you think students are coming into your classroom with the data skills they need to succeed?
- Are you incorporating data literacy instruction into your classroom?
 - *If yes:* What courses? How are you incorporating data literacy into your class? When did you start? Do you feel students are successful in integrating data-literate practices into their work by the end of the course?
 - *If no:* Is there a reason why not?

- What kind of support do you wish you had for data literacy instruction?
- Do you have any questions for us as researchers and librarians?
- Would you like to be notified of the results of the study?

SESSION ENDING

Thank you for participating in our research, taking the time to meet with us, and answering our questions. The information you provided today will be valuable in helping us understand how librarians can support faculty with regard to undergraduate student data literacy. If you think of anything else you would like to add, please feel free to contact us. *[Give your contact information here.]*

Thank you for your time.