College & Research Libraries

CRL

May 2024 • Volume 85 • Number 4

Association of College and Research Libraries

COLLEGE & RESEARCH LIBRARIES

May 2024
VOLUME 85
NUMBER 4
ISSN 0010-0870

Editorial

Connecting with Co-workers to Build Trust

I read Denise Brush's "Trust in Academic Libraries: How to Build Connections between New Co-workers," in the April 2024 issue of *College & Research Libraries News*.[1] Many of her points resonated with my own experiences working in my institution's library. Trust has become a recurrent theme—across many professions' organizational cultures—since the emergence of the COVID-19 pandemic. Trust can erode when there is a lack of connection with other library staff, new or veteran. The library is constantly growing, changing, or adapting to sync with the institutional community's research and instructional needs, current and anticipatory; and new colleagues—across various departments—arrive as steadily long-time library staff retire, or seek new opportunities. Such constant change and growth can make it challenging to build connections with new employees.

Brush writes that:

> A library where a large percentage of the employees are new must build trust. Employees who have worked together for a long time have shared understandings, but when new employees join the organization, they must develop relationships with their co-workers before there can be mutual trust (157).

Connecting with colleagues—getting to know them, to understand their work, and to find ways to work together—is very important, but it can be difficult to do. Library staff working varying schedules—such as working from home and working on-site on different days—impacts when, and how often, colleagues see each other. This can make it difficult to build trust in the library. As Brush writes, libraries must intentionally plan events or other ways for colleagues to connect. Otherwise, the lack of connections can be detrimental to the library's organizational culture. In addition to social events and social media platforms for connecting, I would suggest developing cross-departmental task forces, working groups, or other kinds of committees to meet shared work goals. Working together towards shared goals—whether they are part of a strategic plan goal or objective—is invaluable for many reasons. Not only does working together across departments increase the quality of the work outputs, due to the diversity of perspectives, it also organically allows for opportunities to get to know one's colleagues as people, and to learn about their work, which helps establish trust. In addition, inter-departmental work breaks down silos, and requires colleagues to consider different department's need while collaborating to achieve mutual work goals. Finally, working together across departments fosters future collaborations.

While it is essential to get to know colleagues through library-wide social events, be they virtual or physical, working on projects—across different departments but toward a common goal—can also be a powerful way to build connections and trust between colleagues. When change, evolving services, and having library spaces in flux are the norm, library employees

and leadership cannot take for granted that all coworkers in the organization know, or trust, one other. In addition to building connections socially, libraries should strive to build professional connections, and prioritizing inclusive, goal-oriented projects across departments is a natural way to do this. Through building connections between colleagues, a library also builds trust in its organizational culture.

Note

1. Denise Brush, "Trust in Academic Libraries: How to Build Connections between New Co-workers," *College & Research Libraries News* 85, no. 4 (2024): 157–159, https://doi.org/10.5860/crln.85.4.158

Librarians and Academic Libraries' Role in Promoting Open Access: What Needs to Change?

Shlomit Hadad and Noa Aharony

Profound changes due to Open-Access (OA) publications lead to organizational changes in universities and libraries. This study examines Israeli librarians' perceptions regarding their role and the academic library's role in promoting OA-publications, including the barriers, challenges, needs, and requirements necessary to promote OA publishing. Lack of a budget for OA-agreements, no cooperation from university management, and researchers' unfamiliarity with OA were among the most prominent barriers. Librarians see great importance in their role of advising researchers regarding OA. However, they insist on a regulated OA-policy at the national and institutional levels to strengthen their status as change-leaders of the OA-movement.

Introduction

Open Access (OA) is a term that is used to describe unrestricted online access to scientific articles as part of an effort to "open up" scientific output to the public.[1] The premise is that OA may improve the rigor, validity, replicability, and availability of research.[2] One of the major arguments against the subscription-based model of publishing is that while authors contribute their work to publishers without monetary gain, readers are required to pay a subscription fee to the journal.[3] At the same time, organizations and academic institutions have to pay publishers through mega-agreements, known as the "big deals," to allow researchers and students access to those articles.[4]

Following OA initiatives from the early 2000s, which formed the ideological and practical basis of the movement, countries, funders, and research institutions across the globe commenced to provide OA for their research output, while also attempting to develop a clear OA policy.[5] The evolving form of new business models of academic publishing and the entry of "new players" to this field are among the main reasons for the transformation of academic libraries.[6]

However, open access has its own concerns. These include the costs associated with article processing charges (APCs) for OA journals, which affect the ability of academics from the social science and humanities (SSH) to publish in OA journals. Other problematic issues are implementing and maintaining an institutional repository and the fear of copyright infringement when depositing articles in open repositories.[7] In addition, there is the rise of predatory journals interested in only making quick money that pay little or no attention to

* *Shlomit Hadad is Lecturer in the Department of Digital Learning Technologies at The Israel Academic College in Ramat-Gan, email: hadadshe@biu.ac.il; Noa Aharony is Full Professor at Bar-Ilan University, email: Noa.Aharony@biu.ac.il.*

peer review.[8] These predatory journals negatively influence researchers' attitudes towards OA publishing.[9]

Many previous studies have dealt with the new and emerging roles of librarians resulting from a new digital era and the change from the traditional publishing models to OA.[10] Scholars have explored librarians' role changes and have suggested ways to improve and promote OA in their institutions.[11] However, only a few have explored ways to promote OA from the perspective of librarians, who are considered experts in their field. The present study aims to address this gap by exploring Israeli librarians' perceptions of their role and the academic library's role in promoting OA publication, as well as the barriers, challenges, needs, and requirements needed to promote OA publishing in their institutions and on a national level.

Literature Review

Librarians' and Libraries' Roles in the Scientific Communication System over Time

In the scholarly world, libraries and librarians have always played a central role in the creation, preservation, and dissemination of information.[12] Over the years, academic libraries have evolved alongside the development of higher education institutions, and have adapted to social, political, and technical changes.[13] Thus, the ever-changing research landscape and the relentless advances in technology have significantly influenced the responsibilities of academic librarians.[14]

Traditional functions, such as reference work and collection management, are rapidly losing their status as primary responsibilities of librarians, while new functions related to research support, data management, bibliometrics, and digital initiatives, are increasingly becoming part of the academic librarian's responsibilities.[15] These changes enable librarians to perform new and significant roles, redefine their roles, and provide libraries with the opportunity to remain relevant in the digital age.[16]

Open-access Publication and its Impact on Librarians' Roles

Digital developments are not the only factor that have influenced academic library service. Major changes in scientific communication such as the OA movement affected scientists and publishers and led to organizational changes in universities and libraries.[17]

The OA movement was initiated in the 1990s, as access to the Internet became widely available and online publishing became the norm. It was intensified by three initiatives, known as the BBB declarations: the *Budapest OA Initiative*, the *Bethesda Statement on OA Publishing* in 2002, and the *Berlin Declaration on OA* in 2003. These initiatives represent the most highly regarded definitions of OA, and all agree on the essentials. As stated by Peter Suber, "an OA work is digital, online, free of charge, and free of most copyright and licensing restrictions" (para. 1).[18] The *Budapest OA Initiative* recommended two approaches to providing open access to the research literature: open access journals (known as the "gold" road) and institutional or individual self-archiving in digital repositories (known as the "green" road).[19] Heather Piwowar et al. prefer a less strict definition: "OA articles are free to read online, either on the publisher website or in an OA repository."[20]

Following the BBB declarations, major research institutions across the globe committed to providing OA for their research output.[21] More recently, grant conditions of many funding organizations, including Plan S, Europe PMC Funders' Group, and Horizon Europe, began

requiring peer-reviewed research output to be freely available. These aims can be achieved either by publishing in OA journals, archiving publications in an OA repository, or in some cases, both options are required.[22] Although there is an increased awareness regarding OA over the years, there is still confusion and misunderstanding concerning the various OA models. Moreover, following the rise of the gold OA model (OA journals), many predatory journals have emerged.[23] Hence, researchers' suspicions towards OA journals are understandable. Researchers question the reliability of OA journals and now consider gold journals as providers of lower-quality articles.[24]

Because of this, researchers need to recognize and distinguish between OA publication models and routes, as well as between legitimate and predatory journals. In addition, the increasing costs of toll-access subscriptions, particularly via so-called ''Big Deals'' from publishers, forced libraries and other institutions to initiate large-scale subscription cancellations.[25] As libraries make difficult budgetary decisions, the OA movement allows them to redefine their roles within this emerging publishing model.[26] Further, with the development of OA, there has been an expectation that academic libraries will take on additional responsibilities like managing research data and open access requirements.[27] As a result, OA promoters asked librarians to be the change leaders in their institutions,[28] adding suggestions on how to promote OA publication among researchers and management.[29] Studies have found that academic libraries promote OA in a variety of ways such as: including records for OA journals in their public catalogs and electronic journal lists, collaborating with their institutions to establish institutional repositories, participating in institutional initiatives to encourage faculties to deposit research outputs in the institution's database and more.[30] However, some studies have argued that for experienced librarians as well as those new to the profession, there may be a lack of understanding about potential roles in a changing vision of scholarly communication that includes advocacy for openness.[31] Furthermore, some librarians may not believe that "open access" has relevance to their busy roles in the library and they need clear instructions on how to change their daily work in the library.[32]

In addition, previous studies demonstrated that faculty staff did not perceive librarians as team members for policy development, funding, publishing, or rewards and recognition regarding OA.[33] Faculty members would like to have librarians' assistance and support while keeping the traditional vision of the library as a useful warehouse of information and of librarians as selectors and minders of the inventory.[34] Thus, although academic libraries have the ability to provide services in accordance with OA requirements and information system management, they may encounter a lack of cooperation and support for their initiatives from the institution's leaders and faculty.[35] Moreover, scholars question the ability of librarians to integrate the new requirements into the library's administrative structure.[36] Therefore, the aim of this study was to examine Israeli librarians' perceptions regarding their role and the academic library's role in promoting OA publishing, including barriers, challenges, and difficulties. Further, it explored the factors and requirements needed to promote open access publishing in their libraries and nationally. The research questions that guided the study are:

RQ1. How do librarians perceive their role and the academic library's role in promoting OA publishing?

RQ2. What are the barriers, challenges, and difficulties in implementing open-access?
RQ3. What are the factors and needs that are required to promote open access?

Method

Participants

The study was conducted within the qualitative research paradigm. Qualitative interviews offer ecological validity, provide rich and insightful descriptions, and have the ability to aid in the understanding of complex organizational realities.[37] Monique Hennink et al. found that in order to reach code saturation, the point when no additional issues are identified and the codebook begins to stabilize, qualitative research needs 9 interviews.[38] In this study, the consideration in choosing the number of research participants was to allow one representative participant from each of the ten existing universities in Israel. The participants were 10 librarians and academic library administrators from ten universities in Israel. Regarding gender, 90% were female. Among them: five (50%) were administrators of the library system at their institution, three (30%) were directors of disciplinary libraries and two (20%) were directors of information systems at the libraries within their institution. In terms of seniority, six (60%) had been employed by their institution for over 10 years, and four (40%) had less than 10 years in their current positions.

Procedure and Instruments

An email was sent to Israeli academic librarians having positions of administrators or directors of disciplinary libraries or information systems in academic libraries of universities, with a request to participate in the study. Respondents were provided with full details about the research and were invited to an informal telephone discussion with the interviewer to discuss the research aims and procedures. The researcher then conducted semi-structured interviews via Zoom with the participants who agreed to take part in the project. All interviews were conducted between April to June 2020 and lasted between 30 and 90 minutes. The librarians' interviews were based on items extracted from the "Librarians' attitudes towards open access, principles and related behaviors survey,"[39] and included items that deal with changes in the practice of the academic library based on Perkins and Slowik's (2013) study.[40] During the interviews, librarians were asked: 1) to report how they perceive their role and the role of academic libraries in promoting OA publishing; 2) to address the barriers, challenges and difficulties regarding promoting open access; and 3) to address the factors and needs required to promote open access publication.

The answers were analyzed from the "bottom-up." Researchers categorized the answers using a thematic analysis technique.[41] This analysis allowed researchers to reach the main categories. Further, it enabled researchers to catalog and code the interviewees' quotes and to identify common expressions and recurring themes. In addition, during the analysis process, researchers merged themes and categories and identified the overlap between themes. The richness of responses justified the number of participants, showing data collected are of sufficient depth to provide salient information in relation to the research purpose.

The thematic content analysis of the librarians' narratives yielded 1,264 statements which were classified into three main and broad categories. Each main category included several sub-categories (see Table 1). The unit of analysis in this study was a statement presenting a content unit. The coding was not exclusive, as the same statements could be attributed to

TABLE 1
Research Categories (no. of statements=1,264)

Main Category	Sub-categories
1. The role of the library and librarians in promoting open access (N = 428, 34%)	1. The library is responsible for implementing the change, after setting OA policy 2. Guidance and advising researchers regarding issues in OA publishing 3. Involvement in agreements through "MALMAD" consortium 4. Contacts with university administration and the research authority 5. CRIS (Current research information system) operation 6. Promotion of an institutional repository
2. The barriers, challenges and difficulties in promotion of open access (N = 541, 43%)	1. Lack of budget and OA agreements with publishers 2. Lack of cooperation with management, the research authority, and the rector 3. Researchers' lack of awareness regarding OA publication and fear of predatory journals and copyright infringement 4. The Journal Impact Factor (IF) effect 5. Fear of changes resulting from the transition to OA 6. Lack of information regarding the researchers' publications venues and norms 7. Difficulties in "marketing" the OA publishing and the fear of leading the change 8. Lack of personnel, guidance, and training for the librarians 9. Opposition from stakeholders and researchers to OA promotion
3. Factors and requirements that are needed to promote open access (N = 295, 23%)	1. The need to change policy at the national level and adopt advanced OA policy 2. The need for cooperation between academic institutions to establish new agreements and an institutional repository 3. Collaboration of the university management, research authority, and the library for the advancement of OA 4. The need tutorials and training for library staff regarding OA publishing 5. The need to guide researchers and reward them for OA publications 6. The need to establish the library's status as a central body for OA care

several categories. To ensure inter-rater reliability of the coding, 25% of the statements were analyzed by a second coder (a trained researcher knowledgeable in research methods and the relevant topic, in addition to the study researchers) and the agreement level between them had a Cohen's Kappa of .86. Table 1 describes the final research categories and sub-categories.

Rigor

To ensure reliability in the findings, rigor in interviews was based on "trustworthiness of data,"[42] by adhering to four principles: 1. *Truth-value of data*: Librarians were informed in advance that their perspectives and reports would be confidential. To preserve the privacy of the participants, researchers removed all names and places from data sheets; 2. Applicability of the data: was achieved by selecting librarians from different universities, genders, and seniority; 3. Consistency of the data was assured by verbatim transcriptions of the interviews and keeping records of data collection. However, there are no ages, genders, or names associated with the quotes; and 4. Neutrality of data was assured by recording all steps during data coding of the interviews and trying to present librarians' perceptions concerning the phenomenon.

Results

Role and Contribution of Librarians and the Library in Promoting Open Access

The first research question examined the contribution and role of librarians and the library in promoting open access. The librarians discussed six main aspects of their actual role in promoting OA publication in their institutions. Table 2 presents the sub-categories according to their frequency.

TABLE 2
The Role and Contribution of Librarians and the Library in Promoting OA (N = 428, 34%)

Subcategory	Total Number of Statements Including	
	N = 428	%
1. The library is responsible for implementing the change, after setting OA policy	111	26%
2. Guidance and advising researchers regarding issues in OA publishing	101	24%
3. Involvement in agreements through "MALMAD" consortium	67	16%
4. Contacts with university administration and the research authority	65	15%
5. CRIS (Current research information system) operation	46	11%
6. Promotion of an institutional repository	38	9%

Analysis of the sub-categories and representative quotes are presented below:

1. The library as leader of the change, after setting OA policy.

Librarians expressed their feeling that they are leading the process of promoting open access in their institution. However, they claim that promoting OA is not their sole responsibility; it is the national government's and university administration's role to establish a regulated policy on the subject: "We try to promote OA in every possible way. But still, the library is not the regulator, it is the execution contractor. We can be the ones in the field who encourage policy, recommend, educate and implement the changes" (L10).

2. Guidance and advising researchers regarding issues in OA publishing.

Librarians advise and guide researchers: "We provide information; we have a special page on our library portal that really explains the different routes in OA, what to do not to do, questions and answers. We also conduct individual and group trainings" (L10). In addition, librarians reported that they help researchers obtain research funding for publishing in OA journals: "We have a small fund to support OA. The founder invests a relatively small amount in it, and set very strict criteria for which researchers. The whole process is managed through the library" (L5).

3. Involvement in agreements through MALMAD consortium.

Most librarians mentioned the connection with MALMAD as the body responsible for promoting OA in Israel. MALMAD is the "Inter-University Center for Digital Information Services" and is a consortium for acquiring, licensing, and managing digital information services to Israel's universities and colleges. The director of MALMAD reported that there are significant conflicts with publishers to lower the price of those "big deals," and transfer to models that would incorporate OA: "We are partners in the whole process and try to involve stakeholders in the university" (L4).

4. Contact with university management and the research authority.
Contact with management and the research authority is one of the important roles in promoting OA: "As part of the ongoing process, there are meetings with deans and the research authority, and in every meeting the issue of OA arises, and we are asked to explain why this issue is important and worthwhile to the university" (L7).

5. CRIS system operation.
Librarians referred to CRIS (Current Research Information System), a database that stores and manages data about research activities, as a system that will eventually promote OA: "After a struggle, now the CRIS returns to the library. Finally, the university managers understand that it's the library's role. Perhaps this will promote OA" (L1).

6. Promoting an institutional repository.
Librarians referred to promoting and establishing an institutional repository (IR) at their universities: "There is now a demand among many researchers, due to the funders' requirements—to deposit not only the article but also the research data. We contacted the university administration, showed them researchers' requests, and asked for a budget to promote the construction of IR" (L9).

Barriers, Challenges, and Difficulties

The second research question examined the barriers, challenges, and difficulties of OA as perceived by librarians. Librarians' response to this topic yielded the largest category (541 statements, 43% of all statements). Table 3 presents the sub-categories according to their frequency.

TABLE 3
Librarians' Perceptions: Challenges and Barriers (N = 541, 43%)

Subcategory	Total Number of Statements Including	
	N = 541	%
1. Lack of budget and OA agreements with publishers	112	21%
2. Lack of cooperation with management, the research authority, and the rector	83	15%
3. Researchers' lack of awareness regarding OA publication and their fear concerning predatory journals and copyright infringement	80	15%
4. The Journal Impact Factor (IF) effect	73	13%
5. Fear of changes resulting from the transition to OA	54	10%
6. Lack of information regarding researchers' publications venues and norms	44	8%
7. Difficulties in "marketing" OA publishing and the fear of leading the change	50	9%
8. Lack of personnel, guidance, and training for the librarians	33	6%
9. Opposition from stakeholders and researchers	12	2%

Analysis of the sub-categories and representative quotes are given below.

1. Lack of budget and OA agreements with publishers.

Librarians expressed helplessness in the face of mega-agreements with publishers that leave them no budget for further agreements with OA journals: "Once these mega-agreements with publishers are signed, we have no ability to deal with it. We renew agreements from year-to-year and we need almost the entire operating budget for acquisition" (L4). Thus, there is no budget left to OA journals: "Researchers contact us to request a budget for OA publication, but unfortunately, we have nothing to offer them" (L10).

2. Lack of cooperation with management, the research authority, and the rector.
Lack of cooperation and disconnection is a frequently mentioned barrier by the librarians, as expressed in the following quote: "For many years there has been no contact with the research authority. They did not see or meet us unless they needed the help of the library. Thus, we cannot make such progress in promoting OA" (L9).

3. Researchers' lack of awareness regarding OA publication and fear of predatory journals and copyright infringement.
Researchers, according to librarians, are unaware of OA in general, and in their field of research in particular: "Our feeling is that researchers are not aware of OA. They need someone to explain them both the OA ideology and information relevant to their discipline" (L1). Researchers are also afraid of predatory journals: "We expose researchers to the fact that there are many quality open journals. Many researchers think that all OA journals are predatory journals" (L7). Regarding self-archiving in repositories, according to the librarians, researchers are concerned about copyright infringement and scooping: "Researchers are afraid to deposit a post-print article in an open database because they fear violating the copyrights they have committed to with the publisher. In addition, they are also afraid to deposit a pre-print article, because of the fear of plagiarism" (L1).

4. Journal Impact Factor (IF).
Researchers are evaluated by publishing in high IF journals: "Researchers do not want to publish in an open journal or in an institutional repository that we as a library want to promote. They want to publish in a journal with a high IF, which will improve their CV (curriculum vitae) as their promotion depends on publications. Some OA journals have a high IF, but most have a relatively low IF for their field, so this is one of the main reasons why researchers refuse to publish in OA journals" (L5).

5. Fear of changes resulting from the transition to OA.
Librarians discussed their own concerns. They have difficulties adjusting to the transition to OA, which often reduces the need for library services: "The library is changing, and this world of OA will change the world of access to information. The librarians will still be needed, but everything will look different and that is a cause for concern" (L7). Librarians have also mentioned researchers' concerns. Researchers are also afraid of changing the existing traditional model: "Researchers think; why do we have to change the existing publishing model? What's wrong with what we do today? Why rock the boat?" (L3).

6. Lack of information regarding researchers' publications venues and norms.
Another difficulty is the lack of information regarding researchers' publications: "The infor-

mation is not centralized, we actually make surveys and ask our researchers: Tell us where do you publish? How much money do you pay for publications? Otherwise, how do we know what to offer them?" (L1).

7. Difficulties "marketing" OA publishing and the fear of leading the change.
Librarians reported difficulties marketing OA to researchers and the institution's management. They were worried about being responsible for failure: "Agreements that contain open components are expensive, and if in the end there are not enough publications to justify the investment, it will be our fault. That is why it is very difficult for us to explain, market, and promote OA" (L2). In addition, researchers are not aware of library's activities regarding OA, and it leads to researchers' lack of information: "One of the most difficult problems is that researchers are unaware of how much help they can get from the library, so they don't use the library to publish in OA" (L3).

8. Lack of personnel, guidance, and training for librarians.
A number of statements addressed the lack of manpower and insufficient guidance regarding OA: "In addition to the manpower we lack, we lack professional training to learn about OA, so that we can be professional while conveying the information to researchers and management" (L8).

9. Opposition from stakeholders and researchers to OA promotion.
Librarians have reported opposition regarding OA, resulting in conflicts between researchers and the management at their institution: "Unfortunately, some of the researchers are employees of some of journals' publishers, and they resist promoting open-access journals. This is a blatant intervention by stakeholders in academia" (L10).

Factors and Requirements Needed to Promote Open Access

The third research question examined what is needed to promote OA publication. This category includes 295 statements (23% of the total). The librarians gave six factors and requirements to promote OA.

Table 4 presents the sub-categories according to their frequency.

TABLE 4
Factors and Requirements Needed to Promote Open Access (N = 295, 23%)

Subcategory	Total Number of Statements Including	
	N = 295	%
1. The need to change policy at the national level and adopt advanced OA policy	89	30%
2. The need for cooperation between academic institutions to establish new agreements and an institutional repository	60	20%
3. Collaboration of the University management, research authority, and the library for the advancement of OA	50	17%
4. The need for tutorials and training for library staff regarding OA publishing	49	17%
5. The need to guide researchers and reward them for open access publications	32	11%
6. The need to establish the library's formal status as a central body for open access care	15	5%

Analysis of the sub-categories and representative quotes are listed below.

1. The need to change policy at the national level and adopt an advanced OA policy.
The largest number of statements addressed the need to change policy: "Why would researchers consider publishing in OA? It should be a policy. The breakthrough of OA will come from a national policy, as exists in many other countries in Europe and United States. Currently, we are ' tilting at windmills'" (L1).

2. The need for cooperation between academic institutions to establish new agreements and an institutional repository.
Cooperation between all academic institutions in Israel would create a greater advantage to negotiate with the publishers: "It's too big for each university individually. If all universities are together, they will consist one incorporated group that can negotiate with publishers" (L2). In addition, the librarians mentioned the establishment of a common IR: "If there was a common institutional repository for all institutions, it would give researchers an alternative to the publishers" (L10).

3. Collaboration between the University management, research authority, and the library for advancing OA.
Collaboration between university authorities will help promote OA open access: "The university management and the research authority must cooperate with us (the library) if we really want to promote OA" (L8).

4. The need for tutorials and training for library staff on OA publishing.
Librarians mentioned the need for professional development regarding OA: "The library staff must specialize in OA. We must be ready to guide and advise both researchers and management" (L4).

5. The need to guide researchers and reward them for open access publications.
Librarians argued that researchers need guidance concerning high IF OA journals, and their added value: "Researchers need guidance because they do not understand the value in publishing OA, they wonder why do they need it. And even if they do understand, it is not certain that they will publish on their own initiative" (L6). Therefore, librarians offer to provide funding for researchers: "The Higher Education Council should budget OA publications, and at the same time oblige researchers to self-archive pre-post version in open repository" (L3).

6. The need to establish the library as a central body for Open Access.
Librarians suggested that the library would take a major role and handle everything related to OA: "It would be correct if we as the library centralize the issue of OA publication and not the dean. If researchers need help, they should contact us because we work with publishers" (L9).

Discussion

Academic libraries have the expertise and mindset to be early adopters of new technologies such as digital curation, digital preservation, digital archiving, and more.[43] This study examined Israeli librarians' perceptions regarding their role and the academic library's role

in promoting OA publication, the barriers, challenges, needs, and requirements needed to promote OA publishing in their institutions as well as at the national level. Interpretation of the findings was presented in an integrative way.

It reveals that librarians perceive themselves as being at the forefront of promoting OA in their institutions. However, they emphasized that they are not the first or only persons to be responsible for promoting OA publishing. They assert that since there is not a definite policy towards OA in Israel, it is the national government's and the university administration's role to create a regulated policy towards OA. Once the policy is outlined, they will be the ones to guide, recommend, educate, and implement the changes. A number of studies evaluated the involvement of countries in the international OA movement, and in particular examined the distribution of the number of OA repositories, OA journals, institutional OA policies, and OA articles among selected countries.[44] Studies found a positive relationship between countries' involvement in OA and the proportion of research outputs published in gold/green OA.[45]

Librarians perceive the relationship with the university management as key to promoting OA. However, they emphasized that a lack of cooperation with management, the research authority, and the rector does not enable a substantial advancement in agreements with OA components or promotion of an institutional repository. Librarians also reported that sometimes, due to conflicts of interest, management actively opposes librarians' initiatives. Therefore, and as found in past studies,[46] cooperation with the university administration is a necessary condition for promoting OA publishing and strengthening the library's role in it.

The lack of budget for OA agreements with publishers is the greatest barrier according to librarians in this and previous studies.[47] Librarians feel frustrated by the mega-agreements with publishers that leave them no budget for further agreements with OA journals. Further, they added that they have no budget for researchers who approach them for help in funding APCs to publish in an OA journal. Therefore, to confront the budget barrier, they are involved in supporting activities carried out through the MALMAD consortium aimed at promoting OA, even without the support of university management. Librarians assume that their involvement in two major current projects will lead to OA awareness. The first is the CRIS database. Librarians assert that in order to recommend and advise researchers and university management regarding publishing in OA, they need information about all researchers' publications. CRIS assists them in achieving this goal by centralizing publications and performing data analysis.[48] The second system is the institutional repository (IR), which was mentioned in other studies as contributing greatly to the OA movement.[49] Librarians note that researchers ask them to deposit their work in an institutional repository based on requests from funding agencies. Therefore, they suggest establishing a shared institutional repository for all academic institutions, which would emphasize cooperation between academic institutions. According to the librarians, the collaboration of all academic institutions in Israel will result in a consolidation of forces and a better position to negotiate with the publishers.

Librarians see great importance in their role of guiding and advising researchers regarding OA publishing. One of the biggest barriers in promoting OA is researchers' lack of awareness concerning OA in general, and in their field of study. Further, researchers fail to distinguish between legitimate OA journals and predatory journals. Librarians mention the journal impact factor as one of the main barriers to OA promotion. The journal IF index has a broad

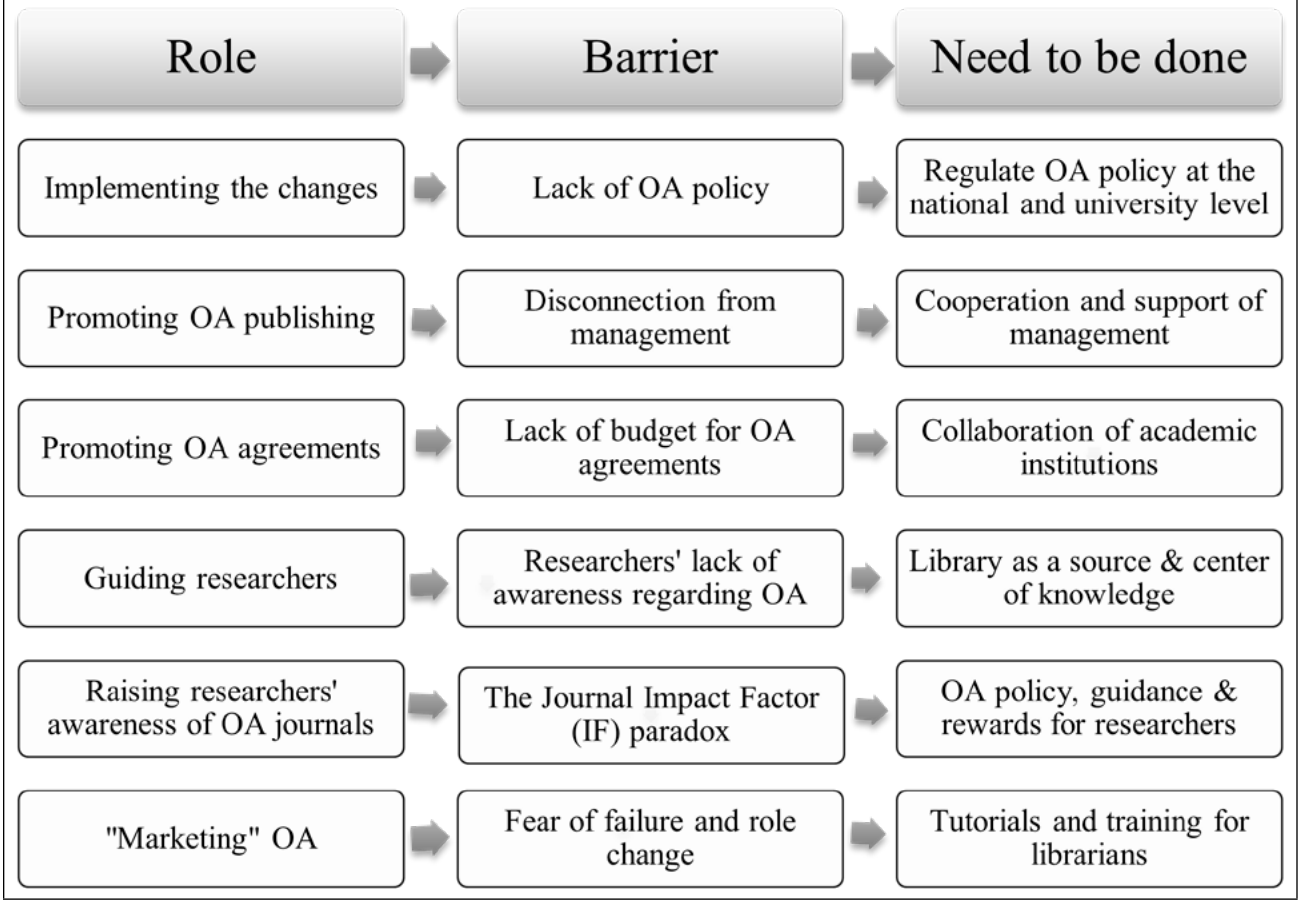
and long-term impact on research institutions and researchers. In most academic disciplines, researchers have to publish in journals with a high IF in order to succeed, especially for those on a tenure track. In many cases, OA journals have a less established IF.[50]

Thus, due to these considerations, junior academics have less experience with OA journals.[51] In addition, researchers, according to librarians, do not know the copyright terms of publishers and therefore avoid self-archiving. They are also concerned about depositing a preprint version for fear of "scooping" (i.e., that someone will steal their research idea). Moreover, librarians add that researchers are afraid of changing the traditional publishing model and need guidance adjusted for their discipline. Therefore, and as found in other studies, librarians consider their role as facilitating and guiding proper publication in OA.[52]

Open access is transforming scholarly communication. Various modes of OA include: gold, hybrid, delayed, bronze, institutional and subject-based repositories, and others, which reflect the complexity of OA.[53] Thus, new challenges emerge for academic library faculty that require investing in developing skills and continuous improvement.[54] With professional development, support and proper guidance, librarians will be able to be real promoters and leaders for OA in their institutions.

To summarize librarians' findings and to outline what can be drawn from this study, Figure 1 shows the roles, barriers, and what is necessary to promote OA:

FIGURE 1
OA Promotion: Roles, Barriers, and What Is Needed to Promote OA



Librarians referred in this study to several stakeholders who can promote open access: at the state level, at the universities management level, and at researcher, and academic library level. Librarians see great importance in their role of advising researchers regarding issues in OA publishing. In order to provide the appropriate training, librarians acknowledge that they need professional development in the various aspects of OA. At the universities management level—the findings indicate the need for support in libraries' activities in relation to OA, but also the need for collaboration between the academic institutions to promote the publication of OA. However, librarians insisted on a regulated OA-policy at the national and institutional levels, which would strengthen their status as change-leaders of the OA-movement. Finally, as Figure 1 indicates, academic libraries have the opportunity to contribute to the adoption of OA and change their traditional roles, provided they get the support they need.

Conclusions and Future Work

Over a decade ago, and 10 years after the BBB declarations, some scholars argued that the growth in OA publication is encouraging. Considering the indicators of progress made by the OA movement against the obstacles in the first decade, there is a reason for great optimism for the next decade.[55] Now, 20 years after the BBB declarations, the struggle continues and OA publishing is not yet the norm in some countries and academic institutions. According to the librarians, the university administration and researchers are not aware of the potential of open access publishing. To encourage the adoption of OA practices—publishing in OA journals, depositing in OA repositories—advocacy is important but insufficient. Librarians require a set of regulated and legal OA policies. Otherwise, they are "tilting at windmills." Much is written in the literature about the difficulty of adapting to the change of roles with the transition to the digital world.[56] The librarians in the current study are willing to lead the change, guide researchers, and support the transition to the OA publishing model. Yet, they need the strengthening that comes from regulated policies at the national level, as well as public support from academic institutions' management of their libraries.

The current study has some recommendations for future research. Future studies may expand the sample and crosscheck librarians' perspectives with other quantitative methods. In addition, future studies should include and examine researchers' and policymakers' perspectives too.

Acknowledgements

The authors thank the academic libraries who volunteered to participate and assist in our study.

Notes

1. Heather Piwowar, Jason Priem, Vincent Larivière, Juan Pablo Alperin, Lisa Matthias, Bree Norlander, Ashley Farley, Jevin West, and Stefanie Haustein, "The State of OA: A Large-Scale Analysis of the Prevalence and Impact of Open Access Articles," *PeerJ* 6 (2018): e4375, https://doi.org/10.7717/peerj.4375.

2. Peter E. Clayson, Scott A. Baldwin, and Michael J. Larson, "The Open Access Advantage for Studies of Human Electrophysiology: Impact on Citations and Altmetrics," *International Journal of Psychophysiology* 164 (2021): 103–11, https://doi.org/10.1016/j.ijpsycho.2021.03.006.

3. Aled Edwards, "Perspective: Science is Still Too Closed," *Nature* 533, no. 7602 (2016): S70, https://doi.org/10.1038/533S70a.

4. Bo-Christer Björk, "The Open Access Movement at a Crossroads—Are the Big Publishers and Academic Social Media Taking Over?," *Learned Publishing* 29, no. 2 (2016): 131–34, https://doi.org/10.1002/leap.1021.

5. Vladimir M. Moskovkin, Tatyana V. Saprykina, Marina V. Sadovski, and Olesya V. Serkina, "International

Movement of Open Access to Scientific Knowledge: A Quantitative Analysis of Country Involvement," *The Journal of Academic Librarianship* 47, no. 1 (2021): 102296.

6. Elizabeth Tait, Konstantina Martzoukou, and Peter Reid, "Libraries for the Future: The Role of IT Utilities in the Transformation of Academic Libraries," *Palgrave Communications* 2.1 (2016):1–9, https://doi.org/10.1016/j.acalib.2020.102296.
7. Jeroen Bosman, and Bianca Kramer, "Open Access Levels: A Quantitative Exploration Using Web of Science and OA DOI Data," e3520v1, *PeerJ Preprints*, 2018, https://doi.org/10.7287/peerj.preprints.3520v1.
8. Jeffrey Beall, ""Predatory" Open-Access Scholarly Publishers," *The Charleston Advisor* 11, no. 4 (2010): 10–17, https://doi.org/10.5260/chara.12.1.50.
9. Tait et al., "Libraries for the Future," 1–9.
10. Mohamed Boufarss, and J. Tuomas Harviainen, "Librarians as Gate-Openers in Open Access Publishing: A Case Study in the United Arab Emirates," *The Journal of Academic Librarianship* 47.5 (2021): 102425, https://doi.org/10.1016/j.acalib.2021.102425.
11. Meghana Manohar Sanjeeva, and Sushama C Powdwal, "Open Access Initiatives: Reframing the Role of Librarians," *Library Herald* 55.4 (2017): 467–487, https://doi.org/10.5958/0976-2469.2017.00037.9; Wm. Joseph Thomas, "The Structure of Scholarly Communications within Academic Libraries," *Serials Review* 39.3 (2013): 167–171, https://doi.org/10.1016/j.serrev.2013.07.003.
12. Boufarss and Harviainen, "Librarians as Gate-Openers in Open Access Publishing," 102425.
13. Tait et al., "Libraries for the Future," 1–9.
14. Ada Ducas, Nicole Michaud-Oystryk, and Marie Speare, "Reinventing Ourselves: New and Emerging Roles of Academic Librarians in Canadian Research-Intensive Universities," *C&RL* 81.1 (2020): 43–65, https://doi.org/10.1016/j.acalib.2021.102425.
15. Ducas et al., "Reinventing Ourselves," 43–65.
16. Tlou Maggie Masenya, and Collence Takaingenhamo Chisita, "Futurizing Library Services in a Technology-Driven Dispensation: Reflections On Selected Academic Libraries in Zimbabwe and South Africa," In Tlou Maggie Masenya (ed.), *Innovative Technologies for Enhancing Knowledge Access in Academic Libraries* (pp. 1–21), IGI-Global, https://doi.org/10.4018/978-1-6684-3364-5; Tait et al., "Libraries for the Future" 1–9.
17. Ducas et al., "Reinventing Ourselves" (2020): 43–65.
18. Peter Suber, "Open Access Overview" (2007), accessed July 10, 2022, https://philpapers.org/rec/SUBOAO.
19. Powdwal, "Open Access Initiatives."
20. Piwowar et al., "The State of OA," 4.
21. Aled, "Perspective," S70.
22. Juliet Sherpa, "About Sherpa Juliet," accessed July 10, 2022, https://v2.sherpa.ac.uk/juliet/about.html.
23. Jeffrey, "Predatory," 10-17.
24. Carol Tenopir, Elizabeth D. Dalton, Lisa Christian, Misty K. Jones, Mark McCabe, MacKenzie Smith, and Allison Fish, "Imagining a Gold Open Access Future: Attitudes, Behaviors, and Funding Scenarios Among Authors of Academic Scholarship," *College and Research Libraries* 78 no. 6 (2017), https://doi.org/10.5860/crl.78.6.824.
25. Piwowar et al., "The State of OA," 2.
26. Tenopir et al., "Imagining a Gold Open Access Future," 6.
27. Tait et al., "Libraries for the Future."
28. Suber, "Open Access Overview"; Wm Joseph Thomas, "The Structure of Scholarly," 167-171.
29. Christine Antiope Daoutis, and Maria de Montserrat Rodriguez-Marquez, "Library-Mediated Deposit: A Gift to Researchers or a Curse on Open Access? Reflections from the Case of Surrey," *Publications* 6.2 (2018): 20, https://doi.org/10.3390/publications6020020; Suber, "Open Access Overview."
30. Emma Cryer, "Incorporating Open Access Into Libraries," *Serials Review* 37, no. 2 (2011): 103-107; Kelemwork Kassahun and Chatiwa Nsala, "The Awareness of Academic Librarians Towards Open Access Resources to Support Reference Services: A Case of Private Institutions of Higher Learning in Gaborone," Botswana, IFLA World Library and Information Congress (2015), 1-11.
31. L. B. Mullen, "Open Access and the Practice of Academic Librarianship: Strategies and Considerations for 'Front Line' Librarians," in *IATUL Proceedings* (2011), IATUL, https://bit.ly/3H8JvOb.
32. Andrew M. Cox, Stephen Pinfield, and Jennifer Smith, "Moving a Brick Building: UK Libraries Coping with Research Data Management As A 'Wicked' Problem," *Journal of Librarianship and Information Science* 48.1 (2016): 3-17.; Mussarat Sultan and Muhammad Rafiq, "Open Access Information Resources and University Libraries: Analysis of Perceived Awareness, Challenges, and Opportunities," *The Journal of Academic Librarianship* 47.4 (2021): 102367.
33. Daoutis and de Montserrat Rodriguez-Marquez, "Library-mediated Deposit," 20.
34. Amy VanScoy, "Bridging the Chasm: Faculty Support Roles for Academic Librarians in the Adoption of

Open Educational Resources," *College & Research Libraries* 80.4 (2019): 426; Daoutis and de Montserrat Rodriguez-Marquez, "Library-mediated Deposit," 20.

35. Justin Fuhr, "Developing Data Services Skills in Academic Libraries," *College & Research Libraries* 83.3 (2022), https://doi.org/10.5860/crl.83.3.474.

36. Joseph, "The Structure of Scholarly Communications," 167-171.

37. Sandy Q. Qu and John Dumay, "The Qualitative Research Interview," *Qualitative Research in Accounting & Management* 8, no. 3 (2011): 238-264.

38. Monique M. Hennink, Bonnie N. Kaiser, and Vincent C. Marconi, "Code Saturation Versus Meaning Saturation: How Many Interviews Are Enough?" *Qualitative Health Research* 27, no. 4 (2017): 591-608.

39. Kristi L. Palmer, Emily Dill, and Charlene Christie, "Where There's A Will There's A Way?: Survey of Academic Librarian Attitudes About Open Access," *College & Research Libraries* 70, no. 4 (2009): 315-335, https://doi.org/10.5860/0700315.

40. Gay Helen Perkins and Amy JW Slowik, "The Value of Research in Academic Libraries," *College & Research Libraries* 74, no. 2 (2013): 143-158, https://doi.org/10.5860/crl-308.

41. Virginia Braun and Victoria Clarke, "Using Thematic Analysis In Psychology," *Qualitative Research in Psychology* 3, no. 2 (2006): 77-101, https://doi.org/10.1191/1478088706qp063oa.

42. Yvonna S. Lincoln and Egon G. Guba, *Naturalistic Inquiry*, Sage, 1985.

43. Gurdish Sandhu, "The Role of Academic Libraries in the Digital Transformation of the Universities," *5th International Symposium on Emerging Trends and Technologies in Libraries and Information Services (ETTLIS*, 2018), IEEE, 2018, https://doi.org/10.1109/ETTLIS.2018.8485258.

44. Moskovkin et al., "International Movement of Open Access to Scientific Knowledge," 102296.

45. Chun-Kai Huang, Cameron Neylon, Richard Hosking, Lucy Montgomery, Katie S. Wilson, Alkim Ozaygen, Chloe Brookes-Kenworthy, "Meta-Research: Evaluating the Impact of Open Access Policies on Research Institutions," *eLife* (2020)9: e57067, https://doi.org/10.7554/eLife.57067.

46. Bhuva Narayan and Edward Luca, "Issues and Challenges in Researchers' Adoption of Open Access and Institutional Repositories: A Contextual Study of a University Repository," in *Proceedings of RAILS (Research Applications, Information and Library Studies),* School of Information Management, Victoria University of Wellington, New Zealand, 6-8 December, 2016; *Information Research* (2017) 22(4), http://InformationR.net/ir/22-4/rails/rails1608.html.

47. Heather Morrison, "Economics of Scholarly Communication in Transition," *First Monday* (2013). https://doi.org/10.5210/fm.v18i6.4370.

48. David Walters, Monique Ritchie, and Megan Kilb, "CRIS Power! Taming the Reporting Requirements of Open Access," *The Serials Librarian* 70 no. 1-4 (2016): 229-235, https://doi.org/10.1080/0361526X.2016.1160306.

49. Teresa Auch Schultz and Elena Azadbakht, "Open but Not for All: A Survey of Open Educational Resource Librarians on Accessibility," *College & Research Libraries*, *82*(5): 755, https://doi.org/10.5860/crl.82.5.755.

50. Cassidy Sugimoto, Liz Allen, Bosman Jeroen, Tindaro Cicero, Stephen Curry, Sarah de Rijcke, Annette Flanagin, et al., "Rethinking Impact Factors: New Pathways In Journal Metrics," *F1000Research* 8, no. 671 (2019): 671, doi/org.10.7490/F1000RESEARCH.1116751.1.

51. Yimei Zhu, "Who Support Open Access Publishing? Gender, Discipline, Seniority and Other Factors Associated with Academics' OA Practice," *Scientometrics* 111, no. 2 (2017): 557-579, https://doi.org/10.1007/s11192-017-2316-z.

52. Christine Antiope and de Montserrat Rodriguez-Marquez, "Library-mediated Deposit," 20.

53. Keiko Kurata, Keiko Yokoi, Tomoko Morioka, Yukiko Minami, and Masashi Kawai, "Monitoring the Transition to Open Access through Its Mode of Implementation: A Principal Component Analysis of Two Surveys," *PloS One* 17.7 (2022): e0271215, https://doi.org/10.1371/journal.pone.0271215.

54. Ducas et al., "Reinventing Ourselves"; Tait et al., "Libraries for the Future."

55. Heather Joseph, "The Open Access Movement Grows Up: Taking Stock of a Revolution," *PloS Biology* 11.10 (2013): e1001686, https://doi.org/10.1371/journal.pbio.1001686.

56. Ducas et al., "Reinventing Ourselves"; Perkins and Slowik, "The Value of Research in Academic Libraries"; Gurdish. "The Role of Academic Libraries" (2018).

Exploring Social Media as an Information Source in IL Instruction

Kathia Salomé Ibacache Oliva, Elizabeth Novosel, and Stacy Gilbert

According to a 2021 Pew Research report, over 80 percent of 18- to 29-year-olds use social media. Studies also show that higher education students use social media in both academic and everyday life. However, there is minimal research about how, or whether, librarians utilize social media in their library instruction as a source of information for students' academic work. We examined 162 responses to a survey sent to an uncounted number of librarians in higher education, asking them about their teaching practices regarding social media to enable us to answer the following research questions: (1) How are teaching librarians using social media as an information source in their instruction?; (2) What are the benefits teaching librarians perceive regarding the use of social media in instruction?; (3) What are the challenges teaching librarians perceive regarding the use of social media in instruction? The survey results showed that most librarians surveyed are not utilizing social media as an information source in their library instruction. Although our results cannot be generalized, our study sheds light on how librarians incorporate social media in information literacy (IL) instruction, the tension between scholarly literature and voices not considered authoritative, and librarians' perceptions of benefits and challenges to incorporating social media in library sessions.

Introduction

In today's world, information moves fast, circulated by digital technologies such as social media. The first social media platforms, as we know them today, appeared in the late 1990s;[1] this led to dozens of sites captivating millions of people by the early 2000s, propelled by the emergence of Web 2.0 technologies. Today, social media is very much present in the lives of higher education students. According to a 2021 Pew Research report, over 80 percent of 18- to 29-year-olds use social media.[2] With people and organizations worldwide sharing information of varied authority, there is both potential—and a need—for librarians to teach students how to use social media critically to discover and access information for academic assignments. Addressing a gap in the literature, this paper explores whether and how teaching librarians utilize

* *Kathia Salomé Ibacache Oliva is Romance Languages Librarian and Assistant Professor at University of Colorado Boulder, email: Kathia.Ibacache@colorado.edu; Elizabeth Novosel is Faculty Fellow and Liaison & Projects Librarian at University of Colorado Boulder, email: elizabeth.novosel@colorado.edu; Stacy Gilbert is Social Sciences Librarian and Assistant Professor at University of Colorado Boulder, email: stacy.gilbert@colorado.edu.*

social media as a pedagogical tool in their library instruction to help learners find sources of information for their academic work. Additionally, our study reports on librarians' perceptions of the benefits and challenges of using social media in library sessions.

Inspired by a faculty member's request for a library session covering hashtags and social media skills to help students find reliable information sources, this exploration also responds to changes we have observed in higher education attitudes toward including underrepresented voices in addition to traditional scholarly perspectives in the academic narrative. First, we will summarize scholarly literature regarding college students' use of social media as an information source, social media in higher education classrooms, and social media within library instruction. Second, we will present data gathered through a survey. Third, we will discuss the data in relation to our research questions and address possible benefits and challenges of using social media in library instruction.

What is Social Media in this Study?

This paper follows Andreas M. Kaplan and Michael Haenlein's definition of *social media*: "a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content."[3] We favored this definition because it emphasizes the notion of user-generated content. The eleven social media platforms utilized in this study were taken from a Pew Research Center article on social media use in 2021. The platforms are: YouTube, Facebook, Instagram, Pinterest, LinkedIn, Snapchat, Twitter, WhatsApp, TikTok, Reddit, and Nextdoor.[4]

How is "Information Source" Defined in this Study?

In the context of social media, we understand an information source to be a primary or secondary source used for academic research or for assignments in a higher education setting. For example, students may access a report from a governmental office's social media account, or may find a scholarly article posted by an author or an academically affiliated account.

Literature Review

There is abundant literature about the use of social media in different academic fields.[5] However, research about teaching librarians' utilization of social media as a source of information within the context of IL instruction is scarce.[6] In a 2011 paper examining social media in the research workflow, David Nicholas expressed concern about librarians' disconnect with social media in the academic setting because scholars were using social media to share their research.[7] His study surveyed 2,414 academics in 215 countries and found that researchers strongly agree that social media enhances their academic work "through the greater visibility it affords them."[8] Nonetheless, a decade later, academic librarians still appear to be unlikely to introduce social media in IL instruction beyond minimal coverage of evaluation, shared content, online collaboration, and observation of informal scholarly interaction. This literature review considers studies addressing how students use social media as an information source for academic work, as well as social media applications in the higher education classroom.

Students' Use of Social Media as an Information Source

Within the last decade, a few studies have noted that students use social media as a source of both background information and news for their academic work.[9] Referring to student use of

social media, both for information seeking and academic purposes, librarians Kyung-Sun Kim, Yuqi He, and Sei-Chin Joanna Sin noted in 2013 that, while undergraduate students utilized media sharing sites such as Social Q&A and Wikis, graduate students preferred blogs.[10] These authors observed that students' majors also determined the social media platforms they utilized.[11]

Similarly, in 2014, Kim, Sin, and Eun Young Yoo-Lee found that students use some social media platforms as information sources, including Wikipedia and social networking sites.[12] Recently, a 2021 literature review examining students' preferences on reading formats and the use of social media information sources for academic and non-academic purposes noted that college students use social media for "class assignments and projects."[13] This review added that students use YouTube as a supplement for textbooks, blogs to follow topics of interest, and LinkedIn to check the authority of authors.[14]

Interestingly, a 2021 paper—in which Kim, Sin, and Yoo-Lee compared two undergraduate cohorts' (2013 and 2019) use of social media as an information source for academic and everyday purposes—found data that contradicts that of the Pew Research Center. Kim et al. found that the 2019 cohort utilized social media as an information source less frequently than the 2013 cohort did, suggesting a decline in use,[15] while a Pew Research Center study showed the opposite, reporting increased use of social media in 18- to 29-year-olds in 2021.[16] Although the Pew Research Center report did not specify data on the use of social media as an information source for academic purposes, the diverging data between these two sources is intriguing.[17] Perhaps future research could address this dichotomy.

Social Media in the Higher Education Classroom

The literature reveals that social media has been incorporated into higher education classrooms for different purposes since the 2000s.[18] Perhaps due to students' familiarity with social media, instructors often report using it to communicate with students, instructors often report using it to communicate with students, increase class participation, promote engagement with course topics, encourage further discussion outside the classroom environment, and build a sense of community among participants.[19] For example, Anatoly Gruzd et al. note that social media promotes "networked learning," in which students connect beyond classmates and instructors to broader academic and social communities with shared interests.[20] Scholars also observe that social media's popularity has grown for "formal and informal learning in many disciplinary areas."[21]

In addition, some studies address the advantages of including social media in the classroom, such as faculty exposing students to sources outside the traditional academic setting, and "promoting learning through social interaction and collaboration."[22] Of particular interest to our study is the idea of expanding the learning environment, which Gruzd et al. align with "discovery," permitting students and instructors to find and access resources outside traditional settings.[23]

Another social media use in the higher education classroom relates to platform data analysis and career skills development. For example, Michele Ewing, et al. write about the importance of social media analytics for students in the public relations field, arguing that practical experience using tools and methods is vital for student success.[24] Similarly, other authors argue in favor of teaching strong social media skills in journalism programs, addressing news coverage and competencies to create content.[25] This disciplinary interest in various aspects of social media suggests there is value in including social media in IL instruction to support course curricula.

Social Media in Library Instruction

Social media literature connected to information literacy emerged in the early 2000s, showing that librarians have utilized social media to help students understand the quality of sources, share content, promote online collaboration, and discuss informal scholarly interaction.[26] However, as mentioned earlier, not many studies provide data on teaching librarians' incorporation of social media as an information source in IL instruction.[27] One study—examining Mississippi academic librarians—found that, even when these librarians concurred that social media was an essential component of IL instruction that would improve students' ability to consume, disseminate, and create information, fewer than half of them attempted or planned to use social media in their library sessions.[28]

Benefits of Using Social Media as an Information Source in Instruction

Some librarians see benefits to using social media in instruction. Sheila Stoeckel and Caroline Sinkinson emphasize the value of social media concerning its, "participatory nature of research and knowledge negotiation," as opposed to the passive collection of resources.[29] Additionally, Natalie Burclaff and Catherine R. Johnson invite librarians to go beyond the demonstration of databases to incorporate social media as a research tool to access different perspectives and voices.[30] Kim et al. also recognize that social media provides information not found in traditional academic sources, offering a broader range of voices and opinions to consider.[31] Other librarians encourage using social media as a source of information based on its ability to disseminate information, the usefulness of hashtags to cover subject headings, and its ability to engage students in the research process.[32]

Students' familiarity with social media is another potential benefit that some authors consider when addressing social media in instruction. Scholars note that students' comfort and previous experience with social media help them to practice critical thinking and become a part of informal scholarly interaction by participating in discussions in library sessions.[33] Critical thinking is crucial for students when evaluating information and sources. In a discussion about the Association of College & Research Libraries' *Framework* and metaliteracy, Valerie J. Hill and Thomas P. Mackey highlight the initial inception of metaliteracy to promote critical thinking and collaboration in social media and other digital communities.[34] The authors add that a primary goal for metaliteracy is for learners to, "actively evaluate content while also evaluating one's own biases."[35] Additionally, Casey Fiesler—addressing information reliability in news disseminated on TikTok—asserts that this platform may provide accurate information, but advises us to, "consider the credibility of sources and information we choose to believe and share."[36]

Challenges of Using Social Media as an Information Source in Instruction

Some authors have expressed concerns about social media as an information source. These concerns include privacy and ethical implications, disinformation, and instructor competency.[37] Alison Hicks and Caroline Sinkinson discuss privacy issues within active learning activities, cautioning that digital tools may expose learners to tracking and surveillance.[38] Regarding personal data shared on social media, Lucy Pangrazio and Neil Selwyn emphasize the importance of including social and ethical considerations in "digital literacy" instruction, noting that students need to know how to engage with social media responsibly.[39]

Other scholars have voiced concerns about the quality of information available on social media. Donna Witek and Teresa Grettano recognized that social media is an environment

where "formal and informal sources of information" meet, and where it is important to teach students how to evaluate information.[40] This concern results from information being generated within a social context that is "directly related to how others in their networks value that information and whether the information is relevant to their friends or to themselves."[41]

This literature review addressed college students' use of social media for academic purposes, social media use in the classroom, and favorable and critical views on integrating social media in instruction. Our study seeks to address the gap in the literature, and to advance the scholarship by shedding light on whether teaching librarians utilize social media as a pedagogical tool in IL instruction, as well as considering how teaching librarians perceive the benefits and challenges of teaching social media in library sessions.

Methodology

This research paper seeks to answer the following questions:

1. How are teaching librarians using social media as an information source in their instruction?
2. What are the benefits teaching librarians perceive regarding the use of social media in instruction?
3. What are the challenges teaching librarians perceive regarding the use of social media in instruction?

Instrument

We used a thirty-five-question Qualtrics survey approved by the University of Colorado Institutional Review Board. The branched survey had thirty-three multiple-choice questions and two free-text questions. The survey was open from August 13th to October 6th, 2021. Although 162 respondents filled out the survey, the number of individuals who completed each question varied. The data collected were quantitative and qualitative. The closed-ended questions gathered data on the participants' institutional affiliations and responsibilities as academic librarians. The multiple-choice questions gathered data on social media platforms used in library instruction sessions, frequency of use, and type of use. The survey had two hybrid questions and two fully qualitative questions. Questions 32 and 33 utilized a hybrid question design, where partially close-ended questions allowed respondents to select from the options provided and write additional comments using the field "Other." Questions 32 and 33 gathered data on the challenges and benefits of using social media in IL instruction. Questions 34 and 35 collected qualitative information in free-text boxes about librarians' perceptions of faculty's reaction to social media in IL instruction. We organized the data in a spreadsheet for analysis. See the Appendix for the complete survey instrument.

Participants

We distributed the survey via email to the following disciplinary listservs: the Association of College and Research Libraries, History Librarians Interest Group, Education and Behavioral Sciences Section, Literature in English Section, the Seminar on the Acquisition of Latin American Library Materials, the Council on East Asian Libraries, the Collaborative Initiative for French and North American Libraries, the Colorado Council of Medical Librarians, and the Engineering Library Division. We chose these organizations due to their connection with library instruction in academic libraries. Due to low levels of participation, we also distributed

the survey directly through email to 139 librarians with teaching responsibilities at eleven universities.[42] We searched the directories of these university libraries, identified all librarians who appeared to have teaching responsibilities, and emailed them directly.

Our Approach to the Data

Our study used a survey research approach. We identified our target audience as librarians with teaching responsibilities. We designed the questionnaire and piloted it with three instruction librarians who provided feedback to improve this instrument. Rather than approaching the data with a pre-existing hypothesis, the two main authors conducted three rounds of coding. First, we individually identified keywords from each response to the four questions with qualitative data and organized these keywords (codes) in a spreadsheet. Second, we compared our keywords and agreed on themes. Third, we reviewed our data again and allocated each response to the thematic categories we created.

Results

Respondent Demographics

Table 1 summarizes the relationship between social media use and respondents' background information. We separated responses into "Yes" and "No" categories of social media use, based on whether the respondent selected the option: "I have not used social media platforms in my instruction." This table shows that 110 of 139 respondents work in a university that confers doctoral degrees, with only 29 respondents at other types of institutions. Table 1 also shows how respondents identified the nature of their positions, with a large representation of subject specialists followed by librarians who categorized their positions as instruction and reference. The largest percentage of librarians who used social media were those who identified their job as involving "outreach." However, this group represents a smaller pool of respondents compared to subject specialists or those who categorized their position as "instruction."

In addition, respondents reported supporting a wide variety of departments and programs, ranging from the sciences and humanities to the social sciences. The small percentage of respondents who supported departments related to communication, journalism, and media studies utilized social media more frequently than those supporting other academic programs. While it is interesting to note some of these trends in the data, we make no statistical claims about the relationship between respondents' background information and their social media use. Additional data collection is necessary to ensure adequate subgroup sample sizes before this analysis is useful.

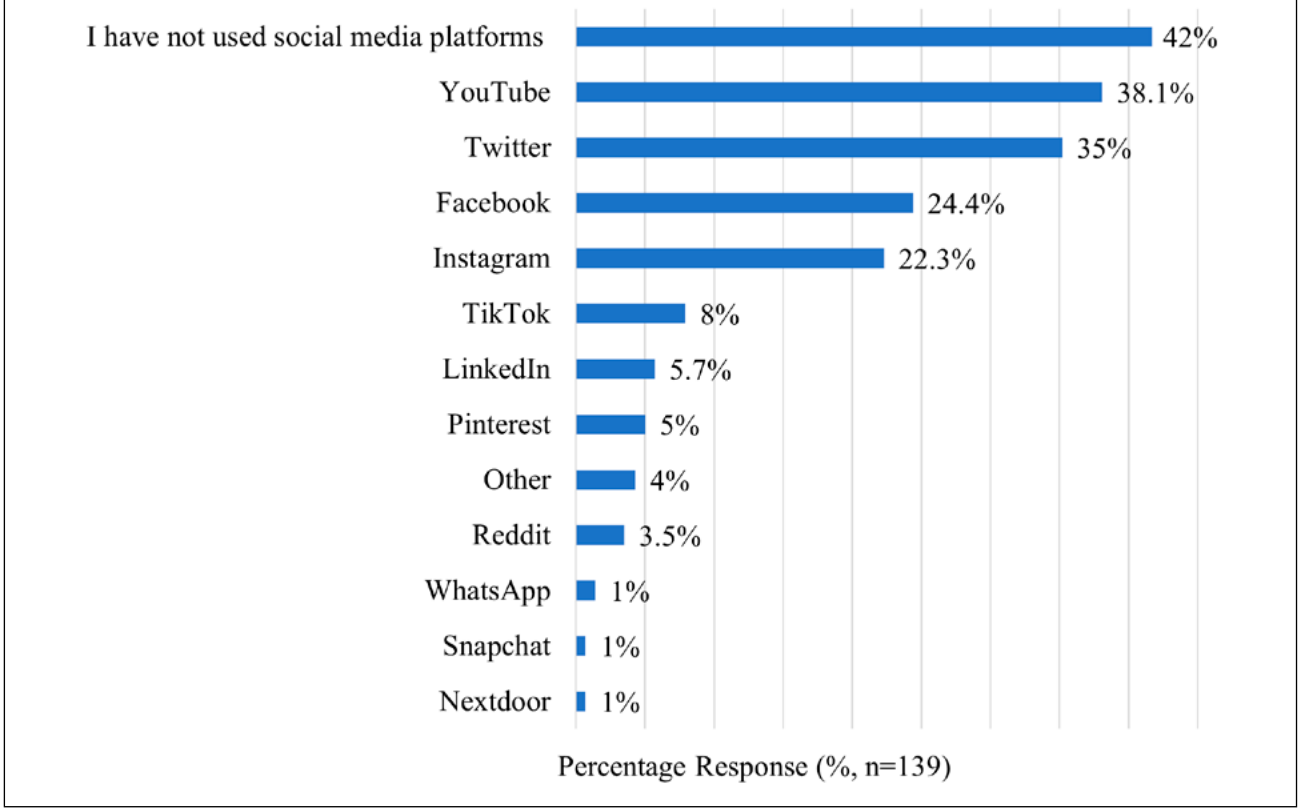
TABLE 1
Summary of Background Information and Social Media Use

Background information		***Do you use social media?***	
Do you have one shot or course instruction responsibilities? (Select all that apply)		*Yes*	*No*
	Yes, courses (multiple related sessions) (n = 55)	42	13
	Yes, one-time sessions (workshops, one-shots, webinars, etc.) (n=84)	39	45
What category best describes your institution?		*Yes*	*No*
	Doctorate University (n = 110)	60	50
	Other (n = 29)	21	8

TABLE 1
Summary of Background Information and Social Media Use

Background information	***Do you use social media?***	
How long have you taught library instruction sessions in a higher education context?	*Yes*	*No*
0-2 years (n = 15)	7	8
3-5 years (n = 32)	23	9
6-10 years (n = 37)	25	12
11 or more years (n = 55)	26	29
How would you categorize your position? (Select all that apply)	*Yes*	*No*
Subject specialist, liaison (n = 113)	65	48
Instruction (n = 89)	53	36
Reference (n = 77)	46	31
Outreach (n = 45)	32	13
Acquisition, collection development, management, strategy (n = 39)	26	13
Scholarly communication (n = 21)	14	7
Digital initiatives, systems information, technology, web development (n = 13)	9	4
Data services, GIS (n = 10)	7	3
Assessment, analytics, user experience (n = 10)	7	3
Archiving, curational, rare books, preservation, conservation (n = 9)	7	2
Access Services (n = 10)	9	1
Other (n = 12)	8	4
Do you support specific academic programs or departments? (Select all that apply)	*Yes*	*No*
Math/ Computer Science/ Statistics/ Technology/ Engineering (n = 38)	21	17
Humanities (n = 31)	18	13
Art/Art History/Theater/Music/Film (n = 27)	20	7
Social Sciences (n = 23)	15	8
Health/Medicine/Nursing (n = 22)	13	9
Natural Sciences (n = 22)	14	8
Literature (n = 21)	13	8
Languages (n = 19)	14	5
Business (n = 13)	10	3
Communication/Journalism/Media Studies (n = 14)	12	2
History (n = 32)	24	8
Undergraduate Writing and Rhetoric (n = 11)	7	4
Education (n = 9)	5	4
Other (n = 37)	20	17
I do not liaise with any departments (n = 12)	9	3

FIGURE 1
Which Social Media Platforms have you Used in your Instruction Sessions? (Select all that Apply)



Use of Social Media in Instruction

Figure 1 shows the social media platforms respondents used in their instruction sessions. Data indicate that 42 percent of 139 respondents do not use social media platforms in their library sessions. Fifty-eight percent of respondents noted that they had used social media platforms in IL instruction. This percentage represents all the purposes of use as listed in table 2, and not just social media as a source of information. Figure 1 also indicates that the online video-sharing platform YouTube, and the microblogging and social networking site Twitter, have the highest use in library instruction, with 38 percent and 35 percent, respectively. Over 20 percent of respondents utilized Facebook and Instagram in their instruction. However, seven of the eleven platforms considered in this study had minimal use in library sessions, with fewer than three respondents selecting Nextdoor (platform for reporting news and events at the neighborhood level,) Snapchat (multimedia instant messaging app,) and WhatsApp (instant messaging and voice-over-IP service).

Table 2 shows how respondents used social media platforms in their instruction sessions (81 respondents). Table 2 indicates that 33 respondents (40 percent) reported that they referred to social media platforms in their instruction, but did not demonstrate their use. In addition, table 2 shows that most respondents (38 respondents) utilized a platform in their instruction session to teach about the context of information and what information is available on social media, although none of the purposes selected were utilized by even half of the respondents. Fewer than 45 percent of respondents (36 respondents) utilized social media for teaching lessons on citing, copyright and fair use, or to teach evaluation skills.

TABLE 2
How Social Media Platforms were Used in Instruction Sessions

Purpose of Use	Number of Respondents n = 81
Teaching context of information and what information is available on social media platforms (e.g., currency of posts, who can post information to this platform, etc.)	38
Used platforms to teach lessons on citing, copyright, fair use, and/or Creative Commons	36
Used platform to teach evaluation skills	36
Taught platform posts using the platform itself	35
Used platform to engage with students with the goal of creating relevant examples	34
Taught platform with references to platform but do not demonstrate/use the platform	33
Used platform to teach how to find posts or other information by authoritative sources	29
Used platform to teach lessons on fake news/misinformation	28
Used platform to engage with students to increase classroom participation by using a familiar platform	28
Used platform to find posts by general public, or to observe social interactions	27
Used platform to teach how to find news and current events	25
Used platform to engage with students with the goal of communicating with students	25
Used platform to teach platform specific skills, such as how to use platforms in future professions	23
Used platform to engage with students to build classroom community	17
Used platform to engage with students to teach analytics and data skills for platform	16
Used platform to teach platform specific skills, such as teaching how to create content for the platform	14
Used platform to teach how to find instructions and solutions to problems	9
Taught platform using third-party platform that makes data available	8

Only 29 respondents (36 percent) used a social media platform to teach how to find information by scholarly sources; 25 respondents (31 percent) utilized a platform to teach how to find news and current events. Fewer than sixteen respondents used platforms to teach data analytics, how to create platform content, or to find solutions to problems. In addition, respondents utilized platforms for other purposes, such as student engagement and classroom community building, as table 2 indicates.

Figure 2 shows the three most utilized platforms (YouTube, Facebook, and Twitter), as well as their purpose of use related to social media as an information source. We chose seven "purpose of use" categories based on the number of respondents and on their relevance to teaching. Respondents most frequently used Twitter to teach: evaluation skills, lessons on fake news, context of information and what information is available, and how to find news and current events. Respondent utilized YouTube the most often to teach lessons on citing, and on how to find posts by scholarly sources. Respondents most often used Facebook as a reference, but respondents who selected this option did not demonstrate or use this platform in instruction.

FIGURE 2
The Three Most Used Platforms and the Purpose of Use Related to Social Media as an Information Source

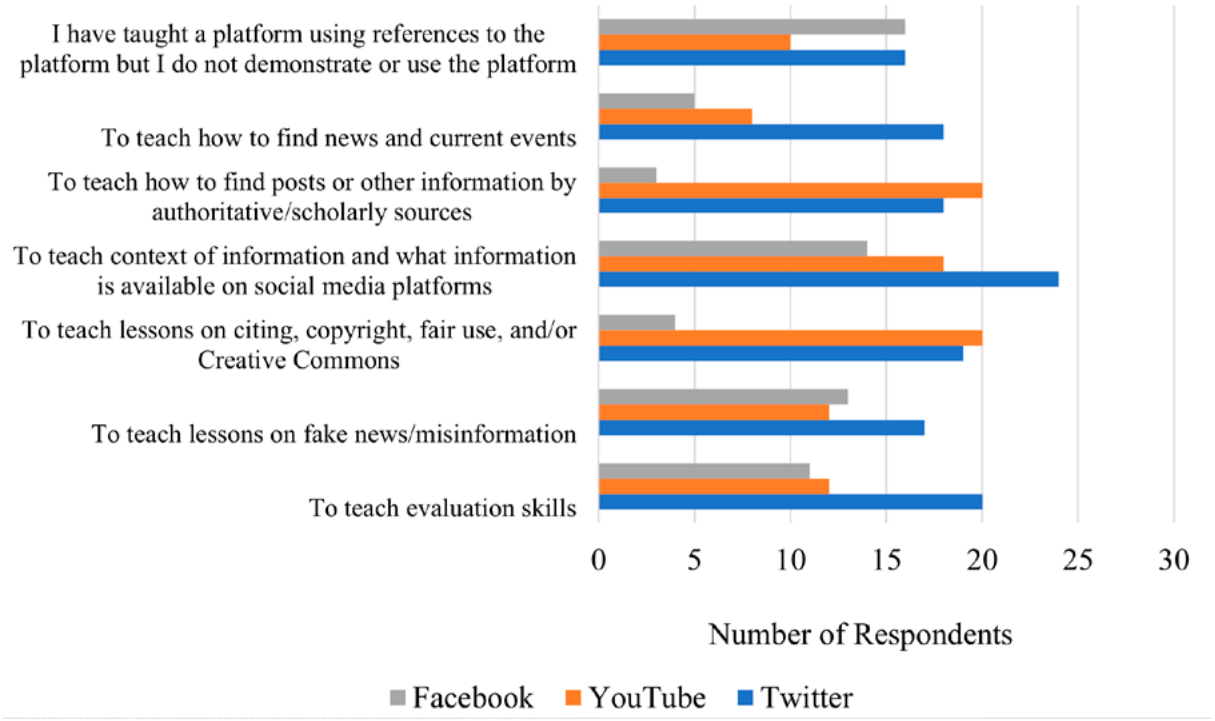
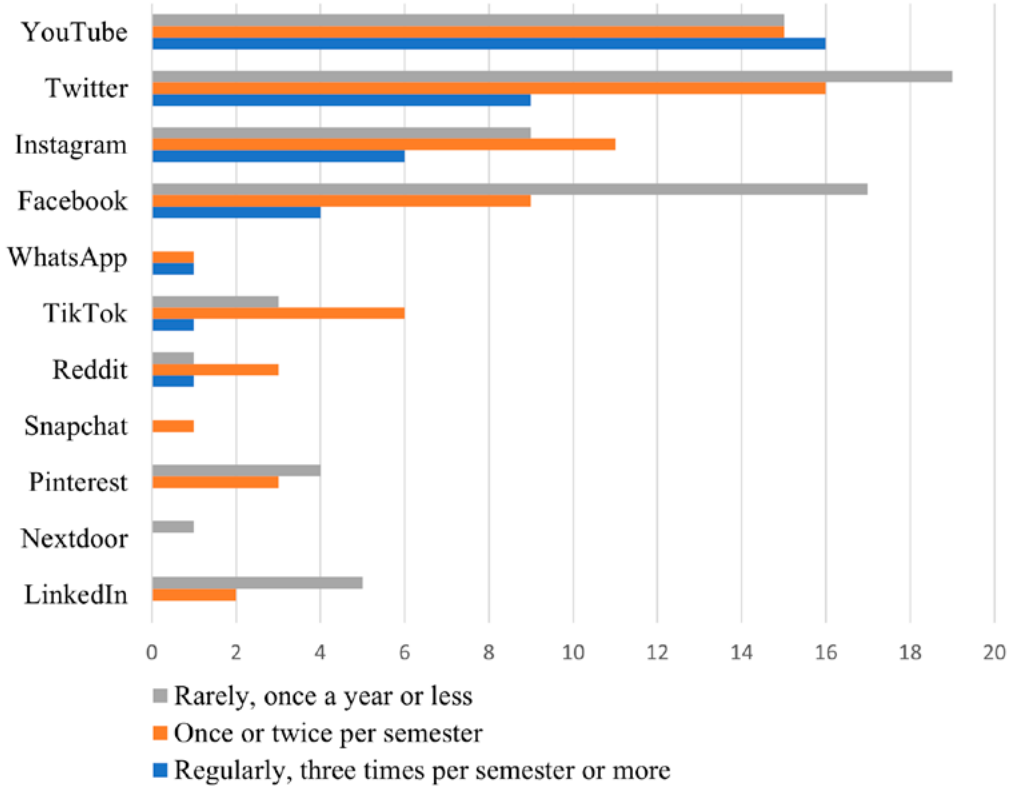


FIGURE 3
Frequency of use of a Platform in Library Instruction



Frequency of Use of the Eleven Platforms

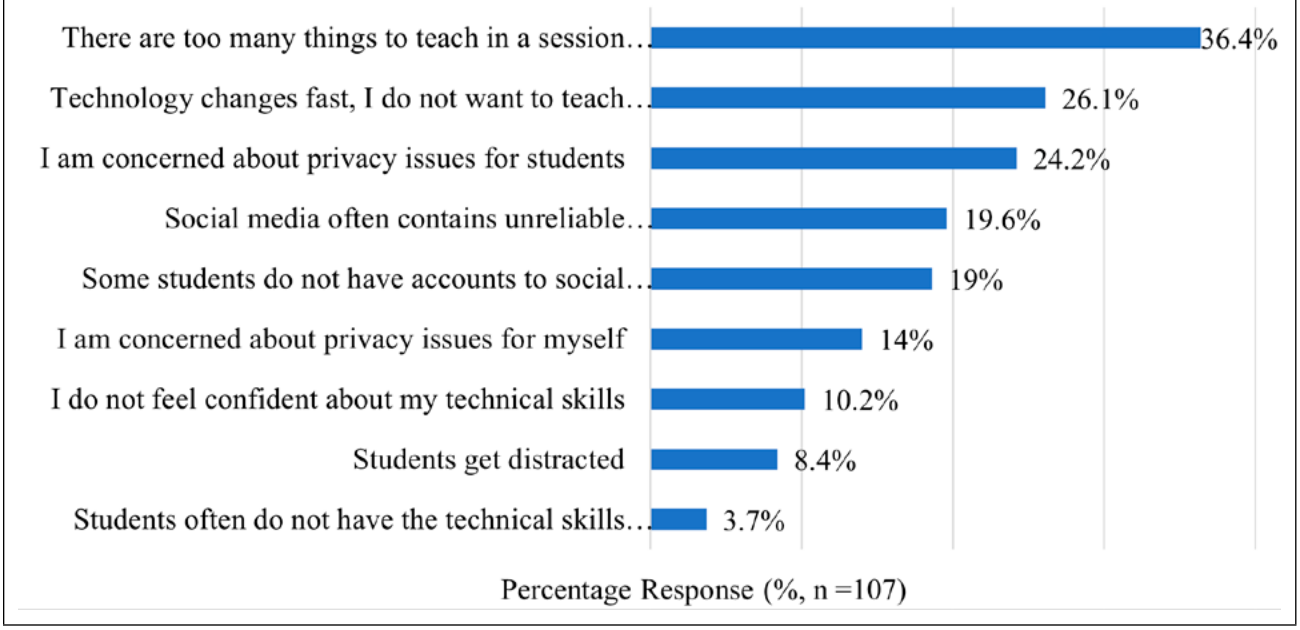
Figure 3 represents how frequently respondents reported using each of the 11 social media platforms in their instruction. Data showed that Twitter and YouTube had the highest frequency of use of the eleven platforms, with nine respondents utilizing Twitter, and 16 utilizing YouTube about three times per semester or more. It is important to note that platform use totals in figure 3 may not match the totals in table 2, as not all respondents answered all survey questions.

Challenges When Using Social Media in Library Instruction

Figure 4 indicates that 36 percent of respondents selected "there are too many things to teach in a session to include social media in library instruction." Nearly 25 percent of respondents indicated concern regarding fast technological changes and privacy issues. Figure 4 also shows that there was no substantial consensus when selecting challenges. In addition, fewer than ten percent of respondents noted that students lack the skills and become distracted. Only 20 percent of respondents agreed with the statement: "social media often contains unreliable information." Ten percent of respondents felt apprehensive about their own technical skills.

Thirty-six percent of respondents selected the option "other," not shown in figure 4. We coded their statements using keywords derived from respondents' answers. The following keywords represent these answers about challenges to using social media in IL instruction: "not useful," "changes in technology access," "lack of support," and "not suitable and not scholarly." Three respondents did not answer this question directly; however, they displayed a pessimistic attitude towards social media, stating: "I am disgusted by social media, which pervades my instruction," "Giving off the 'hello fellow kids' vibe," and "I do not want to support social media platforms in general so I don't want to encourage using them for any purposes…"

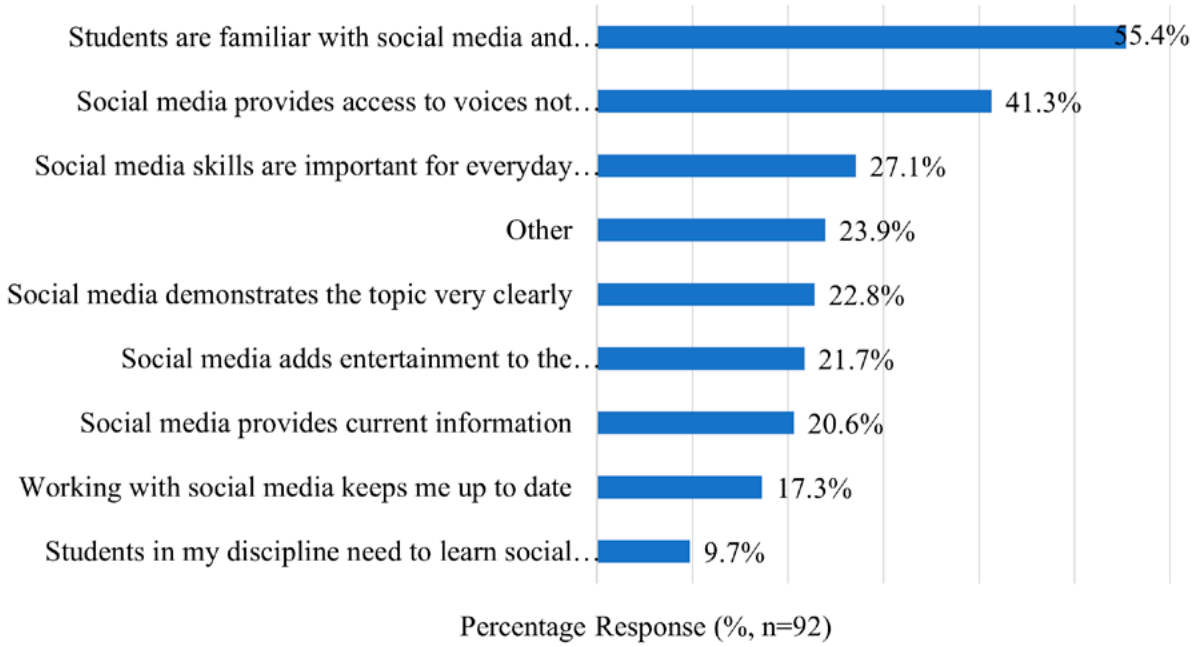
FIGURE 4
Respondents' Selection of Three Challenges when Using Social Media in Library Instruction



Benefits When Using Social Media in Library Instruction

Figure 5 shows that 55 percent of 92 respondents perceived students' familiarity with social media as a benefit of including it in library instruction. Forty-one percent of respondents noted that social media provides access to voices not represented in academic sources, and 27 percent considered social media skills necessary for everyday life. Slightly more than 20 percent indicated that social media provides current information, adds entertainment to the instruction session, and demonstrates a topic well.

FIGURE 5
Respondents' Selection of Three Benefits when using Social Media in Library Instruction



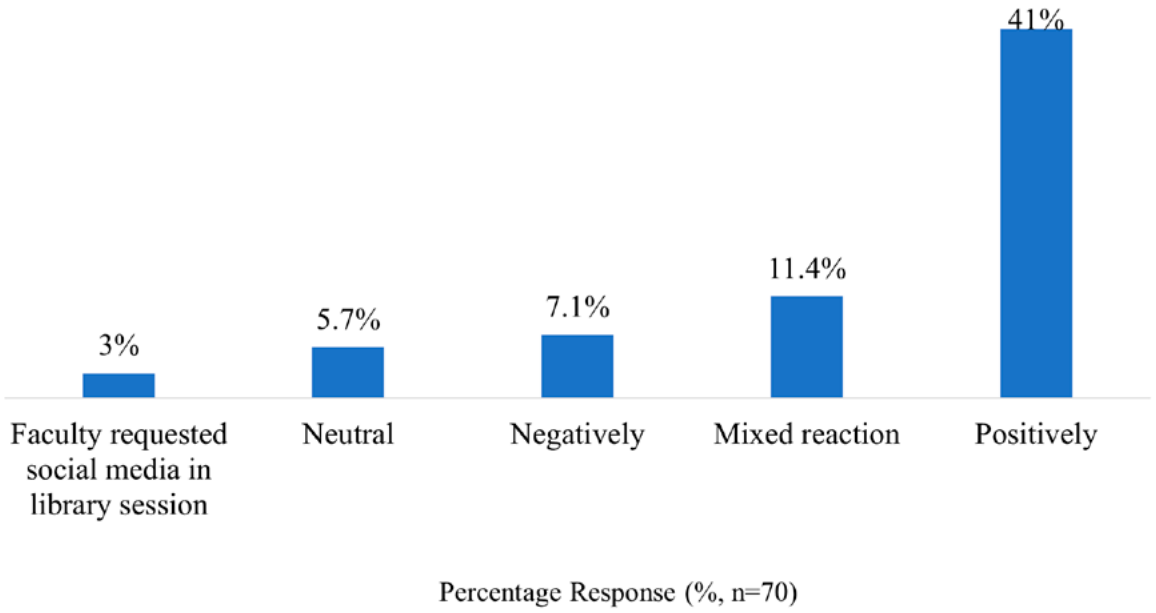
Twenty-two respondents (24 percent) noted additional benefits in the option "Other." This "Other" category revealed a few minor themes in how librarian instructors see teaching social media as beneficial, including that the platforms provide a convenient and helpful medium for instruction. Three respondents reported using social media in this way, two of whom specifically named YouTube. The third remarked that "videos" help them to teach about using the library catalog. Another theme was that social media provided data for students to work with and analyze. Other individuals identified benefits such as social media allowing them to teach concepts in classes, reach broader audiences, and learn to interact with social media in ways essential for their academic field.

Reactions of Faculty in Academic Departments to Social Media in Library Instruction

Seventy respondents shared their perceptions about departmental faculty's reactions to social media in library instruction in a free-text box. Twenty-four respondents selected "not applicable" (N/A), or did not answer the question. Figure 6 shows that forty-one percent of respondents perceived their faculty to react positively to using social media in library instruction. Some librarians commented that professors were curious, supported the evaluation of

contextual information and real-life examples, found social media fun, trusted the librarian, or considered the inclusion of social media inconspicuous because it was used in a limited capacity in IL instruction.

FIGURE 6
Respondents' Perceptions of Faculty Reactions to Social Media in Library Instruction



General Comments from Survey Respondents

The survey's final question asked respondents if they had anything else to share about using social media in instruction. Forty-two people provided comments. Eight respondents advocated for teaching social media in library instruction, suggesting that social media is a part of modern life used by students and researchers, and thus should be addressed in library sessions. Three respondents expressed an interest in learning more about how others included social media in their instruction, or about the support they received from their institutions. Some participants mentioned that social media was helpful to them for teaching concepts and ideas. Other comments included the importance of social media for scholarly communication, engagement and outreach, and research interests.

Discussion

Research Question 1: How are Teaching Librarians Using Social Media in their Instruction?

Our data indicate that, while respondents are using social media platforms in instruction, they mostly use them to provide examples and to help students understand the current information landscape, but not as a source of information that could help students find resources for their academic assignments. Specifically, our findings show that many respondents had either used—or referred to—a social media platform in instruction to teach a concept (e.g., citation, evaluation, fair use) or for classroom community building (table 2). However, fewer librarians use social media platforms as a source of information to find posts by authoritative sources, to teach lessons on fake news, or to find news and current events.

As mentioned above, some librarians utilize social media for lessons about fake news and misinformation; however, most respondents did not prioritize this option. Considering the amount of literature published in library science journals on evaluating sources, metaliteracy, and the challenges of doing research in the age of misinformation, it is intriguing that the librarians who responded to our survey did not report more actively using social media to address fake news and misinformation in IL instruction.[43]

On the other hand, respondents' relatively low use of social media in library instruction to seek information from authoritative or scholarly sources is somewhat predictable. Social media's reputation as a credible or even acceptable source of information has plummeted in response to many scandals, such as when people used social media platforms to spread false news during the 2016 U.S. presidential election, or when publications—such as the *New York Times*—highlighted the negative effects of social media on youth, and published articles related to the Facebook whistle-blower (which prompted Congress to call for increased regulation of this platform).[44] Despite the many critiques of social media, we wonder whether teaching librarians should introduce these platforms as a pedagogical tool for students to find information for academic work in higher education. To this end, we consider how respondents perceived benefits and challenges to including social media in library instruction.

Research Question 2: What are the Benefits Teaching Librarians Perceive Regarding the Use of Social Media in Instruction?

The two greatest benefits to using social media in IL that respondents identified were students' familiarity with social media platforms, and the access social media provides to voices that are often not represented in academic sources (figure 5). Some respondents considered social media to be helpful for the development of skills for everyday life, the ability to demonstrate topics, and for providing access to current information. Additionally, a few respondents mentioned that social media provided data for analysis, and that it was good for instruction using videos, or for teaching a concept.

Since social media is a part of everyday life for many people, including higher education students, it is unsurprising that some librarians would seek to harness students' familiarity and comfort with social media to teach them about library research. The survey data reflects this use: many librarians felt that students' familiarity with social media helped them relate to library instruction topics. One respondent offered the statement: "Great evaluation examples." Similarly, another respondent stated, "I use social media as an example in discussions about how information moves in online spaces…" These statements help us reflect on the possible uses of social media in IL instruction.

Another use of social media is the dissemination of underrepresented voices. Scholarly literature represents one source of information commonly recognized as academic due to the peer review process, but it is not the only source of information. Social media may allow access to additional, different voices—such as those of activists, Indigenous peoples, members of underrepresented communities, and governmental and non-profit organizations—as well as to different contexts for information. Kim et al. acknowledge social media's function as a place where different voices may be found:

> Social media takes[sic] advantage of the wide range of experiences and expertise that people have. While their quality varies, social media information sources can

> provide information that traditional resources cannot efficiently provide. It would therefore be beneficial to acknowledge their usefulness as information sources and to develop IL programs for helping individuals to effectively evaluate and use them.[45]

While we acknowledge the benefit of student familiarity with social media in IL instruction, we consider the more important benefit to be the access to first-hand knowledge and experiences from a wide variety of perspectives it provides, which may broaden students' exposure to different ideas, and to help them understand the world more comprehensively.

Question 3: What are the Challenges Teaching Librarians Perceive Regarding the Use of Social Media in Instruction?

Respondents identified three substantial challenges to using social media in IL instruction: limited time to cover more material in sessions; fear of not being up to date due to fast changes in technology; and concern about privacy issues (figure 4). The biggest challenge—the lack of time to cover social media in IL instruction—suggests that respondents view social media as an additional or "extra" topic, rather than a source of examples for concepts already integrated within their lesson plan. Since time is a major concern for librarians using the one-shot instruction model, one solution could be incorporating social media into the topics they already cover. This model could benefit students by teaching them evaluation skills, and by providing access to a wide range of voices they could include in their academic work.

Another challenge respondents identified relates to the amount of time required to stay current using these technologies. The quickly evolving nature of social media "poses[sic] significant challenges for users," forcing them to develop new skills and different types of competencies constantly.[46] Recognizing that technology changes continually, and that social media will probably continue to exist, we wonder whether teaching librarians would consider planning lessons where social media is seen as a source of information.

Privacy was another challenge respondents identified, including privacy issues for themselves as well as for their students. However, only 41 out of 107 respondents selected privacy concerns as one of their top three challenges for including social media in IL instruction. This data is remarkable, as we anticipated that privacy would be a top concern for many respondents, given the criticism of social media companies who use information from user accounts and social media posts to target users with political propaganda and advertisements. Social media has also been criticized for exposing users to illegal data appropriation and compromising users' accounts. These issues are relevant to our discussion because librarians using social media platforms in instruction may legitimize the use of social media platforms, which subsequently exposes students to these risks.

Another risk associated with the use of social media in IL instruction relates to students' ethical use of platforms. To illustrate: if students decided within a class research activity to search for information about their classmates or instructor instead of the research topic, then students might be infringing on the privacy of others. For this reason, any IL instruction activity involving social media would need to be carefully designed to include a discussion about using platforms in an ethical, responsible, and safe manner.[47]

Another topic that received surprisingly little interest from respondents was that of misinformation. Few librarians indicated that "unreliable information" was a challenge when using social media. Given the degree of attention this topic has received in both scholarly

and popular spaces, it is curious that so few respondents would identify this as an issue. A few respondents included comments such as "inappropriate content," and "not useful," but none specifically referred to "fake news," or similar concerns.

Librarians' Perceptions of Faculty Attitudes Towards Social Media in IL Instruction

Our data suggest that many librarians perceived positive reactions from departmental faculty toward using social media in IL instruction. This positive reaction might encourage librarians to explore using social media in IL instruction for a variety of purposes, including using it as an information source for academic work. However, advocacy for social media could be difficult if teaching librarians have a negative attitude toward it.

A few respondents openly expressed disapproval of social media, suggesting that some librarians may not approve of using social media in IL instruction. As discussed earlier, librarians are often in the position of trying to teach a great deal of content in very limited time frames. Therefore, it is understandable that teaching librarians would not want to spend time on something they consider not useful in the academic context.

However, a negative attitude toward including social media in IL instruction could have consequences. Teaching librarians may miss an opportunity to help students use platforms critically and effectively for scholarly research or other information-gathering purposes. Additionally, ignoring social media and its impact on the world could prevent librarians from understanding how students look for information, generating a disconnect between librarians and students. As teaching librarians ourselves, we recognize the concerns about social media as a source of information for academic work. Nonetheless, we invite our colleagues to investigate further the pedagogical use of these platforms as a source of information in IL instruction.

Limitations to This Study

The data collected in this study represents the responses of a small number of participants within the United States. Therefore, the findings cannot be used to make generalized assumptions, and statistical information is considered exploratory. In addition, it was challenging to create keyword codes for some of the qualitative data due to difficulties in interpreting responses. The survey also presented a limitation concerning some of the multiple-choice questions, which prompted respondents to select from author-created options rather than allowing respondents to generate their own answers.

Conclusion

Our study examines whether and how teaching librarians are using social media in their IL instruction, as well as the benefits and challenges they encounter when they do so. The desire to include other voices and resources in addition to traditional scholarly perspectives in the academic narrative motivated this exploration.

Despite the challenges of utilizing social media in IL instruction, these platforms provide a medium for people to participate in global conversations, to share many types of information—as well as experiences, beliefs, and solutions to problems—that may otherwise go unnoticed by scholars, and that may be useful to students' academic work. Finally, given the popularity of social media, it seems likely that people will continue to use it extensively,

and that voices representing societal changes and trends may be found there. Therefore, we encourage librarians to participate in the conversation about incorporating social media and non-scholarly voices into IL instruction.

Topics for Future Research

Future studies could examine the attitudes of professors toward the use of social media in IL instruction and their expectations about students' inclusion of non-scholarly voices in their academic work. Future research is also needed about the most effective methods of integrating social media into IL instruction.

Acknowledgments

The authors thank Victoria Kravets, Katie Stewart, and Gregor Rumenov Tzinov from the Laboratory for Interdisciplinary Statistical Analysis at the University of Colorado Boulder for their support in reviewing the statistical data in this paper and preparation of Table 1. We also thank Aditya Ranganath, Data Librarian at the University of Colorado Boulder, for his input about survey research. Finally, the authors also thank the University of Colorado Boulder's Writing Center for assisting us during our writing process.

Appendix. The Survey Instrument

1. Are you a librarian in higher education (academic librarian)?
   - □ Yes
   - □ No
2. Do you have one shot or course instruction responsibilities? Select all that apply
   - □ Yes, one time session (workshops, one-shots, webinars, etc.)
   - □ Yes, courses (multiple related sessions)
   - □ No instruction responsibilities.
3. What category best describes your institution?
   - □ Doctorate university
   - □ Master's college/university
   - □ Baccalaureate college/university
   - □ Associate's college/community college
   - □ Special focus institution/technical college
   - □ Tribal college
   - □ None of the above
4. How long have you taught library instruction sessions in a higher education context?
   - □ 0-2 years
   - □ 3-5 years
   - □ 6-10 years
   - □ 11 or more years
5. How would you categorize your position? Select all that apply.
   - □ Subject specialist/liaison
   - □ Instruction
   - □ Data services/GIS
   - □ Reference
   - □ Assessment/analytics/user experience
   - □ Archiving/curatorial/rare books/preservation/conservation
   - □ Digital initiatives/systems/information technology/web development
   - □ Outreach
   - □ Media/multimedia specialist
   - □ Access services
   - □ Cataloging/bibliographic control/metadata
   - □ Scholarly communication
   - □ Press/publishing
   - □ Acquisitions/collection development/management/strategy
6. Do you support specific academic programs/departments? Select all that apply.
   - □ Communication/journalism/media studies
   - □ Business
   - □ Math/computer science/statistics/technology/engineering
   - □ Social sciences
   - □ Natural sciences
   - □ Health/medicine/nursing
   - □ Education
   - □ Art/art history/theater/music/film

- □ Social work
- □ Undergraduate writing and rhetoric
- □ Humanities
- □ Literature
- □ Languages
- □ History
- □ Other, please specify
- □ I do not liaise with any department

7. Which social media platforms have you used (e.g., as an information source, demonstrated, discussed, used data from) in your instruction sessions? Select all that apply.
   - □ Facebook
   - □ Instagram
   - □ LinkedIn
   - □ Nextdoor
   - □ Pinterest
   - □ Reddit
   - □ Snapchat
   - □ TikTok
   - □ Twitter
   - □ WhatsApp
   - □ YouTube
   - □ Other
   - □ I have not used social media platforms in my instruction
8. How do you use Facebook in your instruction sessions? Select all that apply.
   - □ I've used Facebook to teach evaluation skills (accuracy, relevancy, currency, etc.).
   - □ I've used Facebook to teach lessons on fake news/misinformation.
   - □ I've used Facebook to teach lessons on citing, copyright, fair use, and/or Creative Commons.
   - □ I've used Facebook to teach content of information and what information is available in social media platforms (e.g., currency of posts, who can post information to this platform, etc.).
   - □ I've used Facebook to teach how to find posts or other information by the general public, or to observe social interactions.
   - □ I've used Facebook to teach how to find posts or other information by authoritative/scholarly sources.
   - □ I've used Facebook to teach how to find news and current events.
   - □ I've used Facebook to teach how to find instruction and solutions to problems.
   - □ I've used Facebook to engage with students with the goals of increasing classroom participation by using a familiar platform.
   - □ I've used Facebook to engage with students with the goals of communicating with students.
   - □ I've used Facebook to engage with students with the goals of building classroom community.
   - □ I've used Facebook to engage with students with the goals of creating relevant examples that students can relate to the instruction.

- □ I've used Facebook to engage with students with the goals of teaching analytic and data skills for this platform.
- □ I've used Facebook to teach platform-specific skills, such as how to use the platform in their future professions.
- □ I've used Facebook to teach platform-specific skills, such as teaching how to create content for this platform.
- □ I've taught Facebook or Facebook posts using the platform itself.
- □ I've taught Facebook or Facebook posts using a 3rd party platform that makes the data available (e.g. Netlytic, Social Studio, etc.).
- □ I've taught Facebook or Facebook posts using references to the platform, but I do not demonstrate or use the platform.

9. How do you use Instagram in your instruction sessions? Select all that apply.
10. How do you use LinkedIn in your instruction sessions? Select all that apply.
11. How do you use Nextdoork in your instruction sessions? Select all that apply.
12. How do you use Pinterest in your instruction sessions? Select all that apply.
13. How do you use Reddit in your instruction sessions? Select all that apply.
14. How do you use Snapchat in your instruction sessions? Select all that apply.
15. How do you use Tiktok in your instruction sessions? Select all that apply.
16. How do you use Twitter in your instruction sessions? Select all that apply.
17. How do you use WhatsApp in your instruction sessions? Select all that apply.
18. How do you use YouTube in your instruction sessions? Select all that apply.
19. You indicated you use an "other" social media platform. How do you use the "other" social media platform in your instruction sessions? Select all that apply.
20. How often do you use Facebook in your Instruction?
    - □ Regularly, about three times per semester or more.
    - □ Once or twice per semester.
    - □ Rarely, about once a year or less.
21. How often do you use Instagram in your Instruction?
22. How often do you use LinkedIn in your Instruction?
23. How often do you use Nextdoor in your Instruction?
24. How often do you use Pinterest in your Instruction?
25. How often do you use Reddit in your Instruction?
26. How often do you use Snapchat in your Instruction?
27. How often do you use TikTok in your Instruction?
28. How often do you use Twitter in your Instruction?
29. How often do you use WhatsApp in your Instruction?
30. How often do you use YouTube in your Instruction?
31. You indicated to teach an "other" social media platform. How often do you use your identified social media platform in your instruction?
    - □ Regularly, about three times per semester or more.
    - □ Once or twice per semester.
    - □ Rarely, about once a year or less.
32. What are the top three challenges when using social media in your library instruction? Select up to three challenges.
    - □ Students get distracted.

- □ I do not feel confident about my technical skills.
- □ Students often do not have the technical skills needed.
- □ I am concerned about privacy issues for myself.
- □ I am concerned about privacy issues for students.
- □ Technology changes fast, I don't want to teach something outdated.
- □ There are too many things to teach in a session to include social media.
- □ Social media often contains unreliable information.
- □ Some students do not have accounts to social media sites I present.
- □ Other

33. What are the top three benefits of using social media in your instruction? Select up to three benefits.
    - □ Students are familiar with social media and can relate to the instruction.
    - □ Students in my discipline need to learn social media skills for their future careers.
    - □ Working with social media keeps me up to date.
    - □ Social media demonstrate the topic very clearly.
    - □ Social media adds entertainment to the instruction session.
    - □ Social media provides access to voices not represented in academic sources.
    - □ Social media offers current information.
    - □ Social media skills are important for everyday life.
    - □ Other
34. In your experience, how do most faculty/instructors from the departments you support react to social media in your library instruction?
35. Is there anything else you would like to share with us about your use of social media in instruction?

Notes

1. Danah M. Boyd and Nicole B. Ellison, "Social Network Sites: Definition, History, and Scholarship," *Journal of Computer-Meditated Communication* 13 (2008): 214, https//doi.org/10.1111/j.1083-6101.2007.00393.x.

2. Brooke Auxier and Monica Anderson, "Social Media Use in 2021," *Pew Research Center*, (April 2021), https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/.

3. Andreas M. Kaplan and Michael Haenlein, "Users of the World, Unite! The Challenges and Opportunities of Social Media," *Business Horizons* 53, no. 1 (2010): 61, https://doi.org/10.1016/j.bushor.2009.09.003.

4. Auxier and Anderson.

5. For research on incorporating social media into the journalism curriculum, see: Ainara Larrondo Ureta and Simón Peña Fernández "Keeping Pace with Journalism Training in the Age of Social Media and Convergence: How Worthwhile is it to Teach Online Skills?," *Journalism* 19, no. 6 (2018): 877-891, https://doi.org/10.1177/1464884917743174; For research on social media and teaching metaliteracy see: Donna Witek and Teresa Grettano, "Teaching Metaliteracy: A New Paradigm in Action," *Reference Services Review* 42, no. 2 (2014): 188-208, https://doi.org/10.1108/RSR-07-2013-0035. Also, Denny McCorkle and Janice Payan write about the demand for social media skills in business marketing and how students can learn practical job skills in the classroom to help them attain proficiency and confidence. These authors advocate for class projects involving Twitter to teach students how to communicate with consumers about new products, gather customer information, and monitor competitors. See: "Using Twitter in the Marketing and Advertising Classroom to Develop Skills for Social Media Marketing and Personal Branding," *Journal of Advertising Education* 21, no. 1 (2017): 33-43.

6. Carrie Mastley points out the lack of literature about academic librarians' attitudes to including social media in IL instruction in Mississippi. See: "Information Literacy Instruction and Social Media: A Survey of Mississippi Academic Librarian Attitudes," *Mississippi Libraries 83*, no. 3 (2020): 37, http://www.misslib.org/resources/Documents/MLarchive/ML2020Fall.pdf.

7. Jennifer Howard, "Social Media Lure Academics Frustrated by Traditional Publishing," *The Chronicle of Higher Education* (February 2011).

8. Ian Rowlands, David Nicholas, and Bill Russell, "Social Media and Research Workflow," *Information Services & Use* 31 (2011): 72, 82, 83, https://doi.org/10.3233/ISU-2011-0623.

9. L. Pangrazio and N. Selwyn. "It's not Like it's Life or Death or Whatever: Young People's Understanding of Social Media Data," *Social Media + Society* 4, no. 3 (2018): 7.

10. Kyung-Sum Kim, Sei-Ching Joanna Sin, and Yuqi He, "Information Seeking through Social Media: Impact of User Characteristics on Social Media," *Proceedings of the American Society for Information Science and Technology* 50, no. 1 (2013).

11. Ibid.

12. Kyung-Sun Kim, Sei-Ching Joanna Sin, and EunYoung Yoo-Lee, "Undergraduates' Use of Social Media as Information Sources," *College & Research Libraries* 75, no. 4 (2014): 442, 447, https://doi.org/10.5860/crl.75.4.442. In a 2016 paper, Kim and Sin echoed Kim, Sin, and Eun Young Yoo-Lee's 2014 study by stating that students use social media for information-seeking in the academic context. See Kyung-Sun Kim and Sei-Chin Joanna Sin, "Use and Evaluation of Information From Social Media in the Academic Context: Analysis of Gap Between Students and Librarians," *The Journal of Academic Librarianship* 42 (2016): 76.

13. Deepti Khatri, "Use of Social Media Information Sources: A Systematic Literature Review," *Online Information Review* Vol. 45 No. 6, 1039-1063, https://doi.org/10.1108/OIR-04-2020-0152.

14. Ibid.

15. Kim, Sin, and Yoo-Lee, "Use and Evaluation of Information from Social Media: A Longitudinal Cohort Study," *Library and Information Science Research* 43 (2021): 6, https://doi.org/10.1016/j.lisr.2021.101104.

16. Auxier and Anderson.

17. Kim, Sin, and Yoo-Lee also noticed this discrepancy. See: "Use and Evaluation of Information," 6.

18. Martin Rehm, Staphanie Manca, Diana Brandon, and Christine Greenhow, "Beyond Disciplinary Boundaries: Mapping Educational Science in the Discourse on Social Media," *Teachers College Record* 121, no. 14 (2019): 2, https://doi.org/10.1177/016146811912101403.

19. Lilian W. Mina, "Social Media in the FYC Class: The New Digital Divide," in *Social Writing/Social Media: Publics, Presentations, and Pedagogies*, eds. Douglas M. Walls and Stephanie Vie, (The WAC Clearinghouse, 2018), 265, 270, 274, https://doi.org/10.37514/PER-B.2017.0063.2.14, https://wac.colostate.edu/books/perspectives/social/; Sheila Stoeckel and Caroline Sinkinson, "Social Media," *Tips and Trends*: *Instructional Technologies Committee* (Summer 2013): 1-5, https://acrl.ala.org/IS/wp-content/uploads/2014/05/2013summer.pdf; Laurie M. Bridges, "Librarian as Professor of Social Media Literacy," *Library Innovation* 3, no.1 (2012): 50; Jennifer Wright Joe, "Assessment of Social Media in the Library: Guidelines for Administrators," *Journal of Library Administration* 55, no. 8 (2015): 671, https://doi.org/10.1080/01930826.2015.1085251.

20. Anatoly Gruzd et al., "Uses and Gratifications Factors for Social Media Use in Teaching: Instructors' Perspectives," *New Media & Society* 20, no.2 (2018): 478, 489, https://doi.org/10.1177/1461444816662933.

21. Rehm, Manca, Brandon, and Greenhow, 2; Referring to the integration of LinkedIn as a pedagogical tool, Mostafa Hamadi, Jamal El-Den, Sami Azam, and Narumon Sriratanaviriyakul found that collaborative learning could be achieved successfully in social media when instructors implement a well-defined framework for students. See: "Integrating Social Media as a Cooperative Learning Tool in Higher Education Classrooms: An empirical Study," *Journal of King Saud University Computer and Information Sciences* 34 (2022): 3728.

22. Gruzd et al., 479, 488-9.

23. Ibid, 489.

24. Michele E. Ewin et al., "Teaching Digital and Social Media Analytics: Exploring Best Practices and Future Implications for Public Relations Pedagogy," *Journal of Public Relations Education* 4, no. 2 (2018): 51–86.

25. Larrondo Ureta and Peña Fernández, 882.

26. Carrie P. Mastley states that social media connected with IL emerged with the appearance of Web 2.0 technologies. See: "Information Literacy Instruction and Social Media: A Survey of Mississippi Academic Librarian Attitudes," *Mississippi Libraries* 83, 3 (2020): 34–36, https://doi.org/10.1515/opis-2020-0014.

27. Heidi Julien, Melissa Gross, and Don Latham found that 14 percent of 343 librarians recruited from the ILI-L listserv—an American Library Association listserv related to information literacy instruction—use social media as a method of instruction. See "Survey of Information Literacy Instructional Practices in U.S. Academic Libraries," *College & Research Libraries* 79, no. 2 (2018): 186. https://doi.org/10.5860/crl.79.2.179.

28. Mastley, 42.

29. Stoeckel and Sinkinson, "Social Media," 1–5.

30. Natalie Burclaff and Catherine R. Johnson, "Teaching Information Literacy Via Social Media: An Exploration of Connectivism," *Library Philosophy and Practice* (2016): 7.

31. Kim, et al., "Undergraduates' Use," 453.

32. Jennifer Wright Joe refers to the dissemination of information. See "Assessment of Social Media," 674;

Amanda Kraft and Aleck F. Williams Jr., address the use of hashtags and promote the use of social media in library instruction alluding to the familiar environment platforms provide, see "#Shelfies are Encouraged: Simple, Engaging Library Instruction with Hashtags," *College &Research Libraries News* (January 2016): 11, https://doi.org/10.5860/crln.77.1.9425.

33. Kraft and Williams Jr.,"#Shelfies are Encouraged," 11. Bridges suggests that librarians may guide students to apply critical thinking to their use of social media, "Librarian as Professor," 49. Also, addressing participatory pedagogy, Wright Joe notes that social media provides opportunities for students to explore and participate in discussions. See: "Assessment of Social Media," 675-676.

34. Valerie J. Hill and Thomas P. Mackey. "Embracing Metaliteracy Metamodern Libraries and Virtual Learning Communities," *College & Research Libraries News* (May 2021): 2019-220.

35. Ibid, 221. Authors Diane M. Fulkerson, Susan Andriette Ariew, and Trudi E. Jacobson note that metaliteracy has four learning areas known as cognitive, behavioral, affective, and metacognitive, and that learners must be critical as evaluators and creators of information. See: "Revisiting Metacognition and Metaliteracy in the ACRL Framework," *Communications in Information Literacy* 11, no. 1 (2017): 24-26. https://doi.org/10.15760/comminfolit.2017.11.1.45. To review the ACRL framework, see "Framework for Information Literacy for Higher Education," *Association of College & Research Libraries* (2016). http://www.ala.org/acrl/files/issues/infolit/framework.pdf.

36. Ally Dever, "Millions are Turning to TikTok for the Latest on Ukraine, But Can the Platform be Trusted?" *CU Boulder Today* (March 18, 2022). Retrieved March 22, 2022 from https://www.colorado.edu/today/2022/03/18/millions-are-turning-tiktok-latest-ukraine-can-platform-be-trusted.

37. Sarah Hartman-Caverly and Alexandria Chisholm argue that privacy literacy (PL) is minimally present in library instruction, and is more often addressed as an issue in the technology side of librarianship. See: "Privacy Literacy Instruction Practices in Academic Libraries: Past, Present, and Possibilities," *International Federation of Library Associations and Institutions Journal* 46, no.4 (2022): 310, https://doi.org/10.1177/0340035220956804; Daniel G. Krutka et al., "Teaching 'Against' Social Media: Confronting Problems of Profit in the Curriculum," *Teachers College Record* 121 (2019): 11, 16, 26, https://doi.org/10.1177/016146811912101410; Jamie Halliwell, "Applying Social Media Research Methods in Geography teaching: Benefits and Emerging Challenges," *Journal of Geography* 119, no. 3 (2020): 110-111, https://doi.org/10.1080/00221341.2020.1755717; Stoeckel and Sinkinson, "Social Media," 1-5.

38. Alison Hicks and Caroline Sinkinson, "Participation and Presence: Interrogating Active Learning," *portal: Libraries and Academy* 21, no. 4 (2021): 759, https://doi.org/10.1353/pla.2021.0040.

39. L. Pangrazio and N. Selwyn, "It's not Like it's Life or Death or Whatever: Young People's Understanding of Social Media Data," *Social Media + Society* 4, no. 3 (2018), 7. Stefania Manca, Stefania Bocconi, and Benjamin Gleason also refer to the connection between social media skills and digital literacies. See: "Think Globally, Act Locally": A Global Approach to the Development of Social Media Literacy," *Computers & Education* 160 (2021): 2-3.

40. Witek and Grettano, 201.

41. Ibid, 199.

42. We selected the top eleven universities according to a US News and World Report list. See: "Best National University Rankings," *U.S. News & World Report*, accessed September 2021, https://www.usnews.com/best-colleges/rankings/national-universities. The librarians contacted worked for the following institutions: Princeton University, Harvard University, Columbia University, Massachusetts Institute of Technology, Yale University, Northwestern University, John Hopkins University, California Institute of Technology, University of Pennsylvania, University of Chicago, and Stanford University.

43. Citing Wardle and Derakshan, Claire McGuinness defines mis-information as "false information that is shared without meaning to cause harm;" dis-information as "false information that is shared knowing that it will cause harm;" and mal-information as "genuine information" that "is shared to cause harm, often by moving information designed to stay private into the public sphere." See *The Academic Teaching Librarian's Handbook*, (London, UK: Facet Publishing, 2021): 28; Claire Wardle and Hossein Derakhshan, "Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking," *Strasbourg: Council of Europe*, (2017): 5; Ian O'Hara notes that information literacy helps learners to understand misinformation and computational propaganda in social media. See: "Automated Epistemology: Bots, Computational Propaganda & Information Literacy Instruction," *Journal of Academic Librarianship* 48, no. 4 (2022), https://doi.org/10.1016/j.acalib.2022.102540.

44. For reference to the spread of fake news during the 2016 U.S. presidential election, see: Mike Wendling, "The (Almost) Complete History of 'Fake News,'" *BBC News* (22 January 2018), https://www.bbc.com/news/blogs-trending-42724320. For reference to social media effects on youth see: Christina Caron, "Worried about your Teen on Social Media? Here's How to Help," *The New York Times* (Sept. 21, 2021), https://www.nytimes.com/2021/09/21/well/family/teens-social-media-help.html; Erin Woo, "Teenage Girls Say Instagram's Mental Health Impacts are

no Surprise," *The New York Times* (Oct. 5, 2021), https://www.nytimes.com/2021/10/05/technology/teenage-girls-instagram.html; Jessica Grose, "The Messy Truth About Teen Girls and Instagram," *The New York Times* (Oct. 13, 2021), https://www.nytimes.com/2021/10/13/parenting/instagram-teen-girls-body-image.html; Amanda Hess, "How Social Media Turned 'Prioritizing Mental Health' Into a Trap," *The New York Times* (Oct. 27, 2021) https://www.nytimes.com/2021/10/27/magazine/social-media-mental-health.html; Kelly Browning, "Parents Sound Off on Testimony about the Harms of Facebook and Instagram," *The New York Times* (Oct. 5, 2021), https://www.nytimes.com/2021/10/05/technology/parents-sound-off-on-testimony-about-the-harms-of-facebook-and-instagram.html. For reference to the whistle-blower hearing in Congress see: Sheera Frenkel, "Key Takeaways from Facebook's Whistle-blower Hearing," *The New York Times* (Oct. 5, 2021); Frenkel, "Whistleblower Discusses how Instagram May Lead Teenagers to Eating Disorders," *The New York Times* (Oct. 5, 2021), https://www.nytimes.com/live/2021/10/05/technology/facebook-whistleblower-frances-haugen?partner=IFTTT#in-hearing-haugen-discusses-how-instagram-may-lead-teenagers-to-eating-disorders.

45. Kim, et al., "Undergraduates' Use," 453.

46. Hadewijch Vanwynsberghe, Ruben Vanderlinde, Annabel Georges, and Pieter Verdegem, "The Librarian 2.0: identifying a Typology of Librarians' Social Media Literacy," *Journal of Librarianship and Information Science* 47, No. 4 (2015): 284, https://doi.org/10.1177/0961000613520027.

47. For more information about privacy issues in IL in connection to the digital environment, see Hicks and Sinkinson's article "Participation and Presence," 759.

Open Access Workflows for Academic Libraries

Matthew W. Goddard and Curtis Brundy

There is a growing acceptance of open access funding models among academic publishers and a growing adoption of open access publishing agreements among academic libraries. In this context, libraries are taking on new roles and new processes to ensure the successful implementation of open access funding initiatives. This article will examine some of the key issues and considerations in the area of open access workflows, and discuss how one research-intensive library in the United States has approached these new functions.

Introduction

The transition of scholarly journal publishing from its historic subscription business model to new open access models is accelerating. In the wake of Plan S, the international initiative by research funders to advance open access, the European approach to advancing open access through centrally licensed agreements has migrated to North America. All major commercial publishers now have open access agreements in the United States, as do many not-for-profit publishers.[1] While there are alternatives that forgo article processing charges (APCs)—such as Subscribe to Open and the tiered model developed by Association of Computing Machinery (ACM)—the most prevalent open models at this time, such as Read and Publish, are based on the payment of charges per article. These charges are intended to substitute for the revenue generated from subscriptions, allowing the object of exchange between library and publisher to shift from paywalled content to open access publishing services for institutional research. At the time of writing, Cambridge University Press had signed over 300 Read and Publish agreements with US libraries.[2] This indicates a broad willingness on the part of US libraries to enter into open access publishing agreements. Furthermore, new guidance, released in August 2022, from the United States Office of Science and Technology Policy will require publications from US federally-funded research be made freely available and publicly accessible without embargo or delay. While concerns still abound over the sustainability, unintended consequences, and equity of APC-based models,[3] the current trajectory and pace point to even higher levels of future uptake.

In part, the recent success of open access models in displacing the traditional subscription model stems from the growing momentum behind open science.[4] The scientific community's adoption of open research practices in response to the COVID-19 pandemic affirmed the

* *Matthew W. Goddard is Access and Acquisitions Dept. Head at Iowa State University, email: mgoddard@iastate.edu; Curtis Brundy is Associate University Librarian for Scholarly Communication & Collections at Iowa State University, email: cbrundy@iastate.edu.*

importance of sharing research data and open access to research publications.[5] That journal publishers, in the chaotic early months of the pandemic, felt obliged to drop paywalls to the coronavirus literature to help accelerate research discovery, will not soon be forgotten, as it demonstrates their full understanding that paywalls impede science. The lessons learned from the pandemic have found their way into UNESCO's new Recommendation on Open Science, which will be followed by 193 countries, and state, "…the global COVID-19 health crisis has proven worldwide the urgency of and need for fostering equitable access to scientific information…"[6] The renewed commitment to open science, and, by extension, open access, will further grow the number of libraries entering into APC-based agreements.

For libraries that are new to making open access agreements, questions often arise about how they are negotiated and the work involved with their successful implementation. This article will examine some of the key issues and considerations in the area of open access workflows. This includes key agreement clauses, funding verification, invoicing, reporting and analysis, post-publication processes, situating open access workflows in the library, and the role of third party-tools for improved workflow management. Where appropriate, this article offers modest suggestions related to these considerations rooted in the lessons learned by one research-intensive US library from implementing open access agreements across a wide variety of academic publishers. The article will also examine the successful incorporation of open access workflows into an Electronic Resources unit, which is a likely area of consideration for other libraries when locating their own open access workflow processes.

This article will help address the current gap in the literature on open access workflows as practiced within the context of the United States. The topics covered will be of interest to those who are currently responsible for developing and implementing open access publishing agreements, regardless of the size of their institutions or their functional location within the library. While support for the open access publication of institutional research may sometimes be seen as the domain of research-intensive institutions, its proportionally lower cost for institutions at the other end of the research output spectrum should make it an easier sell, with the goal of making all local research open more easily achievable. It is the hope that this article will also be useful to those considering the extent of their involvement with these new initiatives.

Literature Review

Open access workflows, as required by APC-based open access models like Read and Publish, are new to most North American academic libraries, but certain aspects are not necessarily unknown. Lessons have been gleaned, for example, from experiences operating local open access funds and, from a distance, from the experiences of European libraries and consortia that were early adopters of APC-based open access agreements. The limited literature on open access workflows in academic libraries largely runs through these two channels, the deepest being, by far, that originating from European experiences implementing central open access agreements.

For many US libraries, their first experience with open access workflows came through the operation of an open access fund. Open access funds, which cover the cost of affiliated author APCs, offered libraries a direct way to support open access publishing on their campuses. A SPARC survey in 2014 found that 51 libraries in North America were operating open access funds, with nearly 4,000 research articles published.[7] Several articles and reports on open access funds mention the underlying workflows. Greg Tananbaum mentioned the importance of vetting applications for eligibility, verifying article publication, tracking results, and fund

disbursement.[8] Stephen Pinfield mentioned the importance of establishing streamlined procedures and regular performance reviews.[9] Heidi Zuniga and Lilian Hoffecker found that student registration data and employee HR information are useful sources for determining author affiliation.[10] By 2016, interest in library-administered open access funds began to wane over funding concerns,[11] slowing open access fund workflow experimentation and development.

The most significant work to date in establishing, improving, and documenting open access workflows has been done by the Efficiencies Standards for Article Charges (ESAC) initiative, which focuses on optimizing, promoting standards, and good practices for open access workflows. In 2016 and 2017, ESAC organized two largely European workshops around open access workflows that demonstrated the need for workflow improvements in areas such as author identification, metadata exchange, and invoicing.[12] Workflow concerns and issues experienced by those negotiating and implementing open access agreements were surfaced in case studies by the Max Planck Digital Library[13] and the Vienna University Library.[14] To help address issues, ESAC published workflow recommendations covering author and article identification and verification; funding acknowledgement and metadata; and invoicing and reporting.[15] In 2021, the recommendations were updated to include responsibilities of institutions, funders, and consortia; responsibilities of publishers; and relevant metadata.[16] The ESAC standards have been heavily utilized in open access negotiations and conversations with publishers to establish and improve open access workflows.

Other relevant research has been done in the areas of open access workflow tools, metadata, and organizationally locating open access workflows. George Machovec provided an overview of tools and services to manage open access agreements.[17] Publishers were found to not deliver consistent metadata to open access agreement customers in Europe.[18] And Jill Emery, Graham Stone and Peter McCracken put forward ways to envisage open access management as part of their "Techniques for Electronic Resources Management" framework (TERMS), which includes considerations for metadata and reporting and for direct deposit of articles into institutional repositories.[19]

Open Access Workflow Considerations

The first thing to say about the work of managing and implementing open access agreements is that there is a wide range of possible approaches. Publishers are eager to offer agreements that are simple to implement, and libraries have limited staff resources available to take on new tasks. And it is indeed possible to enter into some open access publishing agreements without actually doing anything other than signing an agreement and paying an invoice. Others may also require institutions to approve each article for funding under the agreement. Some institutions will benefit from the possibility of such minimalist approaches, but this article takes a more expansive view of the library's role, considering the full variety of ways library staff can take action to ensure the success of these initiatives and ultimately lay the groundwork for the broader shift of collections budgets from subscriptions to open access publishing. What follows is a general description of these processes, followed by a case study discussion of their implementation at the authors' institution.

The Agreement

Not all library support of open access publishing requires a legal agreement. For example, Subscribe to Open, Diamond open access support, and ad hoc institutional APC payments

are all methods of funding open access publishing that can take place outside of any ongoing legal contract. Read and Publish initiatives generally require signing an agreement by their nature, and it can also be advantageous for libraries to establish agreements with pure open access publishers like PLOS and Frontiers. These agreements establish expectations in a way that is binding for both parties, giving libraries the opportunity to improve the publication process (for example, by requiring deposits to an institutional repository) or to standardize across their open access funding portfolios (for example, by establishing uniform eligibility criteria). More broadly, they provide the framework for a scalable and sustainable transition of library funding from subscriptions to open access publishing.

The actual process of negotiating such an agreement is very similar to that of negotiating a traditional content licensing agreement (and the "read" portion of Read and Publish agreements can be identical), but there are several potential considerations that are unique to these agreements. This is not the place to provide an in-depth description of the wide variety of open access publishing agreements, but some of these unique clauses include:

- **Article cap**: is there a maximum number of articles (or a maximum APC value) that can be published under the agreement? Preferably the answer would be "no", allowing the institution to locally determine its own maximum based on its budget. Alternatively, some agreements do not specify a maximum because they are designed from the outset to cover all eligible articles, regardless of their number.
- **Unused funds/vouchers**: for agreements that rely on a deposit account (or a per-article voucher equivalent), what happens to any unused funds at the end of a term? It is in the library's interest to be able to continue to roll these over from term to term.
- **Affiliation**: how does the publisher commit to identifying eligible authors? Do they take responsibility for this identification, or do they put the burden on researchers or the institution?
- **Eligibility**: which roles are eligible authors under the agreement? This might include both institutional roles (e.g. faculty) and author roles (e.g. corresponding authors). At what stage is eligibility considered—at submission, acceptance, publication? Which publication types are eligible?
- **Content license**: from which Creative Commons licenses may authors choose? Do the authors retain the remaining copyrights?
- **Retroactive conversion**: sometimes eligible publications fall through the cracks of the eligibility process and are only identified as eligible after publication. When this occurs, can these publications be converted to open access after the fact?

Most of these clauses will have a direct impact on the rest of the workflows arising from the agreement, and it will be beneficial if the relevant staff have a working knowledge of each agreement. For this purpose, libraries might use a simple table noting the pertinent details of each clause for each agreement.

Funding Verifications

Generally, a "traditional" Read and Publish agreement requires institutional funders to verify the eligibility of each article before it is applied to the agreement. This is in the interests of the library, ensuring that the limited resources defined by the agreement (an APC deposit, or allotment of vouchers) are not used for ineligible publications.

There are, however, funding models that impose no limits or per-article costs, where such a process is unnecessary. Examples include ACM's tiered model, Flat Fee Agreements and other models from PLOS, and uncapped Read and Publish agreements that are not APC-based. Under these deals, the institution pays an annual fee covering all eligible publications, regardless of their number. Since there are no resource constraints, there is little incentive on the institutional side to ensure that only eligible items are applied to the agreement. However, it should be noted that future renewal costs may be based on the volume of articles published under the agreement. So taking the long view, it may be wise to verify submissions under all agreements. Furthermore, even where there is no local incentive, some agreements require institutions to agree to verify eligibility.

Before library funders can confirm eligibility, the publisher must associate a particular submission with a particular institutional agreement. The methods for establishing this affiliation can vary, including:

- Author e-mail domain
- Institution selected from prepopulated list during submission (these lists may be populated based on organizational identifiers like GRID, ROR or Ringgold)
- Institutional affiliation as stated in the manuscript

It is in the interest of libraries to ask publishers to cast a wide net during this process, potentially using all of these methods in combination, so that submissions matching any one of these criteria would be associated with the agreement. False positives can always be filtered out during the next step, eligibility verification.

The criteria defining an eligible publication can be specified in the agreement, but this legal definition should be as broad as possible, allowing institutions to flexibly define their own local criteria as needed throughout the term of the agreement. In that local context, the library's communications with researchers will likely be the constraining factor. In other words, there can be two sets of "agreements" about what makes a submission eligible—a very broad legal agreement with the publisher, and a potentially more narrow informal set of expectations with one's institutional researchers about what will be published under the agreement.

There is a useful analogy here from the library acquisitions method of patron driven acquisitions (PDA). When use is low and purchase triggers are few, it is safe to define the pool of eligible content as broadly as possible. But when a high level of purchases risks draining the budget, the criteria defining that pool can be tightened. It is at least theoretically possible for libraries to take the same approach with open access agreements. That being said, the necessity of communicating these criteria transparently with researchers does limit their flexibility. Budgets notwithstanding, it is certainly best for libraries to provide consistency and intelligibility in their open access funding eligibility criteria. The publishing process is already confusing, without libraries offering a sort of "APC roulette."

From the library perspective, the actual process of verifying eligibility will depend on the publisher and the tools they have adopted. There are three main methods:

- Proprietary dashboard—some publishers have developed their own user interfaces for viewing metadata and confirming (or denying) eligibility.
- Third-party dashboard—other publishers have opted to use a third-party tool for the same functions, for example RightsLink from the Copyright Clearance Center. Alternatively, libraries can use their own third-party platform, an option that will be discussed in more detail below.

- E-mail—If the publisher doesn't have a purpose-built interface for communicating eligibility decisions (and can't connect to any locally implemented systems), e-mail correspondence is the universal backup option.

Regardless of the tool used, the process requires checking whether the submission is eligible for funding. This can be done manually, automatically, or both. For examples of automation, it may be possible to automatically verify institutional role with an automated lookup to the university's personnel directory, or to configure a maximum APC in the verification dashboard. Even when automations are in place, automatically declined submissions should be manually reviewed for problems with the process or potential exceptions.

Once the eligibility has been confirmed or denied, the publisher will communicate that decision to the authors. In cases of denial, there may have been an opportunity for the library to state the reason for the denial, which will typically be included in the same author communication. The authors will then have a choice of using alternative funding, finding an alternative venue for publication, or (in the case of hybrid journals), publishing "behind the paywall."

Invoicing

Details on invoicing will depend to some extent on the structure of that agreement. Possible scenarios include:

- **Deposit**: one lump sum intended to cover all publishing under the agreement during the term, which may be replenished as needed.
- **Flat fee**: one lump sum, which covers all publishing under the agreement during the term.
- **APC batch payments**: periodic payments, perhaps monthly, to cover all publishing during the period.
- **Per APC**: an invoice for each publication under the agreement. This option is only recommended for agreements with a low publishing volume.

In all cases, invoicing may be done directly with the publisher or through an intermediary. The intermediary may be a library consortium in the case of consortial deals, or a third-party service provider.

Reporting & Analysis

It is important for libraries that are parties to open access publishing agreements to receive timely, accurate, and consistent reports on how the agreement is performing. At minimum, such reports should be able to answer questions such as:

- How many articles have been published under the agreement?
- How much of our deposit has been spent? How much remains?
- What is the average APC paid under this agreement?
- How many submissions have been approved by the library? How many denied?
- How many eligible publications were not published open access (and why)?

The same "dashboards" described above for confirming eligibility can typically also be used for reporting submission-level details on all items published under the agreement. And here again, where purpose-built dashboards aren't available, e-mail fills in the gaps, with some publishers periodically sending a spreadsheet with this data.

These reports can be used for a variety of important purposes, including:

- Identifying problems with the current agreement: will we run out of funds before the end of the term? Are we receiving many fewer submissions than anticipated?

- Assessing the value of each agreement, which can also inform future negotiations. What factors result in a low cost-per-article, and how can these be reproduced across publishers?
- Assessing the value of the library's open access funding agreements as a whole, which may be a powerful tool in continuing to build support for this approach.

Of course, for Read and Publish agreements, data on open access publishing will only be half of the story, and evaluation of traditional institutional COUNTER usage data will need to be considered as well. (As an aside—the need to evaluate two forms of "outputs" (articles published and items used) is an argument for maintaining two separate costs (Read and Publish) to evaluate them against, rather than lumping both "read" and "publish" under one cost.)

Other data points might be considered as well, such as the usage (globally or institutionally) of the open access articles published under the agreement, or citations of the same.

Due to a lack of standardization and industry norms, the synthesis of this article-level metadata across agreements currently stands out as a significant challenge. Kate Amos, Bethany Harris and Amy Devenney detail the difficulties of collecting, cleaning, and analyzing such data at the level of a large national consortium, but even a single institution will face similar obstacles as it enters into more open access agreements.[20] There is great potential for international collaboration to improve this situation. One such initiative, the OA Switchboard, is discussed below.

Post-Publication Processes

There are two tasks that can only be completed after the submission has been accepted for publication and officially published: verifying open access and depositing the version of record to an institutional repository.

A single open access article can represent thousands of dollars of institutional investment and mistakes can happen. So for hybrid publishers in particular, it is important to verify that all of the items that a publisher claims to have published on an open access basis are indeed open. This check can be performed manually or in an automated way by, for example, passing the DOI to the CrossRef API to retrieve the type of license assigned to the publication.

The Creative Commons licenses assigned to open access publications remove any potential legal barriers to ingesting these publications into an institutional repository in a systematic way. By doing so, institutional repository administrators can get closer to providing a comprehensive collection of their institution's research output and thereby stewarding the future of their institutional research. Ingesting every publication from each open access agreement doesn't need to be a burdensome manual process. The SWORD protocol was developed with just such a use case in mind. SWORD was "designed to facilitate the interoperable deposit of resources into systems such as repositories."[21] By implementing SWORD, a publisher can automatically deposit its publications into the institutional repositories of its authors. While most open access publishers have not yet implemented SWORD, academic libraries and other open access funding institutions should continue to press for this functionality.

The Role of Third Party Tools

The growth of open access publishing, and Read and Publish agreements in particular, has spurred the development of a number of tools or systems designed to ease the implementation of open access funding schemes.

Most importantly, the OA Switchboard is a new initiative, developed in Europe but global in scope, to create a centralized data exchange hub that facilitates the automated delivery of open access publication data among publishers, author institutions, and research funders. This hub has been in operation since 2021, and currently sends two basic kinds of "messages"—one message inquires about the eligibility of a submission, the other notifies that a submission has been published. The OA Switchboard represents a collaborative and open solution to the problem of open access workflows—instead of each publisher running its own proprietary system to serve its institutional open access funders, under this option it sends standardized messages via the Switchboard. The great benefit of the OA Switchboard is the standardization of the format and delivery method of these communications between the relevant stakeholders. This not only has the potential to streamline workflows for libraries and institutional funders, but it also potentially reduces the barriers to entry to open access publishing for small and medium academic publishers.

The OA Switchboard is not, primarily, a front-end user interface—it is agnostic about the destination of the messages it relays. E-mail notifications may be the most basic option for delivering OAS messages, but this avenue is technically limited. This is where the third-party dashboards described above, sometimes called "federated OA account management systems" can play a role.

This emerging category of digital service platforms is designed to ease the burdens of institutional open access management. It includes Oable (from Knowledge Unlatched/Wiley), and Chronos Hub. Institutions with the necessary resources, especially large consortial funders, may have the opportunity to develop their own solutions customized to their local contexts. These systems function as a dashboard, a single interface to handle open access workflows across (ideally) all publishers, including those described above: funding verifications, reporting, and invoicing.

A significant challenge faced by these services is integration with publisher data systems. To exchange data in an automated way between these platforms and any given publisher system requires an investment of legal and technical expertise on both sides, investments that increase as the number and variety of these systems increase. The OA Switchboard is thus poised to solve a real problem by offering a centralized neutral hub for publishers and institutional funders to send and receive standardized article-level data.

Open Access Workflows At Iowa State University

Iowa State University is a public land-grant university with approximately 1,500 faculty. According to Dimensions, the bibliographic index from Digital Science, Iowa State authors published 3,945 articles in 2021. The University Library signed its first open access agreements in 2019, and at the time of writing has agreements providing for the open publication of Iowa State research with sixteen publishers. While the foregoing has been a general description of considerations related to open access workflows, this section describes this work in this specific institutional context, focusing in particular on its largest open access agreement, a Read and Publish agreement with Wiley.

Open access workflows at Iowa State have been established and are overseen in the library's Electronic Resources unit. For the library's earliest open access agreements, however, the workflows did not have an organizational home and were handled jointly by the library's Scholarly Communications Librarian and the Collections program staff. This approach be-

came untenable when the volume of articles covered by the increasing number of agreements ballooned. The Electronic Resources unit offered several advantages as a permanent open access workflow home. First, after shifting and prioritizing current responsibilities, the unit offered staff time that could be permanently reassigned. Second, the ER unit staff brought experience and expertise working with publishers and publisher platforms. And finally, the ER unit already worked closely with the Collections program on licensing and access, making the transition of the new responsibilities somewhat seamless.

The Agreement

The first stage of adopting a new open access agreement is typically initiated by Collections program staff, once they have identified prospective publishers with common ground with the library on pricing and models. Once the costs and basic structure of an agreement are in place in a draft, the E-Resources Librarian reviews the agreement against a standardized rubric that includes the library's preferred clauses and requests any necessary changes. As with a standard content licensing process, there can be several rounds of edits and, when successful, the resulting document is inevitably a compromise between the two parties.

Getting Started

Once the agreement is finalized, processes are initiated by both the publisher and library to implement it. The publisher activates their eligibility processes and may create a dashboard account for the library. Both entities will typically take measures to promote the new agreement. Generally this includes, at the very least, mutually listing the other party in a public list of open access agreements. Iowa State University maintains a page on its website listing current open access agreements,[22] with the intended audience of researchers looking for open access publishing options. This page is intentionally simple and glosses over many of the complexities described above. For each agreement we only state who is eligible (typically "Iowa State Corresponding Author") and which publications, with a link to a more detailed list where available.

More can be done in terms of promotion, such as a press release or e-mails to faculty. At Iowa State, to retain a position of neutrality in the publishing industry's competition for the university's research, the library is typically reluctant to do more than a press release. Also note that the incentive structure of an agreement may have a bearing as well. Librarians should carefully reflect before heavily promoting an agreement based on APC payments per article.

The Iowa State agreement with Wiley began with a kickoff meeting where the details of the workflow were reviewed. At this meeting, Wiley shared a series of screenshots showing the author workflow, which have often been referred to since, in order to help answer questions from Iowa State researchers. Whether in the form of screenshots or a video, libraries should ask for detailed information on the author experience. This information can help identify problems encountered by their researchers, which can result in valuable feedback to the publisher on improvements to their author services. For example, one agreement was underperforming, and it was only when the library saw what submitting authors see that the cause was understood. For some journals from this publisher, authors were being asked to select a publishing agreement from a list of three or four options, and there was nothing clearly steering them to the open access option. This information allowed Iowa State to communicate with the publisher about improvements to this process (work which was already ongoing) and know where researchers were likely to encounter a barrier to open access publication.

Funding Verifications & Reporting

In 2021, the first year of the Iowa State agreement with Wiley, open access publishing in both hybrid and fully Gold open access journals was paid from a deposit based on a discounted APC rate that varied per journal. Iowa State is notified by e-mail when a new Iowa State article has been submitted (for Gold open access journals) or accepted (for hybrid journals), and at that point relevant metadata can be reviewed and eligibility can be verified on a proprietary Wiley dashboard. This dashboard also provides a summary of the library's account, including the total amount of the deposit spent thus far.

Because Iowa State maintains very broad criteria for eligibility, the actual eligibility verification process is simple. When a new submission arrives, staff look up its corresponding author in the institutional personnel directory to verify that they are currently affiliated with the university. In those cases where the corresponding author can't be found, a search is done online to attempt to confirm that they are not affiliated before denying the funding.

Beginning in 2020, Iowa State has worked with the OA Switchboard to support its development, as well as with Knowledge Unlatched as beta development partners of their Oable platform. The library's motivation was both to implement a scalable open access management solution, with publishers routing messages through the OA Switchboard that are delivered to their Oable dashboard, but also to support services that can help accelerate the broader open access transition globally. More concretely, the library wanted a single platform for eligibility verification and reporting. That the "single platform" goal hasn't been reached for either system is due to the challenge of publisher buy-in described above. Publishers need to agree to share their data with these systems, and invest staff time in implementing the legal and technical requirements of such a connection. Out of sixteen Iowa State agreements, nine publishers were working at time of writing with one or both systems to some extent, leaving piecemeal workflows for the remaining seven.

Iowa State continues to make the case to the remaining publishers for working with the OA Switchboard, which should allow their data to flow to whichever platform their institutional partners prefer, whether that is Oable or another similar product.

Even without a direct data connection, with staff assistance it is possible to manually upload article-level metadata to Oable for reporting purposes. This method sits uneasily alongside the more direct data connection method, particularly as publishers switch from one method to another, and care must be taken to avoid duplication.

In August 2021, Wiley established a data connection directly with Oable, and in summer 2022 phased out their own proprietary dashboard, in favor of using Oable for all institutional customers. This has allowed Iowa State to verify eligibility, report on publishing, and monitor the deposit from the same platform as (some of) the library's other agreements.

Invoicing

The invoicing methods used at Iowa State are a function of the size and structure of the agreement. Large agreements, as well as Read and Publish agreements, are typically paid annually, either as a deposit or (more commonly) as a flat fee. Smaller agreements, which may only result in one or two open publications each year, are more often handled on a per-APC basis, especially for pure open access publishers.

Post-Publication Processes

Above, two post-publication processes were identified: open access verification and institutional repository deposits. Both processes can theoretically be automated, but at Iowa State the necessary systems have not been in place. It is likely that in the future, automated OA verification will be a feature offered by open access management systems like Oable. Until that is in place, library staff manually check each publication that appears in the reporting to ensure that it is indeed open to all.

Automated institutional repository deposits typically require a SWORD-enabled system. The library's previous commercial repository system did not have this functionality, but a recent migration to an open source SWORD-enabled option now allows set up with publishers who have already agreed to it.

Due to the great variety of models even within the library's sixteen agreements, a valuable reference tool for staff responsible for these workflows has been a simple spreadsheet listing the following values for each agreement:

- Publisher
- Eligible authors
- Eligible publications
- Term begin
- Term end
- Status (active, expired, etc.)
- Total annual cost (current or most recent term)
- Model (Read and Publish, etc.)
- Publishing limit (describes whether agreement includes limit on publishing under the agreement)
- Approvals method (Oable, e-mail, etc.)
- Approvals frequency
- Approvals stage (submission/acceptance)
- Reporting method (Oable, proprietary dashboard, etc.)
- Reporting frequency
- Invoice method (direct, consortium, etc.)
- Invoice frequency
- PO Line
- Contact
- Notes

While the lack of standardization can make this work complex, it should be noted that it is generally not time-intensive. At Iowa State, even with sixteen agreements, the total time spent on the tasks described above does not add up to one full-time position. As the number of articles and agreements continues to grow, it is hoped that greater standardization, aided by important initiatives like the OA Switchboard, will allow this to continue to be true.

Conclusion

Like scholarly communication more broadly, open access workflows are in flux. Besides adapting to future changes and refining the processes described above, librarians involved in this work should explore methods of evaluating it. That means evaluating not only the performance of the agreements themselves, but the work of implementing them described

above. How might managers assess the productivity, efficiency, effectiveness or accuracy of open access agreement implementation? When do we know we are doing well? Is it possible to answer these questions systematically when such wide variations exist between agreements?

These questions raise interesting avenues for future research. Because OA workflows are relatively new, little is known about the attitudes of technical services staff towards them. What is their level of understanding and interest? It would be helpful to know, as well, more broadly how US and North American academic libraries are establishing and approaching OA workflows. Where do they live in the organization? How much staff time do they require? How are staff being trained and supported to be successful? Further exploration of these topics would assist libraries in negotiating and implementing OA agreements.

It is a challenge to describe processes where little uniformity exists, whether between publishers or across time. Open access publishing is in a period of robust experimentation, with changing models cropping up regularly as publishers search for sustainable business strategies. At the same time, new infrastructure, products and services are regularly emerging to help facilitate this relatively new library function. Those responsible for implementing the workflows required by open access agreements must therefore be highly adaptive. They must also be vigilant. While a database outage is likely to be brought to the attention of library staff, a paywalled article that ought to be open might not ever be noticed by anyone. It falls to the staff responsible for this work to hold publishers accountable to their agreements, ensuring that every eligible publication is made freely available to all.

Notes

1. The ESAC Transformative Agreement Registry can be searched by country: https://esac-initiative.org/about/transformative-agreements/agreement-registry/.

2. Cambridge University Press, "OA Agreements in the United States," https://www.cambridge.org/core/services/open-access-policies/read-and-publish-agreements/united-states.

3. John Harle and Verity Warne, *Open Access: Challenges and Opportunities for Low- and Middle-Income Countries and the Potential Impact of UK Policy* (London: Foreign, Commonwealth & Development Office, 2020), https://www.gov.uk/research-for-development-outputs/open-access-challenges-and-opportunities-for-low-and-middle-income-countries-and-the-potential-impact-of-uk-policy; Sara Rouhi, Romy Beard and Curtis Brundy, "Left in the Cold: The Failure of APC Waiver Programs to Provide Author Equity," *Science Editor*, 45 no. 1: 5–13. https://doi.org/10.36591/SE-D-4501-5.

4. It should be noted that while much of the discourse around open access publishing focuses on scientific research outputs, the agreements that are the focus of this article potentially cover all academic disciplines, including those in the social sciences and humanities.

5. Michael M. Crow and Greg Tananbaum, "We Must Tear Down the Barriers That Impede Scientific Progress," *Scientific American* (December 18, 2020), https://www.scientificamerican.com/article/we-must-tear-down-the-barriers-that-impede-scientific-progress.

6. UNESCO, "UNESCO Recommendation on open science" (2021), https://unesdoc.unesco.org/ark:/48223/pf0000379949.

7. Greg Tananbaum, "North American Campus-Based Open Access Funds: A Five-Year Progress Report" (SPARC, 2014), https://sparcopen.org/wp-content/uploads/2016/01/OA-Fund-5-Year-Review.pdf.

8. Greg Tananbaum, "Campus-Based Open-Access Publishing Funds: A Practical Guide to Design and Implementation" (SPARC, 2010), https://sparcopen.org/wp-content/uploads/2016/01/oafunds-v1.pdf .

9. Stephen Pinfield, "Paying for Open Access? Institutional Funding Streams and OA Publication Charges," *Learned Publishing* 23 (2010): 39–52, https://doi.org/10.1087/20100108.

10. Heidi Zuniga and Lilian Hoffecker, "Managing an Open Access Fund: Tips from the Trenches and Questions for the Future," *Journal of Copyright in Education & Librarianship* 1 (2016), https://doi.org/10.17161/jcel.v1i1.5920.

11. Gail McMillan, Leslie O'Brien and Philip Young, "SPEC Kit 353: Funding article processing charges,"

(ARL, 2016), https://publications.arl.org/Funding-Article-Processing-Charges-SPEC-Kit-353/.

12. Kai Geschuhn and Graham Stone, "It's the Workflows, Stupid! What Is Required to Make 'Offsetting' Work for the Open Access Transition," *Insights*, 30 (2017): 103-114, https://doi.org/10.1629/uksg.391.

13. Adriana Sikora and Kai Geschuhn, "Management of Article Processing Charges – Challenges for Libraries," *Insights*, 28 (2015): 87-92, https://doi.org/10.1629/uksg.229.

14. Rita Pinhasi, Guido Blechl and Brigitte Kromp, "The Weakest Link – Workflows in Open Access Agreements: The Experience of the Vienna University Library and Recommendations for Future Negotiations," *Insights* 31 (July 2018): 27, https://doi.org/10.1629/uksg.419.

15. Efficiency and Standards for Article Charges, "Customer recommendations for article workflows and services for offsetting/ open access transformation agreements," (2017), https://esac-initiative.org/wp-content/uploads/2017/04/ESAC_workflow_recommendations_1st_draft20march2017.pdf.

16. Efficiency and Standards for Article Charges, "ESAC Workflow Recommendations for Transformative and Open Access Agreements: 2021 Enhancement to the ESAC Workflow Recommendations," (2021), https://esac-initiative.org/about/oa-workflows.

17. George Machovec, "Selected Tools and Services for Analyzing and Managing Open Access Journal Transformative Agreements," *Journal of Library Administration* 60 (2020): 301–7, https://doi.org/10.1080/01930826.2020.1727280.

18. Mafalda Marques, Saskia Woutersen-Windhouwer, and Arja Tuuliniemi, "Monitoring Agreements with Open Access Elements: Why Article-Level Metadata Are Important," *Insights* 32 (2019): 35, https://doi.org/10.1629/uksg.489.

19. Jill Emery, Graham Stone and Peter McCracken, *Techniques for Electronic Resource Management: TERMS and the Transition to Open* (Chicago: ALA Editions, 2020).

20. Kate Amos, Bethany Harris, and Amy Devenney, "Monitoring Transitional Agreements: The Challenges and Successes of Implementing Article-Level Metadata Collection," in *14th International Conference on Performance Measurements in Libraries, November 2-4, 2021* (Library Performance Measurement Conference), https://library-performance.files.wordpress.com/2021/12/libpmc14_2021_proceedings.pdf.

21. Stuart Lewis, Pablo de Castro, and Richard Jones, "SWORD: Facilitating Deposit Scenarios," *D-Lib Magazine* 18 (January 2012), https://doi.org/10.1045/january2012-lewis.

22. Iowa State University Open Scholarship Services, "Open Access Agreements," https://www.lib.iastate.edu/research-and-teach/publish-and-share/publish-open-access#agreements

Longitudinal Associations between Online Usage of Library-Licensed Content and Undergraduate Student Performance

Felichism Kabo, Annaliese Paulson, Doreen Bradley, Ken Varnum, Stephanie Teasley

Seeking to better understand the longitudinal association between online usage of library-licensed content and short- and long-term student performance, we linked EZproxy logs to institutional university data to study how library usage impacts semester and cumulative GPAs. Panel linear mixed effects regression models indicate online library usage is significantly associated with both semester and cumulative GPAs. The library usage effect is larger for semester GPA, and varies by on- and off-campus residency. The effect on semester GPA is larger for off-campus students, while for cumulative GPA the effect is larger for on-campus students. Longitudinally linked library-institutional data offers key insights on the library's value.

Introduction

Library usage is correlated with important undergraduate student outcomes including academic performance and retention. However, the relationship between library usage and academic performance is better understood over the short term, and for specific subsets of students, such as first-year undergraduate students.[1] We need to develop a better understanding of this relationship both over the long term, and for all undergraduate students. One reason for our currently limited understanding of this relationship is that, in most universities—owing to privacy concerns—libraries either do not collect or retain user data with identifiers. This makes it impossible to link library usage data with other institutional or administrative data from the university, including data regarding academic success and retention. Another limitation is that library usage data are often collected as very large logs (millions and billions of records) that may require the application of methodological approaches, such as Big Data techniques, to structure and store in ways that make them more amenable to analysis. Therefore, there

* *Felichism Kabo is Director of Research, CannonDesign, email: fkabo@cannondesign.com and Research Fellow, Zell Lurie Institute, Ross School of Business, University of Michigan, email: fkabo@umich.edu; Annaliese Paulson is a PhD student at the School of Education, University of Michigan, email: annamp@umich.edu; Doreen Bradley is the Director of Learning Programs and Initiatives at University of Michigan Library, email: dbradley@umich.edu; Ken Varnum is a Senior Program Manager and Discovery Strategist at University of Michigan Library, email: varnum@umich.edu; Stephanie Teasley is a Research Professor at the School of Information, University of Michigan, email: steasley@umich.edu.*

is a need for empirical, longitudinal studies that not only use identifiable library data, but also employ Big Data and statistical methods to advance our understanding of the library's contribution to student success. In this paper, we present the results of a longitudinal study of the association between online library resource usage and student performance for the entire population of undergraduates enrolled at the University of Michigan (U-M) between 2016 and 2019.

The privacy concerns described above are valid; however, other research domains—for which the potential risk of unintended exposure is higher than those of library usage data, such as the type of patient health information covered by the Health Insurance Portability and Accountability Act of 1996 (HIPAA) —have found ways to successfully handle data while maintaining privacy. Yet, these advances in the biomedical and social sciences, which would better serve the privacy requirements of library professional ethics, are still not widely known in libraries. Fortunately, many libraries now adopt the best privacy practices from the social and biomedical sciences. These initiatives make it possible to employ Big Data methods in longitudinal studies of the links from library usage to academic outcomes for the entire student body.

There are two such initiatives critical to the work described in this paper: first, after a multi-year process of engaging with a diverse set of stakeholders including the U-M Learning Analytics Task Force, the U-M Library revised its privacy policy in 2016 to allow the collection and retention of identifiable library usage data;[2] second, the Library Learning Analytics Project (LLAP)—funded by the Institute of Museum and Library Services (IMLS)—examined how libraries impact learning outcomes including in course instruction. Learning processes require that members of the university community engage in activities such as accessing digital data and publication repositories, conducting literature reviews, managing citations, and creating data management plans. These activities often entail interacting with the library virtually, such as when accessing and retrieving library licensed content through the proxy server. This paper reports on analyses performed on the links between off-campus, or off-network, electronic usage of library resources, as well as undergraduate academic performance over the short- and long-term. The best context for work of this nature is one in which library users have agency with how they engage with the library services in question. For library licensed content, individuals can access these resources via computers that are on-campus (physically located in the library or elsewhere in the university), or virtually via the proxy server should they choose to use these resources when off-campus. For this reason, the authors limited the analysis to the relationship between online library usage and student outcomes to the time before the COVID-19 pandemic. That is, the study focuses on when students had the choice of accessing library licensed content through on- or off-campus means.

Literature Review

This work is informed by models of information behavior,[3] which describes how individuals seek and utilize information.[4] Information behavior is contingent on factors such as social contexts, socio-demographics, individual expertise, as well as access to, and ease of use of, technology.[5] The work also builds on two lines of inquiry: 1) research into the associations between college residence and academic performance; 2) work on digital inequalities or the digital divide. We examine the link from library usage to student outcomes in two ways: first, defining library usage in terms of use of licensed online content provided by the library,

and second, evaluating the impacts of on-campus residency for access to library and other resources and reliable internet.

Research on campus residency has examined the issue of whether there are gains in learning and academic performance from living on- versus off-campus. A study of nearly 95,000 first year students in the United States found living on-campus was significantly associated with a range of learning variables, even though the residency effect size was small to medium.[6] An earlier study of first-year students found that the benefits of on-campus residency on academic performance were different across, and within, racial groups. For example, Black students who lived on-campus had significantly higher grade point averages (GPAs) than Black students who lived off-campus.[7] Approaching the issue from a different angle, a study of the causal link between campus residency and academic outcomes found living in university-owned housing had a positive association with student retention.[8] This finding was in line with prior analysis that established an association between on-campus living and academic performance and student retention for first-year students.[9] However, an important caveat is that students who were better prepared academically were more likely to live on-campus as opposed to off-campus.[10] Most studies of the link between on-campus residence and student persistence are based on four-year institutions. One exception is a quasi-experimental analysis of community college students that found that living on-campus was associated with a significant increase in upward transfer (to a four-year institution) and, subsequently, bachelor's degree completion rates.[11] However, the association between on-campus residence and academic outcomes is not always positive. A study conducted at a public four-year university in the southeast United States found that commuter or off-campus students had higher GPAs than residential or on-campus students.[12]

Demographic, geographic, and economic factors all help shape digital disparities in American K-16 education. These disparities are commonly referred to as the "digital divide," or the gap between those who have access to the internet and other information and communication technologies (ICT), and those who do not. Digital inequalities and disparities affect a broad range of life opportunities and outcomes beyond education, such as economic activity and health care.[13] In education, digital inequalities and disparities are a life-course issue and affect disadvantaged students. Their effects are felt from early[14] to late in the K-16 pipeline.[15] The increasing use of technology inside and outside the classroom has significant ramifications for the digital divide and its effect on student performance. Importantly, some groups of students are systematically more likely to experience digital disparities than others. For example, in 2015, higher percentages of students who were White (66%) used the internet at home compared to Black (53%), Hispanic (52%), and American Indian/Alaska Native (49%) students.[16] American Indian/Alaska Native students are more likely than other racial groups to have no internet access, or to have only dial-up internet access at home.[17] The interaction of demography and geography disadvantages some students further. While 18 percent of all students in remote rural areas did not have internet access, or had only dial-up access in 2015, a much larger percentage of Black (41%) students in remote rural areas did not have internet access compared to White (13%) and Asian (11%) students. Having no or low-bandwidth internet is detrimental to any form of online learning. For example, students cannot participate in classes offered via video meeting systems that rely on high-speed internet.[18] The COVID-19 pandemic worsened the effects of the digital divide, such as for rural students.[19] Students of color have been especially impacted by the pandemic and, as noted earlier, are more likely to

lack access to reliable broadband internet, and even computers. The pandemic exacerbated existing educational disparities for minority students and likely widened the achievement gap for students of low socioeconomic status.[20]

In the United States, the effects of the pandemic on the digital divide have demonstrably impacted the entire K-16 pipeline. There were varied institutional responses across the American higher education landscape. Perversely, these varied responses present opportunities for "quasi-experimental" observations regarding the impact of the digital divide on amplifying disparities in student performance. For example, where many colleges and universities stipulated that students residing on-campus leave these residences, some made allowances for students who could not return home, which thus allowed them to still have access to reliable broadband internet via the institution.[21] What was fairly universal, however, was the extent and speed with which university libraries adapted to offering primarily online resources,[22] which can only meaningfully be accessed via reliable internet connections. Thus, not only were students no longer able to access the library's physical collections, but they also no longer had access to the library as a study space, including for group or collaborative activities.[23] By examining how "regular" (pre-pandemic) electronic library usage is associated with academic performance, this study may therefore help us better understand the likely impacts of the worsening of the digital divide during the pandemic. Based on evidence that the digital divide has worsened during the pandemic,[24] we can reasonably assume that the importance of the relationship between online library usage and academic performance has only increased.

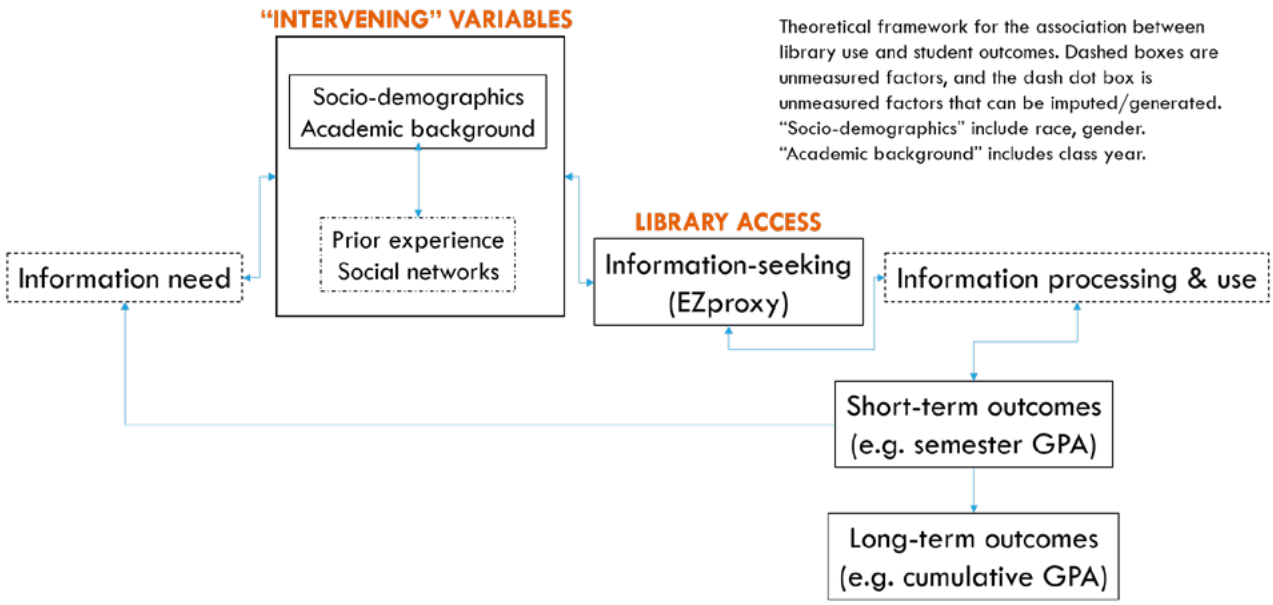
The literature also indicates that models of student performance need to account for other demographic, socioeconomic, and academic factors, including include gender, first-generation status, family or household income, high school GPA, and academic class level. Across national contexts in developed countries, female students are more likely to have both higher work ethics and GPAs than males.[25] First-generation students are more likely to contend with barriers to academic success—such as job and family responsibilities and/or inadequate study skills[26]—and thus tend to have poorer academic outcomes.[27] Students who enter college with higher family or household incomes have significantly higher GPAs than those from lower socioeconomic backgrounds.[28] High school GPA is a strong predictor of college or university GPA as well, especially in the first year.[29] Academic class level is correlated with GPA, as upper class students (e.g. seniors) are more likely to have higher grades, especially in classes that also have lower class students, such as sophomores.[30]

Theoretical Framework

Building on models of information-seeking behavior, we developed a theoretical framework (Figure 1) that correlates student performance with library usage as captured by EZproxy sessions, controlling for factors like socio-demographics and academic background.[31] A key strength of the framework is that it presents testable relationships among demographic and contextual factors, information-seeking behaviors, and academic outcomes.

This paper examines the association between information-seeking behavior (off-campus or off-network electronic library resource use), and both semester and cumulative GPA. However, this relationship must also be understood in the context of contextual factors ("intervening" variables), which contribute to disparities in access to the digital resources that are needed to make effective use of electronic library licensed content. Research shows that access to, and proper use of, digital technology generally has a positive correlation with academic

FIGURE 1
Theoretical Framework for Associations between Library Usage and Student Outcomes Adapted from Models of Information Behavior[32*]



*These models draw on research from multiple fields including information science, psychology, decision-making, innovation, health communication, and consumer research.

performance; this finding is robust across regional and national settings.[33] Based on these findings, we hypothesize that students identified as accessing online library licensed content will have better academic outcomes than those students with no evidence of digital access to these resources. However, there is also evidence that our hypothesized relationship has both short- and long-term implications. While not specific to electronic resources, studies suggest that library usage is positively correlated with student performance both in the short-term,[34] and in the long-term.[35] Therefore:

H1: Students who electronically access library licensed content will have higher semester GPAs.

H2: Students who electronically access library licensed content will have higher cumulative GPAs.

Methodology

The study sample is all undergraduate students (N = 45,254) who were enrolled at the University of Michigan (U-M) from fall 2016 through winter 2019 (or September 2016 through April 2019). We focus on these six semesters before the pandemic because students had more agency with respect to their usage of electronic library licensed content. That is, students could choose to access materials using computers that are physically on-campus, or off-campus access via the proxy server. We sourced library usage data from EZproxy logs (690,300,076 records) stored in a secure repository that the U-M Library managed. We obtained student demographic and outcome data (GPAs) from the research-focused Learning Analytics Data Architecture (LARC) data set maintained by the U-M Office of Enrollment Management. The project team implemented several measures to protect the privacy and confidentiality of the

individuals in the library and LARC data. For example, the library data were classified at the "Restricted" level of data security. This is the highest classification or sensitivity level for U-M institutional data, has the most stringent legal or regulatory requirements, and has the most prescriptive security controls. These controls included restricting access to only two members of the project team, storing and curating the data on a secure enclave, setting up access to the enclave via a terminal in a locked and restricted data room, and requiring that all analyses be performed on the enclave.

Our primary interest in this paper is the relationship between information-seeking behavior (EZproxy sessions) and student performance. EZproxy is proxy server software that many academic libraries use to give authenticated off-campus users access to electronic resources licensed by the library as if they were on campus. After authenticating to a campus system, off-campus users receive an on-campus IP address and are then considered to be a member of the campus community by the information provider. The authors cleaned and normalized raw, unstructured EZproxy logs using Python scripts and regular expressions, and then entered the data into a relational database using structured query language (SQL) scripts. Over 80 percent of the EZproxy data have strong university identifiers which enables merges with other administrative data, such as LARC. It is critical to note that EZproxy logs available to the study: a) did not include any on-campus usage, and b) did not include anyone who used the university's virtual private networks (VPN). Using SQL and R scripts, we merged the data and exported the resultant data set into Stata 16 statistical software for modeling and analysis.[36]

The theoretical framework shown in Figure 1 suggests that student outcomes are a function of factors, such as race and gender, that apply to all the students in the study ("fixed effects"), and factors, such as academic units or schools, that cluster student behaviors and outcomes ("random effects"). We also accounted for student random effects for unobserved, time invariant factors, such as motivation or grit. Thus, we ran panel linear mixed effects (LME) regression models of the association between library usage and student GPA, contingent on students being enrolled in at least four semesters over the study period.

Variables

The two continuous dependent variables are semester GPA ("SEM_GPA") and cumulative GPA ("CUM_GPA"). While SEM_GPA is on a 0 – 4.4 scale and CUM_GPA is on a 0 – 4.314 scale, fewer than 0.5 percent of students have a semester or cumulative GPA that is higher than 4.0. The dichotomous independent variable "EZproxy Session in Term" is coded one if a student is associated with one or more EZproxy sessions during an academic term, and is coded zero otherwise.

We also account, or control, for potential "intervening" variables as follows: the dichotomous variable "On-campus Residence" is coded one if a student was residing in a university residence, and zero otherwise; the variable "High School GPA" is on a continuous 0 – 4 scale and captures a student's academic performance before enrollment at the university; gender is captured by the dichotomous variable "GENDER" (1 = Female, 2 = Male). Note that the LARC data set used for the study does not account for non-binary options. The effects of race, first generation status, family income, and class level were controlled for using the categorical variables "RACE" (1 = White, 2 = Asian, 3 = Black, 4 = Hispanic, 5 = Two or More, 6 = Other, 7 = Not Indicated), "FIRST GENERATION" (1 = First Gen, 2 = Not First Gen, 3 = Don't Know),

"FAMILY INCOME" (1 = More than $100,000; 2 = Less than $25,000; 3 = $25,000 - $49,999; 4 = $50,000 - $74,999; 5 = $75,000 - $99,999; 6 = Don't Know; 7 = Missing), and "CLASS LEVEL" (1 = Freshman, 2 = Sophomore, 3 = Junior, 4 = Senior), respectively.

Statistical Modeling

We ran panel LME regression models with random effects for individuals, as well as by school or academic unit (see Table A.7 in the appendix for a list of the 15 schools that undergraduate students were affiliated with). LME models, an extension of simple linear models, are useful when there is non-independence in the data. This arises from, for example, a hierarchical structure in the data, such as when students are sampled from within academic units. Panel regression approaches are necessary when working with longitudinal study designs, where multiple observations are made on each individual subject. LME models have both fixed effects, which are directly estimated and are analogous to standard regression coefficients, and random effects, which in our case take the form of random intercepts. The fixed effects in our LME models correspond to the "intervening" variables. The random effects account for the fact that student behaviors and outcomes may, instead of being uniform across all undergraduates, be grouped by academic units which map onto disciplinary boundaries that likely affect library usage. The random effects also enable us to account for unobserved, time-invariant individual-level factors, such as motivation or grit. Table A.7 in the appendix shows that there are notable differences across schools with respect to the percentage of students who have at least one EZproxy session during an academic term. After each LME model, we ran a likelihood-ratio comparing this model with a one-level ordinary linear regression. This test was highly significant for each of the LME models in our study, supporting the decision to use the LME model.

Findings and Discussion

Descriptive Statistics

Over half of enrolled undergraduates had at least one EZproxy session during an academic term over the study period (Table 1).

There are some notable differences in library usage among enrolled undergraduates. Table 2 below illustrates differences in library usage by demographic, academic, and residency factors for the winter 2019 term (see the appendix for similar statistics on all semesters). Off-campus students are more likely to have at least one EZproxy session in the academic term than are on-campus students. This makes sense because students who are on-campus are more likely to access electronic library resources on the university's network, in which case authentication is not required. Recall that students are identifiable in the EZproxy logs only when authentication is required. An example of this is when a student accesses electronic library resources outside

TABLE 1
Percentage of Students Associated with EZproxy Sessions by Semester, Fall 2016 – Winter 2019

Academic Term	Enrolled Students	EZproxy Session	% ≥ 1 EZproxy Session
FA 2016	28,682	16,605	58%
WN 2017	27,408	13,434	49%
FA 2017	29,161	16,034	55%
WN 2018	27,852	14,855	53%
FA 2018	29,726	16,191	54%
WN 2019	28,355	16,299	57%
TOTAL*	171,184	94,418	55%

*This is a tally of unique student-term combinations, as there were 45,254 enrolled undergraduates over the study period.

TABLE 2
Percentage of Undergraduate Students Associated with EZproxy Sessions by Socio-Demographics and Academic Background, Winter 2019

Variable	Category	Enrolled Students	EZproxy Session	% ≥ 1 EZproxy Session
First Gen Status	First Gen	3,890	2,310	59%
	Not First Gen	24,418	13,957	57%
	Don't Know	47	32	68%
Family Income	Less than $25,000	1,507	923	61%
	$25,000–$49,999	2,212	1,269	57%
	$50,000–$74,999	2,009	1,217	61%
	$75,000–$99,999	2,074	1,213	58%
	More than $100,000	13,951	7,892	57%
	Don't Know	515	278	54%
	Missing Income Information	6,087	3,507	58%
Class Level	Freshman	2,557	1,300	51%
	Sophomore	6,397	3,373	53%
	Junior	7,132	4,114	58%
	Senior	12,269	7,512	61%
Race	Asian	5,829	3,137	54%
	Black	1,268	766	60%
	Hispanic	1,899	1,099	58%
	White	16,604	9,738	59%
	2 or More	1,302	745	57%
	Other	46	22	48%
	Not Indic	1,407	792	56%
Gender	Female	14,204	9,219	65%
	Male	14,151	7,080	50%
Residency	On-campus	9,261	4,540	49%
	Off-campus	19,110	11,765	62%
Academic Unit	Architecture	181	124	69%
	Art and Design	524	381	73%
	Business Administration	1,799	740	41%
	Dental Hygiene	101	70	69%
	Education	126	54	43%
	Engineering	6,313	2,847	45%
	Information	260	122	47%
	Joined Degree Program	10	7	70%
	Kinesiology	954	678	71%
	Literature, Science and the Arts	16,409	10,030	61%
	Music, Theare, & Dance	717	515	72%
	Nursing	607	475	78%
	Pharmacy	55	36	65%
	Public Health	157	116	74%
	Public Policy	142	104	73%

the university's network such as from an off-campus residence, coffee shop, etc. There is a significant gender difference, with females much more likely than males to have an EZproxy session, despite more males (69%) than females (66%) residing off-campus in winter 2019. Note that the likelihood of having at least one EZproxy session increases with each class level. Perhaps this is because students are more likely to move or reside off-campus as they progress from freshman to seniors. However, a factor that weakens this explanation is U-M does not require freshmen and sophomores to live on-campus, as is the case in some colleges and universities. An alternative explanation is that lower-level classes are less research-intensive and students may not need library-provided resources to complete research and writing projects.

Finally, there are noteworthy differences between academic units. Additional work would be needed to clarify the factors that account for these differences. For example, 45 percent of engineering undergraduates had at least one EZproxy session compared to 73 percent of art and design undergraduates, even though both academic units are co-located at the university. A potential explanation could be that these differences reflect disciplinary differences (STEM versus arts and humanities). Another plausible explanation could be that the differences reflect gaps in technological expertise between the two groups of students, with engineering students being more likely to access electronic library resources using the university's VPN which bypasses the authentication process on the library's proxy server. We should also keep in mind factors such as the interplay between residency and socioeconomic statuses. It is more expensive to live on- rather than off-campus, implying that students in the former group may tend to be from wealthier families. For example, 78 percent of nursing undergraduates had at least one EZproxy session, compared to 41 percent of business administration undergraduates. Tabulations of residency for the two academic units showed that 32 percent of business undergraduates resided on-campus in winter 2019, compared to 20 percent of nursing undergraduates. Similarly, tabulations of family income for these two academic units showed that 58 percent of business undergraduates had a family income of more than $100,000, compared to 48 percent of nursing undergraduates. These findings suggest that library usage data have the potential to reveal disparities and inequalities, and could therefore help libraries make significant analytical contributions of interest to their institutions.

Regression Models

The results from the regression modeling are summarized in Tables 3 (semester GPA) and 4 (cumulative GPA). The regression models showed positive and statistically significant associations between having at least one EZproxy session in an academic term, and both semester and cumulative GPAs, controlling or accounting for residency, race, gender, high school GPA, family income, first generation status, and class level.

Overall, the results from the regression models for semester GPA provide strong support for hypothesis **H1**. That is, students that use electronic library licensed content have higher semester GPAs. Having an EZproxy session during an academic term was correlated with a 0.14 point increase in semester GPA (model 1). To further examine the impact of campus residency, considering the link between authentication requirements and a student's presence in the EZproxy logs, we ran separate models for on-campus (model 2) and off-campus (model 3) students. For off-campus students, having an EZproxy session in an academic term is correlated with a 0.17 point increase in semester GPA. In comparison, for on-campus students, having an EZproxy session in an academic term is correlated with a

TABLE 3
Panel LME Regressions for Association between Library Usage and Semester GPA, FA 2016–WN 2019 (Four or More Semesters)

	(1: All Students)	(2: On-campus)	(3: Off-Campus)
VARIABLES	**SEM_GPA**	**SEM_GPA**	**SEM_GPA**
EZproxy Session in Term	0.138***	0.0837***	0.171***
	(0.00304)	(0.00415)	(0.00419)
On-campus Residence	0.0967***		
	(0.00471)		
High School GPA	0.0273***	0.0435***	0.0211***
	(0.00194)	(0.00345)	(0.00235)
GENDER (Reference = Female)			
Male	–0.0908***	–0.0616***	–0.108***
	(0.00529)	(0.00662)	(0.00685)
RACE (reference = White)			
Asian	0.0499***	0.0534***	0.0404***
	(0.00660)	(0.00838)	(0.00851)
Black	–0.376***	–0.374***	–0.400***
	(0.0128)	(0.0145)	(0.0181)
Hispanic	–0.164***	–0.181***	–0.143***
	(0.0107)	(0.0126)	(0.0145)
Two or More	–0.101***	–0.0812***	–0.121***
	(0.0126)	(0.0150)	(0.0167)
Other	–0.239***	–0.209**	–0.255**
	(0.0631)	(0.0781)	(0.0784)
Not Indic	–0.00568	0.0168	–0.0188
	(0.0121)	(0.0160)	(0.0155)
FIRST GENERATION (reference = First Gen)			
Not First Gen	0.119***	0.138***	0.112***
	(0.00851)	(0.0106)	(0.0112)
Don't Know	–0.166**	–0.0157	–0.202**
	(0.0525)	(0.0845)	(0.0640)
FAMILY INCOME (reference = More than $100,000)			
Less than $25,000	–0.150***	–0.129***	–0.166***
	(0.0127)	(0.0159)	(0.0167)
$25,000 – $49,999	–0.101***	–0.115***	–0.102***
	(0.0106)	(0.0131)	(0.0141)
$50,000 – $74,999	–0.0557***	–0.0719***	–0.0581***
	(0.0104)	(0.0133)	(0.0134)
$75,000 – $99,999	–0.0545***	–0.0528***	–0.0572***
	(0.0100)	(0.0129)	(0.0128)
Don't Know	–0.0505*	–0.0385	–0.0688**
	(0.0196)	(0.0238)	(0.0260)
Missing Income Information	–0.00505	–0.0117	–0.00127
	(0.00652)	(0.00827)	(0.00831)

TABLE 3
Panel LME Regressions for Association between Library Usage and Semester GPA, FA 2016–WN 2019 (Four or More Semesters)

	(1: All Students)	(2: On-campus)	(3: Off-Campus)
VARIABLES	**SEM_GPA**	**SEM_GPA**	**SEM_GPA**
CLASS LEVEL (reference = Freshman)			
Sophomore	0.0176***	0.0237***	0.0184
	(0.00498)	(0.00455)	(0.0229)
Junior	0.0326***	0.00259	0.0704**
	(0.00605)	(0.00680)	(0.0229)
Senior	0.0815***	0.0403***	0.116***
	(0.00662)	(0.0112)	(0.0230)
Constant	3.207***	3.242***	3.174***
	(0.0357)	(0.0444)	(0.0448)
Observations	151,049	53,896	97,153
Standard errors in parentheses			
*** p<0.001, ** p<0.01, * p<0.05			

0.09 point increase in semester GPA. For the other "intervening" variables, it is noteworthy that the GPA gender gap in favor of females is smaller for on-campus students compared to their off-campus peers. Interestingly, notwithstanding the small sizes of the effects, the first-generation disadvantage of lower GPAs is more pronounced for on-campus students relative to their off-campus peers.

Overall, the results from the regression models for cumulative GPA provide strong support for hypothesis **H2**. That is, students that use electronic library licensed content have higher cumulative GPAs. However, the effect of having at least one EZproxy session in an academic term is smaller for cumulative GPA than it is for semester GPA. Model 4 shows that having an EZproxy session in an academic term was correlated with a 0.02 point increase in cumulative GPA. To examine the effect of being on- or off-campus, we ran separate models for on- (model 5) and off-campus (model 6) students, which show differences between the two groups of students—although in ways that are opposite to semester GPA. Having an EZproxy session in an academic term has a larger effect on cumulative GPA for on-campus students

TABLE 4
Panel LME Regressions for Association between Library Usage and Cumulative GPA, FA 2016–WN 2019 (Four or More Semesters)

	(4: All Students)	(5: On-Campus)	(6: Off-Campus)
VARIABLES	**CUM_GPA**	**CUM_GPA**	**CUM_GPA**
EZproxy Session in Term	0.0201***	0.0242***	0.0144***
	(0.000896)	(0.00190)	(0.000871)
On-campus Residence	0.0216***		
	(0.00149)		
High School GPA	0.0222***	0.0364***	0.0141***
	(0.00162)	(0.00313)	(0.00182)

TABLE 4
Panel LME Regressions for Association between Library Usage and Cumulative GPA, FA 2016–WN 2019 (Four or More Semesters)

	(4: All Students)	(5: On-Campus)	(6: Off-Campus)
VARIABLES	**CUM_GPA**	**CUM_GPA**	**CUM_GPA**
GENDER (Reference = Female)			
Male	–0.0735***	–0.0573***	–0.0841***
	(0.00447)	(0.00603)	(0.00528)
RACE (reference = White)			
Asian	0.0655***	0.0654***	0.0559***
	(0.00558)	(0.00763)	(0.00658)
Black	–0.330***	–0.328***	–0.364***
	(0.0108)	(0.0134)	(0.0139)
Hispanic	–0.157***	–0.168***	–0.150***
	(0.00904)	(0.0116)	(0.0112)
Two or More	–0.0769***	–0.0648***	–0.0885***
	(0.0107)	(0.0137)	(0.0129)
Other	–0.197***	–0.159*	–0.192**
	(0.0549)	(0.0717)	(0.0611)
Not Indic	0.0120	0.0296*	–0.00456
	(0.0102)	(0.0145)	(0.0121)
FIRST GENERATION (reference = First Gen)			
Not First Gen	0.105***	0.118***	0.102***
	(0.00721)	(0.00971)	(0.00867)
Don't Know	–0.209***	–0.0901	–0.233***
	(0.0451)	(0.0786)	(0.0503)
FAMILY INCOME (reference = More than $100,000)			
Less than $25,000	–0.113***	–0.101***	–0.126***
	(0.0107)	(0.0147)	(0.0129)
$25,000 – $49,999	–0.0806***	–0.0963***	–0.0837***
	(0.00901)	(0.0120)	(0.0109)
$50,000 – $74,999	–0.0342***	–0.0543***	–0.0364***
	(0.00883)	(0.0122)	(0.0104)
$75,000 – $99,999	–0.0438***	–0.0416***	–0.0460***
	(0.00850)	(0.0118)	(0.00990)
Don't Know	–0.0326*	–0.0317	–0.0454*
	(0.0165)	(0.0217)	(0.0200)
Missing Income Information	–0.00391	–0.00968	–0.00127
	(0.00553)	(0.00753)	(0.00644)
CLASS LEVEL (reference = Freshman)			
Sophomore	–0.00343*	–0.00615**	0.00310
	(0.00150)	(0.00209)	(0.00513)
Junior	–0.00137	–0.0235***	0.0217***
	(0.00187)	(0.00322)	(0.00517)
Senior	0.0241***	–0.0114*	0.0483***
	(0.00209)	(0.00538)	(0.00520)

TABLE 4
Panel LME Regressions for Association between Library Usage and Cumulative GPA, FA 2016–WN 2019 (Four or More Semesters)

	(4: All Students)	(5: On-Campus)	(6: Off-Campus)
VARIABLES	**CUM_GPA**	**CUM_GPA**	**CUM_GPA**
Constant	3.430***	3.376***	3.456***
	(0.0275)	(0.0358)	(0.0304)
Observations	151,049	53,896	97,153
Standard errors in parentheses			
*** p<0.001, ** p<0.01, * p<0.05			

compared to their off-campus peers. However, the magnitude of both effects is very small. Also note that, like semester GPA, the female advantage in cumulative GPA was smaller for on-campus students relative to off-campus students. The first-generation disadvantage with respect to lower cumulative GPAs is more pronounced for on-campus students compared to those that are off-campus.

The study findings suggest that using library resources positively effects academic performance. These effects were larger in magnitude for semester GPA relative to cumulative GPA. For example, regarding semester GPA, first-generation students had a lower GPA (-0.119) than non-first-generation students. Further, males had a lower semester GPA (-0.091) than females. Thus, the impacts of gender and first-generation status on semester GPA were smaller in magnitude than the impact of having at least one EZproxy session during an academic term.

Conclusion

Because library data are often not integrated into other university data, there are major obstacles in demonstrating the richness and complexity of the value of academic library usage for the students who use these resources. We show that merging library usage and student outcome data yields valuable insights on the value of the academic library. Understanding patterns of off-campus use of library resources offers an additional point of insight into potential gaps in use by certain groups of students, such as those living off campus, which may correlate with lower academic success and retention. If students in particular programs tend to live off campus, yet their programs are library-research intensive, what could this mean for those students? For example, 80 percent of undergraduate nursing students live off campus, yet the nursing program integrates the library heavily in its curriculum. We could explore off-campus use by students in this program to potentially identify students at risk of lower academic performance, or to provide indicators to faculty advisors if a student's GPA in research-intensive courses falls below a certain threshold. As additional data from other library services is collected in the future, libraries can develop models to explore other questions around library usage, student success, and curricular integration. Libraries could use the work by the LLAP and allied initiatives to identify opportunities for mitigating educational disparities. Library usage data adds depth of perspective of the student experience, and student engagement broadly, during undergraduate study, and can therefore be a valuable addition to institutions of higher education as they continue to make data-informed

decisions to improve undergraduate education. Further, in the process of doing this work, we have created shareable scripts and tools that could be used to replicate our work in other institutional settings. These and other resources can be downloaded for free from the LLAP project's GitHub site (https://github.com/Learning-Library-Analytics-Project) and website (https://libraryanalytics.org/).

Libraries are often new participants within campus learning analytics efforts. The research described here could lead to new partnerships between libraries and other institutional organizations. Much as traditional academic advisors and partners have great insight into the specific needs and capabilities of their students, so could libraries better tailor their services to those needs. By being better informed about both the kinds of assignments and the needs of the individual students, along with a more granular conceptualization of the technologies they have access to, library staff could be better situated to deliver information services tailored to individual needs. As noted by researcher Megan Oakleaf, designing library services and instruction for the average student harms almost everyone (Oakleaf et al. 2020).[37]

Future work could build on our findings by disentangling the effects of students who are off-campus and not using the VPN (and thus need authentication), versus those who are on-campus but choose to access library licensed content via non-university devices, and hence the library proxy server. Undoubtedly there are economic, technical, and experiential factors contributing to these types of differences in accessing library licensed content. Unfortunately, we were not able to capture them in our study. In addition to multiple socioeconomic factors that could impact student use of library licensed content, there are other factors that could account for these differences, such as the varying nature and demands of curricula across programs and colleges. While there is a healthy demand for library curriculum-integrated instruction (CII) at U-M, programs and instructors may require CII at different times in the progression of a student's academic career. For example, some programs require library CII in first-year experience courses, while other programs may only require CII in the third- or fourth-year. This suggests several lines of future inquiry, such as how course selection affects the need and motivation to use library-licensed resources, or even how the level of study (such as first-year, third-year, and so on) correlates to use of licensed resources and, subsequently, to academic outcomes.

Acknowledgement

The work described in this paper is primarily supported by funding from the Institute of Museum and Library Services (IMLS, LG-96-18-0040-18), and secondarily by the University of Michigan Library.

Appendix

Tables A.1 – A.7 show the percentages of students who had at least one EZproxy session in an academic term by various sociodemographic and academic factors.

TABLE A.1
Percentage of Undergraduate Students Associated with EZproxy Sessions by First-Gen Status, FA16–WN19

Academic Term	First-Gen Status	Enrolled Students	EZproxy Session	% ≥ 1 EZproxy Session
FA 2016	First-Gen	3,520	2,062	59%
	Not First-Gen	24,903	14,372	58%
	Don't Know	259	171	66%
WN 2017	First-Gen	3,364	1,664	49%
	Not First-Gen	23,818	11,631	49%
	Don't Know	226	139	62%
FA 2017	First-Gen	3,753	2,054	55%
	Not First-Gen	25,316	13,928	55%
	Don't Know	92	52	57%
WN 2018	First-Gen	3,605	2,025	56%
	Not First-Gen	24,162	12,788	53%
	Don't Know	85	42	49%
FA 2018	First-Gen	4,091	2,308	56%
	Not First-Gen	25,582	13,855	54%
	Don't Know	53	28	53%
WN 2019	First-Gen	3,890	2,310	59%
	Not First-Gen	24,418	13,957	57%
	Don't Know	47	32	68%

TABLE A.2
Percentage of Undergraduate Students Associated with EZproxy Sessions by On-Campus, FA16–WN19

Academic Term	Residency	Enrolled Students	EZproxy Session	% ≥ 1 EZproxy Session
FA 2016	Off-campus	19,130	11,554	60%
	On-campus	9,552	5,051	53%
WN 2017	Off-campus	17,971	10,353	58%
	On-campus	9,437	3,081	33%
FA 2017	Off-campus	19,993	12,049	60%
	On-campus	9,168	3,985	43%
WN 2018	Off-campus	18,793	11,043	59%
	On-campus	9,059	3,812	42%
FA 2018	Off-campus	20,357	12,014	59%
	On-campus	9,386	4,187	45%
WN 2019	Off-campus	19,110	11,765	62%
	On-campus	9,261	4,540	49%

TABLE A.3
Percentage of Undergraduate Students Associated with EZproxy Sessions by Gender, FA16–WN19

Academic Term	Gender	Enrolled Students	EZproxy Session	% ≥ 1 EZproxy Session
FA 2016	Female	14,296	9,510	67%
	Male	14,386	7,095	49%
WN 2017	Female	13,630	7,817	57%
	Male	13,778	5,617	41%
FA 2017	Female	14,599	9,227	63%
	Male	14,562	6,807	47%
WN 2018	Female	13,910	8,589	62%
	Male	13,942	6,266	45%
FA 2018	Female	14,833	9,304	63%
	Male	14,893	6,887	46%
WN 2019	Female	14,204	9,219	65%
	Male	14,151	7,080	50%

TABLE A.4
Percentage of Undergraduate Students Associated with Ezproxy Sessions by Class Level, FA16–WN19

Academic Term	Class Level	Enrolled Students	EZproxy Session	% ≥ 1 EZproxy Session
FA 2016	Freshman	5,665	2,982	53%
	Sophomore	6,621	3,724	56%
	Junior	7,035	3,979	57%
	Senior	9,361	5,920	63%
WN 2017	Freshman	2,727	874	32%
	Sophomore	6,296	2,383	38%
	Junior	6,489	3,291	51%
	Senior	11,896	6,886	58%
FA 2017	Freshman	5,387	2,391	44%
	Sophomore	7,043	3,704	53%
	Junior	7,084	3,918	55%
	Senior	9,647	6,021	62%
WN 2018	Freshman	2,511	1,088	43%
	Sophomore	6,407	2,911	45%
	Junior	6,949	3,785	54%
	Senior	11,985	7,071	59%
FA 2018	Freshman	5,440	2,477	46%
	Sophomore	6,957	3,601	52%
	Junior	7,666	4,257	56%
	Senior	9,663	5,856	61%
WN 2019	Freshman	2,557	1,300	51%
	Sophomore	6,397	3,373	53%
	Junior	7,132	4,114	58%
	Senior	12,269	7,512	61%

TABLE A.5
Percentage of Undergraduate Students Associated with EZproxy Sessions by Family Income, FA16–WN19

Academic Term	Family Income	Enrolled Students	EZproxy Session	% ≥ 1 EZproxy Session
FA 2016	Less than $25,000	1,470	896	61%
	$25,000 – $49,999	2,073	1,206	58%
	$50,000 – $74,999	2,190	1,294	59%
	$75,000 – $99,999	2,356	1,372	58%
	More than $100,000	14,246	8,256	58%
	Don't Know	935	558	60%
	Missing Income Information	5,412	3,023	56%
WN 2017	Less than $25,000	1,417	724	51%
	$25,000 – $49,999	1,973	953	48%
	$50,000 – $74,999	2,114	1,069	51%
	$75,000 – $99,999	2,249	1,145	51%
	More than $100,000	13,636	6,683	49%
	Don't Know	851	435	51%
	Missing Income Information	5,168	2,425	47%
FA 2017	Less than $25,000	1,486	855	58%
	$25,000 – $49,999	2,091	1,139	54%
	$50,000 – $74,999	2,090	1,134	54%
	$75,000 – $99,999	2,210	1,263	57%
	More than $100,000	14,336	7,749	54%
	Don't Know	476	263	55%
	Missing Income Information	6,472	3,631	56%
WN 2018	Less than $25,000	1,441	795	55%
	$25,000 – $49,999	2,026	1,119	55%
	$50,000 – $74,999	2,024	1,123	55%
	$75,000 – $99,999	2,080	1,167	56%
	More than $100,000	13,689	7,095	52%
	Don't Know	430	222	52%
	Missing Income Information	6,162	3,334	54%
FA 2018	Less than $25,000	1,586	911	57%
	$25,000 – $49,999	2,307	1,299	56%
	$50,000 – $74,999	2,066	1,150	56%
	$75,000 – $99,999	2,161	1,204	56%
	More than $100,000	14,632	7,760	53%
	Don't Know	540	285	53%
	Missing Income Information	6,434	3,582	56%
WN 2019	Less than $25,000	1,507	923	61%
	$25,000 – $49,999	2,212	1,269	57%
	$50,000 – $74,999	2,009	1,217	61%
	$75,000 – $99,999	2,074	1,213	58%
	More than $100,000	13,951	7,892	57%
	Don't Know	515	278	54%
	Missing Income Information	6,087	3,507	58%

TABLE A.6
Percentage of Undergraduate Students Associated with EZproxy Sessions by Race, FA16 – WN19

Academic Term	Race	Enrolled Students	EZproxy Session	% ≥ 1 EZproxy Session
FA 2016	Asian	5,460	3,019	55%
	Black	1,268	730	58%
	Hispanic	1,564	916	59%
	White	17,743	10,439	59%
	2 or More	1,111	642	58%
	Other	53	30	57%
	Not Indic	1,483	829	56%
WN 2017	Asian	5,282	2,425	46%
	Black	1,213	574	47%
	Hispanic	1,500	747	50%
	White	16,876	8,438	50%
	2 or More	1,084	515	48%
	Other	53	23	43%
	Not Indic	1,400	712	51%
FA 2017	Asian	5,685	2,941	52%
	Black	1,291	698	54%
	Hispanic	1,762	955	54%
	White	17,803	10,053	56%
	2 or More	1,206	631	52%
	Other	56	29	52%
	Not Indic	1,358	727	54%
WN 2018	Asian	5,501	2,746	50%
	Black	1,252	683	55%
	Hispanic	1,698	908	53%
	White	16,924	9,220	54%
	2 or More	1,155	599	52%
	Other	54	26	48%
	Not Indic	1,268	673	53%
FA 2018	Asian	6,047	3,063	51%
	Black	1,315	748	57%
	Hispanic	1,972	1,051	53%
	White	17,525	9,794	56%
	2 or More	1,346	702	52%
	Other	49	23	47%
	Not Indic	1,472	810	55%
WN 2019	Asian	5,829	3,137	54%
	Black	1,268	766	60%
	Hispanic	1,899	1,099	58%
	White	16,604	9,738	59%
	2 or More	1,302	745	57%
	Other	46	22	48%
	Not Indic	1,407	792	56%

TABLE A.7
Percentage of Undergraduate Students Associated with EZproxy Sessions by School, FA16–WN19

Academic Term	School	Enrolled Students	EZproxy Session	% ≥ 1 EZproxy Session
FA 2016	Architecture	145	65	45%
	Art and Design	495	356	72%
	Business Administration	1,673	890	53%
	Dental Hygiene	111	77	69%
	Education	112	66	59%
	Engineering	6,078	2,736	45%
	Information	208	123	59%
	Joined Degree Program	10	7	70%
	Kinesiology	946	698	74%
	Literature, Science & the Arts	17,306	10,395	60%
	Music, Theater & Dance	732	447	61%
	Nursing	705	626	89%
	Pharmacy	14	11	79%
	Public Policy	147	108	73%
WN 2017	Architecture	140	71	51%
	Art and Design	462	249	54%
	Business Administration	1,639	746	46%
	Dental Hygiene	107	63	59%
	Education	112	53	47%
	Engineering	5,909	1,958	33%
	Information	186	89	48%
	Joined Degree Program	8	5	63%
	Kinesiology	918	576	63%
	Literature, Science & the Arts	16,400	8,614	53%
	Music, Theater & Dance	700	402	57%
	Nursing	685	512	75%
	Pharmacy	14	7	50%
	Public Policy	128	89	70%
FA 2017	Architecture	155	82	53%
	Art and Design	497	363	73%
	Business Administration	1,773	869	49%
	Dental Hygiene	112	79	71%
	Education	120	46	38%
	Engineering	6,409	2,666	42%
	Information	253	147	58%
	Joined Degree Program	12	8	67%
	Kinesiology	976	627	64%

TABLE A.7
Percentage of Undergraduate Students Associated with EZproxy Sessions by School, FA16–WN19

Academic Term	School	Enrolled Students	EZproxy Session	% ≥ 1 EZproxy Session
	Literature, Science & the Arts	17,160	9,942	58%
	Music, Theater & Dance	747	495	66%
	Nursing	667	516	77%
	Pharmacy	42	19	45%
	Public Health	85	72	85%
	Public Policy	153	103	67%
WN 2018	Architecture	153	107	70%
	Art and Design	481	356	74%
	Business Administration	1,757	760	43%
	Dental Hygiene	109	82	75%
	Education	118	40	34%
	Engineering	6,150	2,571	42%
	Information	214	122	57%
	Joined Degree Program	12	7	58%
	Kinesiology	951	594	62%
	Literature, Science & the Arts	16,294	9,034	55%
	Music, Theater & Dance	715	509	71%
	Nursing	636	487	77%
	Pharmacy	42	25	60%
	Public Health	84	72	86%
	Public Policy	136	89	65%
FA 2018	Architecture	181	119	66%
	Art and Design	556	396	71%
	Business Administration	1,826	753	41%
	Dental Hygiene	103	71	69%
	Education	131	60	46%
	Engineering	6,649	2,755	41%
	Information	302	135	45%
	Joined Degree Program	11	9	82%
	Kinesiology	962	617	64%
	Literature, Science & the Arts	17,262	9,918	57%
	Music, Theater & Dance	743	524	71%
	Nursing	632	543	86%
	Pharmacy	56	33	59%
	Public Health	158	130	82%
	Public Policy	154	128	83%
WN 2019	Architecture	181	124	69%

TABLE A.7
Percentage of Undergraduate Students Associated with EZproxy Sessions by School, FA16–WN19

Academic Term	School	Enrolled Students	EZproxy Session	% ≥ 1 EZproxy Session
	Art and Design	524	381	73%
	Business Administration	1,799	740	41%
	Dental Hygiene	101	70	69%
	Education	126	54	43%
	Engineering	6,313	2,847	45%
	Information	260	122	47%
	Joined Degree Program	10	7	70%
	Kinesiology	954	678	71%
	Literature, Science & the Arts	16,409	10,030	61%
	Music, Theater & Dance	717	515	72%
	Nursing	607	475	78%
	Pharmacy	55	36	65%
	Public Health	157	116	74%
	Public Policy	142	104	73%

Notes

1. Krista M. Soria, Jan Fransen, and Shane Nackerud, "Library Use and Undergraduate Student Outcomes: New Evidence for Students' Retention and Academic Success," *portal: Libraries and the Academy* 13, no. 2 (2013): 147–64, https://doi.org/10.1353/pla.2013.0010.

2. Laurie Alexander, Doreen R. Bradley, and Kenneth J. Varnum, "On the Road to Learning Analytics: The University of Michigan Library's Experience with Privacy and Library Data," in *Using Digital Analytics for Smart Assessment*, ed. Tabatha Farney (Chicago: American Library Association, 2018), 83–93.

3. T. D. Wilson, "Models in Information Behaviour Research," *Journal of Documentation* 55, no. 3 (1999): 249–70, https://doi.org/10.1108/EUM0000000007145; J. David Johnson, "Cancer-related information seeking." *Health Communication* (Cresskill, N.J.: Hampton Press, 1997); Thomas D. Wilson, "Information Behavior Models," in *Encyclopedia of Library and Information Science, 4th Edition* (CRC Press, 2017), 2086-2093.

4. Marcia J. Bates, "Information Behavior," in *Encyclopedia of Library and Information Science, 4th Edition* (CRC Press, 2017), 2074-2085.

5. Bates, "Information Behavior,"; Lotta Haglund and Per Olsson, "The Impact on University Libraries of Changes in Information Behavior among Academic Researchers: A Multiple Case Study," *The Journal of Academic Librarianship* 34, no. 1 (2008): 52–9, https://doi.org/10.1016/j.acalib.2007.11.010; Xi Niu and Bradley M. Hemminger. "A Study of Factors that Affect the Information-seeking Behavior of Academic Scientists," *Journal of the American Society for Information Science and Technology* 63, no. 2 (2012): 336–53, https://doi.org/10.1002/asi.21669.

6. Polly A. Graham, Sarah Socorro Hurtado, and Robert M. Gonyea, "The Benefits of Living on Campus: Do Residence Halls Provide Distinctive Environments of Engagement?" *Journal of Student Affairs Research and Practice* 55, no. 3 (2018): 255–69, https://doi.org//10.1080/19496591.2018.1474752.

7. Ruth N. López Turley and Geoffrey Wodtke. "College Residence and Academic Performance: Who Benefits from Living on Campus?" *Urban Education* 45, no. 4 (2010): 506–32, https://doi.org/10.1177/0042085910372351.

8. Lauren T. Schudde, "The Causal Effect of Campus Residency on College Student Retention," *Review of Higher Education* 34, no. 4 (2011): 581–610, https://doi.org/10.1353/rhe.2011.0023.

9. Clare Huhn, "The "Housing Effect" on First-Year Outcomes," (Madison, WI: Academic Planning and Analysis, Office of the Provost, University of Wisconsin-Madison, 2006).

10. Ibid.

11. Jonathan M. Turk and Manuel S. González Canché, "On-Campus Housing's Impact on Degree Comple-

tion and Upward Transfer in the Community College Sector: A Comprehensive Quasi-Experimental Analysis," *The Journal of Higher Education* 90, no. 2 (2019): 244–71, https://doi.org/10.1080/00221546.2018.1487755.

12. Denise Balfour Simpson and Dana Burnett, "Commuters Versus Residents: The Effects of Living Arrangement and Student Engagement on Academic Performance," *Journal of College Student Retention: Research, Theory & Practice* 21, no. 3 (2017): 286–304, https://doi.org/10.1177/1521025117707516.

13. Laura Robinson et al., "Digital Inequalities and Why They Matter." *Information, Communication & Society* 18, no. 5 (2015): 569–82, https://doi.org/10.1080/1369118X.2015.1012532; Xinzhi Zhang et al. "Big Data Science: Opportunities and Challenges to Address Minority Health and Health Disparities in the 21st Century," *Ethnicity & disease* 27, no. 2 (2017): 95–106, https://doi.org/10.18865/ed.27.2.95.

14. Paul F. Cleary, Glenn Pierce, and Eileen M. Trauth, "Closing the Digital Divide: Understanding Racial, Ethnic, Social Class, Gender and Geographic Disparities in Internet Use among School Age Children in the United States," *Universal Access in the Information Society* 4, no. 4 (2006): 354–73, https://doi.org/10.1007/s10209-005-0001-0; Linda A. Jackson et al., "Does Home Internet Use Influence the Academic Performance of Low-income Children?" *Developmental Psychology* 42, no. 3 (2006): 429–35, https://doi.org/10.1037/0012-1649.42.3.429.

15. Elizabeth F. Farrell, "Among Freshmen, a Growing Digital Divide," *The Chronicle of Higher Education* February 2, 2005, A32-; Steve Jones, Camille Johnson-Yale, Sarah Millermaier, and Francisco Seoane Pérez, "U.S. College Students' Internet Use: Race, Gender and Digital Divides," *Journal of Computer-Mediated Communication* 14, no. 2 (2009): 244–64, https://doi.org/10.1111/j.1083-6101.2009.01439.x.

16. Angelina KewalRamani et al., *Student Access to Digital Learning Resources outside of the Classroom. NCES 2017-098*. Edited by American Institutes for Research. *National Center for Education Statistics*. Washington, DC: U.S. Department of Education, 2018.

17. Lauren Musu, "The Digital Divide: Differences in Home Internet Access," in *NCES Blog*, edited by NCES Blog Editor. Washington, DC: National Center for Education Statistics, October 18, 2018, https://nces.ed.gov/blogs/nces/post/the-digital-divide-differences-in-home-internet-access.

18. Ana-Paula Correia, "Healing the Digital Divide During the COVID-19 Pandemic," *Quarterly Review of Distance Education* 21, no. 1 (2020): 13–21.

19. John Lai and Nicole O. Widmar, "Revisiting the Digital Divide in the COVID-19 Era," *Applied Economic Perspectives and Policy* 43, no. 1 (2021): 458–64, https://doi.org/10.1002/aepp.13104.

20. Fawzia Reza, "COVID-19 and Disparities in Education: Collective Responsibility Can Address Inequities," *Knowledge Cultures* 8, no. 3 (2020): 68–75, https://doi.org/10.22381/KC83202010; Megan Kuhfeld et al., "Projecting the Potential Impact of COVID-19 School Closures on Academic Achievement," *Educational Researcher* 49, no. 8 (2020): 549–65, https://doi.org/10.3102/0013189X20965918.

21. Terence Day et al. "The Immediate Impact of COVID-19 on Postsecondary Teaching and Learning," *The Professional Geographer* 73, no. 1 (2021): 1–13, https://doi.org/10.1080/00330124.2020.1823864.

22. Ibid.

23. Ibid.; Dipti Mehta and Xiaocan Wang, "COVID-19 and Digital Library Services – A Case Study of a University Library," *Digital Library Perspectives* 36, no. 4 (2020): 351–63, https://doi.org/10.1108/DLP-05-2020-0030.

24. Ana-Paula Correia, "Healing the Digital Divide During the COVID-19 Pandemic," *Quarterly Review of Distance Education* 21, no. 1 (2020): 13–21.

25. Kyong Hee Chee, Nathan W. Pino, and William L. Smith, "Gender Differences in the Academic Ethic and Academic Achievement *," *College Student Journal* 39, no. 3 (September 2005): 604+; Michael Sheard, "Hardiness Commitment, Gender, and Age Differentiate University Academic Performance," *British Journal of Educational Psychology* 79, no. 1 (2009): 189–204. https://doi.org/10.1348/000709908X304406; Arna Kristín Harðardóttir, Sigurður Guðjónsson, Inga Minelgaite, and Kári Kristinsson, "Ethics as Usual? Gender Differences in Work Ethic and Grades," *Management: Journal of Contemporary Management Issues* 24, no. 2 (2019): 11–21, https://doi.org/10.30924/mjcmi.24.2.2.

26. Michael J. Stebleton and Krista M. Soria, "Breaking Down Barriers: Academic Obstacles of First Generation Students at Research Universities," *The Learning Assistance Review* 17, no. 2 (2012): 7+.

27. Xi Wang, Minhao Dai, and Robin Mathis, "The Influences of Student- and School-level Factors on Engineering Undergraduate Student Success Outcomes: A Multi-level Multi-school Study," *International Journal of STEM Education* 9, no. 1 (2022): 23, https://doi.org/10.1186/s40594-022-00338-y.

28. John M. Trussel and Lisa Burke-Smalley, "Demography and Student Success: Early Warning Tools to Drive Intervention," *Journal of Education for Business* 93, no. 8 (2018): 363–72, https://doi.org/10.1080/08832323.2018.1496893; Julian R. Betts and Darlene Morell, "The Determinants of Undergraduate Grade Point Average: The Relative Importance of Family Background, High School Resources, and Peer Group Effects," *The Journal of Human Resources* 34, no. 2 (1999): 268–93, https://doi.org/10.2307/146346.

29. Siu-Man Raymond Ting and Tracy L. Robinson, "First-year Academic Success: A Prediction Combining

Cognitive and Psychosocial Variables for Caucasian and African American Students," *Journal of College Student Development* 39, no. 6 (1998): 599–610; Betts and Morell, "The Determiniants of Undergraduate Grade Point Average"; Andrew P. Barkley and Jerry J. Forst, "The Determinants of First-Year Academic Performance in the College of Agriculture at Kansas State University, 1990–1999," *Journal of Agricultural and Applied Economics* 36, no. 2 (2004): 437–48, https://doi.org/10.1017/S1074070800026729.

30. Forrest E. Huffman, "Student Performance in an Undergraduate Advanced Real Estate Course: Real Estate Majors vs. Finance Majors," *Journal of Real Estate Practice and Education* 14, no. 2 (2011): 111–23, https://doi.org/10.1080/10835547.2011.12091693.

31. Wilson, "Models in Information Behaviour"; Johnson, *Cancer-related information seeking.*

32. Ibid.; Johnson, J. David. *Cancer-related information seeking*. Cresskill, N.J.: *Health Communication*. Cresskill, N.J.: Hampton Press, 1997.

33. Jerry Chih-Yuan Sun and Susan E. Metros, "The Digital Divide and Its Impact on Academic Performance," *US-China Education Review A* 1, no. 2 (2011): 153–61; Joanna Goode, "The Digital Identity Divide: How Technology Knowledge Impacts College Students," *New Media & Society* 12, no. 3 (2010): 497–513, https://doi.org/10.1177/1461444809343560; Laura Robinson, Øyvind Wiborg, and Jeremy Schulz, "Interlocking Inequalities: Digital Stratification Meets Academic Stratification," *American Behavioral Scientist* 62, no. 9 (2018): 1251–72, https://doi.org/10.1177/0002764218773826; KewalRamani et al., *Student Access to Digital Learning Resources*.

34. Penny Beile, Kanak Choudhury, Rachel Mulvihill, and Morgan Wang, "Aligning Library Assessment with Institutional Priorities: A Study of Student Academic Performance and Use of Five Library Services," *College & Research Libraries* 81, no. 3 (2020): 435–58, https://doi.org/10.5860/crl.81.3.435; Soria, Fransen, and Nackerud, "Library Use and Undergraduate Student Outcomes."

35. (Soria, Fransen, and Nackerud 2014, 2016; Wong and Webb 2011); Krista M. Soria, Jan Fransen, and Shane Nackerud, "Stacks, Serials, Search Engines, and Students' Success: First-Year Undergraduate Students' Library Use, Academic Achievement, and Retention," *The Journal of Academic Librarianship* 40, no. 1 (2014): 84–91, https://doi.org/10.1016/j.acalib.2013.12.002; Krista M. Soria, Jan Fransen, and Shane Nackerud, "Beyond Books: The Extended Academic Benefits of Library Use for First-Year College Students," *College & Research Libraries* 78, no. 1 (2016): 8–22, https://doi.org/10.5860/crl.78.1.8; Shun Han Rebekah Wong and T. D. Webb, "Uncovering Meaningful Correlation between Student Academic Performance and Library Material Usage," *College & Research Libraries* 72, no. 4 (2011): 361–70, https://doi.org/10.5860/crl-129.

36. Stata Statistical Software: Release 16. StataCorp LLC, College Station, TX.

37. Megan Oakleaf et al., *Connecting Libraries and Learning Analytics for Student Success*, (Syracuse, NY: Syracuse University, 2020), https://library.educause.edu/-/media/files/library/2020/12/cllassfinalwhitepaper.pdf.

Leaning Into the Future, Together: Applying Business Process Management to Increase Efficiency and Manage Change in Archives and Special Collections

Jodi Allison-Bunnell, Anne Jenner, and Emily Dominick

The time and resources required to prepare archival collections for use by researchers is a source of constant frustration in archives and libraries. Almost always, aspirations and collections exceed limited resources. The last fifteen to twenty years have seen archivists and librarians putting great effort into increasing standardization and efficiency. However, there are few examples of applying techniques from other fields that are proven to increase productivity. This dual case study shows that applying Lean techniques, which were originally developed for automobile manufacturing, yields significant results: measurable reductions in processing time and resource use; increased adherence to standards; increased engagement in and willingness to change by staff; effective coordination across departments; and increased ability to meet the needs of stakeholders.

Introduction

The time and resources required to prepare archival collections for use by researchers, usually referred to as "processing,"[1] is a source of constant frustration in archives and libraries. Nearly every repository contends with unprocessed backlogs and struggles to meet administrator and donor expectations. All have many aspirations; All have limited resources. Compounding these challenges, they may struggle with staff who resist changes to processes, standards, technologies, and the workplace. Over the last fifteen to twenty years, archives have put great effort into reconsidering processing and making it more efficient. However, the profession has few examples of applying techniques from other fields like business, manufacturing, and engineering that are proven to increase productivity and better match aspirations to resources.

At the University of Washington's Special Collections, starting in 2014, and Montana State University Library's Archives and Special Collections, starting in 2021, they applied a suite of

* *Jodi Allison-Bunnell is Head of Archives and Special Collections and Assistant Professor at Montana State University Library, email: jodi.allisonbunnell@montana.edu; Anne Jenner is Curator of the Pacific Northwest Collection at University of Washington Libraries, email: ajenner@washington.edu; Emily Dominick is Manager, Puget Sound Branch, Washington State Archives, email: emily.dominick@sos.wa.gov.*

techniques from manufacturing called Lean to revise their approaches to processing. They found that this technique applied to their organizations yielded significant results: measurable reductions in processing time and resource use; increased adherence to standards; increased engagement in and willingness to change by staff; effective coordination across departments; and increased ability to meet the needs of their stakeholders. As there are no other published examples of applying Lean in archives, and few on applying related techniques, the authors aim to share the results and to suggest that these techniques can yield similar results for other institutions.

Literature Review and Background

Because Lean is not well known in archives or libraries, a few key concepts and terms will be defined. Lean is a sub-discipline of Business Process Management (BPM),[2] which is "the art and science of overseeing how work is performed in an organization to ensure consistent outcomes and to take advantage of improvement opportunities." BPM focuses on managing "entire chains of events, activities and decisions that ultimately add value to the organization and its customers"; those "chains" are collectively known as "processes."[3] Nearly everything an organization does is a process that enables it to provide products or services for customers or clients. The way that processes are designed and performed in an organization affects both the quality of products or services and the speed at which they are delivered. Lean itself emerged from the 1930 "Toyota Way," which has two pillars for organizational excellence: 1. continuous improvement and 2. respect for people.[4] From the "Toyota Way," James Womack and Daniel Jones defined "Lean" in 1996 as five principles that focus on specifying value for each project, identify how value is created, avoid any interruptions in creating value, let customers pull (identify) value from the producer, and pursue perfection.[5] Lean focuses on the elimination of anything that does not add value to the customer, which is termed "waste."

Lean in Academic Libraries

Even though other methodologies for evaluating workflows from a user-centered perspective are quite common, there are few articles describing applications of BPM in libraries. In their article on implementing Six Sigma (a specific sub-discipline of BPM closely related to Lean) at Sungkyunkwan University Library in South Korea, Dong-Suk Kim observed that at that time (2010), not many academic libraries had applied Six Sigma or similar frameworks to improving their processes. They reported that the process was successful and well suited to the work of academic libraries.[6] Around the same time, Sarah Anne Murphy concluded that libraries benefit significantly from the structures of BPM:

> Libraries can customize and borrow a number of quality management systems and tools from the business community to both assess their service process and continuously improve their operations. By adopting an approach like Lean Six Sigma, a library can respond better to changing customer needs and desires by creating an infrastructure that supports, nurtures, and sustains a culture of assessment and change.[7]

In her 2015 article, Elizabeth Nelson articulates how Lean Six Sigma can be applied in academic libraries and suggests that the most strategic use may be in reducing errors in service and in

increasing service satisfaction.[8] She also observes that these tools have been used to improve technical services workflows, including purchasing and processing books and reducing time needed to re-shelve materials.

Some examples focus on transformative outcomes in customer-facing operations. A case study from the Columbus Metropolitan Library in Ohio has compelling examples of reducing wait times for telephone reference (detailed inquiries of process improvement revealed that using a single button on the phone, not increasing staff, vastly reduced caller wait times) and increasing on-time delivery for internal duplication orders to 100%.[9] The Columbus study, together with another case study from the University of Arizona's interlibrary loan service, provides examples of the ways in which Lean aids in identifying and addressing the root cause(s) of quality deficits and delays.[10] Other examples include using Lean to improve email reference and for general process improvement for a large library system facing dwindling resources.[11]

A book-length treatment of a case study from the University of Maryland's University College examines in detail the application of Lean to managing electronic resources.[12] In that study, Nelson observes that rather than adopting an all-or-nothing approach across the library, these techniques and their many tools and approaches work well when applied to specific processes and scenarios. In most cases, library staff make significant discoveries about the actual origins of waste; challenging assumptions proves to be very powerful. However, Nelson also notes that in some cases, libraries have referred to "process improvement" rather than naming specific methodologies because library employees are naturally suspicious of "managerial names" of techniques from the business world.[13]

The archives literature has just two articles related to BPM, both case studies from Brigham Young University: Gordon Daines' 2014 article (and a closely related 2009 article) on applying BPM to processing workflows. Daines describes applying process modeling to help department staff adopt new practices and workflows around preparing collections. The group had used a project management model but moved to process analysis and revision based on a fundamentally important insight: processing archival collections is not a one-off, unique activity that must be defined anew for each project. Instead, there are so many strong similarities across projects that many aspects of the work should be standardized and done the same way for every collection.[14] Process mapping helped the department visualize and understand how these processes needed to work. It helped the department implement systems well and has driven continued change and adaptation in the organization in the years since.[15]

Why Lean for Archives?

Despite the dearth of specific applications in archives, Lean is closely related to changes in processing workflows and a systematic re-thinking of archival description over the last fifteen to twenty years. That re-thinking, in turn, has two threads: an increasing emphasis on user needs and the reality of scarce resources for processing.

The call to focus on user needs arose in the 1980s with Mary Jo Pugh and Elsie Freeman calling for a reorientation of description toward users.[16] Paul Conway amplified those ideas, suggesting that archives should and could seek out information on user needs.[17] In subsequent decades, the emergence of the term "hidden collections" focused on the most basic user need: easily discovering where collections were held. In discussions between

1998 and 2008, stakeholders asserted that outdated practices for cataloging and processing collections for use were a major factor in creating unacceptable backlogs and lack of access to collections.[18] Following on these findings, the Council on Library and Information Resources (CLIR) launched the Cataloging Hidden Collections regrant program in 2008. That program, which continued until 2014, focused on developing and implementing efficient practices—including collection-level description, re-use of finding aid data, using slightly augmented accession records for public access—to challenge traditional notions of processing.[19] It formed a strong underpinning for the CLIR regrant program that continues today: Digitizing Hidden Collections.[20]

Closely related, and driving the promulgation of revised practices, was Greene and Meissner's seminal 2005 article "More Product, Less Process."[21] Their work was transformative because it redefined processes based on documented needs of end users, eliminating work that was tangential to the needs of those users. Their work is thus consonant with the Lean concept of letting the customer pull value from the provider.

Other essential works and standards are consistent with this strategic focus on end users. In its introductory principles, Describing Archives: A Content Standard (DACS) version 2019.0.3.1 states:

> Because it facilitates use, archival description is a user-centered product and process.… It is imperative that repositories identify, engage, and seek to understand the motivations and needs of their users, which may include but are not limited to scholarly production, collection care and control, institutional knowledge, connection to family ties, artistic endeavors, government accountability, justice-seeking endeavors, and symbolic purposes of holding records.[22]

DACS also states that description beyond the minimum should at all times be user-driven. At this point, the profession can leave behind any idea of slavishly following old procedures and the notion of "the right way," without giving thought to the functional and real requirements of users.

More recently, OCLC Research's Total Cost of Stewardship report (TCS) moves beyond the focus on reducing backlogs that is part of the "hidden collections" concept to provide means to potentially prevent the accumulation of backlogs in the first place. By addressing the organizational gap that often exists between collection development and collection stewardship, it is part of an overall trend to strategically re-focus the work of archivists, librarians, and other cultural heritage professionals.[23] It provides a toolkit for estimating and articulating the total cost of stewarding collections throughout their lifecycle so that repositories can better match ambitions with available resources.

Even with all of these advances, institutions and their staff continue to struggle with processes and tools that may not be producing the results they and their users need. It can be difficult to help individuals and teams understand where established practices are or are not beneficial. As a grassroots process that fundamentally respects the individuals that do the work, Lean and other structures like it can help individuals (and thus organizations) to see past personalities, territories, and incidents or conditions long past to develop new and more effective approaches to collection preparation.

Case Studies

University of Washington

Background

The University of Washington is the flagship institution of six public universities in Washington state and includes one of the largest library systems in the world. Special Collections is situated on the Seattle campus of the University in the Suzzallo & Allen Library. Holdings in Special Collections include over 70,000 cubic feet of archival collections and nearly 200,000 non-circulating titles.[24]

The division has seen a steady reduction in the number of staff over the last two decades, but the volume of acquisitions has increased, creating a notable imbalance of acquisitions and resources. Their current staff structure is the result of a merger of Manuscripts & University Archives with Special Collections in 2000, forming a single division named Special Collections. Merging the two service desks into one resulted in reduced staffing. The 2008 recession further reduced staff numbers, including three key technical services positions, the Head of Special Collections Technical Services, the Processing Specialist, and eventually an Accessioning Specialist.[25] The backlog of un-accessioned and unprocessed collections grew. When the economy recovered, Special Collections hired more curatorial staff positions but did not fill vacant Technical Services positions.[26] Seven curators acquired collections without shared collection development goals or awareness of Technical Services capacity to handle the quantity of acquisitions. The backlog continued to grow and conversations regularly ended in frustration because of ineffective accessioning and processing systems. Technical Services staffing declined through attrition and by 2012 it consisted of two part-time accessioning staff members with on-the-job training in collections management and one full-time computer/database support staff member. In the absence of a Head of Technical Services, the University Archivist supervised this three member team. Collections work—accessioning, processing, collection records maintenance, and finding aid[27] changes—was activated and tracked with paper forms, but each curator used the forms differently. The lack of standard practices was confusing and caused rifts among staff and between departments. While they agreed on the key steps of the accessioning process (acquisition, logging collections information, preparing materials for the storage and access, and creating and uploading the Encoded Archival Description [*EAD*] finding aid), they could not agree on who was responsible for each step. Curators were aware that they had developed distinct accessioning workflows and standards but were unsure of a way to resolve the differences.

Special Collections learned that the University Records Management Services department went through a Lean process improvement exercise and successfully developed more efficient work processes. Hoping that they could have similar success, the curators and Special Collections leadership turned to the University of Washington's Lean Process, a program managed loosely through the university's Finance and Facilities Department. The university supported the Lean Process by providing a gathering space for Lean launches, a Lean facilitator, and organizational assistance. They were assigned a facilitator from a corps of University of Washington volunteers who guide other departments through their Lean redesigns. Their facilitator was a staff member from the University Records Center who held a library degree and understood some aspects of their work. Her primary role was to give them direction, keep them focused on particular tasks, lead them to key milestones, and then step back so they could do the work.

The Lean Redesign

PROJECT SCOPE

Once they decided to take the Lean leap, key team members met with their facilitator to begin working on a scope statement. That statement stated the problem they wanted to solve:

> There is a large and steadily growing backlog of un-accessioned acquisitions, and the current accessioning process is cumbersome and prone to stalling. Staff have differing understanding and expectations of what work is to be done during accessioning, by whom, and when.

The Statement connected the Lean effort to the unit's organizational goals: to make materials discoverable by and accessible to the researchers, and to be accountable to its donors.

The Statement focused on three areas of accessioning:

1. Identifying necessary, unnecessary, and desirable elements in the accessioning process;
2. Eliminating stall points to make time-to-completion predictable;
3. Reducing the accessioning backlog month to month.

Success was defined as:

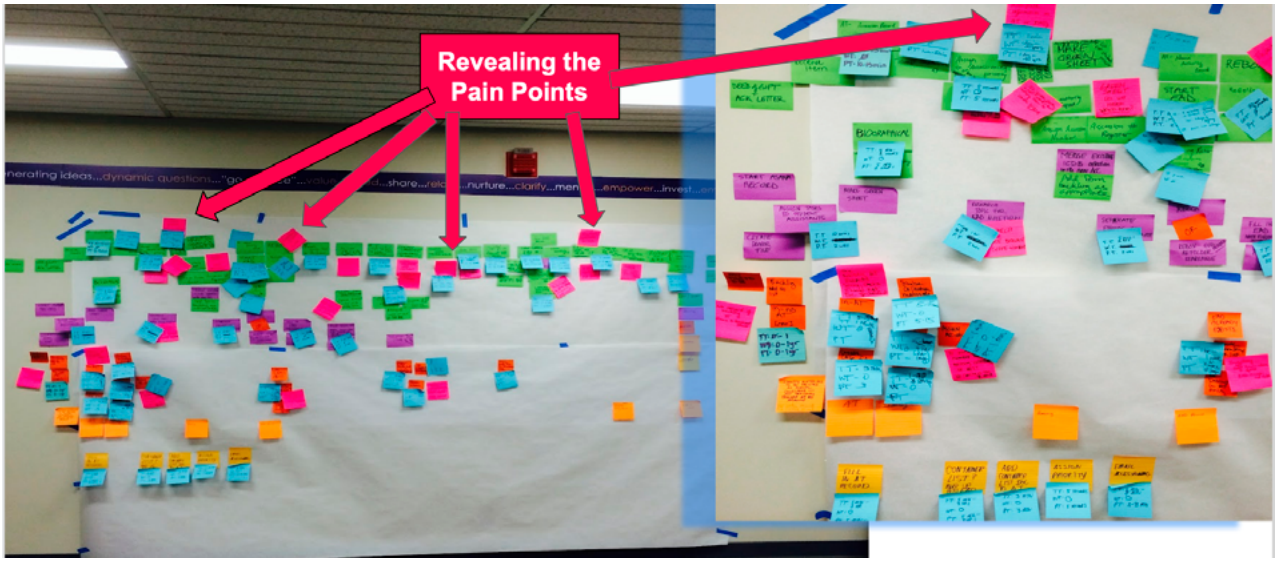
1. Producing a revised accessioning manual;
2. Reducing the time needed to move materials from intake to completion;
3. Regular monitoring of backlog numbers shows a reduction in size.

The scope statement also listed the team members who would gather for a collaborative three-day Lean launch. The ten-member Lean team consisted of curators (archivists and librarians), Technical Services staff, and the director of Special Collections. Aside from the director, every person on the team was directly involved in some way with the accessioning process.

CURRENT STATE

The University of Washington provided an ideal space for their Lean launch: It was across campus and away from the library; had rolling office furniture to quickly gather, disperse,

FIGURE 1
Documenting the Current State



and re-convene; and had blank walls to cover with butcher paper and sticky notes of all sizes and colors. The team began documenting the current state with sticky notes on the wall, mapping the accessioning process from the beginning (when the new acquisition is delivered to the workroom) to the end, (when the collection is staged for shelving). The intervening steps included logging preliminary information about the new accession into the collection management database, rehousing the collection if needed, creating or updating an EAD finding aid, creating or updating catalog records and authority records, and locating an appropriate storage location.

FINDING PAIN POINTS

By mapping their processes in detail, they revealed a workflow with 86 steps, 26 stall points, process times ranging from 77 to 466 hours of touch time; 88 to 244,131 hours of process time; and 33 to 130,482 hours of wait time. They identified and addressed 31 ideas for improvement as they tweaked the process. They also found four shortcuts or "ghost processes" that staff developed to overcome accessioning barriers. One example: incoming visual materials collections were temporarily staged on shelves apart from other materials awaiting accessioning. These collections were recorded and tracked in a spreadsheet accessible only to the visual materials curator and kept in that state for an undetermined period of time because the accessioning staff were unable to allocate sufficient time to address the materials.

FUTURE STATE AND INITIAL REDESIGN

The team's next focus was envisioning a future state. Here, they faced a challenge: focusing on the "happy path" of finding a workflow for the 80% (the *most common* situations) while accepting that *less common* situations will make up the remaining 20%. Their second challenge was to set a goal for improvement. The Lean facilitator urged them to make an audacious goal: a 50% improvement. To fulfill this goal, they would need to dismantle current processes and build a common workflow that worked for every member of the team. The process stopped from time to time for the team to step back and resolve disputed perspectives or settle on a shared definition of terms, like, "what is the definition of accessioning or hand-off?" Unresolvable issues were added to a list of projects to address back at the shop.

They also developed a larger goal: to continue to pursue the ideal future state. They understood that the team would remain empowered to implement improvements and make their work lives better and their communication smoother, while also receiving support from leadership for their efforts. The Lean process is more than a set of tools and techniques. It is meant to have lasting organizational impact by building a culture where staff identify and fix problems collectively, work with a sense of urgency, purpose and teamwork, think creatively, learn, grow, and share lessons learned with others.

NEW PROCESSES THROUGH KAIZENS

The Lean launch concluded with the framework of a single common accessioning workflow and four kaizens. Kaizens are simply short-term projects undertaken by a subset of the team with the aim of improving one element or aspect of the overall process. They are designed with an expectation of group-wide report-outs at 30, 60, and 90 days. The team agreed to complete four kaizens:

1. Flesh out Future Process: The entire team would work together toward a goal to resolve 26 pain points and test the new common workflow.

2. Create Queue Management: Three team members would test electronic tools to replace paper forms and make a recommendation to the team.
3. Address Backlog: Three team members researched and documented backlogs and established new norms for recording incoming acquisitions to facilitate discoverability and prevent the backlog from increasing.
4. Space Management: Three team members examined storage capacity to prepare for space challenges.

After committing to a new common workflow, the team overhauled their pre-Lean accessioning manual. The new manual was designed as a living document to be continually updated with efficiencies as they were developed and adopted. Eighty-six steps in the workflow were reduced as practices were consolidated using shared project management software, while paper accessioning forms and other redundancies were eliminated. Upon eliminating paper forms, they turned to managing the accessioning queue with Asana project management software. As a web-based system accessible to every member of the staff, Asana allowed them to track each collection through the workflow with transparency and improved command over the workflow. Like many project management tools, Asana's templates and task lists can include mandatory and optional steps as the workflow requires. The templates ensured consistent processes for all curators while allowing each person to assign tasks to colleagues or to student assistants on their teams. Using Asana created consistency but also still allowed for flexibility.

To address the backlog and prevent new backlogs from accumulating, Kaizen 3 examined each curator's undocumented or alternately documented collections. They identified the minimum level of work needed to record materials in the collections management and project management databases and provided training and coaching to adopt the new practices. They suggested approaches to continuously monitor activities aimed at reducing and eliminating the backlog.

Kaizen 4 members researched the current spaces and space management databases. The work prepared the team for an eventual shelf reading project and a revamp of the existing database that tracks space availability.

Beyond the four kaizens, the team immediately adopted different and more productive ways to work together. Daily "huddles" (also known as stand-up meetings) began the day they returned to the office after the Lean launch. They committed to weekly one-hour meetings to track progress and followed through on kaizens and their 30, 60, and 90 report outs. During daily huddles and weekly meetings, they refined their approaches to complete tasks and negotiated changes in their collective workflow. Rather than selecting one or more test projects, they simply applied the new processes to all the work. As a relatively large organization, they receive new collections weekly if not daily. Incoming collections, along with the identified backlog, were excellent test cases for the improved process.

Results

Quality

Team communication improved right away with daily huddles, weekly meetings, and shared goals and continued to improve with the addition of project management software and a revised accessioning manual. Having one queue for accessioning that was available and visible to all team members immediately cut out the problems created by shadow systems. Curators could trust that their collections were being addressed and successfully churned through the

workflow. Accessioning was now commonly understood to be a series of discrete steps agreed upon by the entire team, Bottlenecks that inevitably cropped up were brought to a huddle or the longer weekly meeting.

Quantity

Measuring quantitative improvements is often the driving force behind implementing the Lean process, because better numbers can be equated with cost savings by management. Common metrics to consider are the time-to-completion, the number of completed accessions per week/month, the number of workflow stall points, and quantifying and reducing the backlog of unaccessioned collections. The scope document was vague on some of these points. The team aspired to "improve the accessioning process to eliminate stall points and make time to completion predictable," and "to continuously reduce the backlog month to month." Even with this lack of clarity, it did not take long for the new accessioning system to show marked quantitative improvements in productivity. The number of accessions completed doubled in the first year, from about 150 accessions completed to more than 300.

Relationships

Although the quantitative improvements were impressive, the team improved qualitatively, as well. They found it easier to communicate and build stronger working relationships. Because Lean focuses on the process instead of the person doing the work, Lean was a key factor in that change. Each team member's voice and perspective had equal value in building workflows, tracking progress, and implementing changes. Each curator had equal access to technical services, common communication tools, and a commitment to using shared standards.

Lean requires willing participants across the board. Some team members were not ready or willing to make changes. Although the team determined it necessary for Technical Services staff to work full-time, two part-time accessioning staff members did not want to increase their hours. Feeling no longer suited to the work, both opted to retire. The team adapted and hired a full-time Technical Services Archivist a year after the Lean launch. She was charged with managing the workflow, leading daily huddles and weekly hour-long meetings. With a functioning common workflow, she could move collections from the hands of curators through the accessioning process and make them ready for researchers. The Technical Services Archivist was a neutral position on equal footing with all curators.

The success of this new approach to queue management prompted them to use tracking systems within other areas of their library work, such as using the library management system Alma to barcode archival materials. Barcoding allowed for location information on a per-item basis to be centrally tracked, which prepared the collections for eventual circulation to the reading room. It also enabled them to track materials that were routed from the department to Preservation for treatments or to curators for exhibits. They added partners from the Preservation and Cataloging departments to their Asana workspace so they could also participate in the workflow.

They grew confident enough with the Lean process that they held subsequent Lean launches to redesign their digital collections accessioning and processing and to manage the cross departmental work of Special Collections library materials acquisitions and cataloging. While these Lean processes were helpful, they did not work with the outside facilitator or do the full three-day redesign, and therefore saw less success and impact than they did with

the first Lean process. Neutral facilitation, support, and substantial time away from the shop to build new processes are important elements of success in the Lean redesign experience.

Montana State University

Background

The Archives and Special Collections (ASC) department of the Montana State University Library has significant primary and secondary source materials in its focus areas.[28] With significant recent acquisitions, the department has a strong mandate from library administration to continue to build its collections.[29] The library hired a new department head in 2020 to lead two and one-half paraprofessionals (Curator, Archives Technician, and Digital Production Manager) and four full-time faculty librarians (Special Collections Librarian, Data Librarian, Outreach and Humanities Librarian, and Archivist).

Like their peer institutions, they face constraints on their ability to efficiently and promptly prepare new acquisitions of unique materials—archival, bibliographic, digital, or a mix—for use by researchers. Before 2020, processing was done primarily by the Curator, the Head of Special Collections, and occasional temporary faculty, staff, or interns. Workflows developed within each of those positions with little coordination or knowledge of those workflows by the remainder of the department members. Processing techniques, though very solid from a traditional standpoint, were centered on manually producing HTML and EAD finding aids and MARC records. Accession records and locations were managed in an aging ProCite database that was inaccessible to all but the department head and the curator.[30] In general, the department operated very separately from the rest of the organization.

Preparation of collections for use in analog or digital form also involves not only ASC, but also Digital Library Initiatives (DLI) and Cataloging, Access and Technical Services (CATS). CATS does subject analysis and name authority work for EAD and MARC records; creates metadata for digital collections; and advises on metadata structures for description and management. DLI builds and maintains the library's digital collections; manages in-house and outsourced digitization with ASC; and oversees technical infrastructure for the library that includes the digital collections system. While the working relationships among the three departments were reasonably good, there were few routine processes established for collection processing. Instead, each project was treated as unique, with little clarity about who initiated and oversaw projects; whether digital projects and metadata were part of routine work or were an "extra"; and who was empowered to determine or adjust timelines and deadlines. Each time a project transitioned from one department to another, individuals had to schedule meetings in order to discuss next steps. Frequently, projects would stall for weeks or months because they were handed off to the wrong person or because individuals lacked adequate information to do the next step in a project. Project documentation was uneven, requiring repeated decision making and making it difficult to declare successful outcomes because not everyone agreed on end products. Re-starting stalled projects took time and fueled frustration between departments and individuals, and library administration often had to get involved to satisfy promises made to stakeholders.

The combination of all of these factors resulted in great difficulty predicting processing times and introduced difficulties for both staff and collection donors. Without predictability, making promises to donors about the availability of their materials was risky and tended to create unsustainable timelines and overwork. For instance, processing and digitizing the

initial deposit of the Ivan Doig papers resulted in significant stress and overwork even as it produced an opportunity for library staff to meet a new challenge and exposed the library's premier literary collection.[31]

ASC staff, their colleagues in CATS and DLI, and library administration were dissatisfied with the status quo and were ready to find a way to distribute processing among existing positions, promote high-quality standards compliant work, support coordinated teamwork within ASC and across the organization, and enable better awareness of capacity. Luckily, they had several key components in place that made success more likely based on Lean in Higher Education expert William Balzer's measures: Desire to change that aligned with the library's strategic plan and its support for employee development and continuous improvement; department leadership change that brought new communication practices; and an acknowledgement of and support for needed changes in position descriptions.[32]

Inspired by their colleagues at the University of Washington, who presented about their Lean redesign at a Northwest Archivists meeting in 2016,[33] they conducted a Lean redesign in 2020–2021. The staff of ASC, along with eight colleagues from DLI and CATS, planned and carried out the redesign, and the Head of Archives and Special Collections was the project director. A Faculty Excellence Grant from Montana State University's Office of the Provost funded Lean consultants Irene Mauch (The Mauch Group) and Megan Mozina (Cresta Solutions) (hereafter Mauch/Mozina) to facilitate the process.[34]

The Lean Redesign

PROJECT SCOPE

Mauch/Mozina facilitated a series of whole-group and small-group workshops in March-June 2021 to introduce the team to the Lean methodology and to define the breadth and depth of the project, the roles and responsibilities, and the end user's value proposition; to analyze current state, gaps, and opportunities; to create a high-level implementation plan; to frame the new process tools; and to provide post event coaching and support. They worked with library administration, the project director, and the CATS and DLI department heads to develop a high-level understanding of the need for the project, define the project charter, and help the redesign team engage with and commit to the project.[35] Because this redesign was conducted during the COVID-19 pandemic, it was done entirely online.

CURRENT STATE AND PAIN POINTS

Like their Washington colleagues, they began by defining the customer, the customer's needs, and then mapping the current state. Mauch/Mozina led the entire group in creating an overview of all the steps involved in preparing a collection (archival, bibliographic, digital, or all three) for use by researchers, from intake to availability. In additional small group sessions, team members described each part of the process in detail. Other team members asked questions to clarify meaning, challenged some assertions, and looked for missing pieces. During these sessions, individuals were more able to describe a single process in detail than they were to participate in a group description of a series of processes, so a rewarding by-product was that participants and departments learned a great deal from one another. Team members struggled to describe all the processes at a similar level (e.g. how do we label boxes? How do we find space on the shelves? How do we decide what level of processing to apply?). Individuals were challenged to describe the most common

scenarios—the 80% rather than the 20%—and required support to redirect their thinking from unusual situations to more routine ones. The project director checked in with individuals throughout the current state workshops and found that individuals reported that the workshops were positive and that they were learning a great deal from one another. For most team members, the current state map was the first time they saw the process from end to end. The current state was documented, and seventy-five pain points were identified where the current processes were not working well. Because this was an overwhelming list that could not be addressed in a single redesign, the list was narrowed to sixteen items that ranged in scope from identifying department roles and responsibilities to a standardized process for labeling boxes.

SMALL GROUPS ON PAIN POINTS

The next phase of the redesign was to address the identified problems. They created sixteen small project groups of varying sizes according to topic and group members' strengths and expertise. (These projects were kaizen events, though that term was not used.) Over the course of two weeks, each group examined the pain point, identified an approach to relieve the pain, outlined the steps needed, and produced a draft of the essential elements of a new approach. Each group met with Mauch/Cresta for support, coaching, and coordination with the project as a whole.[36] They also formed a coordination group (composed of the project director, CATS and DLI department heads, and the Digital Production Manager) that met frequently during the two weeks that the small groups were working to integrate that work into the main flowchart, to observe where groups or individuals needed additional support, and to consider new issues as they emerged.

During the future state planning OCLC Research released its TCS report and tools at precisely the moment that the team needed them. Specifically, they adopted the project plans for archival, bibliographic, and digital as well as the quick cost calculator. With well-considered and field-tested models for collections consideration, project plans, and calculating capacity, the team was able to create stronger processes.

REDESIGN CLOSE

At the end of the small group work, the project had drafts of new practices that addressed critical pain points, a revised overall flowchart, and a clear sense of where to adopt the existing tools from the TCS Toolkit. Mauch/Cresta met with the entire team to show the revised flowchart and how the mini-projects fit in and how they laid the groundwork for continuously testing and revising processes. They also met with library administration to show the results to date, to emphasize the critical role that executive support plays in a successful redesign, and to clarify that the team needed to focus considerable energy on implementation for the coming four months with minimal distractions.[37] This represented the end of Mauch/Cresta's direct work on the redesign. The redesign coordination group became the Cross-Functional Group (CFG), responsible for identifying and managing capacity and the preparation of all unique collections—archival, bibliographic, digital, or a mix—for use by researchers. This group now meets twice a week for 10-30 minutes to provide consistent oversight for all current and upcoming projects.[38] They also established the Technical Review Committee (TRC), which meets weekly to plan and carry out digital projects specifically. The Digital Production Manager serves on the CFG and leads the TRC.

PROTOTYPING TOOLS

Lean relies heavily on trying new processes, rapidly assessing how well they work, and making revisions—all of which enable continuous improvement. The team emerged from future state planning with workflows and tools that were deliberately in draft form so they could make changes before they were over-invested in particular approaches. The first month of implementation consisted of taking the tools that were in draft form and refining them to a usable state. Members of the CFG met with the individuals who worked on the mini-projects by workflow (archival, bibliographic, digital) to step through the workflows and tools in detail so that the group members could clearly understand how each tool functions in the greater whole and could see the gaps between the draft form of each and a form ready for implementation. The CFG prioritized the tools most critical to all processes along with those in the most incomplete form.

The whole team met again for a workflow-specific review of the flowchart, a view of the draft project tracker, and articulation of plans for training for both small groups and the whole team. They celebrated the work done during the first month of implementation and recognized groups and individuals for their contributions. During that meeting, they reinforced the idea of prototyping the new processes and documentation and making notes on needed changes in a shared document rather than getting caught up in details. This is an essential underpinning for continuous improvement in the short and long term.

IMPLEMENTATION: TEST PROJECTS

With those tools and mindset in hand, the team moved to test projects, each of which used one or more specific components created during the redesign. All of them also tested the overall workflows and structure for project planning and tracking and were the focus of all the work; other collection preparation stopped in order to fully embrace the new processes. Over the course of three months, the team completed one archival, one bibliographic, and the end of two digital projects.[39] The two digital projects proceeded smoothly with clear project plans, templates for metadata, and project tracking. The bibliographic project enabled student assistants to accurately search the library's catalog to identify duplicate and unique items which addressed a major pain point for the CATS department. The archival project tested the use of a number of components, including inventory spreadsheets for import into ArchivesSpace, exporting MARC records from ArchivesSpace, and using digital-on-arrival processes.

Results

After these test projects were completed, the project director and all involved staff conducted a review of the initial implementation. Persons involved in each workflow met as a group to step through the process and note where there was insufficient information, missing steps, or flaws in the tools. The review processes uncovered elements that needed improvement, specifically the project tracker, specific uses of ArchivesSpace, and the review of digital collections. Also marked for discussion was initial curation of bibliographic collections and whether unnecessarily complex processes had been implemented. Small groups then made changes and improvements to the tools and processes.

During the review period, the project director also met individually with all involved individuals for 30-minute structured interviews. Interview questions addressed the experiences of each individual and areas of change or continuity in both their work and the work of

the organization (See Appendix B for questions). The project director took notes during each interview, invited every participant to review and correct those notes, summarized both the interview notes and the workflow-specific sessions, and performed analysis of keywords and themes. The summary was shared with and discussed by all Lean participants in a celebration, and with the entire organization in an all-staff meeting.

FIGURE 2
Word Cloud of Summary Notes, 2021 December

Quality

Participants reported that the work completed during the test projects was more consistent and of higher quality. Project plans meant that the overall information on a project was easily available, and decisions were made just once rather than several times. Because individuals had clear roles in the processes, their work felt more relevant and appreciated. And even though work stopped on other projects in order to focus on and measure the results of the test projects, the quality of the test projects meant that the overall impacts were positive.

Quantity

At this stage of the project, there was still insufficient data to assess whether the new processes increased the quantity of collections that could be prepared for researchers. Considering they went from having no idea how long work took to having some means to estimate it through the TCS toolkit, it is an improvement, nonetheless. With the new processes, participants stated that the work felt faster because of having increased clarity and fewer bottlenecks. Continued data collection will show whether this perception is quantitatively supportable.

Relationships

Improvements in quantity and quality were important, but somewhat unsurprising. The most significant change seen in the initial implementation was in relationships. The keywords

used most frequently during the interviews described increased understanding, collaboration, clarity, communication, and transparency. Participants reported an increased sense of responsibility toward one another, and an increased sense that all roles are clear and valued. They reflected that ASC is much more connected and integrated with the rest of the organization, rather than operating as a distinct and siloed entity. Several participants discussed the experience of talking openly about the pain points, which they felt were previously taboo in the organization, and expressed gratitude and support for being encouraged to have these discussions and to seek solutions together

Discussion

The Lean redesign processes at the University of Washington and Montana State University occurred at very different organizations and more than five years apart. However, both processes share some common themes: collections stewardship, the function of standards compliance, and fomenting and sustaining organizational change. For both organizations, having permission to discuss pain points was not only needed, but was an entryway to creating solutions. In both cases, the Lean redesign resulted in significant positive outcomes along with challenges that should inform similar projects at other institutions.

Stewardship

Both institutions faced slightly different versions of the same challenge: responsible stewardship. For the University of Washington, a backlog of material lacking even basic metadata meant that a large number of collections were completely inaccessible. Montana State University lacked the same degree of backlog, but was challenged to meet donor and administrative expectations without either making promises that could not be fulfilled or keeping promises while significantly overtaxing library staff. As the TCS report observes,

> Archives and special collections are charged with collecting materials that document our society and its institutions as well as with the ongoing, responsible stewardship of these records. Yet many archives and special collections struggle to manage the volume of materials under their care. Accumulations of inaccessible, poorly described collections and inadequately preserved materials can create a breach of the trust we hold with collection donors and users. Matching collecting activities to resources is fundamental to stewardship.[40]

The TCS authors also observe that archives and special collections have given significant focus to reducing backlogs and preparing collections more efficiently. They propose a new approach that uses a constraint model for collecting so that libraries collect within their capacity and make promises to stakeholders that they can keep.[41] That model builds on existing efficiencies and adds a measured approach to calculating and communicating capacity. These experiences suggest that processes redesigned using Lean support both approaches. The Lean redesign at the University of Washington fit, chronologically and theoretically, into the "increase efficiency" trend. It fulfilled that goal: quantitative measures showed a marked increase in the number of collections accessioned each year. The redesign at Montana State University, with its twin focus on increasing efficiency *and* understanding and articulating capacity for preparing new collections, was parallel with the framework of TCS and "responsible stewardship." While

the quantitative data is still emerging, the improved communication lays firm groundwork for responsible stewardship.

However, the calculation and communication tools provided by the TCS toolkit are just one part of responsible stewardship. The other part, which both institutions confronted, is creating and supporting a team rather than an individualistic approach to special collections work. University of Washington Lean team members were not unanimous in their support for the process. Some reluctantly participated, convinced that their voice would not be heard. The Lean process is designed to address this issue: team members incrementally and collectively document the current state and imagine the future state point by point and moving forward only when all are in agreement, turning the skeptic into a big supporter of the process. For Montana State University, departments and individuals had no real understanding of how the library prepared unique collections for use. Lean current state process mapping helped every member of the team understand where their work fit in. One team member observed, "every piece of this flowchart represents a promise that we make to one another."[42]

Standards Compliance

The "special" in "special collections" meant, for decades, unique approaches for processing every collection; it is a legacy of the profession that is difficult to move past.[43] The emergence of DACS as a descriptive standard just eighteen years ago counters that legacy. With the developments of the last twenty years—EAD, DACS, standardized rights statements, and all of the advances in name authority work including the Virtual International Authority File (VIAF) and Social Networks and Archival Context (SNAC)—practitioners in unique collections increasingly understand and see both the advantages of standardization and the limitations of customization. Both institutions used Lean processes to significantly enhance their ability to create consistent metadata for use and re-use and to move that metadata through systems.

For MSU, standardization was one approach to relieving pain points uncovered during the Lean redesign. These changes—metadata templates for digital collections, ensuring that bibliographic searching was done consistently by student assistants, clearly defining levels of physical processing for archival collections—made immediate and tangible differences in how the work was done. They represent a move from a completely customized process to one with the right level of routine and room for appropriate customization. However, this change was difficult: for some staff, standardization felt like moving toward a "cookie cutter" process. The current state mapping process provided an opportunity for the project leader and the consultants to key in on individuals' expertise and perspectives and to carefully consider how to draw on it to create balanced approaches.

For both institutions, revising processes also included ensuring that the metadata created during them was both consistent across the organization and compliant with national standards. Inherent to the process was engaging the role of accessioning in establishing consistent bibliographic control. For the University of Washington, the Lean process forced a reckoning about the meaning of accessioning—once everyone agreed on the definition and what they wanted out of the process, it was easier to move forward and engineer the process to check all of the boxes. They realized they wanted each accession to have at least a minimum level of description. They identified all of the fields in their collection management software that were essential at each level of the process. In some cases these fields corresponded to fields that would be displayed in their finding aid and their MARC record (the outward facing mani-

festations of the invisible accessioning work). In this way they baked standards compliance into their everyday expectations of what accessioning is and always strove for the "golden minimum." They agreed that all accessions would be made discoverable by the end of the accessioning process whether or not any detailed processing had been done or was planned. They embraced the "accessioning as processing" framework out of necessity.[44]

For Montana State University, the Lean process established a clear practice of doing fully standards-compliant minimal description on accessioning, and in ArchivesSpace, so that the metadata could be re-used for as many outputs as possible: EAD finding aid, MARC record, digital collections metadata, and a variety of administrative reports, including those for donors. This eliminated both tedious cut-and-paste work across multiple systems and the re-creation (often in different forms) of metadata for those outputs. For ASC staff members, learning and practicing standards-compliant basic description was a hurdle but ultimately achievable for most.

Managing Change

Fundamentally, Lean redesign is about people embracing the changes they identified, implementing the solution they have formulated, and tracking progress toward goals they have established. It is a bottom-up, grassroots process that respects the people who do the work. This makes it a powerful approach to creating and sustaining change in an organization, so long as there is also continuous support and expectations for sustaining new processes. Absent that support, new processes will lose steam.

At the University of Washington, the Lean process was supported through their department administration and encouraged in general at the University.[45] Within this context, each team member was expected to be motivated to participate and commit to new and more efficient practices. They learned from their facilitator that it is common for team members to not find themselves in the new work, so the retirements of the accessioning staff were unsurprising. The changes also set up the new Technical Services Archivist for success: they joined an already motivated group that had adopted new processes and wanted to continue to improve them.

However, over time enthusiasm waned for the idea of continuous improvement at the University of Washington. Some staff were, if not resistant then, more likely to fixate on the extra effort to master new tools and less committed to a common workflow. In the end, a uniform workflow was not adopted by all curators. The director did not manage the process, failing to insist on the need for adherence to common/shared practices and workflows and failing to assign the authority to enforce it to another person in the department.

At Montana State University, the project director and Mauch/Cresta closely monitored the level of engagement and skills evidenced by members of the team throughout the process. Some individuals were expected to be strong contributors to this work, either directly by contributing knowledge of specific processes or more indirectly by asking questions of colleagues. Some individuals emerged as strong contributors unexpectedly, proving themselves strong systems thinkers and reliable collaborators. Some demonstrated that they were less engaged by not following agreed-on project norms or contributing to discussions. And some individuals had struggled more than others to articulate how work might be done differently and why. Sharing these observations allowed the project director and Mauch/Mozina to shape individuals' places in the project and identify both those who could contribute a great deal, those who were likely to contribute more to very specific topics, and those who would need

additional support and encouragement to contribute well.

One of the roles the strong contributors began to play was to provide peer support. A few team members expressed concern that the new processes were overkill, and that they were documenting and planning at the expense of actually accomplishing the work. This is a reasonable concern and a reflection of what a profound change this was. This type of skepticism can emerge because it is hard for people to see how the work they do fits into the process: they are simply doing the work in the way that seems best to them, sometimes in the way that is most expedient or convenient. Stepping through the Lean process enhanced everyone's understanding of how and why their work impacts others. For example, a new process at Montana State University ensured that donated bibliographic materials were accurately searched in the ILS before being selected for the collection was one of the most impactful. When one team member expressed their reluctance about changing processes, another team member described how inaccurate searching wasted their time and made it harder to do their job. Yet another team member stated to the resistant one, "[name], [they] need you to do this in order to do [their] work!"[46] After this discussion, one of the test projects that used the new processes yielded significant change. While CATS had in the past received 100% of bibliographic materials with unclear selection criteria or that lacked other key information, the new processes meant that none of the materials had those issues.[47] This, in turn, made the CATS team members whose work was improved even more enthusiastic about the new processes. The individual who was resistant is willing to continue to follow the new processes as a courtesy to their colleagues.

Managing change and supporting individuals is also an ongoing process that takes considerable time and attention from leadership. At Montana State University, Mauch/Cresta provided both intensive support during the redesign and supported the project leader in sustaining support for managing change. Mozina observed that when people are used to doing something a particular way, the easy path will be to do it the same way. Deviating from the old way can arouse emotion and stress, and people will naturally resist the additional effort associated with the new practice.[48] Countering this natural tendency requires significant effort by individuals, peers, and organizations.[49] Team members needed to acknowledge the urge to fall back into old habits, remind themselves of the rational path, and use the new practice. Manager expectations and peer support play essential roles in supporting these changes. Balzer acknowledges this reality in his observations about an institution's readiness to undertake a Lean redesign: "Major institutional change is an ambitious undertaking, and university leaders should be fully cognizant of the sustained commitment needed to implement LHE [Lean in Higher Education]."[50] Indeed, some observe that many Lean redesigns fail in the long term because of lack of leadership support and because employees just don't want to change.[51] Additionally, staff turnover and changes in leadership introduce both challenges and opportunities. Onboarding new employees, or even new student assistants, gives the team a chance to introduce the improved workflows to colleagues unfettered by the "old way" of doing things. Furthermore, new team members are good at exposing points in the process that need more clarity or refinement.

The University of Washington Special Collections is a standalone department in a relatively large organization where, historically, most of the work has been done within the department rather than across the organization. Montana State University had the same situation despite its much smaller size and sought better integration across the three departments. For both organizations, the Lean process resulted in better integration both within the department and

across the organization. For instance, during the qualitative assessment at Montana State, one of the most frequently cited advantages was a better understanding of others' work. The process built better relationships across departments. For the University of Washington, attempting standardized work has necessitated partnerships with other departments. With the realization that SC can't do the work alone, they looked to other experts including Acquisitions, Cataloging, and Preservation.

Advocacy and Administration

Process redesign, continuous improvement, and managing change may be grassroots in nature, but they also require substantial support from leadership in order to be successful.[52] Responsible stewardship is not in itself the focus that upper administration may gravitate toward! But it enables the higher visibility things that administrators want, including exposing high profile collections, enabling impactful scholarship, and supporting transformative learning experiences for students.

For MSU, library administration enthusiastically supported the Lean process, including advocating for the grant that supported it and providing encouragement throughout the initial redesign. The results of the redesign increased library administration's understanding of responsible stewardship. Acquisition of prestigious collections that raise the library's profile and support transformative learning experiences for students was already a priority, but an understanding of what is required for responsible stewardship was less so. By articulating the steps involved, using well-formed tools from the TCS kit, and showing a determination to improve both quality and quantity, they were able to increase the understanding of those requirements. Over time, they anticipate they will learn more about how well they are able to match aspirations and resources to complete collection preparation within predicted timeframes.

Assessing Success

For both institutions, carrying out a Lean redesign advanced both efficiency and responsible stewardship. The most immediate results were improved communication and relationships both within the departments and across the organizations. For the University of Washington, the redesign measurably increased the amount of work done. For both organizations, the redesign increased standards compliance and the ability to re-use metadata.

For the University of Washington, with eight years passed since the Lean redesign, they can now see both success and failure. The director was thrilled with the initial outcomes and because the process had improved so dramatically, it seemed by all accounts that "accessioning was fixed." Everyone involved in the Lean launch and the subsequent activities–including daily huddles that continued for several years and weekly meetings that continue to this day—understands how crucial an improved accessioning process is to every other part of our work. In addition to permanent "fixes," they hired a temporary accessioning assistant for backlog management for 22 months. Using the new processes and concentrated staff time to attack the backlog, they moved 545 collections through the accessioning process, effectively cutting their "hidden collections" in half. It was due to the Lean process mapping that they were able to get a handle on the steps needed to daylight collections that hadn't yet gone through the accessioning process.

In spite of these successes, they still have a substantial backlog. Why haven't they obliterated it by now? It turns out that they aren't alone. TCS reasons that continuing backlogs

"cannot be addressed solely through increased efficiency in technical services and infusions of extra labor"; capacity constraint must be considered as well.[53] Capacity constraints, defined as "factors that limit production, performance, or output," are at the heart of the TCS framework and surrounding tools.[54] At the University of Washington, while accessioning was improved and continued to evolve, no amount of efficiency in that realm could get at the underlying issues surrounding the lack of sustainable stewardship of collections. Leadership (deans and directors) are encouraged to engage in process improvement by providing leadership support and developing an overarching collection management policy that embraces the TCS framework and thoughtfully considers staffing and space constraints.

A Lean redesign also requires significant time and energy to support it. In both cases, the departments needed the help of skilled consultants to design and carry out the redesign. For the University of Washington, that expertise was available in a unit on campus; for Montana State that expertise required an internal grant to support Mauch and Cresta's work. Both entities found that after conducting a successful redesign, it was challenging to sustain change and make continuous improvement part of the culture. Successfully sustaining change requires substantial support that begins at the highest level in the organization and continues throughout the hierarchy, down to individuals who are invested in the work and can also contribute to peer support. Balzer elaborates on all of these points.[55] Last, the constant change in any organization, including gaining or losing positions, key staff moving to other organizations, and special projects that take precedence over day-to-day work, makes continuous improvement difficult to sustain as a single unit. Instead, the enthusiasm for and commitment to sustaining change needs to be supported across the organization. Middle managers (e.g. department heads for archives and special collections) must be prepared to not only continuously support their and other department members, but to engage administration and solicit ongoing support.

Lean Principles for Archives Collection Preparation

Lean Principle[56]	Lean for Archives Collection Preparation
Customer defines value for each project	Prioritize what users value most over what archivists value most.
Identify (map) how value is created	Map the work in detail so that collections are prepared through a clearly identified sequence of actions
Focus on the flow of work and avoid any interruptions	Understand how work moves from one person to another and provide transparency in those processes.
Let the customer pull value from the producer	Provide predictability for preparation times.
Pursue perfection	Enable continuous improvement of processes that integrate the expertise of all personnel involved. Integrate new practices and innovations. Don't get stuck on just one way to do things!

Conclusion

For the special collections units in the libraries at the University of Washington and Montana State University, a Lean redesign for portions or all of collection preparation processes had

transformative impacts on quality and quantity of work. Both institutions learned valuable lessons about what is necessary to initiate and sustain change, to support team- and standards-based approaches, the key benefit of responsibility to each other, and the value of a framework that is not commonly used in special collections and archives. Both institutions increased their capacity for responsible stewardship. Most importantly, Lean's grassroots approach and respect for people allowed all staff, particularly those resistant to change, to engage with both the process and to carry through on changes to the work. With sustained administrative support and attention, archives and libraries can benefit from implementing Lean or other BPM approaches. Doing so is consistent with recent moves toward standardization, efficiency, and collecting within constraints. It is a substantial effort and an ongoing investment, but the ongoing results are worth it!

Acknowledgment

The work at Montana State University described in this paper would not have been possible without the work of Irene Mauch and Megan Mozina, of The Mauch Group and Cresta Solutions respectively, a dynamic duo who facilitated a really transformative project despite a global pandemic! The project was also a success because of the engagement and thoughtful contributions from the project team: Doralyn Rossmann, Amy Foster, Jan Zauha, Sophia Phillips, Sara Mannheimer, James Thull, Gary Barnhart, Brandon Watson, Anne Stefani, Molly Arrandale, Rhonda Borland, Jim Espeland, Jakob Schultz, and Anika-Anzum Prima. The author additionally appreciates the steadfast support of Associate Dean Brian Rossmann and Dean Kenning Arlitsch.

The University of Washington Libraries Special Collections Lean efforts were coordinated within the greater campus Lean program and led by volunteer launch facilitator and coach Cara Ball, compliance analyst at UW Records Management. The Special Collections Lean team was composed of Technical Services and Curatorial staff members: John Bolcer, Nicolette Bromberg, Mark Carlson, Conor Casey, Nan Cohen, Anne Jenner, Susan Kemp, and Sandra Kroupa. The team was sponsored by Paul Constantine, Associate Dean of UW Libraries and Director of Special Collections.

Appendix A: Lean Background and Terms

The following are some common terms used in Lean. Montana State University and the University of Washington did not use these terms in precisely the same ways; these definitions reflect the common and variant uses.

Current state: The process "as is" today. Articulated and analyzed through a variety of means, including flowcharts, spaghetti diagrams, and others.

Future state: What processes will look like in the future, after redesign. Articulated through a flowchart or value stream diagram.[57]

Huddle or cross-functional group: A short, frequent, meeting of people who play roles in a workflow. Also known as a standup, these can be daily or another interval.

Lean launch: A cycle with a defined beginning and end devoted to designing or redesigning a process[58]

Process Redesign: Re-engineering a business process so that it delivers greater value to the customer(s).[59]

Lean Redesign: The active work of using Lean tools to reduce waste, increase value to the customer, and improve efficiency and quality.[60]

Respect for People: A focus on valuing the contribution and opinion of those involved with the work. A fundamental value of Lean.[61]

Report Out: Describing expected results and how the process must operate to deliver them.[62]

Kaizen: the Japanese philosophy of continuous improvement, applied in a short, intense, focused workshop that redesigns a sub-process.[63]

Appendix B: Montana State University Project Assessment Questions

1. Describe your role in processes (e.g. which new or revised processes did you "touch" the most?)
2. What do you think has changed about your work as the result of having new processes?
3. What hasn't changed about your work?
4. What do you think has changed about us as an organization as the result of having new processes?
5. What hasn't changed about us as an organization?
6. What do you think needs to change?
7. What are you most pleased about?

As part of planning for this assessment, the Project Director confirmed requirements for IRB review. She determined that, as an internal improvement/quality assessment initiative, IRB approval is not required (https://www.montana.edu/orc/irb/human_subjects_research.html). In reporting on this assessment, she is following best practices, e.g. not identifying individuals and seeking permission for any quotation.

Notes

1. See https://dictionary.archivists.org/entry/processing.html.
2. Other sub-disciplines include Total Quality Management, Operations Management, and Six Sigma.
3. Marlon Dumas, Marcello La Rosa, Jan Mendling, and Hajo A. Reijers, "Introduction to Business Process Management" in *Fundamentals of Business Process Management* (Springer Berlin, Heidelberg 2013): 1.
4. William K. Balzer, *Lean and Higher Education: Increasing the Value and Performance of University Processes*, 2nd Edition (New York: Routledge, 2020): 15.
5. James P. Womack and Daniel T. Jones, "Beyond Toyota: How to Root out Waste and Pursue Perfection," *Harvard Business Review*, vol. 74, no. 5 (1996): 140.
6. Kim Dong-Suk, "Eliciting Success Factors of Applying Six Sigma in an Academic Library: A Case Study," *Performance Measurement and Metrics* 11, no. 1 (2010): 25–38, https://doi.org/10.1108/14678041011026847.
7. Sarah Anne Murphy, "Leveraging Lean Six Sigma to Culture, Nurture, and Sustain Assessment and Change in the Academic Library Environment," *College & Research Libraries* 70, no. 3 (May 2009): 215.
8. Elizabeth Nelson, "Using Six Sigma and Lean in the Library," *College & Undergraduate Libraries*, 22: 3–4 (2015): 312–324, https://doi.org/10.1080/10691316.2015.1070701.
9. Shaunessy Everett and Marihelen Hatcher, "Do More, Better, for Less," *Library Journal* 131, no. 15 (September 15, 2006): 28.
10. Jeanne F. Voyles, Linda Dols, and Ellen Knight, "Interlibrary Loan Meets Six Sigma: The University of Arizona Library's Success Applying Process Improvement," *Journal of Interlibrary Loan, Document Delivery & Electronic Reserve* 19, no. 1 (January 2009): 75.
11. Sarah Anne Murphy, "Leveraging Lean Six Sigma to Culture, Nurture, and Sustain Assessment and Change in the Academic Library Environment," *College & Research Libraries* 70, no. 3 (May 2009): 218.
12. Lenore A. England and Stephen D. Miller, *Maximizing Electronic Resources Management in Libraries: Applying Business Process Management* (Witney: Elsevier Science & Technology, 2015).
13. Elizabeth Nelson, "Using Six Sigma and Lean in the Library," 319.
14. J. Gordon Daines III, "Aligning Customer Needs: Business Process Management (BPM) and Successful Change Management in the L. Tom Perry Special Collections," *Library Leadership & Management* 29 no. 1, http://dx.doi.org/10.5860/llm.v29i1.7101; J. Gordon Daines III and Cory L. Nimer, "Integrating Process Management with Archival Management Systems: Lessons Learned," *The Code4Lib Journal* 6 (2009), http://journal.code4lib.org/articles/1016 .
15. Daines, personal communication with Jodi Allison-Bunnell, 2020 October 2.

16. Mary Jo Pugh, "The Illusion of Omniscience: Subject Access and the Reference Archivist," *The American Archivist* 45, no. 1 (1982): 33–44; Elsie T. Freeman, "In the Eye of the Beholder: Archives Administration from the User's Point of View," *American Archivist*, 47 (Spring 1984): 111–123.

17. Paul Conway, "Facts and Frameworks: An Approach to Studying the Users of Archives," *The American Archivist* 49, no. 4 (1986): 393–407.

18. ARL Research, Teaching, and Learning. Exposing Hidden Collections Conference Summary, 2003. http://www.arl.org/rtl/speccoll/hidden/EHC_conference_summary.shtml; ARL Special Collections Task Force Final Status Report July 2006, https://www.arl.org/wp-content/uploads/2006/07/special-collections-task-force-final-status-report-july2006.pdf; both accessed 2021 January 11.

19. "Program History," https://www.clir.org/hiddencollections/program-history/, accessed 2021 January 11.

20. "Digitizing Hidden Special Collections & Archives," https://www.clir.org/hiddencollections/, accessed 2021 January 11.

21. Mark A. Greene and Dennis Meissner, "More Product, Less Process: Revamping Traditional Archival Processing," *The American Archivist* 68, no. 2 (2005): 208–63.

22. "User-centered Archival Description," in *Describing Archives: A Content Standard*, https://saa-ts-dacs.github.io/dacs/04_statement_of_principles.html#user-centered-archival-description, accessed 2020 December 11.

23. Chela Scott Weber, Martha O'Hara Conway, Nicholas Martin, Gioia Stevens, and Brigette Kamsler, *Total Cost of Stewardship: Responsible Collection Building in Archives and Special Collections*, (Dublin, OH: OCLC Research, 2021), https://doi.org/10.25333/zbh0-a044, accessed 2021 September 23.

24. Collecting areas include a rich Pacific Northwest Collection and other special collections including architecture, 19th-century and 20th-century American and English Literature, Book Arts Collection, the Labor Archives of Washington, and the University Archives.

25. In archives, "accessioning" refers to the steps taken by staff to establish preliminary physical and bibliographic control over an analog or digital acquisition. Although a standardized process is only just emerging, accessioning usually involves a minimal-level DACS description, rehousing, and a shelf or digital location. See https://dictionary.archivists.org/entry/accession.html.

26. Traditionally at University of Washington, curatorial staff are primarily responsible for acquiring materials whereas technical services staff accession, arrange, describe, and make materials discoverable using library systems. Curators are more public-facing while technical services staff are involved in hidden labor.

27. See https://dictionary.archivists.org/entry/finding-aid.html.

28. Collecting areas include agriculture; architecture and engineering; Montana and western history; Native Americans; Montana State University; politics and government; Montana history; and regional writers.

29. Since 2015, these have included collections from renowned writer Ivan Doig, scientist Frank Craighead, writer and filmmaker John Heminway, and modern artists Bob and Gennie deWeese.

30. ProCite was first released in 1983 and was popular in many industries, including libraries. It has been unsupported since 2013. https://en.wikipedia.org/wiki/ProCite, accessed 2021 October 25.

31. Kenning Arlitsch et al., "Digitizing the Ivan Doig Archive at Montana State University: A Rise to the Challenge Illustrates Creative Tension," *Journal of Library Administration* 57, no. 1 (January 2, 2017): 99–113, https://doi.org/10.1080/01930826.2016.1251251.

32. William K. Balzer, *Lean Higher Education*; Montana Library Strategic Plan, https://www.lib.montana.edu/about/strategic-plan/index.html, accessed 2021 June 7.

33. Cara Ball, Nicolette Bromberg, and Emily Hughes Dominick, "Accessioning All-Stars: Turning our Dysfunctional Process around Using Lean, A Process Improvement," Northwest Archivists Annual Meeting, Seattle, Washington, May 2016.

34. Center for Faculty Excellence, Montana State University, https://www.montana.edu/facultyexcellence/grantsawards/feg/2020feg-Rd2.html, accessed 2021 October 25. The Lean Redesign for Archives and Special Collections at Montana State University was facilitated by the partnership between Mauch Group LLC and Cresta Solutions LLC. Irene Mauch, principal of the Mauch Group, has over thirty years of experience with Lean and Six Sigma in a variety of industries, including higher education. Megan Mozina, principal of Cresta Solutions, has fifteen years of developing leaders and supporting strategic changes in higher education. In combination, they bring strong experience in process improvement in higher education, complementary skills, and extensive experience in online workshop facilitation. https://mauchgroup.com; https://www.crestasolutions.com, both accessed 2021 June 14.

35. As part of this process, the team members also agreed to a set of ground rules for the process that described shared expectations for attendance, engagement, timelines, video conferencing, and mindset.

36. In order to support the work of many small groups, the project director created an additional page for the main flowchart to document metadata and information flow among all the processes the group was develop-

ing. For instance, the overall description of a collection that is created in the process of working with a donor or seller becomes the basis for both the preliminary scope and content note in the accession record, which in turn appears in final descriptions like the finding aid. This allowed groups to see where information could be collected once and re-used, with or without modifications, and where standards-compliant forms needed to emerge.

37. Montana State University Library, Executive Team minutes, 2021 April 19 (not publicly available).

38. Adopting daily stand-up meetings with this group, or even with the entire group involved in the redesign, might be more typical. These twice-weekly meetings were more organizationally viable.

39. The two digital projects were completed by a vendor but needed to be integrated into the library's systems and made available to researchers. Two additional projects anticipated had unexpected issues with previous processing that made them unsuitable as test projects.

40. Chela Scott Weber, et al., *Total Cost of Stewardship*.

41. Ibid., 3–4.

42. Paraphrase, 2021 April. Used by permission.

43. Jodi Allison-Bunnell, Megan Mozina, Irene Mauch, "Your Work is Special! Lean Makes it Repeatable," Lean in Higher Education Conference, Glasgow, Scotland, 2021 October 21.

44. Christine Weideman, "Accessioning as Processing," *American Archivist* 69 (2): 274–283.

45. Washington Lean is an official administrative program offered through the University of Washington Finance division. A case study, "Transforming Higher Ed: Implementing a Culture of Continuous Improvement at the University of Washington," was winner of the 2019 Shingo Research and Publication Award, https://finance.Washington.edu/lean accessed 4/13/2022.

46. Paraphrase of team member, 2021 March. Used with permission.

47. Interview with KC and MA, 2021 November 4. Used with permission.

48. Mauch and Mozina, Full Team Meeting presentation, 2021 April 28.

49. https://dictionary.apa.org/plan-continuation-bias, accessed 2022 January 7.

50. William K. Balzer, *Lean Higher Education*, 102.

51. Lynn Kelley, "The 'Playbook' of Sustaining Change." https://www.lean.org/the-lean-post/articles/the-playbook-of-sustaining-change/, accessed 2022 October 21.

52. In the kickoff meeting with the library's administration, Mauch and Cresta stated that leadership support is the single biggest factor in the success of a redesign.

53. Chela Scott Weber, et al., *Total Cost of Stewardship*.

54. Ibid, 15.

55. William K. Balzer, *Lean Higher Education*, 259–262

56. As per Womack et al., 1996.

57. https://www.leansixsigmadefinition.com/glossary/value-stream-map/, accessed 2021 December 10.

58. 10 Steps to a Lean Launch, https://www.lean-labs.com/lean-launch, accessed 2022 February 25.

59. https://www.leanblog.org/2017/01/lean-redesign-continuous-improvement-organizations-blog/, accessed 2021 December 10.

60. https://www.leanblog.org/2017/01/lean-redesign-continuous-improvement-organizations-blog/, accessed 2022 February 25.

61. Glossary of Lean Production Related Terms, All About Lean, https://www.allaboutlean.com/lean-glossary/#r, accessed 2021 October 4.

62. http://theleanthinker.com/2011/03/04/the-report-out/, accessed 2022 February 14.

63. Glossary of Lean Production Related Terms, All About Lean, https://www.allaboutlean.com/lean-glossary/#k, accessed 2021 October 4.

Transfiguring the Library as Digital Research Infrastructure: Making KBLab at the National Library of Sweden

Love Börjeson, Chris Haffenden, Martin Malmsten, Fredrik Klingwall, Emma Rende, Robin Kurtz, Faton Rekathati, Hillevi Hägglöf and Justyna Sikora

This article provides an account of the making of KBLab, the data lab at the National Library of Sweden (KB). The first part discusses the work involved in establishing a lab as both a physical and a digital site for researchers to use digital collections at previously unimaginable scales. The second part explains how the lab has deployed the library's collections as data to produce high quality Swedish AI models, which constitute a significant new form of digital research infrastructure. We situate this work in the context of uneven AI coverage for smaller languages, and consider how the lab's models have contributed to the making of important AI infrastructure for the Swedish language. The conclusion raises the possibilities and challenges involved in continuing this type of library-based AI development.

Introduction

In an era of big data, significant new demands are being placed upon libraries.[1] As the world becomes increasingly amenable to processes of datafication, and more and more previously unquantified aspects of life are rendered into data, the library as a cultural heritage institution has been forced into a period of creative transformation.[2] This is partly a matter of developing collecting practices for the vast amount of material being produced online and exploring sustainable ways to describe and store these web archive collections for future users.[3] But it also involves strategies to meet the needs of users in the present, especially the novel requirements of digital scholarship.[4] Researchers in the humanities and social sciences pursuing

* *Love Börjeson is Director of KBLab at the National Library of Sweden, love.borjeson@kb.se; Chris Haffenden is Research Co-ordinator at KBLab, email: chris.haffenden@kb.se; Martin Malmsten is Head Data Scientist at KBLab and an IT Architect at the National Library of Sweden, email: martin.malmsten@kb.se; Fredrik Klingwall is a Developer at KBLab and the National Library of Sweden, email: fredrik.klingwall@kb.se; Emma Rende is a Product Manager at KBLab and the National Library of Sweden, email: emma.rende@kb.se; Robin Kurtz is a Senior Data Scientist at KBLab, email: robin.kurtz@kb.se; Faton Rekathati is a Data Scientist at KBLab, email: faton.rekathati@kb.se; Hillevi Hägglöf is a former Data Scientist at KBLab, email: hillevi.hagglof@gmail.com; and Justyna Sikora is a Data Scientist at KBLab, email: justyna.sikora@kb.se.*

digital approaches now routinely expect to be able to conduct analysis of library collections at previously unimaginable scales.[5] Such an expectation is particularly evident at research and national libraries with legal deposit material, where it creates distinctive challenges for information systems that have historically favored the analogue object and single item use. How do these libraries go about providing access to their collections as data, when so much of their underpinning socio-technical imaginaries have been centered upon the individual book?[6]

This article explores this question via the organizational form of the *data lab*. Faced with increasing demands for computational access to collections over the past decade, university and national libraries have responded by instituting such labs—with LC Labs at Library of Congress, British Library Labs and Yale Digital Humanities Lab as characteristic examples. Broadly speaking, these amount to the creation of an internal platform where the professional expertise of data scientists can be harnessed towards the informational complexities of digitization and facilitating new forms of digital research. Here we use the example of KBLab at the National Library of Sweden (*Kungliga biblioteket*, hereafter KB) to discuss what is involved in creating such a lab in a library setting. The first part details the infrastructural work required to make KB's digital collections available for large-scale analysis, as well as the practical and technical setup established at KBLab. The second part moves on to explain how the use of collections as the basis for development work with artificial intelligence (AI) has proved foundational in transforming the library into a digital research infrastructure. Though the particular details of KBLab are specific to the Swedish context, we raise broader arguments relevant for a wider international audience of library professionals, digital researchers and policy makers. In sum, the article elaborates on the value of such data labs in the heritage sector, while offering a principal justification for the project of library-centered AI development as a public good.

Literature Overview: AI in the Library

Until very recently there existed a surprising "absence of scholarly research on AI-related technologies in libraries."[7] Yet—and at least in part due to the rapid increase in public discussion of AI prompted by the release of ChatGPT—a greater body of studies exploring the possible roles and applications of AI in the context of academic and research libraries has now started to appear. One strand of such work has approached this subject from a broad perspective, looking at the question of what AI and machine learning could offer libraries in general: this includes Ryan Cordell's recent state-of-the-field report for the Library of Congress and the various publications of the "collections as data" movement propagated by Thomas Padilla and others.[8] An alternative type of study has been those focused more particularly on the AI awareness levels of library staff and the expectations that these professionals have about potential future adoption and application of such techniques.[9] Various studies have also discussed both some of the challenges in enabling a data-driven approach to digital research at scale, and some of the ways in which new forms of AI applications could be integrated in the work processes of academic libraries to make digital collections more amenable for such research.[10]

However, there are two significant dimensions of AI applications in libraries that have received less attention. The first of these is the practical and organizational efforts involved in making it possible for new insights from the field of data science to be explored and experimented with in the information-intensive environment of academic and research libraries in general, and at legal deposit libraries in particular. Here we build upon a recent attempt

to address this lacuna, *Open a GLAM Lab*, the manifesto encouraging the growth of further labs in the heritage sector, by providing a specific case study of such institutional work.[11] The second dimension is the potential for libraries not only to integrate AI techniques developed elsewhere, but also to serve as a site of experimentation for the making and testing of more democratically-inclined AI tools that are transparent and open for scrutiny. In pursuing this aspect, we are reinforcing the suggestion that "data labs at libraries—and especially national libraries—can have a significant role to play in the future of AI.

KBLab as Entrance Point to the Collections for Research

In this opening section, we sketch the practical and organizational conditions that shaped the making of a data lab at the National Library of Sweden. How does KBLab align with and form part of KB's broader mission as a national library? What is it about a national library's collections that is particularly well-suited to the type of work that is possible in such a lab? And what is involved in creating access to this material in a lab environment? Addressing these questions provides the contextual detail necessary to make sense of the subsequent discussion of AI development in the library presented towards the end of the article.

Library Collections as Data

The concern with making collections available is entirely central to KB's purpose as a publicly-funded heritage institution. The library's obligations to the research community in this regard are highlighted in the legal act defining its principal mission, where the opening paragraph describes KB as both a "national library" and a "national research infrastructure."[13] While the concrete tasks that pertain to this—i.e. to collect, describe, preserve, and make accessible material—are outlined in relation to the general good of aiding democratic development, the act specifically connects these activities to the end of safeguarding Swedish research quality. In this sense, KB is bound by law to maintain a close relationship with the shifting and dynamic needs of researchers. In practice, and given today's increasingly digitalized media ecology, this means incorporating the digital into a national research infrastructure and in turn becoming a digital research infrastructure. As we will demonstrate, KBLab comprises an essential component in both of these dimensions.

KB's collections can be characterized in terms of their considerable breadth and scale. While first introduced as a form of official censorship in 1661, with publishers forced to submit a copy of each work to the state for approval prior to public circulation, Sweden's legal deposit act has long since served to make the national library a guarantor of future cultural heritage. The law dictates that a copy of every publication issued in Swedish must be submitted to the library for preservation; since 1979 this has included audio-visual material as well as print, and since 2015 at least a degree of electronic publications.[14] Beyond their historical depth and continuity over time, the collections thus also encompass cultural production from a diverse and shifting media landscape: ranging across newspapers, magazines and books through scientific journals and governmental reports to radio broadcasts, television shows and computer games. To give a sense of the scale involved, KB's physical collections alone now number over 18 million items in the archives.

Although only a small part of these collections has yet been digitized, sufficient volumes of digital material exist to make the "collections as data" perspective highly pertinent.[15] Such a framework entails exciting possibilities but also significant infrastructural challenges for the

GLAM sector. In terms of the former, the creation of high quality, language-specific humanities data opens up new potential for researchers to be able to analyze the contents of digital collections at previously unimaginable scale, and often in previously inconceivable ways. The existence of such data is also especially valuable for the development of AI tools for smaller languages—a point revisted below.

Yet producing and providing access to humanities data is far from a trivial task. The library's collections have a particular history that has shaped their form in the archives, producing data artefacts that need to be managed. To take one key example, we can consider the effects of optical character recognition (OCR) software within the production process of digitized heritage material.[16] Since the particular terms of the Swedish legal deposit law have previously prioritized physical over digital examples, physical newspapers have been submitted to KB and then subsequently digitized. Beyond certain OCR errors with specific Swedish words, an effect of this digitization is the loss of various aspects of metadata that we as humans take for granted. A digital copy is gained from this process, but what is left is a mishmash of text blocks with no clear indication of which blocks belong together, which articles comprise part of the same section, and which texts are editorial content rather than adverts. It is certainly possible to use machine learning to attempt to put Humpty together again and reconstruct the newspaper, but this is a complex and laborious undertaking.[17] Making collections amenable to computational analysis is a qualified task that often demands considerable labor in terms of data cleaning and curation; humanities data is far from ready-made.[18]

It is within this particular context that KBLab came into being. On the one hand, there has been growing demand from scholars within the humanities and social sciences using digital approaches, who wish to be able to access digital collections to conduct large-scale analysis. As the pilot study that laid the ground for the founding of the lab suggested, "researchers, funding agencies and governmental research propositions are also increasingly pushing scholarship in a data intensive direction in order to promote digital scholarship."[19] On the other hand, there is the technical complexity involved in creating an infrastructure capable of providing access to these collections as data, when little has previously existed. Despite any suggestion to the contrary, enabling the production of high-quality datasets fit for research can be a complicated and messy undertaking. We now turn to consider how we sought to address this challenge through the making of a lab at KB.

Designing Technical Infrastructure

When the library formally initiated the project to establish a data lab in 2019, two particular user groups—and purposes—were specified. Internally, KBLab was conceived of as a resource for method development and AI innovation at KB: a means of providing staff and leadership with improved knowledge about the potential for automating various library processes. Externally, the lab was intended to position itself among existing institutions and environments to become an established infrastructure for facilitating and supporting digital research. In the short term, it was to meet the needs of two major projects within the digital humanities and social sciences funded by the Swedish Research Council: "Welfare State Analytics. Text Mining and Modeling Swedish Politics, Media & Culture, 1945–1989" (based at Umeå University)[20] and "Mining for Meaning: The Dynamics of Public Discourse on Migration" (based at Linköping University).[21] That both of these wished to conduct large-scale analysis of mid to later twentieth-century material from KB's collections—principally newspapers, but also

fiction and periodicals—had an important effect on the technical and organizational development of the lab. Since the projects were concerned with the analysis of material still protected by copyright, our initial task was to design a computing infrastructure to provide local access within the library itself.

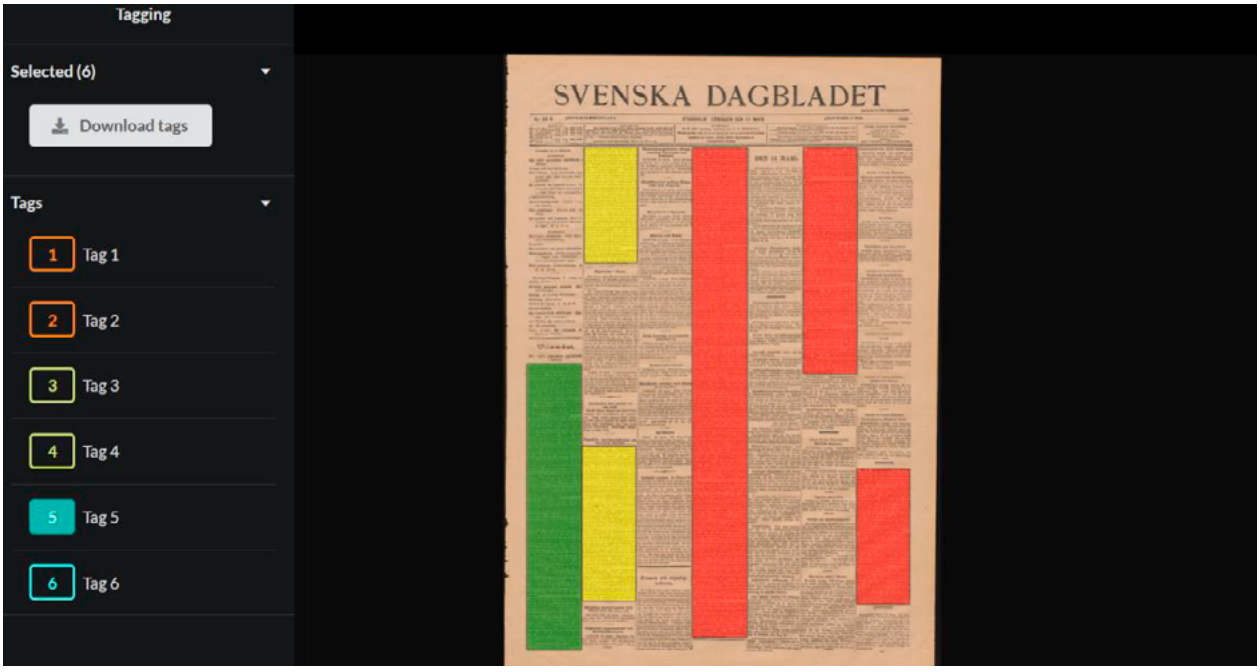
A key starting point in approaching this task was the conviction that a data lab at a library should offer experimental access to the collections. This required a laboratory space that enabled and encouraged researchers to explore. Of course, one way of making digital collections available for further research is to produce predetermined datasets that can then be released to be analyzed and used in diverse ways. However, and beyond the fact that copyright restrictions prevented such an approach in this instance, providing already defined datasets tends towards restricting rather than supporting open-ended critical investigation. Considering the research process as something not necessarily linear—indeed, as often tangential and shaped by serendipitous findings far beyond the initial remit of enquiry[22]— we opted to design an infrastructure where an exploratory working method would be possible, if not inevitable, for the researchers who come to use it. Once a research project is onboarded at the lab, it is granted *unlimited* access to KB's digital collections so researchers can explore and design their own datasets as a result of contact with these collections.

The technical challenge this involved was in how to enable exploration without compromising security. Our solution was to offer indirect access though an Application Program Interface (API). Researchers can conduct searches of KB's digital collections via the lab's API that will give results in the form of a representation of the original data, rather than the unnecessary risk of exposing the library's databases through direct access to the files themselves.[23] These representations take the shape of JSON files, a data form reflecting the longer history of KB's engagement with linked data and one which is particularly apposite for a digital research infrastructure since it is machine-readable.[24] Another key aspect of the linked data model underpinning the lab's design is the presence of Uniform Resource Identifiers (URIs) in the lab environment. By providing stable and persistent URIs for the archival material, researchers are able to find their way back to the same point in the collections and ultimately to demonstrate that their results are reproducible. Through establishing an API and an information model that makes linked data of the digital archive, we could create programmatic access to the library's collections that offers researchers the autonomy to steer their own exploratory processes.

In addition to making the material searchable through the API, we also created a graphical user interface (GUI) for the lab environment (see figure 1). This serves to strengthen the lab's functionality as a research infrastructure in various, overlapping ways. Firstly, it provides a means of validation: in accessing the material in visual form, researchers can verify and navigate among their results. Secondly, it provides a way for scholars within the humanities and social sciences without programming skills to access and interact with the material in the lab environment. This can be particularly important for multi-disciplinary projects that seek to combine the perspectives of data science with more traditional forms of expertise in close reading and analysis within various humanities disciplines. It is also pertinent for mixed-methods analyses that, beyond conducting large-scale computational analysis, wish to incorporate investigation of visual aspects of the material (and therefore also need to be able to see the individual object rather than an aggregation of its textual contents). Thirdly, it allows for the annotation of the material, which can prove a significant element in projects that utilize machine learning by training models based on the collections. There is a specific

FIGURE 1
Visualizing a Lab Environment for Explorative Research, KBLab's GUI

FIGURE 2
Annotating and Extracting Research Data in the Lab Environment



function within the interface that allows users to annotate according to their own chosen labels and then extract the particular text that has been annotated (see figure 2).

We made the lab's GUI available to researchers outside the lab itself through a prototype service called betalab. On the one hand, this forms part of the onboarding process for research projects that have been granted access to use KBLab. Prior to gaining access to the physical premises of the lab, researchers can use betalab to test and accustom themselves to the lab environment—in certain cases, they can even design and prepare scripts to be run on-site once they have access (which can prove a significant time-saving approach for geographically

disparate projects). On the other hand, betalab is also used as an access point for those parts of the digital collections that are available at the lab but not subject to copyright restrictions. This open data includes historical newspaper material up to 1906, the Swedish Government Official Reports (SOU's) and various parliamentary data. If, for example, a research project wished to access newspaper material in machine-readable, structured form to conduct an investigation into nineteenth-century crime reporting, KBLab would provide them with access to this data via betalab. In this sense, it provides an important complement to the primary part of the lab that can be accessed on-site at KBLab's premises.

For research projects that need to use KBLab for large-scale computational analysis of those parts of the collections protected by copyright, we established a computer lab at KB's locale at Karlavägen in Stockholm. This physical manifestation of the lab as research infrastructure is significantly a matter of computing power: since the terms of Swedish copyright legislation mean it is not currently possible for these projects to move the data outside the lab for external processing, we needed to ensure that there were sufficient computational resources in-house to meet the researchers' needs. To this end, we built a local computing infrastructure with three levels: a) powerful workstations at the computer lab; b) a server environment for computation and access to the material via an API; and c) two NVidia DGX A100 servers for more computationally heavy analysis. (We have also since been granted access to the EU's supercomputing infrastructure for our own development work, a point we return to below.)

The guiding principles that shaped the technical work to establish this solution have been pragmatism, flexibility, and a desire to create autonomy for the researchers who use it. The workstations in the computer lab, for instance, use the Linux-based system Ubuntu, as this allows researchers to create and control their own software environments according to their particular needs and preferences. Likewise, to enable researchers to manage their own back-up for code and work-in-progress, we created gitlab, an internal, server-based git function. We elected to start by acquiring consumer rather than enterprise hardware for the lab: in part due to the (relatively) limited resources we had at our disposal, but also because it allowed us to move quickly and adapt according to the shifting needs of researchers as these emerged. The work involved in establishing this setup has depended upon KB's existing staff expertise within IT-architecture and systems design; without the input of an experienced and creative IT-architect, the making of KBLab would not have been possible.

Research Coordination

With the lab established as an entrance point to the collections, another important organizational matter to be dealt with was research coordination. A significant aspect of this involved determining the principles and procedures via which access to the computer lab should be granted. In particular, and given that demand to use KBLab among researchers has consistently been greater than our on-site capacity, how should places at the limited number of workstations be allocated?[25] To address this in a fair and transparent manner that aligns with KB's values and missions as a public authority, we made applying to the lab part of the library's broader process for managing research and development applications.[26] Researchers who are interested in collaborating with KBLab therefore begin by submitting a brief project outline describing what they would like to do in their proposed research. This application is then subjected to an initial screening to confirm that the project actually involves research elements—i.e. that there are questions and hypotheses amenable to further exploration—before any decision is

made about the specific terms of collaboration that might be possible. Important to note in this context is that, apart from confirming the presence of a research question and determining its essential feasibility, we make no judgement upon the substantive content of the research proposal.

The question of sustainable funding is central to the existence of any lab, and this also impacts how new applications to KBLab are handled.[27] While the initial outlay for the lab was financed through a combination of internal funding from the library and external funding from the projects mentioned above, our working assumption is that research projects based at the lab should be self-financing—i.e. that they pay an overhead fee to cover the running costs (technical and administrative) in utilizing the lab, in line with a general Swedish praxis for the use of research infrastructure. Given the configuration of funding for academic research in Sweden, this means researchers have to include a budget post for use of the lab in their applications to research funding organizations such as the Swedish Research Council (VR) and Riksbankens Jubileumsfond (RJ). It also means researchers need to coordinate applying for a place at the lab with the process of submitting a funding application to these organizations.

The advantage of this approach is that it serves as a mechanism for quality control: by granting access to projects that have been awarded funding following a competitive, peer-reviewed process, we can ensure that the research allocated a place at the lab is of the highest caliber. However, a potential disadvantage is that it can favor larger projects proposed by established researchers at the expense of smaller initiatives by less well-established scholars. To counter this, a pragmatic cost-benefit analysis is adopted when considering each potential project, which can allow the overhead fee for use of the lab to be waived in certain cases. For example, if a project involves significant infrastructural gains for the library beyond the particulars of the project itself, then such a solution might be possible. A typical instance where infrastructural positives outweigh any overhead costs is the various Masters projects in machine learning that have been hosted at the lab, which have explored how AI models can help make the library's collections more accessible.[28]

A further dimension that affects how the overhead costs for a potential project at the lab are assessed is the level of data science competence in the project team in relation to the complexity of the proposed research. The underlying issue here is finding productive forms of collaboration between expertise in AI and machine learning, on the one hand, and more traditional qualitative competences in the humanities and social sciences, on the other.[29] Based on our experience, outsourcing the requisite expertise for large-scale data analysis to technical staff outside the project is the *least* effective way of dealing with this question. Such an approach tends to be problematic, partly since it risks making vital technical labor invisible and uncredited, and partly because it lends itself to a situation where researchers in the humanities and social sciences are publishing work where they do not properly understand either the methods used or their subsequent results.

On this basis, we recommend that projects based at KBLab incorporate data science competence within their project team, so that this perspective is represented and accountable at all stages of the research process. In practice, this means we are reluctant to grant lab space to proposals lacking the necessary technical skills, instead referring these to other infrastructural organizations such as the various Swedish centers for digital humanities who can provide greater levels of support. To proposals that have included the necessary expertise, we offer an overhead fee that is adjusted according to the technical complexity and demands of the

specific project: ranging from a standard rate that includes initial support and advice in using the lab, to higher levels when a greater degree of development work is required from the lab's staff to make the project possible. In each case, ascertaining the particular needs and requirements of a proposal presumes an ongoing dialogue with the researcher and deliberation from several of the lab's staff.

Once a project has been offered a place at the lab and received research funding, it is ready for onboarding. This process was designed in accordance with the particular model of explorative research practices mentioned above in the discussion of the lab's technical setup. A thorough introductory phase clarifies the formal terms for using the lab, where researchers sign a personal user agreement stipulating the legal conditions for accessing and using the data available at KBLab, as well as receiving a copy of the code of conduct (see appendix 1). This is followed by a hands-on guide where the researcher(s) will be shown how to access data via the lab's API, how to manage ongoing results and which among the lab's various tools might be of interest. After this introduction, researchers are ready to work autonomously at the lab: beyond consulting with lab staff in the event of problems, they are free to begin interacting with the collections at KBLab according to their particular interests.

Collections-based Models as Digital Infrastructure

Having discussed the making of the lab as a physical site for researchers to access the collections, we now turn to discuss how we have harnessed the collections as the basis for new digital tools that in themselves constitute a significant form of research infrastructure. Whereas the number of researchers who can use the on-site lab is necessarily limited by practical constraints, the creation of such tools that can be distributed beyond the library has enabled the lab to have a far greater reach. In the remainder of this article, we outline our work in producing and releasing collections-based models at KBLab: how have KB's digital collections enabled the emergence of a library-based form of AI development?

Library Collections as Training Data

The past five years have witnessed rapid rates of development within the field of AI and machine learning. For instance, the release of transformer-based language models like BERT has proved the basis for unprecedented performance in many natural language processing tasks.[30] However, the emergence of such AI tools has occurred according to existing global hierarchies of power and resources: they are far from being equally available to all languages and actors. While Google AI developed dedicated BERT models with cutting-edge capabilities for major languages like English and Chinese, other languages had to make do with a less powerful multilingual model. Where big tech companies lacked the commercial interest to train these tools for particular languages, actors within the academy and beyond have tended to take the initiative to produce state-of-the-art monolingual models.[31] For so-called lesser-resourced languages, a significant bottleneck to doing so was the availability of sufficient computational resources and training data. In the instance of Sweden, the first monolingual BERT model was created by the Public Employment Agency using solely data from Swedish Wikipedia which, while better than Google's multilingual model, was still considerably less effective than the English BERT.[32]

Yet the prevailing paradigm for producing state-of-the-art AI models enables national libraries and other heritage institutions to contribute to development in novel ways, especially

in the case of smaller languages. With the emphasis on unsupervised learning in current AI development—i.e. when vast algorithms called artificial neural networks learn through being exposed to huge volumes of unlabeled training data, rather than, as previously, from smaller amounts of (expensively) annotated data[33]—new opportunities have emerged for the custodians of large amounts of high quality, language-specific data.

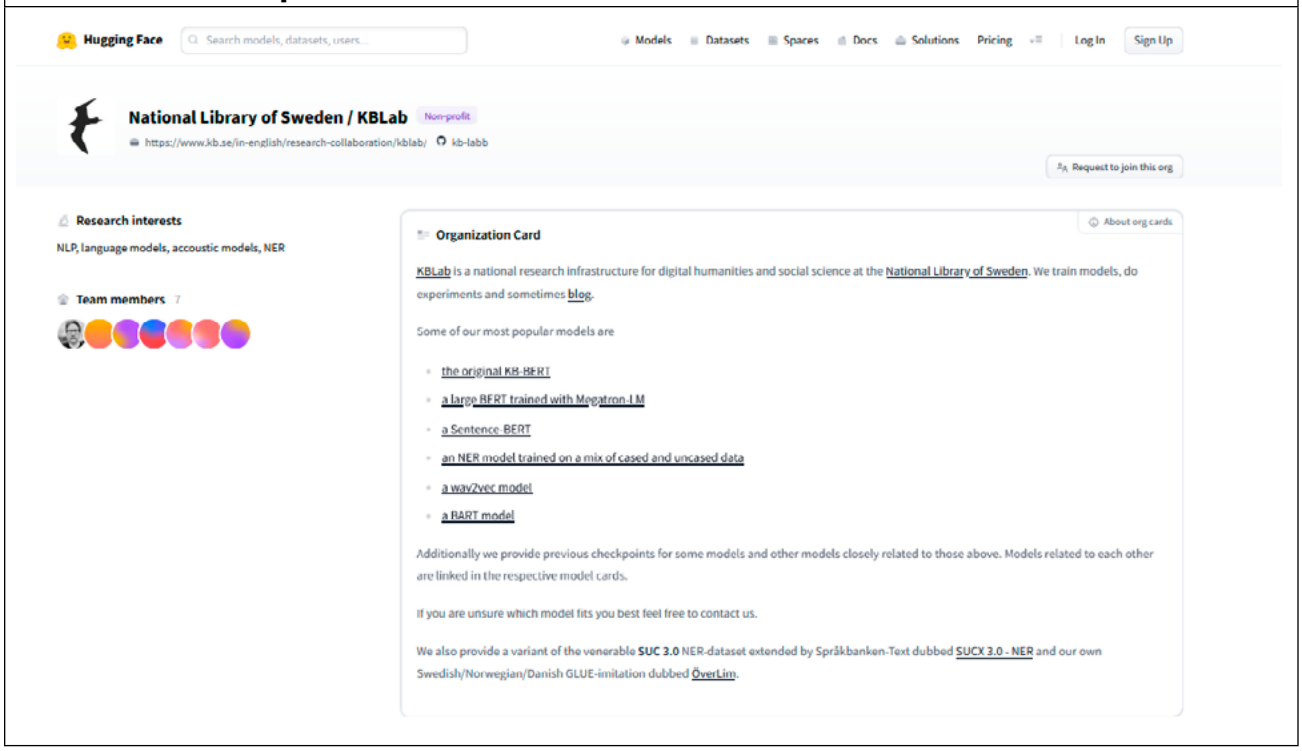
In such a context, the breadth and depth of KB's collections mentioned above becomes a uniquely valuable resource for the making of cutting-edge tools for Swedish AI. Indeed, the fact that legal deposit provides KB with something approaching population data for the language means there is an important *democratic* dimension to harnessing the library's collections as training data. With recourse to a broader and more representative range of data than that available to other actors (who have access chiefly to Swedish data that can be scraped from the web), KB has the potential to produce AI models of greater quality and effect. Given that this data cannot be shared beyond the library due to copyright and GDPR legislation, this creates a powerful rationale for the training of models in-house at KBLab.[34]

Making and Distributing Collections-based AI

Against this backdrop of enhancing the quality of Swedish AI infrastructure towards global state-of-the-art, we have been using the library's digital collections to train new AI tools since the lab was established in 2019. The first phase of this development work focused specifically upon text, with the aim of improving the capabilities available for automated analysis of Swedish text in light of recent innovations with transformer models. Here we turned to the breadth and depth of KB's collections to train a BERT model for Swedish capable of processing "the living language of the national community."[35] To create such data representativity, we produced a large and diverse training corpus that made substantial use of the library's digitized newspaper archives dating back to 1945, as well as more recent online material and social media to capture more colloquial language use. Making this material amenable to machine learning so it could be used as training data also involved painstaking and laborious processes of data curation, which in turn depended upon the specialized competence in data science and programming of the lab's staff. The language model that this enabled, KB-BERT, proved significantly more effective than existing models and has since become the standard model to use for Swedish language processing.[36]

In line with the increasingly multi-modal direction of current AI innovation and the multimedia inclinations of recent humanities scholarship, our development efforts at KBLab have also moved beyond solely text. Here we have been able to take advantage of the diversity of media forms stored in the archive: ranging across a variety of different modes, KB is guardian of unparalleled collections of Swedish text, images, sound, and film, which equates to a considerable range of possibilities for training new models. A pertinent example is the work at the lab in producing improved tools for automated sound recognition (ASR). This involved using the library's enormous, and often largely unexplored, holdings of audio-visual material from the twentieth century. More specifically, we utilized KB's digitized national and local radio programs from the past two decades to produce a corpus of over 1.4 million hours of spoken Swedish, including dialects from all the regions in the country.[37] This was then used as training data for Swedish versions of the wav2vec 2.0 model developed by Facebook (now Meta) AI.[38] As was the case with KB-BERT, the collections-based models that this produced, entitled VoxRex, outperformed existing multilingual and monolingual models for speech-to-

FIGURE 3
Open Access Swedish AI Models Available for Download



text tasks.[39] As we explain below, the existence of cutting-edge tools for Swedish speech-to-text creates a range of synergy effects, both within and beyond heritage institutions.

To ensure the AI tools produced at KBLab can benefit as many as possible, we release our models open access via the data science community platform Hugging Face (see figure 3).[40] To date, we have made available 46 models in this way so that users are able to download and experiment according to their particular interests. In addition to the KB-BERT and VoxRex models mentioned above, these include a canonical Swedish SpaCy, a zero-shot classifier, Sentence-BERT and BERT models fine-tuned for named-entity recognition (NER), as well as Swedish versions of the latest Whisper models for ASR released by OpenAI.[41] Beyond training our own models for Swedish text and sound, we have also collaborated with other actors in developing multimodal tools that connect image and text to enable new forms of image search.[42] As part of the transparent and accountable approach to AI development that we pursue at the lab, we make sure to document the data we have used to train our models through a combination of Hugging Face descriptions, blog posts and research articles.[43] We also share our code via GitHub.[44] In this way, we seek to make it possible for users of the lab's models to understand how these tools have been made, and to consider how, in light of the particular values and emphases contained within the data in KB's collections, the models might need to be adjusted for use in specific applications.[45]

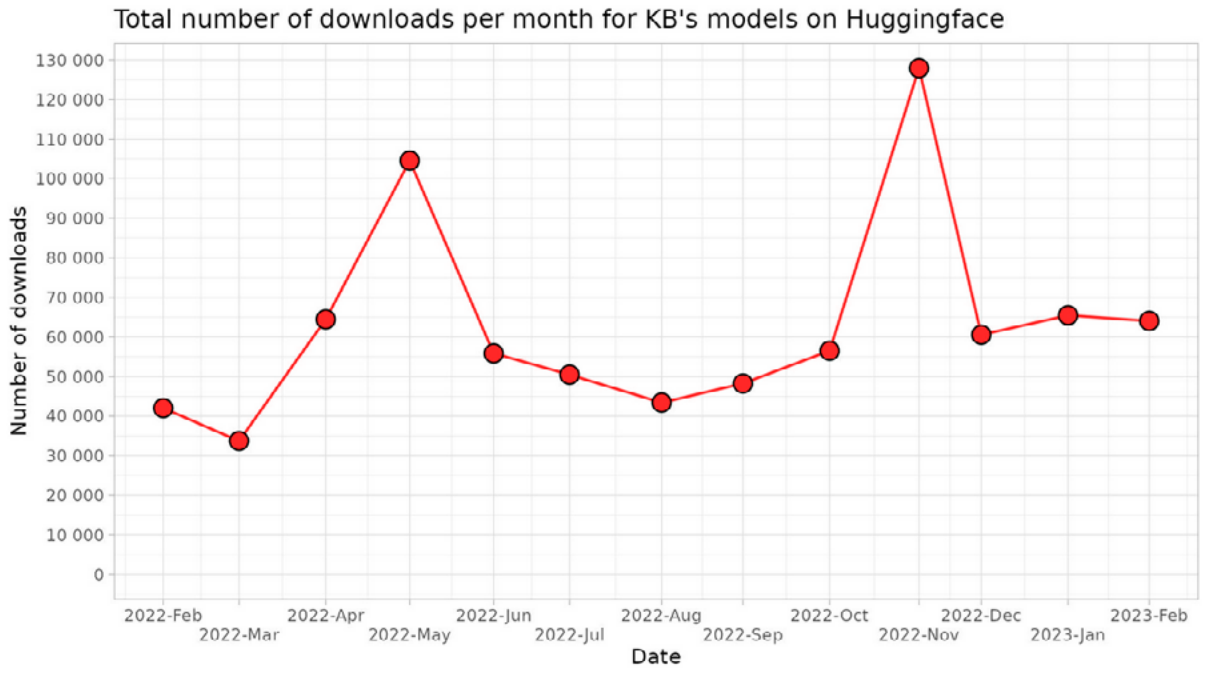
The Value of Collections-based Models in Practice

Turning to tracing the value of the models trained at the lab, these are being put to use in a diverse range of contexts. The initial impetus towards producing these models was as a means for making the library's huge, but often largely uncharted holdings of digital material more accessible; through creating such tools we hoped to help the library better understand and describe

its own collections, while simultaneously improving research access to this material. That this has come to fruition is demonstrated, for instance, in the various Master's projects at KBLab that have shown how KB-BERT could be used for the automated enrichment of metadata in the digital newspaper archive.[46] It is also evident in a pilot project exploring how a topic-modelling approach built upon our Swedish Sentence-BERT, BERTopic, might be used to provide a sort of automated subject headings that offer more fine-grained navigation of the collections.[47] Perhaps most striking, though, is the positive feedback loop created by the lab's development work with sound data mentioned above: first, KB's collections enabled the production of state-of-the-art ASR models; these models can then be used for speech-to-text to make these collections amenable to text searching; and text transcriptions of the radio and television material can be used as new training data for yet another generation of new and better models at the lab. AI development and improved metadata thus work hand-in-hand to improve the accessibility of the material, thereby enhancing the library's function as a research infrastructure.

Beyond the library and research projects based at the lab, KBLab's models have proved valuable for both a surprising array of academic research and for information-intensive organizations outside the academy, in public and private sectors alike. In terms of the former, KB-BERT has now been utilized by medical researchers seeking to develop new lifestyle treatments for diabetes patients; in attempts to automatically identify the presence of implants (i.e. pacemakers or stents) in heart patients prior to MRI scans; and for the classification of legal documents.[48] In terms of the latter, the lab's models have been put to work in automating and streamlining the information handling processes of various public authorities, including local councils, the Swedish Tax Agency (*Skatteverket*), the Swedish Courts (*Domstolsverket*) and most recently, the support function of State administration (*Statens servicecenter*).[49] As a growing number of Swedish organizations and companies start to become aware of the pos-

FIGURE 4
Download Statistics for KBLab's Open Access AI Models for February 2022–February 2023



sibilities presented by AI, they are increasingly turning to the lab's models for easy to access and state-of-the-art performance.[50]

However, the most striking evidence of the scale of the impact of the lab's collections-based models is quantitative in nature, with statistics showing over a million downloads since they were made available on our Hugging Face page (see figure 4).[51] Of course, these figures need to be contextualized: this does not refer to the number of discrete users, but rather the total number of times the models have been used (within, say, a particular application). While there is no way of knowing further details about such usage—beyond what can be traced from citation of the lab's publications, and cases where specific developers have contacted the lab—these statistics can still be taken as a forceful demonstration of the reach of our work in producing new AI tools. As both of the independent evaluations of the lab's first two years have highlighted, the widespread uptake of models trained at the library indicates a pertinent way in which the lab's development work reinforces and furthers KB's democratic commitments.[53] In making and releasing AI models using the library's data, KBLab thus offers a powerful new way of sharing the value of the collections outside the library itself.

Library-based AI Development as a Public Good

The principal merit of this type of AI development is the way it can simultaneously enhance the library's functions as national research infrastructure and guarantor of democratic values.[54] As a form of digital infrastructure, open access collection-based models enable a wide range of AI applications for Swedish, which would have proved difficult, if not impossible, without such infrastructural tools in place.[55] As a means of encouraging and enabling democratic social development, there are distinctive, yet mutually reinforcing aspects, of these models that are worth accentuating.

That we are facilitating an expansion of Swedish AI implementation by making the models freely available is intensified by the particular logic of the models' architecture. More specifically, the relative resource allocation between the pre-training and fine-tuning phases of a Transformer model lends itself to effective dissemination: while pre-training for a general purpose model like KB-BERT is computationally expensive and presumes large amounts of data, subsequent fine-tuning can be carried out with but a fraction of the data and computational resources.[56] This means that a far wider range of social actors outside resource-intensive environments like the university can consider downloading the models, experimenting with them, and applying them to their own particular use cases. Our work at the lab is thereby contributing to a democratizing of both the technology and the library's data.[57] In this sense, KBLab is helping to share the newly-found value of the collections as data—understood as a form of publicly funded and maintained *commons*[58]—to new groups of users than those traditionally reached by the library and beyond.

This form of library-based development can even counter some of the more problematic aspects of an AI future driven purely by private sector actors, particularly the growing deficit of data accountability. As emergent AI technologies become more mature and increasingly governed by commercial concerns, there has been a concurrent move towards treating training data and methods as trade secrets to be protected from competitors. This was exemplified in recent discussions about the lack of transparency surrounding OpenAI's release of GPT-4, with one scientific researcher, Sasha Luccioni, suggesting "it's just completely impossible to do science with a model like this," given the lack of access to details about the data used to

make it.[59] Such opaque practices can also be connected to a wider culture of silence in the tech industry that precludes critical voices about, for example, the overreliance on vast, unaccounted-for web materials in training new models, as the case of Timnit Gebru amply illustrates.[60] By contrast, in adhering to careful practices of documentation, scrutinizing the workings of data representativity, and pursuing more representative models based on the breadth of the library's collections, we are engaged in an accountable form of AI work at the lab that can variously complement and challenge that of private tech companies.[61] Insofar as it is more open and operates according to other imperatives than commercial interest, library-based AI development can constitute one of the much-needed "alternatives to the hugely concentrated power of a few large tech companies and the elite universities closely intertwined with them."[62]

Yet using publicly-funded heritage data as the basis for a more ethical AI development is dependent upon sourcing new forms of resources. Although it was possible to produce cutting-edge tools at the lab when these were of a proportion of a BERT model, the pace of recent AI innovation has led to new models at a scale that makes this far more challenging. To give a sense of the leap in scale: where BERT had hundreds of millions of parameters, GPT-3 has over 175 billion and GPT-4 is suspected to have far more—though, of course, this latter figure remains shrouded in secrecy as yet. While we have the prerequisite training data and specialized expertise in data science to produce larger models, a significant bottleneck has been in locating sufficient computational resources. To solve this, and to be able to further our work in producing state-of-the-art models for Swedish, we sought the help of ENCCS (EuroCC National Competence Centre Sweden) to apply to use the EU's infrastructure for supercomputers, EuroHPC.[63] Gaining access to first HPC Vega (in Slovenia, with 240 GPUs) and now HPC Meluxina (in Luxembourg, with 800 GPUs), has enabled development work of a different scale at the lab.[64] In becoming the first public authority to use these EU-funded development resources, KBLab is furthering the prospect of a Swedish AI that is open, accountable, and democratically inclined.

Finally, contributing to the making of a national AI infrastructure in this way also demands novel collaboration. With the release of KB-BERT establishing the lab as a key player in Swedish language technology, we have since become involved in national and international networks that include a diversity of actors who are engaged in AI questions: from researchers and university departments, to coordinating organizations, public authorities and private companies. Forming new relationships and collaborating with this configuration of groups beyond those that the library has traditionally cooperated with is an important step in trying to work effectively in the rapidly evolving space of AI development. A recent example was the lab's role in a project, together with the National Language Bank of Sweden at Gothenburg University, the Swedish Research Institute (RISE), and AI Sweden, to provide a set of benchmarks for evaluating Swedish language models.[65] By working together to make it easier for users of Swedish AI to determine which models might best fit their purpose, we are helping to make recent innovations more widely accessible. In this way, the lab's research collaboration with external actors is also leading to improved infrastructure.

Conclusion

The establishment of a data lab at the National Library of Sweden has enhanced the library as a digital research infrastructure. As we have explained, various practical and technical considerations shaped the making of KBLab as a physical site where researchers can now access

the collections at unprecedented scales. The library's digital collections have enabled the lab to play an important role in contributing to the development of a national AI infrastructure for the Swedish language. As a closing note, we offer some reflections on the possibilities and challenges facing the lab as a node for library-based AI development.

One of the key justifications for library labs in particular, and GLAM labs more generally, is that they provide new ways of sharing the value of cultural heritage material. Establishing such a lab can lead to snowball effects with various positive, if often unforeseen, consequences. In particular, the consolidation of in-house expertise within data science and machine learning opens up significant possibilities for heritage institutions that are increasingly becoming custodians of large volumes of digital material. Through working in tandem with domain specialists (i.e. librarians, archivists, curators, etc.), such labs can make these collections available at scale to researchers and other users so they can pursue new lines of inquiry. Adopting a collections as data approach also creates significant opportunities to contribute to AI development, especially for lesser-resourced languages that have not been prioritized by major commercial actors. By using high quality, language-specific heritage data to contribute to national infrastructure, and engaging in novel collaboration with external actors, these labs can play a role in democratizing this data, while laying claim to a powerful new form of social relevance in the process. In short, GLAM labs create new and unexpected lives for collections far beyond the heritage sector itself.

Conversely, while it might seem a platitude, it is far easier to start a lab, with all the start-up energy and buzz this entails, than it is to entrench one as a given part of a heritage organization. In part, this is about the thorny question of funding and a systematic tendency to underinvest in digital research infrastructure.[66] But it is also connected more specifically to the difficulty of attracting and retaining highly-qualified staff within publicly-financed AI development, when the demand for this data science expertise in the private sector is intensifying. There are even complexities to be addressed concerning how this expertise should be integrated within the wider organization: should data scientists be centralized within a lab, as is the case with KBLab, or are there arguments for distributing this competence across the organization as a whole? How might fruitful interactions between data scientists and domain experts best be encouraged?[67] In dealing with such questions and seeking to navigate a way forwards through the rapidly shifting terrains of digitalization and AI innovation, there is a compelling need for strategic leadership and direction.

Based on our conviction that the future interactions of AI and the library can be mutually beneficial, we conclude by offering some concrete pointers to any research library considering establishing a data lab or investing in digital transformation connected to AI. The first pointer underlines the importance of *people* and the centrality of an open, interdisciplinary outlook among the project team. Whether recruited internally or externally, the library professionals engaged in this work should be both driven by infrastructural questions and curious about other perspectives, be this data scientists interested in the design and use of research infrastructure or humanities researchers interested in collections as data. The second pointer is about the provision of a realistic time frame to give the project space to experiment before being expected to deliver. This is partly a matter of financial support and ensuring there is sufficient continuity beyond the short-term possibilities of external research grants, but it is also connected to the question of timing for strategic evaluation. (In the case of KBLab, for instance, the external assessment reviewing the lab's establishment took place after two years

before the decision to make the lab a permanent part of the library was taken.) The third pointer is to allow for specialized legal support to assist the project in navigating relevant national and international legislation about data use and sharing. The final pointer is the importance of continued dialogue with various stakeholders, within and beyond the library, concerning the possibilities and risks with ongoing AI development. Working responsibly and transparently to mitigate these risks, library-based AI can be a synergizing venture that significantly enhances the availability, usability, and value of heritage collections for present and future users.

Appendix 1. KBLab's Code of Conduct

KBLab is an open and friendly working environment, where collaboration is encouraged, questions are welcomed, and the commitment to critical, open-ended and independent enquiry is foundational. This openness is essential for new research ideas and projects to flourish.

A prerequisite for such an environment is being kind to one another. To maintain this space for open enquiry, we expect all people connected to KBLab to behave according to the principles of mutual respect and decency. Not honouring these principles can lead to access to KBLab being withdrawn.

If you have any questions about this code, please contact us at kblabb@kb.se.

Notes

1. M.B. Hoy, "Big Data: An Introduction for Librarians," *Medical Reference Services Quarterly* 33:3 (2014): 320-326, https://doi.org/10.1080/02763869.2014.925709.

2. K. Cukier & V. Mayer-Schoenberger, "The Rise of Big Data: How It's Changing the Way We Think About the World," *Foreign Affairs* 92:3 (2013): 28-40, http://www.jstor.org/stable/23526834.

3. N.J. Bingham & H. Byrne, "Archival Strategies for Contemporary Collecting in a World of Big Data: Challenges and Opportunities with Curating the UK Web Archive," *Big Data & Society* (2021), https://doi.org/10.1177/2053951721990409.

4. S. Ames & S. Lewis, "Disrupting the Library: Digital Scholarship and Big Data at the National Library of Scotland," *Big Data & Society* (2020), https://doi.org/10.1177/2053951720970576.

5. For example, T. Underwood, *Distant Horizons: Digital Evidence and Literary Change* (Chicago: The University of Chicago Press, 2019).

6. For the notion of "socio-technical imaginaries," see S. Jasanoff & S. Kim, "Containing the Atom: Socio-technical Imaginaries and Nuclear Power in the United States and South Korea," *Minerva* 47:2 (2009): 119-146, http://www.jstor.org/stable/41821489. For "collections as data," see, for example, T. Padilla, "Humanities Data in the Library: Integrity, Form, Access," *D-Lib Magazine* (2016), https://www.dlib.org/dlib/march16/padilla/03padilla.html [accessed March 8, 2023].

7. A. Wheatley & S. Hervieux, "Artificial Intelligence in Academic Libraries: An Environmental Scan," *Information Services & Use* 39 (2019): 348, https://doi.org/10.3233/ISU-190065.

8. R. Cordell, "Machine Learning + Libraries: A Report on the State of the Field," LC Labs, Library of Congress (14 July 2020), https://labs.loc.gov/static/labs/work/reports/Cordell-LOC-ML-report.pdf [accessed March 8, 2023]; T. Padilla, *Responsible Operations: Data Science, Machine Learning, and AI in Libraries* (Dublin, OH: OCLC Research, 2019), https://doi.org/10.25333/xk7z-9g97 [accessed March 8, 2023]; T. Padilla, L. Allen, H. Frost, S. Potvin, E. Russey Roke & S. Varner, "Final Report - Always Already Computational: Collections as Data," *Zenodo* (2019), https://doi.org/10.5281/zenodo.3152935.

9. A.M. Cox, S. Pinfield & S. Rutter, "The Intelligent Library: Thought Leaders' Views on the Likely Impact of Artificial Intelligence on Academic Libraries," *Library Hi Tech* 37:3 (2019): 418-435, https://doi.org/10.1108/LHT-08-2018-0105; D. Harisanty, N.E.V. Anna, T.E. Putri, A.A. Firdaus & N.A. Noor Azizi, "Leaders, Practitioners

and Scientists' Awareness of Artificial Intelligence in Libraries: A Pilot Study," *Library Hi Tech* pre-print (2022), https://doi.org/10.1108/LHT-10-2021-0356.

10. Ames & Lewis, "Disrupting the Library"; C. Haffenden, E. Fano, M. Malmsten & L. Börjeson, "Making and Using AI in the Library: Creating a BERT Model at the National Library of Sweden," *College & Research Libraries* 84:1 (2023): 30-48, https://doi.org/10.5860/crl.84.1.30.

11. The acronym GLAM stands for Galleries, Libraries, Archives and Museums. M. Mahey, A. Al-Abdulla, S. Ames, P. Bray, G. Candela, S. Chambers, C. Derven, M. Dobreva-McPherson, K. Gasser, S. Karner, K. Kokegei, D. Laursen, A. Potter, A. Straube, S-C. Wagner & L. Wilms with forewords by: T. A. Al-Emadi, J. Broady-Preston, P. Landry & G. Papaioannou, *Open a GLAM Lab. Digital Cultural Heritage Innovation Labs*, (23-27 September, 2019), https://glam-labs.s3.amazonaws.com/media/dd/documents/Open_a_GLAM_Lab-10-screen.9c4c9c7.pdf [accessed March 8, 2023].

12. Haffenden, et al., "Making and Using AI in the Library," 45.

13. The text of this legal act is available (in Swedish) here: "Förordning (2008:1421) med instruktion för Kungl. Biblioteket" (2008), https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/forordning-20081421-med-instruktion-for-kungl_sfs-2008-1421 [accessed March 8, 2023].

14. G. Konstenius, "Plikten under lupp! En studie av pliktlagstiftningens roll, utformning och relevans i förhållande till medielandskapets utveckling" [Eng. Legal Deposit in Focus! A Study of the Role, Design and Relevance of Legal Deposit Legislation in Relation to the Development of the Media Landscape], *Kungl. Biblioteket* (2017),
https://urn.kb.se/resolve?urn=urn:nbn:se:kb:publ-539 [accessed March 8, 2023].

15. See Padilla, *Responsible Operations* and Padilla et al., "Final Report".

16. M.C. Traub, J. van Ossenbruggen & L. Hardman, "Impact Analysis of OCR Quality on Research Tasks in Digital Archives" in *Research and Advanced Technology for Digital Libraries*, ed. S. Kapidakis, C. Mazurek, C., and M. Werla (Springer, Cham, 2015),
https://doi.org/10.1007/978-3-319-24592-8_19 [accessed March 8, 2023].

17. See F. Rekathati, "A Multimodal Approach to Advertisement Classification in Digitized Newspapers," *The KBLab Blog* (2021), https://kb-labb.github.io/posts/2021-03-28-ad-classification/ [accessed March 8, 2023].

18. Cf. L. Gitelman (ed), *"Raw Data" is an Oxymoron* (Cambridge, Mass.: The MIT Press, 2013), https://doi.org/10.7551/mitpress/9302.001.0001.

19. P. Snickars, "Datalabb på KB: En förstudie" [Eng. Data Lab at KB: A Pre-study], *Kungl. Biblioteket* (2018), 29, https://urn.kb.se/resolve?urn=urn:nbn:se:kb:publ-339 [accessed March 8, 2023].

20. For more details, see: https://www.westac.se/en/.

21. For more on which, see: https://liu.se/en/research/computational-text-analysis.

22. Cf. J. Law, *After Method: Mess in Social Science Research* (London: Routledge, 2004).

23. A description of the lab's API is available here: M. Malmsten, "KB Data Lab," *GitHub* (2020), https://github.com/Kungbib/kblab [accessed March 8, 2023].

24. For KB's longer engagement with linked data, see, for instance, M. Malmsten, "Exposing Library Data as Linked Data," *IFLA Satellite Preconference Sponsored by the Information Technology Section* (2009),
https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=2e0791d88a65cb2517e284c2bfca02b7c6660f30 [accessed March 8, 2023].

25. For further details about demand for KBLab among researchers, see M. Fridlund, "Utvärdering av KB-labb" [Eng. Evaluation of KBLab], Gothenburg University, Centre for Digital Humanities (September 2021), 9, https://urn.kb.se/resolve?urn=urn:nbn:se:kb:publ-97 [accessed March 8, 2023].

26. For more details about this process, see: https://kb.se/in-english/research-collaboration/criteria-for-collaboration.html.

27. For further discussion of funding a GLAM lab, see Mahey et al., *Open a GLAM Lab*, 29.

28. Examples of these projects include F. Rekathati, "Curating News Sections in a Historical Swedish News Corpus," Independent Master's Thesis, Linköping University, Department of Computer and Information Science (2020), http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-166313 and G. Henning, "News Article Segmentation Using Multimodal Input: Using Mask R-Cnn and Sentence Transformers," Independent Master's Thesis, KTH, School of Electrical Engineering and Computer Science (2022), http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-309527.

29. Further consideration of these issues can be found in M. Kemman, *Trading Zones of Digital History* (Berlin: De Gruyter Oldenbourg, 2021), https://doi.org/10.1515/9783110682106 and E. Fano & C. Haffenden, "Digital humaniora eller humanistisk datavetenskap?" [Eng. Digital Humanities or Humanistic Computer Science?] *Samlingsbloggen* (2022), https://www.kb.se/hitta-och-bestall/samlingsbloggen/blogginlagg/2022-04-14-digital-humaniora-eller-humanistisk-datavetenskap.html [accessed March 8, 2023].

30. J. Devlin, M.-W. Chang, K. Lee & K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv* (2019), https://arxiv.org/abs/1810.04805v2.

31. For instance, A. Virtanen, J. Kanerva, R. Ilo, J. Luoma, J. Luotolahti, T. Salakoski, F. Ginter & S. Pyysalo, "Multilingual Is Not Enough: BERT for Finnish," *arXiv* (2019), https://arxiv.org/abs/1912.07076; L. Martin, B. Muller, P.J.O. Súarez, Y. Dupont, L. Romary, E.V. de la Clergerie, D. Seddah & B. Sagot, "CamemBERT: a Tasty French Language Model," *arXiv* (2020), https://arxiv.org/abs/1911.03894; and P.E. Kummervold, J. De la Rosa, F. Wetjen & S.A. Brygfjeld, "Operationalizing a National Digital Library: The Case for a Norwegian Transformer Model," *arXiv* (2021), https://arxiv.org/abs/2104.09617.

32. Haffenden, et al., "Making and Using AI in the Library," 35-36.

33. Cf. A. Radford, K. Narasimhan, T. Salimans & I. Sutskever, "Improving Language Understanding by Generative Pre-Training," [pre-print] (2018), https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf [accessed March 8, 2023].

34. Haffenden, et al., "Making and Using AI in the Library," 35.

35. M. Malmsten, L. Börjeson & C. Haffenden, "Playing with Words at the National Library of Sweden: Making a Swedish BERT," *arXiv* (2020), https://arxiv.org/abs/2007.01658.

36. This model is available here: https://huggingface.co/KBLab/bert-base-swedish-cased.

37. M. Malmsten, C. Haffenden & L. Börjeson, "Hearing Voices at the National Library: A Speech Corpus and Acoustic Model for the Swedish Language," *arXiv* (2022), https://arxiv.org/abs/2007.01658.

38. A. Baevski, H. Zhou, A. Mohamed & M. Auli, "Wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations," *arXiv* (2020), https://arxiv.org/abs/2006.11477.

39. Malmsten et al., "Hearing Voices at the National Library." This model is available here: https://huggingface.co/KBLab/wav2vec2-large-voxrex-swedish.

40. This has been the policy for distribution for all of the models released so far, and will continue to remain so for predictive models. The situation with generative models is more complex, since there is a risk for such tools to be used in ways that counter the democratic goals of the library's mission—i.e. in creating misinformation and fake news. On this basis, a more restrictive distribution policy will most likely be pursued in relation to generative models, should they be produced in the future.

41. See organization page here: https://huggingface.co/KBLab.

42. F. Carlsson, P. Eisen, F. Rekathati & M. Sahlgren, "Cross-lingual and Multilingual CLIP," *Proceedings of the Thirteenth Language Resources and Evaluation Conference* (2022), https://aclanthology.org/2022.lrec-1.739/ [accessed March 8, 2023].

43. See, for instance, F. Rekathati, "Swedish Sentence Transformer 2.0," *The KBLab Blog* (2023), https://kb-labb.github.io/posts/2023-01-16-sentence-transformer-20/ [accessed March 8, 2023].

44. See: https://github.com/kb-labb.

45. Fridlund, "Utvärdering av KB-labb," 14.

46. For instance, A. Ekmark, "Text Block Prediction and Article Reconstruction Using BERT," Independent Master's Thesis, Uppsala University, Department of Statistics (2021), http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-447248.

47. Cf. E. Fano & C. Haffenden, "BERTopic for Swedish: Topic Modeling Made Easier via KB-BERT," *The KBLab Blog* (2022), https://kb-labb.github.io/posts/2022-06-14-bertopic/ [accessed March 8, 2023].

48. C. Dwibedi, E. Mellergård, A.C. Gyllensten, K. Nilsson, A.S. Axelsson, M. Bäckman, M. Sahlgren, S.H. Friend, S. Persson, S. Franzén, B. Abrahamsson, K. Steen Carlsson & A.H. Rosengren, "Effect of Self-Managed Lifestyle Treatment on Glycemic Control in Patients With Type 2 Diabetes," *NPJ Digital Medicine* 5:60 (2022), https://doi.org/10.1038/s41746-022-00606-9; O. Jerdhaf, M. Santini, M. Lundberg & A. Karlsson, "Implant Terms: Focused Terminology Extraction with Swedish BERT - Preliminary Results," *Eighth Swedish Language Technology Conference* (2020), http://urn.kb.se/resolve?urn=urn:nbn:se:ri:diva-52378; A. Avram, V. Pais & D Tufis, "PyEuroVoc: A Tool for Multilingual Legal Document Classification with EuroVoc Descriptors," *arXiv* (2021), https://arxiv.org/abs/2108.01139.

49. For the work of the Swedish Tax Agency and the Swedish Courts in using lab models, see M. Juhlin, "De samhällsekonomiska effekterna kopplat till Kungliga bibliotekets AI-baserade språkmodeller" [Eng. The Socioeconomic Effects of the National Library of Sweden's AI Language Models], *Policy Impact* report (2022), 36-41, https://urn.kb.se/resolve?urn=urn:nbn:se:kb:publ-692; for KB-BERT being used as the basis of a new search application for precedent in state bureaucracy, see: https://www.statenssc.se/nyheter/nyhetsarkiv/2023-03-14-ai-baserad-soktjanst-ska-underlatta-remisshanteringen-i-staten [accessed March 8, 2023].

50. Fridlund, "Utvärdering av KB-labb," 14.
51. These download statistics are also available here: https://github.com/kb-labb/huggingface_stats [accessed March 8, 2023].
52. Cf. Juhlin, "De samhällsekonomiska effekterna," 20-22.
53. Fridlund, "Utvärdering av KB-labb," 14; Juhlin, "De samhällsekonomiska effekterna".
54. "Förordning (2008:1421) med instruktion för Kungl. Biblioteket".
55. Cf. P.N. Edwards, G.C. Bowker, S.J. Jackson & R. Williams, "Introduction: An Agenda for Infrastructure Studies," *Journal of the Association for Information Systems* 10:5 (2009), https://aisel.aisnet.org/jais/vol10/iss5/6 [accessed March 8, 2023].
56. Haffenden, et al., "Making and Using AI in the Library," 33.
57. Fridlund, "Utvärdering av KB-labb," 13.
58. Cf. D. Harvey, "The Future of the Commons," *Radical History Review* 109 (2009): 101-107, https://doi.org/10.1215/01636545-2010-017.
59. K. Sanderson, "GPT-4 Is Here: What Scientists Think," *Nature* March 16 (2023), https://doi.org/10.1038/d41586-023-00816-5.
60. E.M. Bender, T. Gebru, A. McMillan-Major & S. Shmitchell, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (2021),
https://doi.org/10.1145/3442188.3445922; T. Gebru, "For Truly Ethical AI, Its Research Must Be Independent From Big Tech" *The Guardian*, December 6 (2021),
https://www.theguardian.com/commentisfree/2021/dec/06/google-silicon-valley-ai-timnit-gebru [accessed March 8, 2023].
61. Fridlund, "Utvärdering av KB-labb," 14.
62. Gebru, "For Truly Ethical AI".
63. For further details about the granting of such access to EuroHPC, see: https://www.eurocc-access.eu/success-stories/success-story-national-library-of-sweden-has-now-access-to-vega/ [accessed March 8, 2023] and https://enccs.se/news/2022/10/national-library-of-sweden-accesses-meluxina/ [accessed March 8, 2023].
64. For instance, R. Kurtz & J. Öhman, "SUCX 3.0," *The KBLab Blog* (2022), https://kb-labb.github.io/posts/2022-02-07-sucx3_ner/ [accessed March 8, 2023].
65. R. Kurtz, "Evaluating Swedish Language Models," *The KBLab Blog* (2022), https://kb-labb.github.io/posts/2022-03-16-evaluating-swedish-language-models/ [accessed March 8, 2023].
See also: https://www.ai.se/en/node/81535/superlim [accessed March 8, 2023].
66. R. Knowles, B.A. Mateen & Y. Yehudi, "We Need to Talk About the Lack of Investment in Digital Research Infrastructure," *Nature Computational Science* 1 (2021): 169–171,
https://doi.org/10.1038/s43588-021-00048-5
67. Haffenden, et al., "Making and Using AI in the Library," 45.

Inviting Knowledge: Enhancing Archival Discovery through Information Design

David J. Williams

Information design—incorporating research in graphic design, typography, visualization, and usability—is a user experience practice directly applicable to contemporary museums, libraries, and archives. Information design principles and guidelines improve engagement at every point of service, effectively and efficiently complementing the mission of knowledge organizations. This historical survey explores information design in the context of contemporary user experience design, and provides an overview of information design principles and guidelines developed over 40 years of research and professional application. Applying elements of information design to archival finding aids enhances usability, while also preserving descriptive and contextual structures.

Introduction

Librarians and archivists are often called upon to practice design. Such activities can take various forms, from creating maps and signs to arranging services and spaces.[1] From an organizational perspective, every policy or service created in a library can be regarded as a design decision.[2] These design activities often occur in environments subject to budgetary constraints and data-driven accountability.[3] When dedicated creative professionals are unavailable, librarians and archivists—regardless of training—are frequently called upon to contribute to the usability of their services, facilities, and instructional materials.

Concurrently, growth in the distribution of goods and services over the internet fueled corresponding growth in the field of user experience (UX) design, a comprehensive, holistic, and iterative practice devised to ameliorate every interaction with a product or service. As organizations with multiple avenues of interaction—both online and in-person—libraries and archives are ideally suited for designing and testing the full range of experiences they provide. However, a rigorous UX program is not always convenient for libraries because it requires commitment of resources and time.

In the absence of UX specialists, integrating UX into libraries and archives can begin with a review and evaluation of the language, text, and visual elements found throughout the organization, following the principles and guidelines of information design. Information design, a component of UX, features well-established principles and guidelines, low barriers to entry, and enhanced usability. This historical survey outlines the evolution and integration of several UX topics, revealing their practical and conceptual interrelationships, and offering a model for

*David J. Williams, MA, MLS, is Digital Initiatives and Special Collections Librarian at William Paterson University of New Jersey, email: williamsd110@wpunj.edu.

expanding the practice of information design into the development of archival finding aids. Familiarity with information design and its relationship to other design practices will inspire measurable improvements in usability while paving the way for future UX activities.

The User Experience Ecology

Although the term "user experience" may have previously appeared in different contexts, engineer and cognitive psychologist Donald Norman is widely credited with popularizing the term in 1993 during his tenure with Apple Computer,[4] where he advanced the philosophy of human-centered design and later adopted the broader, usability-focused concept of user-centered design.[5] Concurrently, Jakob Nielsen, an engineer and human-computer interaction specialist with Sun Microsystems, became one of the leading researchers in the emerging field of web usability. Human-computer interaction, derived from human factors engineering and cognitive psychology, grew in influence following the widespread availability of personal computing technologies and the rapid growth of the internet as a fundamental channel for exchanging goods, services, and information. Nielsen brought his interface design proficiency and interactivity expertise to the web, emphasizing the need for repeated and dedicated usability testing.

Nielsen defined usability as the condition of being easy to learn, remember, and use, resulting in few errors and high user satisfaction. The International Organization for Standardization (ISO) regards usability as a quality of successful goal achievement. Norman added the concepts of familiarity and obviousness, through which a designed object suggests its usage by offering "perceived affordances." These definitions were expanded to include the essential qualities of user experience design, in which individual goals are part of a larger, holistic process.[6] In 1998, Nielsen and Norman formed a consulting partnership, the Nielsen Norman Group, positioning themselves as the leading authorities in research-based UX. Their discoveries resulted in a concise definition of the term: "'User Experience' encompasses all aspects of the end-user's interaction with the company, its services, and its products."[7] Additional definitions—from concise to broad—were advanced, with the consensus centering on the "holistic" nature of user experience. As web usability expert Steve Krug observed, "UX sees its role as taking the users' needs into account at every stage of the product life cycle, from the time they see an ad on TV, through purchasing it and tracking its delivery online, and even returning it to a local branch store."[8]

With roots in product development and web design, UX rapidly evolved into both a dynamic professional practice and a comprehensive research discipline.[9] Like many conceptual fields, UX reflects the convergence of multiple practices, applying them to every aspect of an organization's products and services.[10] Today, UX specialists are increasingly contributing to public and academic libraries as core members of public service departments. The success of UX design has led to the development of numerous accredited and highly regarded academic programs, as shown in Table 1.

Information Design Practices

Although designing informational displays is a fundamental communication activity, as a field of research and professional practice, the formal origins of information design can be traced to the 1970s. Academic researchers in Europe and the United Kingdon envisioned a supportive discipline, with professional government and business sector practitioners apply-

TABLE 1
User Experience Degrees and Programs Currently Offered in the United States

School	Degrees Offered	Relevant Courses
Academy of Art University	Master of Fine Arts	Interaction Design, Product Design, User Experience, Visual Design
Art Center College of Design	Bachelor of Science, Master of Fine Arts	Design, Interaction Design, Human Factors, Design Psychology, Data Visualization, Human Computer Interaction, Product Design, Design Research, Writing for Interaction
Bentley University	Master of Science	Human Factors, Information Architecture, Testing and Assessment, User-Centered Design, Visualizing Information
California College of the Arts	Certificate in Interaction Design	Cognitive Science, Human Computer Interaction, Graphic Design, User Interface Design
California State University, Fullerton	Certificate of User Experience and Customer-Centered Design	Design Thinking, Product Design
Carnegie Mellon University	Bachelor of Science, Master of Science, Master of Professionals Studies, PhD	Cognitive Science, Communications, Design, Human Computer Interaction, Human Factors, Interaction Design, Usability
DePaul University	Master of Science	Content Strategy, Human-Computer Interaction, Information/Data Visualization, Interaction Design, Information Architecture, Usability Evaluation, UX Strategy
Drexel University	Bachelor of User Experience and Interaction Design	Cognitive Psychology, Content Management, Design, Design Thinking, Digital Media, Interaction Design, Human Factors, User-Centered Design, User Experience Design, User Interface Design, User Research
George Mason University	Master of Arts, PhD	Cognitive Science, Human Factors, Psychology, Statistics
Georgia Tech	Master of Science	Communications, Human Computer Interaction, Industrial Design, Psychology
Indiana University	Master of Science	Human Computer Interaction, Interaction Design
Kennesaw State University	Bachelor of Science	Design Thinking, Human-Centered Design, Interaction Design, User Interface Design, Visual Design
Kent University	Master of Science	Accessibility and Universal Design, Information Architecture, Interaction, Usability, User Experience Design
Michigan State University	Bachelor of Arts	Content Strategy, Digital Rhetoric, Experience Architecture, Graphic Design, Information Architecture, Interaction Design
New Jersey Institute of Technology	Bachelor of Science	Human Factors, Psychology, Usability, User Experience Design, Visual Design
New York University	Bachelor of Integrated Digital Media, Master of Integrated Digital Media	Digital Audio Production, Digital Media, Visual Design

TABLE 1
User Experience Degrees and Programs Currently Offered in the United States

School	Degrees Offered	Relevant Courses
Parsons School of Design	Bachelor of Fine Arts	Design
Philadelphia University	Master of Science	Cognitive Psychology, Digital Experience Design, Information Architecture, Interaction Design
Pratt Institute	Certificate in UX/UI Mobile Design	Accessibility, Content Strategy, Digital Analytics, Information Architecture, Information Visualization, Usability, User Experience Design
Purdue University	Bachelor of Computer Graphics Technology, Master of Computer Graphics Technology	Human Factors, Interaction Design, Usability, User Analysis, User-Centered Design, User Experience Design, Visual Design
Rutgers University	Master of Information, Master of Business and Science	Communications, Visual Design, User Experience Design, Usability, Information Architecture, Interaction Design, Informatics, Information Visualization
San Jose State University	Master of Science	Cognitive Psychology, Ergonomics, Human Computer Interaction, Interaction Design, User Interface Design
Santa Monica College	Bachelor of Interaction Design	Cognitive Psychology, Interaction Design, Product Design
Savannah College of Art and Design	Bachelor of Fine Arts	Communications, Graphic Design, Human Computer Interaction, Information Architecture, Interface Design, Product Design, Typography, User Experience Design
The University of Baltimore	Master of Science	Computers and Cognition, Information Architecture, Interaction, Interface Design
The University of Texas at Austin	Master of Science	User Experience Design
Touro College	Master of Arts	Design Thinking, Interactive Design, UI/UX Design
Tufts University	Master of Science	Computer Graphics, Computer Interface Design, Human Computer Interaction, Human Factors, Visualization
University of California	Bachelor of Cognitive Science	Cognitive Science, Communications, Design
University of California, Los Angeles	Certificate in User Experience Design	Accessibility Design, Design Thinking, User Experience Design, User Interface Design
University of Maryland, Baltimore County	Master of Science, PhD	Graphic Design, Human-Centered Computing, Systems Analysis and Design, User Interface Design
University of Miami	Master of Fine Arts	Human Centered Design, Human-Computer Interaction, Interaction Design, UX Research Methods
University of Utah	Certificate in Human Factors	Cognitive Psychology, Human Factors

TABLE 1
User Experience Degrees and Programs Currently Offered in the United States

School	Degrees Offered	Relevant Courses
University of Washington	Bachelor of Science, Master of Science, Master of Human-Computer Interaction and Design, PhD	Accessibility, Communications, Data Visualization, Human Centered Design, Human Computer Interaction, Information Visualization, Interaction Design, Product Design, Usability, User Centered Design, User Experience Design, Visual Literacy
University of Wisconsin	Bachelor of Science, Master of Science	Content Management, Data Science, Data Visualization, Human Factors, Technical Communications
Utah Valley University	Bachelor of Digital Media	Digital Product Design, Interaction Design
Winthrop University	Bachelor of Science	Information Systems, User Experience Design

ing their results to practical requirements such as traffic symbols and product labels.[11] In the United States, technical communications researchers—developing repair manuals and product documentation—began integrating elements from other disciplines and practices, particularly typography and graphic design.[12] These early practitioners referred to their work as "document design," commonly regarding the textual components as scaffolding through which graphics and other visual disciplines could be expressed.[13] Some writers and designers favored this term to distinguish their work from the information model developed by Claude Shannon, describing techniques for encoding messages within noisy communications channels.[14]

As the practice evolved, the concept of messages as fundamental units of information grew increasingly prominent, with content, language, and visual elements integral to message composition.[15] Information, as conceptualized in the practice of message design, was consistent with definitions developed in the fields of information science and communications.[16] Information was understood as being composed of facts and data that can be managed, transmitted, and imparted with meaning in the form of patterns and relationships possessing the capacity to inform.[17] Groundbreaking engineer and schematic designer Per Mollerup defined information design as "explanation design," the art of explaining facts with the aim of producing knowledge.[18] Combining elements of both practical theory and theoretical practice, information design incorporates ideas from many external disciplines.[19] As an academic discipline, information design is cross-disciplinary and integrative, characteristics shared with user experience design.[20] It is regarded by practitioners as an *information discipline*, comparable to Library and Information Science,[21] and is considered complementary to information technology (just as building technology is complementary to architecture).[22] Theoretically, information design is the process of facilitating understanding to help people achieve their goals,[23] and the task of the information designer is to transform data into high-quality information.[24] Functionally, information design is the way information is presented on a page or screen.[25] Essentially, information design adds *seeing* to reading.[26] Engineer and technical communications researcher George Hayhoe grouped information design practices according to technique:

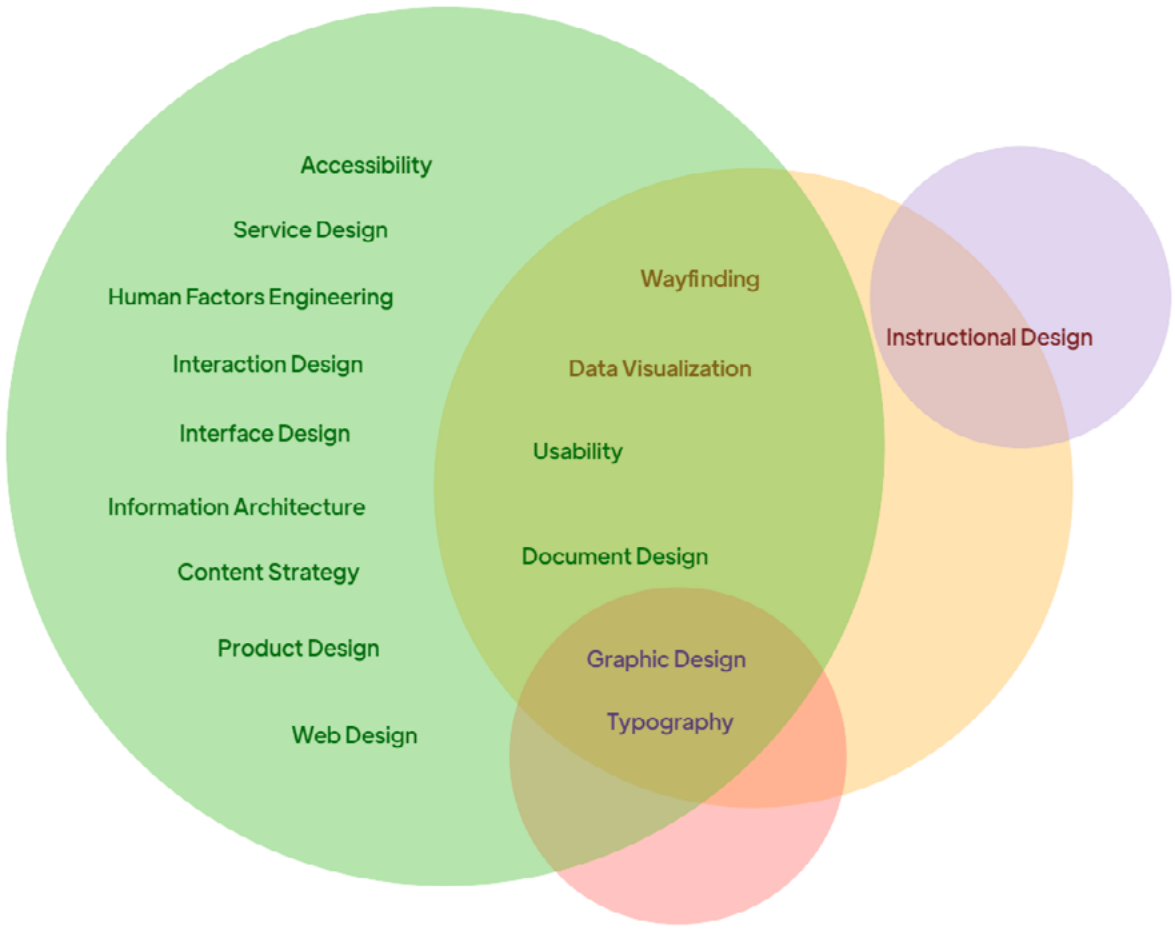
1. Information design is a "design area," similar to interior design,[27] applying graphic design principles to information in order to communicate effectively.
2. Information design is a process, identifying, organizing, and composing information messages to achieve a goal.

3. Information design is a strategy, producing and improving internal communications, products, and services as part of an organizational mission.

All three interpretations are considered equally valid.[28] Pioneering information design researcher Rune Pettersson favored the second, empirical definition: "In order to satisfy the information needs of the intended receivers, information design comprises analysis, planning, presentation, and understanding of a message—its content, language, and form."[29]

As information design evolved beyond technical communications, practitioners began to assimilate research from other disciplines and fields. The practice expanded to include typography, color theory, and graphics.[30] Graphic design—incorporating layout, color, and visual elements—effectively amplifies information design. Visual design elements—increasingly regarded as building blocks of contemporary web design—improve the effectiveness of information by aesthetically augmenting content.[31] Symbols, pictures, and words communicate ideas and express visual relationships.[32] The similarities and shared goals of these varied practices and specialties contribute to the effectiveness of both information design and user experience design, becoming information design "sub-fields,"[33] and collectively forming the larger UX ecology illustrated in Figure 1. The *usability*, *instructional design*, and *wayfinding* domains are particularly useful to library and archives professionals.

FIGURE 1
The UX Ecology*



*green = user experience design; yellow = shared information design practices; red = specialties employed in commercial design; purple = instructional and learning systems development.

Usability

Usability is one of the more prominent component elements of user experience, and the principles and methods developed by web usability experts, including Nielsen and Krug, are generally applicable to all information design products and library communications.[34] Emphasis is placed on the ongoing nature of usability practice, and testing is considered a cyclical, iterative activity informing decision-making.[35] The optimal number of participants for qualitative usability testing was established as 15, but statistically meaningful qualitative testing was accomplished with as few as five participants, producing usable results and facilitating frequent and ongoing test iterations.[36] In business terms, these practices offer the added benefits of a low barrier to entry, as well as high return on investment.

In addition to classifying and developing information resources and services, the academic disciplines of library and information science have a long history of researching their effectiveness. However, libraries and archives with limited resources may find themselves unable to fully integrate UX into strategic planning. Usability, although a valuable assessment metric, is only a single element of the user experience.[37] Considering the scope of the practice, assessing the effectiveness of UX design can involve complex and specialized survey instruments. As media and technology supporting text has evolved beyond traditional concerns for print layout and typography, the criteria for evaluating visual communication usability has correspondingly increased in complexity.[38] Having evolved in parallel, both usability and usability testing are considered essential facets of information design and of UX.[39] Information design, embraced by practitioners as a component of UX, produced a collection of principles, guidelines, and best practices for enhancing the usability of information products and resources. Information design limits complexity to the effective and efficient delivery of discrete messages, thus facilitating assessment. In the field of technical communications, usability has always been part of an integrated program,[40] and information artifacts—such as physical products or services—can be measured in terms of usability.

Instructional Design

Although instructional design, as a research domain, can be traced back to cognitive and behavioral psychology experiments conducted during World War II, it is increasingly regarded as an information design specialization.[41] Many instructional design practices have been integrated directly into message design, including Smith and Ragan's three-phase model of analysis, strategy, and evaluation.[42] Both instructional and information design are concerned with discovering evidence-based principles for presenting verbal and visual information effectively. The primary distinction between the two practices is the emphasis on long-term outcomes for instruction compared to the immediate, short-term application of information. Either goal can be successfully accomplished by applying cognitive techniques that limit irrelevant materials and signals in favor of essential message processing, thus resulting in shared design methodologies.[43]

Wayfinding

UX researchers considered the benefits of expanding the practice to include both in-person services and physical spaces. This expansion proved particularly suitable for libraries, regardless of mission or membership, due in large part to the number of interaction opportunities present in a typical library environment.[44] Wayfinding is a specialized subset of environmental

graphic design (EGD) —a practice rooted in architecture and urban planning—that is applied to built environments. EGD incorporates signs, symbols (usually in the form of pictographs), and other elements of information design.[45] Wayfinding is a common undertaking in libraries, where the goal is to make resources easy to find and use.[46] Signs and symbols assist people both in accomplishing tasks effectively and achieving goals, grouped into four categories: identification, direction, orientation, and regulation.[47] The effective application of information design in signage complements EGD, showing users what they are seeking, and revealing things they do not know.[48] In instances of signs being governed by legislation supporting people with disabilities, typography and visual composition become important information design elements of EGD.[49]

Information Design Guidelines

Information designers, such as Pettersson, have concluded that identifying consistent and firm rules for information design is essentially impossible, since information design is a combined discipline that incorporates research from a variety of fields.[50] Despite this unavoidable complexity, the statistician and data visualization pioneer Edward Tufte considers the underlying principles of information design to be universal.[51] Either way, decades of research and application have, at least, helped to establish substantial guidelines for information design. Such guidelines reflect a general consensus among practitioners regarding the universal principles that are applicable to information design projects, and that are common to their many specialized applications.

Messaging

Understanding the intended audience is necessary for an information design solution to succeed.[52] An information set must be clearly understood by its user;[53] badly designed information artifacts—those that do not consider the intended audience and/or its requirements—frequently fail to reach the target user.[54] In information design, communication is not complete until the intended receiver understands the message.[55] The first step in message design is to identify the audience and define a message. Message design requires both gaining, and holding viewer, attention.[56] Both a user's understanding of a message, and the message itself, require evaluation when assessing its effectiveness.[57] Establishing and emphasizing the essential, decision-relevant meaning of information, referred to as the "gist" of the message, significantly improves evaluation and decision making.[58]

Plain Language

The main goal of information design is clarity of communication.[59] To achieve this goal, early information designers approached text from the perspective of *legibility*. However, legibility is a difficult term to define. Legibility could describe comprehensibility, or could refer to the ability to clearly distinguish characters and words, independent of their meaning.[60] Previous attempts to classify legibility resulted in numeric "readability" scores, using sentence length and syllables per word to measure the educational level required for text comprehension. These assessments proved popular—and were frequently integrated into word processing software—but eventually came to be regarded as antiquated, and based on flawed assumptions.[61] Abstract words, jargon, long and complex sentences, passive constructions, and stilted language obstruct reading.[62] Applying usability practices, document and information designers

developed guidelines for improving communication through written text, a practice labeled "plain language." Popularized in web design and formalized into government legislation,[63] plain language soon became a standard information design assumption, and has been accepted as a fundamental component of practice.[64]

Typography

Typography clearly affects the comprehension and transmission of written information. Understanding this impact requires a deep understanding of visual perception and cognition, with results that can be difficult to interpret.[65] Although the mechanics are unclear, researchers at Johns Hopkins University School of Medicine recently concluded that people interpret typefaces as having an emotional "tone," or personality, affecting how message content is processed. Precisely which characteristics impart these tones is unclear, but their effect is measurable.[66] Potentially impacting this quality are factors such as font design characteristics (e.g. the width and spacing of individual letters) which can influence reading acuity or text legibility.[67] Balancing readability and legibility is as much art as science. Typography is a design art that requires creativity and skill, and it may not be easy to measure or analyze. Web designer Jeffrey Zeldman is credited with the statement, "Ninety percent of design is typography. The other 90 percent is whitespace."[68] The growth of digital typography brings tools and techniques to a wider audience, creating numerous opportunities for research and experimentation.

The principal goal of typography, however, is effective communication, making it an essential information design practice. Many helpful guidelines exist to support this goal. For example, the observation that the brightness contrast between the text and background is—independent of color—central to legibility.[69] Additionally, text entirely in capital letters is difficult to read, because it provides fewer visual "cues" for identifying words, as individuals take in both the individual letters and overall "shape" of a word when reading.[70] The ornamental shaped elements designed into serif fonts, on the other hand, make them better for body text, rendering the words and letters easier to distinguish.[71] Beyond the discrete influence of typeface, the length of a text line affects readability. People read words in clusters, and move from one cluster to another. Lines of text that are too long slow the reading process.[72] Lines that are too close together lead to the reader's vision drifting to adjacent lines, breaking concentration.[73] Lists are easy to parse.[74] They are easier both to recall and to evaluate, and thus have a generally positive effect on conveying information.[75] Mathematical signs and symbols, however, are not as easy to parse as alphabetical characters.[76] The contrast achieved by different sizes, shapes, forms, and weights of text content creates a visual hierarchy that organizes a document, improves comprehension, and helps the author manage how readers will view it.[77] Chunking text elements into small, discrete pieces, and using clear, visible headings and graphics are key elements of technical communication accessibility, particularly for small displays.[78]

Visualization

Data visualization, the art of representing abstract data visually, is a popular practice in the fields of statistical analysis and the digital humanities. Tables outperform graphics when presenting small data sets, but visualization is an effective means for rendering large data sets into comprehensible and practical information.[79] Infographics, in the form of transit

maps and network diagrams, are essential tools for the expeditious interpretation of complex information. These practices create visual relationships and patterns among data, leading to knowledge acquisition.[80] Data visualization applied to message design places it solidly within the conventions of information design.[81] Converting text into visual presentations—such as lists, tables, and maps—is considered an important skill for information designers.[82] Well-designed statistical graphics, expressed using data visualization techniques, satisfy the goal of communicating complexity with clarity.[83]

Graphics

Aesthetic considerations for enhancing the experience and perception of a message may improve its effectiveness.[84] Information design does not ignore aesthetic principles, but they are not the primary focus. Nonetheless, well-structured content and visualizations are often deemed aesthetically pleasing.[85] Introducing aesthetics into a systematic visual style does communicate meaning, and applying style standards cohesively contributes to a harmonious experience. Even small stylistic decisions, such as text justification, can alter the meaning and message conveyed.[86] Carefully integrating words and pictures engages people more effectively than words or pictures alone.[87] Color, images, lines, symbols, and text should be integrated into a meaningful whole, instead of being treated as individual elements.[88] An example of this integration are pictographs—standardized visual elements commonly used in signage and EGD features—which are recognized and understood quickly.[89] However, animated pictographs, and animations in general, decrease recognition accuracy compared to static images.[90]

Images are easier to remember than text,[91] and we perceive the entirety of an image, its gestalt, beyond its individual parts.[92] Visual design applies the graphic design principles of unity, gestalt, space, hierarchy, balance, contrast, scale, and continuity to information displays.[93] Visual language, combining graphics and text, speaks to people holistically and emotionally.[94] Visual messages are superior when the content is emotional, immediate, and spatial.[95] The combination of an effectively articulated message—expressed clearly and presented with a harmonious consideration of visual and textual elements—is essential to contemporary information design.

Archival Finding Aids

Libraries and archives produce many information artifacts that support their operations and users. One example specific to archives and special collections is the finding aid. Finding aids are a specialized form of collection inventory designed to facilitate access to primary source materials, while maintaining intellectual control over their arrangement and representation. The U.S. National Archives and Records Administration defines finding aids as, "tools that help a user find information in a specific record group, collection, or series of archival materials." Examples include inventories, container and folder lists, indexes, registers, and institutional guides, which are formally and informally published.[96] The Society of American Archivists concisely defines a finding aid as "a description that typically consists of contextual and structural information about an archival source."[97]

In contrast to traditional library print collections, archival holdings often contain diverse materials that require structural flexibility. Finding aids reveal a collection's arrangement, which reflects how the contents are grouped and ordered. Depending on the source, materials can be organized based on the original order of accessioning, the subject areas covered, the material and media types, or other characteristics. Where the original order or provenance

is not provided, alphabetical, chronological, and subject-based arrangements are customary. The amount of detail recorded reflects the descriptive activities undertaken during initial and subsequent archival processing.

An effective finding aid should assist researchers in expeditiously locating materials, regardless of complexity, and ideally without requiring assistance or intervention.[98]

Finding Aid Evolution

Until recently, archivists regarded finding aids as supplemental to their practice, assuming that their professional intervention and assistance would be necessary for most users.[99] Although intended for all researchers, historians and genealogists are the primary consumers of archival collections. Traditional archivists brought knowledge of historical research methods to their practice, and historians are among those who find the structure and format of finding aids intuitive.[100] As the practice of developing finding aids evolved, the need for standards became apparent, starting with standardized archival descriptions. After surveying the materials and determining their arrangement, archivists developed shared terminologies to identify the elements, features, and relationships within a collection. These standards formed guidelines for composing finding aids, and were defined both by the International Council on Archives as the General International Standard Archival Description (ISAD(G)), and, in the United States, the Describing Archives: A Content Standard (DACS). However, these standards were not designed for display formatting, or as content guidelines. Their primary audience was the archival community and, as a result, they could be regarded as input, not output, standards.[101] Another way to interpret these earlier standards is that they describe collections, but do not provide access. Intellectual control is achieved by documenting provenance and providing context.[102] This was the finding aid's dual inheritance: standards for internal management and preservation were adapted and applied without modification for use by external users. Many archives did not initially consider the practice of adapting materials for outside visitors, leading to inconsistent labeling and terminology.[103] Consequently, even when institutions apply the same archival standards, the output of archival descriptions across institutions vary widely.[104]

Encoded Archival Description

By the mid-1990s, standards for describing and arranging archives and records were further codified and enhanced by the development of Encoded Archival Description (EAD), which was an initiative intended to provide similar control and discoverability to archival collections as Machine-Readable Cataloging (MARC) brought to library collections in electronic form.[105] EAD encourages coherence around established standards, and facilitates the exchange of collection information data between search and discovery systems.[106] However, adopting EAD is not sufficient if the implementation does not offer information that is understandable or helpful to users.[107] Surveys conducted in 2007 indicated that research communities using archival collections still preferred traditional print finding aids.[108] Ten years later, many researchers still expected access to printable copies of finding aids.[109] As online access to information expanded, users of archival and records collections began expecting a corresponding amount of self-directed item-level searching, regardless of arrangement.[110] Current trends suggest that item-level description and representation is presumed from digital archival technologies.[111] Increasingly, researchers anticipate seeing digital surrogates of documents and photographs associated with finding aid descriptions.[112]

Expectations are evolving, however, and enhanced online finding aids, which support remote discovery, may become the exclusive point of access to collections.[113] Unfortunately, the descriptive data archives provide are rarely user-friendly.[114] If finding aids are intended to provide end-users with efficient and effective access to collections without archival assistance, then adopting information design practices is highly recommended.

Finding Aid Usability

Early efforts to evaluate the usability of archival tools and resources, as recently as 2002, were considered inadequate.[115] At that time, usability training was rare even in archives that implemented EAD. The finding aids produced online using EAD were indistinguishable from earlier print versions.[116] These preliminary tests revealed that end-users were not interested in systems that reproduced archival structures.[117] Online users also demonstrated little patience for reading dense blocks of text typical of print finding aids.[118] In 2006, user-centered design—a flourishing practice popularized by Nielsen and Norman—was still uncommon in online archival resources.[119] Users increasingly expected web usability features in online archives, and they considered the simple transfer of print content to a screen interface suboptimal.[120] By 2008, online finding aids continued to mirror their print counterparts in both content and appearance, offering limited functionality, as well as minimal browsing and searching features.[121] Usability and information design received little attention. Novice users encountered dense paragraphs, few graphics, little white space or visual organization, confusing language, and professional jargon.[122] Institutions that implemented even modest user interface improvements, including online navigational "wayfinding" indicators, yielded significantly better user experiences.[123]

Novice and Nonexpert Finding Aid Interaction

The development of EAD was initially expected to facilitate access to archival collections for inexperienced users.[124] Online distribution granted access to more researchers, and collection information was no longer limited to expert archives users.[125] Educators at every grade level, and in a variety of disciplines, increasingly encouraged and expected novice student researchers to incorporate primary source archival materials into their assignments.[126] Non-experts, including secondary school students, are challenged to learn both new knowledge domains and the domain-specific metacognitive skills needed to analyze and internalize this knowledge.[127] Unlike subject-matter experts visiting an online finding aid, novice users seldom know exactly what they are looking for when initially approaching a collection.[128] Non-experts quickly discover that finding aids generally reflect the archivists' perspective on collections, which often differ substantially from their own.[129] By 2010, novice archives users were increasingly conversant with internet technologies, and expected the availability of typical web-usability features—including sidebar menus, descriptive hover text, and supplementary help guides—when visiting archives websites.[130] For collections that offered such web-usability features, not only did keyword searching with controlled vocabularies grow in popularity, item-level search interfaces proved significantly more popular than engaging with and traversing the collection hierarchy.[131] Today, for both expert and non-expert users, commercial search engines are the preferred tools for locating known items.[132] With the growth of unmediated digital access to archival repositories, information design has become an effective means for providing access and promoting the value of these collections.

Information Design and Finding Aids

In October 2005, William Paterson University of New Jersey (WPU) hosted the Nicholas Martini Conference on Local Government, a series of presentations and panels featuring historians and elected public officials discussing regional political issues. For the conference, the university's Cheng Library accessioned the personal papers and artifacts of Nicholas Martini, the former commissioner of Passaic County, and mayor of Passaic, New Jersey. Processing archivist Trudi Van Dyke organized, described, and compiled the guide, which provides a detailed representation of the collection. Because the collection lacked an original order, Van Dyke arranged the materials by subject into series and grouped chronologically, reflecting the practice of folder-level description. The result is a professional information product following best practices in archival description.[133] At WPU, 45 percent of enrolled students represent the first generation attending university.[134] Undergraduate students are often unfamiliar with primary source materials. Historians and political scientists experience little difficulty in interpreting the Martini Collection finding aid; however novice users, including many WPU freshmen, require guidance and direction. With the increasing availability of hybrid and distance education courses requiring self-directed access to online resources, such interventions are

FIGURE 2
Print Collection Summary Description Page

Guide to the Nicholas Martini Collection, 1931-1991

Archives & Special Collections, Cheng Library
William Paterson University

Descriptive Summary

Creator:	Martini, Nicholas, 1904-1991	
Title:	Guide to the Nicholas Martini Collection	
Dates:	1931-1991	
Quantity:	Campaign Files:	5 Boxes
	Alberta Central:	1 Box
	Scrapbooks:	33 Volumes
	Microfilm:	3 Reels
	Photographs:	6 Boxes
	Personal:	5 Boxes

Abstract: Collection supports research on the life of Passaic, NJ attorney and politician Nicholas Martini, who served as Freeholder, Commissioner and Mayor of Passaic at various times in his career. Personal papers and extensive newspaper clippings document Martini's campaigns for office and present a political history of Passaic in the 1930s through the 1950s. Numerous photos and other personal papers provide information about the Martini family, his business ventures, and involvement in community activities.

FIGURE 3
Print Series Description Page

Series Descriptions

CAMPAIGN FILES

NICHOLAS MARTINI'S CAMPAIGNS:
1936 & 1937 CAMPAIGNS FOR PASSAIC COUNTY FREEHOLDER
1935 & 1939 CAMPAIGNS FOR PASSAIC COUNTY COMMISSIONER
MISCELLANEOUS CAMPAIGNS

Folders arranged topically and chronologically; documents arranged in chronological order (wherever possible) in folders as well.

This series mainly contains documentation pertaining to Nicholas Martini's 1936 and 1937 bids for Passaic County Freeholder and his 1935 & 1939 campaigns for Passaic County Commissioner. However, there are also some documents pertaining to other campaigns, including Republican campaigns of 1931 and 1938. These files contain correspondence pertaining to Martini's political affiliations and invitations to various functions, such as benefits and dedications. These files also contain correspondence related to Martini's duties as a City Commissioner of Passaic, as he was often acting in that capacity during his campaigns for Freeholder as well as when he was running for re-election for Commissioner. Among the documents are letters to/from the Passaic County Clerk's Office, the Central Italian-American Committee, the New American, Robert Wardle Associates, the Dante Political Club, as well as letters from citizens. These files also include letters from Martini asking for support in the elections and documents pertaining to the Martini League.

In addition to the correspondence, these files contain political paraphernalia, including campaign cards, brochures, posters and receipts pertaining to the campaigns. Election ballots and voting tallies of the elections can also be found in these files. Among the documents are offers of condolences to Martini on his 1936 loss and congratulations on his 1937 victory, and Martini's acknowledgement of these materials. These documents are in the form of cards, letters and telegrams.

not guaranteed. Information design offers a means to overcome these limitations. A sample illustration of information design principles and guidelines applied to the Martini Collection finding aid demonstrates their effectiveness.

Two essential components of a finding aid are the collection summary (Figure 2) and series description (Figure 3). Many details provided in a standard summary, such as the author's name and primary language, are administrative and not essential to novice users. Other elements can be contextually introduced. Displaying arrangement descriptions within an individual series and associating access terms with visitor information are methods for conditionally presenting these elements.

For non-experts, the goal is to efficiently determine a collection's subject areas and content. Figure 4 simplifies the presentation of the online summary. The collection creator, title, and repository name are integrated into the page template. The extent and abstract are reformulated using plain language, and the dates—a frequent source of confusion for inexperienced readers[135]—are expressed as a specific subset of coverage reflecting the collection's central topic.

Figure 5 provides an alternative series description, eliminating long blocks of text in favor of brief messages. Plain language intended for specific audiences facilitates the quick scanning and evaluation of key details.

FIGURE 4
Custom Online Summary Description Demonstrating Information Design

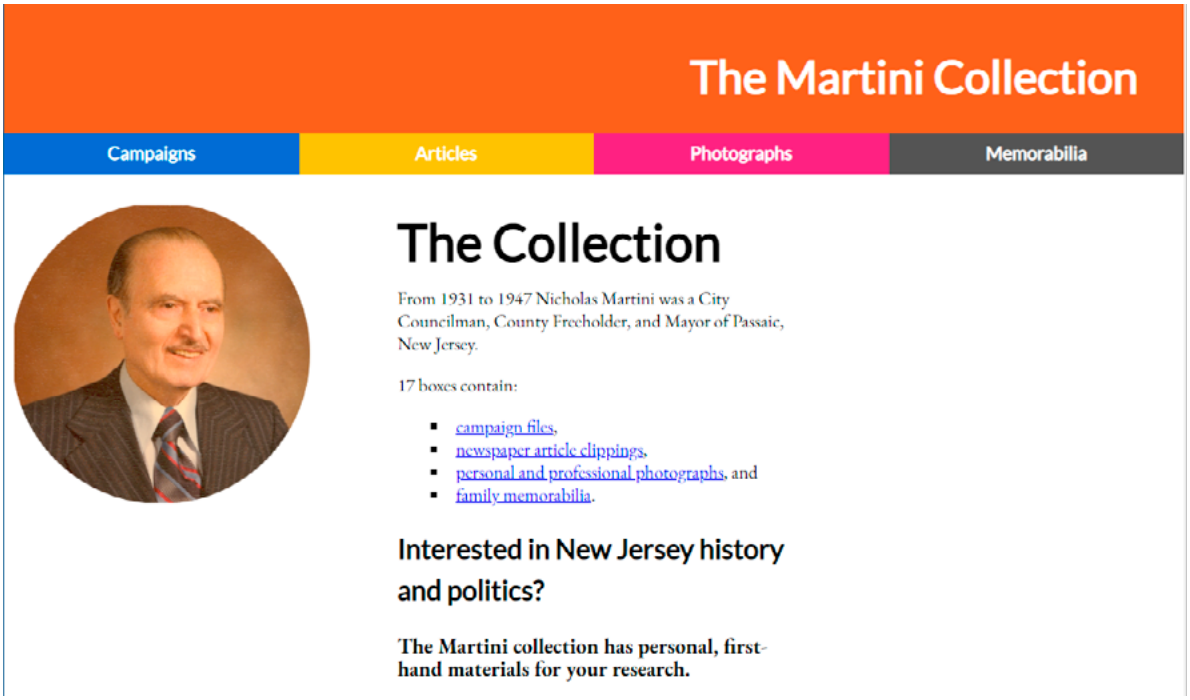


FIGURE 5
Custom Online Series Description Demonstrating Information Design

These sample documents introduce four additional information design practices:

1. Consistent visual style, highlighting the collection title, repository name, search feature, print format finding aid, and other recurring page elements.
2. Application of color for both aesthetic and functional purposes, with a high contrast scheme optionally mirroring wayfinding practices.

3. Inclusion of visual elements to supplement and complement text elements, promoting effective engagement.
4. Implementation of useful typography, in the form of mixed case, varying font sizes and weights, and structured headings prioritizing content. Usability, readability, and comprehension improves significantly when distinct headings are applied.[136]

Conclusion

Digital archives and online archival resources, including electronic finding aids, are valuable contemporary developments that support information management and access. Developments in electronic standards and online finding aid dissemination have improved education. Data visualization offers further educational value to archives by increasing engagement and adding interactivity to primary source collections.[137] Correspondingly, standards-based markup languages and related technologies innovate workflows by electronically extracting and transforming item-level metadata. Reformatting this content into a variety of representations provides multiple customized access points, while also preserving traditional archival description.[138] Future enhancements to electronic standards in the form of integrated audience attributes can establish elements intended for alternative user communities, including surrogate images and plain-language messages clarifying the structure and components of a collection.

Where item-level description is limited or missing, collection search features will be correspondingly deficient.[139] Scanning and processing documents using optical character recognition technologies, although time-consuming and subject to preservation considerations, increases the volume of item-level content. The additional metadata produced by these activities greatly enhance collection discovery, and offset the negative effects of a limited, or absent, search text index.[140] Information design provides the final step of the process, ensuring that the data are tested and distributed in a usable, user-friendly manner. In the absence of comprehensive digitization initiatives, skillfully designed information products will increase entry-level researcher comfort with primary sources and improve educational outcomes.

Information design is an established and economical practice offering significant enhancements to the overall usability and effectiveness of both physical and online environments. Proficiency in information design is readily attainable with immediate, measurable benefits. Applying information design guidelines and best practices to complex and confusing information artifacts, in addition to positively impacting usability, has proven to be a relatively inexpensive investment that is easy to implement.[141] As demonstrated using finding aid content from the Martini collection, librarians, archivists, and library users benefit substantially when incorporating information design into their professional activities.

Bibliography

Beier, Sofie, Chiron A.T. Oderkerk, Birte Bay, and Michael Larsen. "Increased Letter Spacing and Greater Letter Width Improve Reading Acuity in Low Vision Readers." *Information Design Journal* 26, no. 1 (April 28, 2021): 73–88. https://doi.org/10.1075/idj.19033.bei.

Berg, Magnus. "A 'Major Technological Challenge': Multi-Level Description and Online Archival Databases." *Emerging Library & Information Perspectives* 4, no. 1 (July 2, 2021): 62–87. https://doi.org/10.5206/elip.v4i1.12529.

Chapman, Joyce Celeste. "Observing Users: An Empirical Analysis of User Interaction with Online Finding Aids." *Journal of Archival Organization* 8, no. 1 (June 10, 2010): 4–30. https://doi.org/10.1080/15332748.2010.484361.

Chen, Yu-Hui, Carol Anne Germain, and Abebe Rorissa. "Defining Usability: How Library Practice Differs from Published Research." *Portal: Libraries and the Academy* 11, no. 2 (2011): 599–628. https://doi.org/10.1353/pla.2011.0020.

Cole, Tiffany. "Characteristics of Successful Finding Aids." Mid-Atlantic Regional Archives Conference, October 31, 2011. https://www.marac.info/index.php?option=com_content&view=article&id=64.

Daines III, J. Gordon, and Cory L. Nimer. "Re-Imagining Archival Display: Creating User-Friendly Finding Aids." *Journal of Archival Organization* 9, no. 1 (May 20, 2011): 4–31. https://doi.org/10.1080/15332748.2011.574019.

DeRidder, Jody L., Amanda Axley Presnell, and Kevin W. Walker. "Leveraging Encoded Archival Description for Access to Digital Content: A Cost and Usability Analysis." *The American Archivist* 75, no. 1 (April 1, 2012): 143–70. https://doi.org/10.17723/aarc.75.1.5641v61p422u0u90.

Emerson, John. "Visualizing Information for Advocacy: An Introduction to Information Design." Open Society Institute Information Program, January 2008.

Ericson, Jonathan D., William S. Albert, Benjamin P. Bernard, and Elizabeth Brown. "End-User License Agreements (EULAs): Investigating the Impact of Human-Centered Design on Perceived Usability, Attitudes, and Anticipated Behavior." *Information Design Journal* 26, no. 3 (May 23, 2022): 190–208. https://doi.org/10.1075/idj.20018.eri.

Ferriero, David. "Finding Aid Type." Lifecycle Data Requirements Guide (LCDRG). Washington, DC: National Archives and Records Administration, August 15, 2016. https://www.archives.gov/research/catalog/lcdrg/elements/findingtype.html.

Freund, Luanne, and Elaine G. Toms. "Interacting With Archival Finding Aids." *Journal of the Association for Information Science and Technology* 67, no. 4 (April 27, 2015): 994–1008. https://doi.org/10.1002/asi.23436.

Gibson, David. *The Wayfinding Handbook: Information Design for Public Places*. 1st ed. New York, NY: Princeton Architectural Press, 2009.

Godfrey, Krista. "Creating a Culture of Usability." *Weave: Journal of Library User Experience* 1, no. 3 (2015): 1–10. https://doi.org/10.3998/weave.12535642.0001.301.

Hayhoe, George F. "Telling the Future of Information Design." *Communication Design Quarterly* 1, no. 1 (September 2012): 23–26.

Hazlett, Richard L., Kevin Larson, A. Dawn Shaikh, and Barbara S. Chaparo. "Two Studies on How a Typeface Congruent with Content Can Enhance Onscreen Communication." *Information Design Journal* 20, no. 3 (September 1, 2013): 207–19. https://doi.org/10.1075/idj.20.3.02haz.

Jansen, Frank. "How Bulleted Lists and Enumerations in Formatted Paragraphs Affect Recall and Evaluation of Functional Text." *Information Design Journal* 21, no. 2 (November 3, 2015): 146–62. https://doi.org/10.1075/idj.21.2.06jan.

Johnston, Rita D. "A Qualitative Study of the Experiences of Novice Undergraduate Students with Online Finding Aids." Master's thesis. University of North Carolina at Chapel Hill, 2008. https://cdr.lib.unc.edu/concern/masters_papers/12579x16p.

Krug, Steve. *Don't Make Me Think, Revisited: A Common Sense Approach to Web Usability*. 3rd ed. San Francisco, CA: New Riders, Peachpit Press, 2013.

Laugwitz, Bettina, Theo Held, and Martin Schrepp. "Construction and Evaluation of a User Experience Questionnaire," 63–76. Graz, Austria: Springer-Verlag, 2008. https://doi.org/10.1007/978-3-540-89350-9_6.

Lonsdale, Maria dos Santos, David Lonsdale, and Hye-Won Lim. "The Impact of Delivering Online Information Neglecting User-Centered Information Design Principles." *Information Design Journal* 24, no. 2 (April 18, 2019): 151–77. https://doi.org/10.1075/idj.00005.san.

Luca, Edward, and Bhuva Narayan. "Signage by Design: A Design-Thinking Approach to Library User Experience." *Weave: Journal of Library User Experience* 1, no. 5 (June 2016): 1–25. https://doi.org/10.3998/weave.12535642.0001.501.

Marchionini, Gary. "Information-seeking Strategies of Novices Using a Full-text Electronic Encyclopedia." *Journal of the American Society for Information Science* 40, no. 1 (January 1989): 54–66. https://dl.acm.org/

doi/10.5555/65193.65199.
Mayer, Richard E. "Instructional Design as a Form of Information Design." *Information Design Journal* 25, no. 3 (October 22, 2020): 258–63. https://doi.org/10.1075/idj.25.3.03may.
McGahee, Clayton. "Making Special Collections Accessible to Users: Finding Aids." Lecture presented at the International Association of Law Libraries 36th Annual Course on International Law and Legal Information, Civil Rights, Human Rights, and Other Critical Issues in U.S. Law, Atlanta, GA, October 26, 2017. https://doi.org/10.1017/jli.2018.2.
Meloncon, Lisa. "Embodied Personas for a Mobile World." *Technical Communication* 64, no. 1 (February 2017): 50–65.
Mollerup, Per. *Data Design: Visualising Quantities, Locations, Connections*. London, UK: Bloomsbury Academic Publishing, 2015.
Nielsen, Jakob, and Thomas K. Landauer. "A Mathematical Model of the Finding of Usability Problems." In *CHI '93: Proceedings of the INTERACT '93*, 206–13. Amsterdam, NL: Association for Computing Machinery, 1993. https://doi.org/10.1145/169059.169166.
Nielsen, Jakob, and Donald A. Norman. "User Experience," n.d. https://www.nngroup.com/articles/definition-user-experience/.
Norman, Donald A. *The Design of Everyday Things: Revised and Expanded Edition*. 2nd ed. New York: Basic Books, 2013.
Norman, Donald A., Jim Miller, and Austin Henderson. "What You See, Some of What's in the Future, And How We Go About Doing It: HI at Apple Computer." In *Conference Companion on Human Factors in Computing Systems*, 155. Denver, CO: ACM, 1995. https://doi.org/10.1145/223355.223477.
North, Alvin J., and L. B. Jenkins. "Reading Speed and Comprehension as a Function of Typography." *Journal of Applied Psychology* 35, no. 4 (August 1951): 225–28. https://doi.org/10.1037/h0063094.
Oberbichler, Sarah, Katharina Gallner-Holzmann, and Theo Hug. "Generous and Inviting Interfaces Revisited: Examples of Designing Visual Structures for Digital Archives." *Information Design Journal* 26, no. 2 (December 7, 2021): 157–74. https://doi.org/10.1075/idj.20028.obe.
Orna, Elizabeth. "Information Science and Information Design: Have They Anything to Communicate to One Another?" *Information Design Journal* 1, no. 4 (January 1, 1980): 271–80. https://doi.org/10.1075/idj.1.4.07orn.
Pearce-Moses, Richard. "Finding Aid." In *SAA Dictionary of Archives Terminology*. Chicago, IL: Society of American Archivists, April 29, 2020. https://dictionary.archivists.org/entry/finding-aid.html.
Pettersson, Rune. "Information Design–Principles and Guidelines." *Journal of Visual Literacy* 29, no. 2 (2010): 167–82. https://doi.org/10.1080/23796529.2010.11674679.
———. *It Depends*. 4th ed. Stockholm, SE: Institute for Infology, 2012. https://www.researchgate.net/profile/Rune-Pettersson/publication/281810734_11_It_Depends/links/6267908e1b747d19c2a89281/11-It-Depends.pdf.
———. "Seven Theories for ID." In *Proceedings of the 7th Information Design International Conference*, 2:819–30. São Paulo, Brazil: Blucher Design Proceedings, 2015. https://doi.org/10.5151/designpro-CIDI2015-cidi_ZZ.
Pirolli, Peter, and Daniel M. Russell. "Introduction to This Special Issue on Sensemaking." *Human-Computer Interaction* 26, no. 1–2 (March 16, 2011): 1–8. https://doi.org/10.1080/07370024.2011.556557.
Pontis, Sheila, and Michael Babwahsingh. "Improving Information Design Practice: A Closer Look at Conceptual Design Methods." *Information Design Journal* 22, no. 3 (January 19, 2017): 249–65. https://doi.org/10.1075/idj.22.3.06pon.
Prom, Christopher J. "The EAD Cookbook: A Survey and Usability Study." *The American Archivist* 65, no. 2 (September 1, 2002): 257–75. https://doi.org/10.17723/aarc.65.2.b3783jr052731588.
Redish, Janice C. *Letting Go of Words: Writing Web Content That Works*. 2nd ed. Waltham, MA: Morgan Kaufmann Publishers, 2014. https://www.sciencedirect.com/science/article/pii/B9780123859303000327.
———. "Readability Formulas Have Even More Limitations than Klare Discusses." *ACM Journal of Computer Documentation (JCD)* 24, no. 3 (August 1, 2000): 132–37. https://dl.acm.org/doi/pdf/10.1145/344599.344637.
———. "Technical Communication and Usability: Intertwined Strands and Mutual Influences." *IEEE Transactions on Professional Communication* 53, no. 3 (September 2010): 191–201. https://doi.org/0.1109/TPC.2010.2052861.
———. "What Is Information Design?" *Technical Communication: Journal of the Society for Technical Communication* 47, no. 2 (May 1, 2000): 163–66. https://www.researchgate.net/profile/Ginny-Redish/publication/233710886_What_Is_Information_Design/links/56e4494e08aedb4cc8ac232d/What-Is-Information-Design.pdf.
Roth, James M. "Serving Up EAD: An Exploratory Study on the Deployment and Utilization of Encoded Archival Description Finding Aids." Master's, University of North Carolina at Chapel Hill, 2001. https://doi.org/10.17723/aarc.64.2.e687471v304k0u66.
Scales, Alice Y. "Improving Instructional Materials by Improving Document Formatting," 1–12. American Society for Engineering Education, 2011.
Scheir, Wendy. "First Entry: Report on a Qualitative Exploratory Study of Novice User Experience with On-

line Finding Aids." *Journal of Archival Organization* 3, no. 4 (July 18, 2006): 49–85. https://doi.org/10.1300/J201v03n04_04.
Schmidt, Aaron, and Amanda Etches. *Useful, Usable, Desirable: Applying User Experience Design to Your Library*. Chicago, IL: American Library Association, 2014. https://www.alastore.ala.org/content/useful-usable-desirable-applying-user-experience-design-your-library.
Schriver, Karen A. *Dynamics in Document Design: Creating Texts for Readers*. New York: John Wiley & Sons, Inc., 1996.
Shedroff, Nathan. "Information Interaction Design: A Unified Field Theory of Design," 1994.
Stapleton, Katina Rae, and Katherine Spivey. "Federal Plain Language Guidelines." The Plain Language Action and Information Network (PLAIN), May 2011. https://www.plainlanguage.gov/guidelines/.
Strikwerda, Jelle, Bregje Holleman, and Hans Hoeken. "Designing Pension Communication: Lessons from the Medical Domain." *Information Design Journal* 26, no. 3 (August 4, 2022): 1–21. https://doi.org/doi.org/10.1075/idj.21011.str.
Tinker, Miles A. "The Relative Legibility of the Letters, the Digits, and of Certain Mathematical Signs." *The Journal of General Psychology* 1, no. 3–4 (1928): 472–96. https://doi.org/10.1080/00221309.1928.9918022.
Tufte, Edward R. *Envisioning Information*. 4th ed. Cheshire, CT: Graphics Press, 1990.
———. *The Visual Display of Quantitative Information*. 2nd ed. Cheshire, CT: Graphics Press, 2001.
U.S. Dept. of Health and Human Services. "Visual Design Basics." Usability.gov. Department of Health and Human Services, October 8, 2013. https://www.usability.gov/what-and-why/visual-design.html.
Van Dyke, Trudi. "The Nicholas Martini Collection 1931-1991." William Paterson University, October 14, 2005. https://www.wpunj.edu/library/pdf/MartiniFindingAidUpdated060407.pdf.
Wakimoto, Diana K. "Librarians and Graphic Design: Preparation, Roles, and Desired Support." *Public Service Quarterly* 11, no. 3 (July 3, 2015): 171–82. https://doi.org/10.1080/15228959.1054545.
Waller, Rob. "Functional Information Design: Research and Practice." *Information Design Journal* 1, no. 1 (January 1, 1979): 43–50. https://doi.org/10.1075/idj.1.1.06wal.
———. "The Origins of the Information Design Association." University of Reading Department of Typography & Graphic Communication, 2008.
Walton, Rachel. "Looking for Answers: A Usability Study of Online Finding Aid Navigation." *The American Archivist* 80, no. 1 (June 2017): 30–52. https://doi.org/10.17723/0360-9081.80.1.30.
William Paterson University Office of Institutional Effectiveness. "University Facts." William Paterson University, n.d. https://www.wpunj.edu/about-us/university-facts.html.
Yakel, Elizabeth, Seth Shaw, and Polly Reynolds. "Creating the Next Generation of Archival Finding Aids." *D-Lib Magazine* 13, no. 5/6 (May 13, 2007): 7. https://doi.org/10.1045/may2007-yakel.
York, Jeremy. "Legibility and Large-Scale Digitization." Hathi Trust Digital Library, November 2008.
Zeldman, Jeffrey. "Design Is Typography," December 23, 2015. https://twitter.com/zeldman/status/679727437198929921?s=20&t=RoHX_FVinW-yv3TODE8YYA.
Zhang, Jane. "Archival Representation in the Digital Age." *Journal of Archival Organization* 10, no. 1 (May 29, 2012): 45–68. https://doi.org/10.1080/15332748.2012.677671.

Notes

1. Diana K. Wakimoto, "Librarians and Graphic Design: Preparation, Roles, and Desired Support," *Public Service Quarterly* 11, no. 3 (July 3, 2015): 171, https://doi.org/10.1080/15228959.1054545.

2. Aaron Schmidt and Amanda Etches, *Useful, Usable, Desirable: Applying User Experience Design to Your Library* (Chicago, IL: American Library Association, 2014), 7, https://www.alastore.ala.org/content/useful-usable-desirable-applying-user-experience-design-your-library.

3. Krista Godfrey, "Creating a Culture of Usability," *Weave: Journal of Library User Experience* 1, no. 3 (2015): 1–10, https://doi.org/10.3998/weave.12535642.0001.301.

4. Donald A. Norman, Jim Miller, and Austin Henderson, "What You See, Some of What's in the Future, And How We Go About Doing It: HI at Apple Computer," in *Conference Companion on Human Factors in Computing Systems* (Human Factors in Computing Systems, Denver, CO: ACM, 1995), 155, https://doi.org/10.1145/223355.223477.

5. Donald A. Norman, *The Design of Everyday Things: Revised and Expanded Edition*, 2nd ed. (New York: Basic Books, 2013), 9.

6. Yu-Hui Chen, Carol Anne Germain, and Abebe Rorissa, "Defining Usability: How Library Practice Differs from Published Research," *Portal: Libraries and the Academy* 11, no. 2 (2011): 622, https://doi.org/10.1353/pla.2011.0020.

7. Jakob Nielsen and Donald A. Norman, "User Experience," n.d., https://www.nngroup.com/articles/

definition-user-experience/.

8. Steve Krug, *Don't Make Me Think, Revisited: A Common Sense Approach to Web Usability*, 3rd ed. (San Francisco, CA: New Riders, Peachpit Press, 2013), 152.

9. Janice C. Redish, "Technical Communication and Usability: Intertwined Strands and Mutual Influences," *IEEE Transactions on Professional Communication* 53, no. 3 (September 2010): 191, https://doi.org/0.1109/TPC.2010.2052861.

10. Janice C. Redish, *Letting Go of Words: Writing Web Content That Works*, 2nd ed. (Waltham, MA: Morgan Kaufmann Publishers, 2014), xxvi, https://www.sciencedirect.com/science/article/pii/B9780123859303000327.

11. Rob Waller, "The Origins of the Information Design Association" (University of Reading Department of Typography & Graphic Communication, 2008), 4.

12. Janice C. Redish, "What Is Information Design?," *Technical Communication: Journal of the Society for Technical Communication* 47, no. 2 (May 1, 2000): 163, https://www.researchgate.net/profile/Ginny-Redish/publication/233710886_What_Is_Information_Design/links/56e4494e08aedb4cc8ac232d/What-Is-Information-Design.pdf.

13. Redish, "What Is Information Design," 165.

14. Karen A. Schriver, Dynamics in Document Design: Creating Texts for Readers (New York: John Wiley & Sons, Inc., 1996), 6–7.

15. Rune Pettersson, *It Depends*, 4th ed. (Stockholm, SE: Institute for infology, 2012), 15, https://www.researchgate.net/profile/Rune-Pettersson/publication/281810734_11_It_Depends/links/6267908e1b747d19c2a89281/11-It-Depends.pdf.

16. Elizabeth Orna, "Information Science and Information Design: Have They Anything to Communicate to One Another?," *Information Design Journal* 1, no. 4 (January 1, 1980): 271, https://doi.org/10.1075/idj.1.4.07orn.

17. Nathan Shedroff, "Information Interaction Design: A Unified Field Theory of Design," 1994: 4.

18. Per Mollerup, *Data Design: Visualising Quantities, Locations, Connections* (London, UK: Bloomsbury Academic Publishing, 2015), 5.

19. Rune Pettersson, "Seven Theories for ID," in *Proceedings of the 7th Information Design International Conference*, vol. 2 (CIDI 2015: The 7th Information Design International Conference, São Paulo, Brazil: Blucher Design Proceedings, 2015), 1, https://doi.org/10.5151/designpro-CIDI2015-cidi_ZZ.

20. Waller, "Origins of the IDA," 4.

21. Pettersson, *It Depends*, 26.

22. Pettersson, "Seven Theories for ID," 821.

23. Sheila Pontis and Michael Babwahsingh, "Improving Information Design Practice: A Closer Look at Conceptual Design Methods," *Information Design Journal* 22, no. 3 (January 19, 2017): 250, https://doi.org/10.1075/idj.22.3.06pon.

24. Pettersson, "Seven Theories for ID," 822.

25. Redish, "What Is Information Design?," 164.

26. John Emerson, "Visualizing Information for Advocacy: An Introduction to Information Design" (Open Society Institute Information Program, January 2008), 3.

27. Pettersson, *It Depends*, 13.

28. George F. Hayhoe, "Telling the Future of Information Design," *Communication Design Quarterly* 1, no. 1 (September 2012): 23.

29. Pettersson, "Seven Theories for ID," 819.

30. Redish, "What Is Information Design?," 163.

31. U.S. Dept. of Health and Human Services, "Visual Design Basics," Usability.gov (Department of Health and Human Services, October 8, 2013), https://www.usability.gov/what-and-why/visual-design.html.

32. Emerson, "Visualizing Information for Advocacy," 3.

33. Pontis and Babwahsingh, "Improving Information Design Practice," 250.

34. Godfrey, "Creating a Culture of Usability."

35. Krug, *Don't Make Me Think, Revisited*, 135.

36. Jakob Nielsen and Thomas K. Landauer, "A Mathematical Model of the Finding of Usability Problems," in *CHI '93: Proceedings of the INTERACT '93* (CHI '93 Conference on Human Factors in Computing Systems, Amsterdam, NL: Association for Computing Machinery, 1993), 213, https://doi.org/10.1145/169059.169166.

37. Bettina Laugwitz, Theo Held, and Martin Schrepp, "Construction and Evaluation of a User Experience Questionnaire" (Symposium of the Austrian HCI and Usability Engineering Group, Graz, Austria: Springer-Verlag, 2008), 64, https://doi.org/10.1007/978-3-540-89350-9_6.

38. Rob Waller, "Functional Information Design: Research and Practice," *Information Design Journal* 1, no. 1 (January 1, 1979): 45, https://doi.org/10.1075/idj.1.1.06wal.

39. Redish, *Letting Go of Words*, xxvi.
40. Lisa Meloncon, "Embodied Personas for a Mobile World," *Technical Communication* 64, no. 1 (February 2017): 59.
41. Richard E. Mayer, "Instructional Design as a Form of Information Design," *Information Design Journal* 25, no. 3 (October 22, 2020): 258, https://doi.org/10.1075/idj.25.3.03may.
42. Pettersson, *It Depends*, 16.
43. Mayer, "Instructional Design as Information Design," 258.
44. Schmidt and Etches, *Useful, Usable, Desirable*, 2.
45. David Gibson, *The Wayfinding Handbook: Information Design for Public Places*, 1st ed. (New York, NY: Princeton Architectural Press, 2009), 46.
46. Wakimoto, "Librarians and Graphic Design," 173.
47. Gibson, *The Wayfinding Handbook*, 46.
48. Edward Luca and Bhuva Narayan, "Signage by Design: A Design-Thinking Approach to Library User Experience," *Weave: Journal of Library User Experience* 1, no. 5 (June 2016): 1–25, https://doi.org/10.3998/weave.12535642.0001.501.
49. Gibson, *The Wayfinding Handbook*, 79.
50. Pettersson, *It Depends*, 42.
51. Edward R. Tufte, *Envisioning Information*, 4th ed. (Cheshire, CT: Graphics Press, 1990), 10.
52. Pontis and Babwahsingh, "Improving Information Design Practice," 250.
53. Pettersson, *It Depends*, 17.
54. Maria dos Santos Lonsdale, David Lonsdale, and Hye-Won Lim, "The Impact of Delivering Online Information Neglecting User-Centered Information Design Principles," *Information Design Journal* 24, no. 2 (April 18, 2019): 29, https://doi.org/10.1075/idj.00005.san.
55. Pettersson, "Information Design Principles and Guidelines," 169.
56. Pettersson, "Information Design Principles and Guidelines," 177.
57. Pettersson, *It Depends*, 19.
58. Jelle Strikwerda, Bregje Holleman, and Hans Hoeken, "Designing Pension Communication: Lessons from the Medical Domain," *Information Design Journal* 26, no. 3 (August 4, 2022): 1, https://doi.org/doi.org/10.1075/idj.21011.str.
59. Rune Pettersson, "Information Design–Principles and Guidelines," *Journal of Visual Literacy* 29, no. 2 (2010): 168, https://doi.org/10.1080/23796529.2010.11674679.
60. Jeremy York, "Legibility and Large-Scale Digitization" (Hathi Trust Digital Library, November 2008), 14.
61. Janice C. Redish, "Readability Formulas Have Even More Limitations than Klare Discusses," *ACM Journal of Computer Documentation (JCD)* 24, no. 3 (August 1, 2000): 132, https://dl.acm.org/doi/pdf/10.1145/344599.344637.
62. Pettersson, "Information Design Principles and Guidelines," 173.
63. Katina Rae Stapleton and Katherine Spivey, "Federal Plain Language Guidelines" (The Plain Language Action and Information Network (PLAIN), May 2011), i, https://www.plainlanguage.gov/guidelines/.
64. Redish, "What Is Information Design?," 165.
65. Alvin J. North and L. B. Jenkins, "Reading Speed and Comprehension as a Function of Typography," *Journal of Applied Psychology* 35, no. 4 (August 1951): 228, https://doi.org/10.1037/h0063094.
66. Richard L. Hazlett et al., "Two Studies on How a Typeface Congruent with Content Can Enhance Onscreen Communication," *Information Design Journal* 20, no. 3 (September 1, 2013): 207, https://doi.org/10.1075/idj.20.3.02haz.
67. Sofie Beier et al., "Increased Letter Spacing and Greater Letter Width Improve Reading Acuity in Low Vision Readers," *Information Design Journal* 26, no. 1 (April 28, 2021): 85, https://doi.org/10.1075/idj.19033.bei.
68. Jeffrey Zeldman, "Design Is Typography," December 23, 2015, https://twitter.com/zeldman/status/679727437198929921?s=20&t=RoHX_FVinW-yv3TODE8YYA.
69. York, "Legibility and Digitization," 46.
70. Alice Y. Scales, "Improving Instructional Materials by Improving Document Formatting" (ASEE Southeast Section Conference, American Society for Engineering Education, 2011), 2.
71. Scales, 4.
72. Scales, 6.
73. Scales, 6.
74. Gary Marchionini, "Information-seeking Strategies of Novices Using a Full-text Electronic Encyclopedia," *Journal of the American Society for Information Science* 40, no. 1 (January 1989): 56, https://dl.acm.org/doi/10.5555/65193.65199.
75. Frank Jansen, "How Bulleted Lists and Enumerations in Formatted Paragraphs Affect Recall and Evaluation of Functional Text," *Information Design Journal* 21, no. 2 (November 3, 2015): 159, https://doi.org/10.1075/

idj.21.2.06jan.

76. Miles A. Tinker, "The Relative Legibility of the Letters, the Digits, and of Certain Mathematical Signs," *The Journal of General Psychology* 1, no. 3–4 (1928): 492, https://doi.org/10.1080/00221309.1928.9918022.
77. Scales, "Improving Instructional Materials," 7.
78. Redish, "What Is Information Design?," 166.
79. Tufte, *Visual Display of Quantitative Information*, 56.
80. Emerson, "Visualizing Information for Advocacy," 4.
81. Pettersson, *It Depends*, 30.
82. Redish, "What Is Information Design?," 166.
83. Edward R. Tufte, *The Visual Display of Quantitative Information*, 2nd ed. (Cheshire, CT: Graphics Press, 2001), 191.
84. Pettersson, "Information Design Principles and Guidelines," 176.
85. Shedroff, "Information Interaction Design," 2.
86. Shedroff, 14.
87. Pettersson, 825.
88. Pettersson, "Information Design Principles and Guidelines," 178.
89. Strikwerda, Holleman, and Hoeken, "Designing Pension Communication," 14.
90. Strikwerda, Holleman, and Hoeken, 16.
91. Pettersson, 180.
92. Pettersson, 178.
93. U.S. Dept. of Health and Human Services, "Visual Design Basics."
94. Pettersson, "Information Design Principles and Guidelines," 179.
95. Pettersson, "Seven Theories for ID," 827.
96. David Ferriero, "Finding Aid Type," Lifecycle Data Requirements Guide (LCDRG) (Washington, DC: National Archives and Records Administration, August 15, 2016), https://www.archives.gov/research/catalog/lcdrg/elements/findingtype.html.
97. Richard Pearce-Moses, "Finding Aid," in *SAA Dictionary of Archives Terminology* (Chicago, IL: Society of American Archivists, April 29, 2020), https://dictionary.archivists.org/entry/finding-aid.html.
98. Tiffany Cole, "Characteristics of Successful Finding Aids" (Mid-Atlantic Regional Archives Conference, October 31, 2011), https://www.marac.info/index.php?option=com_content&view=article&id=64.
99. James M. Roth, "Serving Up EAD: An Exploratory Study on the Deployment and Utilization of Encoded Archival Description Finding Aids" (Master's, Chapel Hill, NC, University of North Carolina at Chapel Hill, 2001), 228, https://doi.org/10.17723/aarc.64.2.e687471v304k0u66.
100. Luanne Freund and Elaine G. Toms, "Interacting With Archival Finding Aids," *Journal of the Association for Information Science and Technology* 67, no. 4 (April 27, 2015): 995, https://doi.org/10.1002/asi.23436.
101. Wendy Scheir, "First Entry: Report on a Qualitative Exploratory Study of Novice User Experience with Online Finding Aids," *Journal of Archival Organization* 3, no. 4 (July 18, 2006): 50, https://doi.org/10.1300/J201v03n04_04.
102. Freund and Toms, "Interacting With Archival Finding Aids," 995.
103. J. Gordon Daines III and Cory L. Nimer, "Re-Imagining Archival Display: Creating User-Friendly Finding Aids," *Journal of Archival Organization* 9, no. 1 (May 20, 2011): 6, https://doi.org/10.1080/15332748.2011.574019.
104. Magnus Berg, "A 'Major Technological Challenge': Multi-Level Description and Online Archival Databases," *Emerging Library & Information Perspectives* 4, no. 1 (July 2, 2021): 63, https://doi.org/10.5206/elip.v4i1.12529.
105. Christopher J. Prom, "The EAD Cookbook: A Survey and Usability Study," *The American Archivist* 65, no. 2 (September 1, 2002): 258, https://doi.org/10.17723/aarc.65.2.b3783jr052731588.
106. Prom, 274.
107. Prom, 258.
108. Elizabeth Yakel, Seth Shaw, and Polly Reynolds, "Creating the Next Generation of Archival Finding Aids," *D-Lib Magazine* 13, no. 5/6 (May 13, 2007): 7, https://doi.org/10.1045/may2007-yakel.
109. Clayton McGahee, "Making Special Collections Accessible to Users: Finding Aids" (Lecture, International Association of Law Libraries 36th Annual Course on International Law and Legal Information, Civil Rights, Human Rights, and Other Critical Issues in U.S. Law, Atlanta, GA, October 26, 2017), 32, https://doi.org/10.1017/jli.2018.2.
110. Jane Zhang, "Archival Representation in the Digital Age," *Journal of Archival Organization* 10, no. 1 (May 29, 2012): 45, https://doi.org/10.1080/15332748.2012.677671.
111. Zhang, 49.
112. McGahee, "Making Special Collections Accessible," 37.

113. Rachel Walton, "Looking for Answers: A Usability Study of Online Finding Aid Navigation," *The American Archivist* 80, no. 1 (June 2017): 31, https://doi.org/10.17723/0360-9081.80.1.30.
114. Jody L. DeRidder, Amanda Axley Presnell, and Kevin W. Walker, "Leveraging Encoded Archival Description for Access to Digital Content: A Cost and Usability Analysis," *The American Archivist* 75, no. 1 (April 1, 2012): 146, https://doi.org/10.17723/aarc.75.1.5641v61p422u0u90.
115. Roth, "Serving Up EAD," 216.
116. Prom, "EAD Cookbook Survey," 264.
117. Roth, "Serving Up EAD," 234.
118. Scheir, "Novice User Experience with Online Finding Aids," 60.
119. Scheir, 60.
120. Cole, "Successful Finding Aids."
121. Yakel, Shaw, and Reynolds, "Creating the Next Generation," 7.
122. Rita D. Johnston, "A Qualitative Study of the Experiences of Novice Undergraduate Students with Online Finding Aids" (Master's, Chapel Hill, NC, University of North Carolina at Chapel Hill, 2008), 11, https://cdr.lib.unc.edu/concern/masters_papers/12579x16p.
123. Walton, "Looking for Answers," 45.
124. Roth, "Serving Up EAD," 219.
125. Scheir, "Novice User Experience with Online Finding Aids," 52.
126. Johnston, "Experiences of Novice Undergraduate Students," 2.
127. Peter Pirolli and Daniel M. Russell, "Introduction to This Special Issue on Sensemaking," *Human-Computer Interaction* 26, no. 1–2 (March 16, 2011): 7, https://doi.org/10.1080/07370024.2011.556557.
128. Johnston, "Experiences of Novice Undergraduate Students," 45.
129. Freund and Toms, "Interacting With Archival Finding Aids," 1005.
130. Joyce Celeste Chapman, "Observing Users: An Empirical Analysis of User Interaction with Online Finding Aids," *Journal of Archival Organization* 8, no. 1 (June 10, 2010): 5, https://doi.org/10.1080/15332748.2010.484361.
131. DeRidder, Presnell, and Walker, "Leveraging Encoded Archival Description," 148.
132. Freund and Toms, "Interacting With Archival Finding Aids," 1003.
133. Trudi Van Dyke, "The Nicholas Martini Collection 1931-1991" (William Paterson University, October 14, 2005), https://www.wpunj.edu/library/pdf/MartiniFindingAidUpdated060407.pdf.
134. William Paterson University Office of Institutional Effectiveness, "University Facts" (William Paterson University, n.d.), https://www.wpunj.edu/about-us/university-facts.html.
135. Scheir, "Novice User Experience with Online Finding Aids," 58.
136. Jonathan D. Ericson et al., "End-User License Agreements (EULAs): Investigating the Impact of Human-Centered Design on Perceived Usability, Attitudes, and Anticipated Behavior," *Information Design Journal* 26, no. 3 (May 23, 2022): 200, https://doi.org/10.1075/idj.20018.eri.
137. Sarah Oberbichler, Katharina Gallner-Holzmann, and Theo Hug, "Generous and Inviting Interfaces Revisited: Examples of Designing Visual Structures for Digital Archives," *Information Design Journal* 26, no. 2 (December 7, 2021): 167, https://doi.org/10.1075/idj.20028.obe.
138. DeRidder, Presnell, and Walker, "Leveraging Encoded Archival Description," 146.
139. Freund and Toms, "Interacting With Archival Finding Aids," 996.
140. Freund and Toms, 1006.
141. Ericson et al., "End-User License Agreements," 203.

Benchmarking Librarian Support of Systematic Reviews in the Sciences, Humanities, and Social Sciences

Mê-Linh Lê, Christine J. Neilson, and Janice Winkler

Systematic reviews, along with other types of knowledge synthesis, are a type of research methodology that attempt to find all available evidence on a topic to help answer specific questions. Librarian involvement in systematic reviews is well established in the health sciences, and in recent years there has been growing awareness of, and literature about, librarians outside of health supporting systematic reviews. This study benchmarks librarian support of systematic reviews in the sciences, humanities, and social sciences (SHSS) by looking at the growth of demand for support, the disciplines requesting this kind of librarian support, and the specific types of support needed. It also examines what SHSS librarians need to be successful in this type of work, including administrative support and workload adjustments.

Background

Knowledge synthesis is a collection of secondary research methods that use the systematic collection, evaluation, and integration of previous research to answer a research question. Over 40 types of reviews which go by a variety of names (sometimes used interchangeably by researchers) fall under the knowledge synthesis category, including systematic reviews, scoping reviews, integrative reviews, and meta-analyses.[1] For simplicity, we use the term systematic review (SR) in this article to refer to all types of knowledge synthesis. Different areas of research adopted SR methods at different rates. We often hear anecdotally that SRs originated in the health sciences and have since spread to other disciplines, but this is not accurate. SRs took hold and spread widely in health along with the Evidence-Based Practice Movement in the mid-1990s, but the social science disciplines of education, psychology, and business and economics have continuously used the SR methodology since the mid to late 1970's, albeit without the same fervor seen in the health disciplines.[2]

Library literature has discussed library support for SRs since at least the mid 1990s.[3] Health librarians found roles on SR research teams, likely due to well-established methodological guidelines from the Cochrane Collaboration and the Institute of Medicine, that advise review-

*Mê-Linh Lê is Acting Head, Health Sciences Division at the University of Manitoba, email: me-linh.le@umanitoba.ca; Christine J. Neilson is a Health Sciences Librarian at the University of Manitoba, email: christine.neilson@umanitoba.ca; and Janice Winkler is a Social Sciences Librarian at the University of Manitoba, email: janice.winkler@umanitoba.ca.

ers to consult with an experienced health sciences librarian to ensure a high-quality literature search for the project.[4] As secondary research, the literature search that identifies the existing knowledge base forms the foundation for a SR's data collection. Research has demonstrated that health librarian participation on SR teams is associated with better quality search strategies, lower risk of bias, and better reporting of search methods.[5] Health sciences librarians increasingly contribute to SRs in roles that extend beyond searching, including protocol development, source selection, and teaching.[6] The number of SRs published by health researchers is high and continues to grow. Hoffman and colleagues used the PubMed database to estimate that 80 SRs alone—that is, not including other types of knowledge synthesis reviews like scoping reviews, rapid reviews, etc.,—were published *per day* in 2019. According to their analysis, this publication rate is 20 times greater than it was 20 years earlier, in 2000.[7] Unsurprisingly, the ever-increasing popularity of SRs, as well as the demand for health librarian involvement, have led to concerns over librarian workload and the capacity to provide SR support along with other library services.[8]

While SRs and librarians' role in that research have long occupied a large amount of the health librarianship discourse, SR research—and library support for it—within the humanities, social sciences, and sciences has only started to gain more attention in the library literature relatively recently.[9] Outside of health, the reported involvement of librarians in published SRs in the social sciences and sciences is low. Premji found that two percent of SRs on business topics published from 2014 to 2019 mentioned consulting a librarian, and only one percent credited the librarian with running the search itself.[10] Similarly, Slebodnik et al. found that 3.3 percent of the science and social science SRs they examined reported that a librarian was consulted for the review.[11] Given this evidence, one might assume that even though SRs are being conducted in these fields, librarians are typically not involved in the process. However, we know that librarians regularly work on SRs without receiving credit for their contribution in a resulting publication.[12] Indeed, Kogut and colleagues reported in 2020 that the number of librarian consults at their institution for SRs in education increased over a period of six years, from fewer than 20 consults per year to more than 100 per year. This increase in demand threatened to exceed the library's capacity to provide support services and required training additional education librarians to provide SR support to maintain the service.[13]

We know that a variety of disciplines use SRs, but existing library literature has not provided an overview of librarian involvement in SRs outside of health. This article starts to fill this gap by benchmarking Association of Research Libraries (ARL) and Canadian Association of Research Libraries (CARL) SHSS librarians' involvement in SRs. We collected the data presented here as part of a larger survey regarding SHSS librarian involvement in SRs, their comfort and competence with SR processes, and their perception of library administrators' level of support for SHSS librarian participation in SRs.

Methods

Survey Design

We created an open survey using SurveyMonkey software. The survey included 29 potential questions in total, over 14 "screens" or pages; 23 questions were closed-ended and six open-ended. The survey used conditional logic so that respondents were only asked applicable questions. Responses to all questions were optional. Participants could return to previous questions using the back button on their browser to revisit their responses, if desired. The survey began with background questions about participants' subject responsibilities and years

of experience. Next, it asked respondents details about involvement or requests for involvement with SRs. The next set of questions asked about supports available to respondents in their provision of SR support. The final set of questions related to training involvement and preferences, which we will explore in a forthcoming C&RL article. We piloted the survey with five colleagues, using their feedback to refine questions for clarity. The University of Manitoba Research Ethics Board approved the final questions and study design (JFREB J2020:062). We translated the recruitment materials and survey instrument into French to obtain responses from librarians who speak both official Canadian languages.

Study Population and Recruitment

Academic librarians working at ARL and CARL institutions were invited to participate in the survey in a convenience sample using 22 listservs maintained by Canadian and American library associations. Due to restrictions from our research ethics board, we could not contact libraries or librarians directly and were only permitted to use listservs. We obtained consent through a form at the beginning of the survey. We offered no incentives in exchange for completing the survey.

To be included in the study, respondents had to be librarians at an ARL or CARL institution, who currently, or within the last 5 years, provided direct library services and support to faculty, staff, or students within the SHSS disciplines. The survey did not provide definitions for which subject areas fell under these categories because programs can be interdisciplinary and their points of focus can vary, leaving this open to interpretation. Instead, it provided a definition of health sciences to allow respondents to determine whether their subjects fell under the category of health sciences and allowed them to use their judgment as to the categorization of their liaison areas. The provided definition was as follows:

> For the purposes of this study, the Health Sciences is deemed to include programs or disciplines where health or health care is the primary focus and includes: Allied Health, Dentistry, Dental Hygiene, Nursing, Medicine (including Public Health), Pharmacy, and Rehabilitation Sciences (Physical Therapy, Occupational Therapy, Rehabilitation Therapy). We urge those librarians working with other disciplines that may include strong health components (e.g., kinesiology, psychology) to fill out this survey.

We excluded responses from librarians serving the health sciences, except in cases where respondents supported SHSS disciplines as well as the health sciences.

Librarians who identified themselves as not meeting these inclusion criteria were rejected prior to beginning the survey. The survey was open for seven weeks, with an initial recruitment email sent out in November 2020 and a reminder email sent three weeks later in December 2020. Supplementary materials, including the survey instrument, a list of listservs the survey was sent to, and anonymized data, are available via Open Science Framework (OSF) at https://osf.io/mqxf2.

Data Analysis

We downloaded all responses from SurveyMonkey to a private group in Microsoft Teams, which our institution approved as a secure location for research data. We discarded responses that did not meet the inclusion criteria.

Original data files are stored in password-protected files, accessible only to the authors. We anonymized responses prior to data sharing via OSF in the following manner. We categorized reported liaison areas into four broad disciplinary areas for analysis: Humanities, Sciences, Social Sciences, and Health Sciences (see the OSF page (https://osf.io/mqxf2) for the categorization scheme). As there is no definitive list of liaison areas assigned to a specific discipline, our standard for those liaison areas (e.g., kinesiology) that could be part of different disciplines depending on the institution was to categorize based on where they sat within our home institution. We separated references to specific employers from the data set that we analyzed, and we completely removed them from the shared dataset to maintain participant anonymity. Open-ended responses also presented the possibility of including identifiable information, so we removed them from the shared dataset. The statistical consultant who conducted data analysis signed the required oath of confidentiality, as per research ethics board requirements.

We conducted descriptive analysis for each close-ended question included in the survey. Because of the conditional logic used in the survey and the optional nature of the questions, percentages reported below are based on the number of responses received for individual questions, rather than the total number of individual respondents. We coded responses provided in the free-text questions into broad themes that indicated common issues brought up by multiple respondents.

TABLE 1
Relationships Between Variables Explored through Cross-Tabulations

Variable 1	Variable 2
Did librarians discuss SR support with their administrators?	Number of SRs supported in the last 5 years
Administrators' expressed attitudes towards SR support	Number of SRs supported in the last 5 years
Administrators' expressed attitudes towards SR support	Respondents' interest in participating in SRs
Would librarians be provided with time away from regular duties	Respondents' interest in participating in SRs
Would librarians be provided with a reduction in workload	Respondents' interest in participating in SRs
Librarians' years of experience	Confidence in all aspects of SR support
Librarians' years of experience	Number of SRs supported in the last 5 years

Because we were interested in relationships between specific variables, we created a list of these variables and hired a statistical consultant to complete more sophisticated statistical analysis. Depending on the type of data gathered and quantity of responses, the consultant determined whether and how to best complete the analyses. In most cases, this involved cross-tabulations (shown in Table 1). Spearman's correlation coefficient could be calculated to measure the correlation between variables in two cases: the correlation between librarians' years of experience and their confidence in all aspects of SR support; and their years of experience and the number of SRs they had supported in the last 5 years.

Results

The survey received 379 responses. After filtering for the inclusion criteria, the total number of usable responses was 161. Not all respondents answered every question, so the true num-

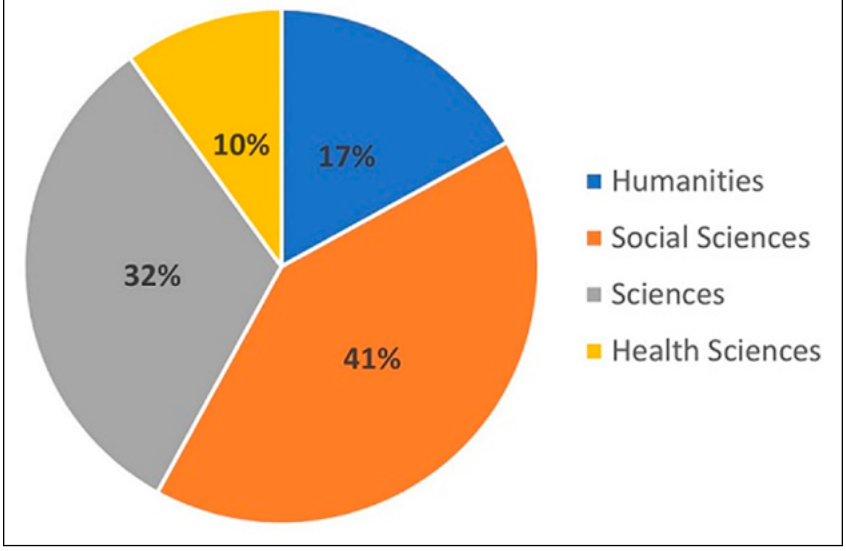
ber of responses is indicated for each individual question below. Anonymized study data is freely available online at OSF.

Due to our recruitment strategy of library listservs, calculating a response rate was not possible, however, respondents were asked to provide the name of their CARL/ARL institution. There were 98 total responses for this question, but multiple respondents came from the same institution. In total, 42 out of 108 US-based ARL institutions (38.8 percent) and 20 out of 31 CARL institutions (64.5 percent) are represented. Of the 42 ARL institutions represented, 38 (90.4 percent) are R1 institutions, three (seven percent) are R2 institutions, and one (two percent) is an M1 institution. Of the CARL institutions, 13 (87 percent) are U15 institutions.[14]

Liaison Area

As outlined above, we categorized liaison areas into four broad disciplinary categories: Humanities, Sciences, Social Sciences, and Health Sciences. Individual respondents may support more than one discipline, faculty, or department. As a result, a total of 521 liaison support areas were indicated; 17 percent were classified as the Humanities; 41 percent were classified as the Social Sciences; 32 percent belonged to the Sciences; and ten percent of liaison areas were classified as the Health Sciences (see Figure 1). As noted in the inclusion criteria above, all included respondents who had liaison responsibilities in the Health Sciences also had liaison responsibilities in non-health disciplines. Of the 151 respondents, 84 (52 percent) had liaison responsibilities within a single category; 58 (36 percent) had liaison responsibilities across two categories; and nine (5.5 percent) had liaison responsibilities across three categories.

FIGURE 1
Respondents' Areas of Responsibility



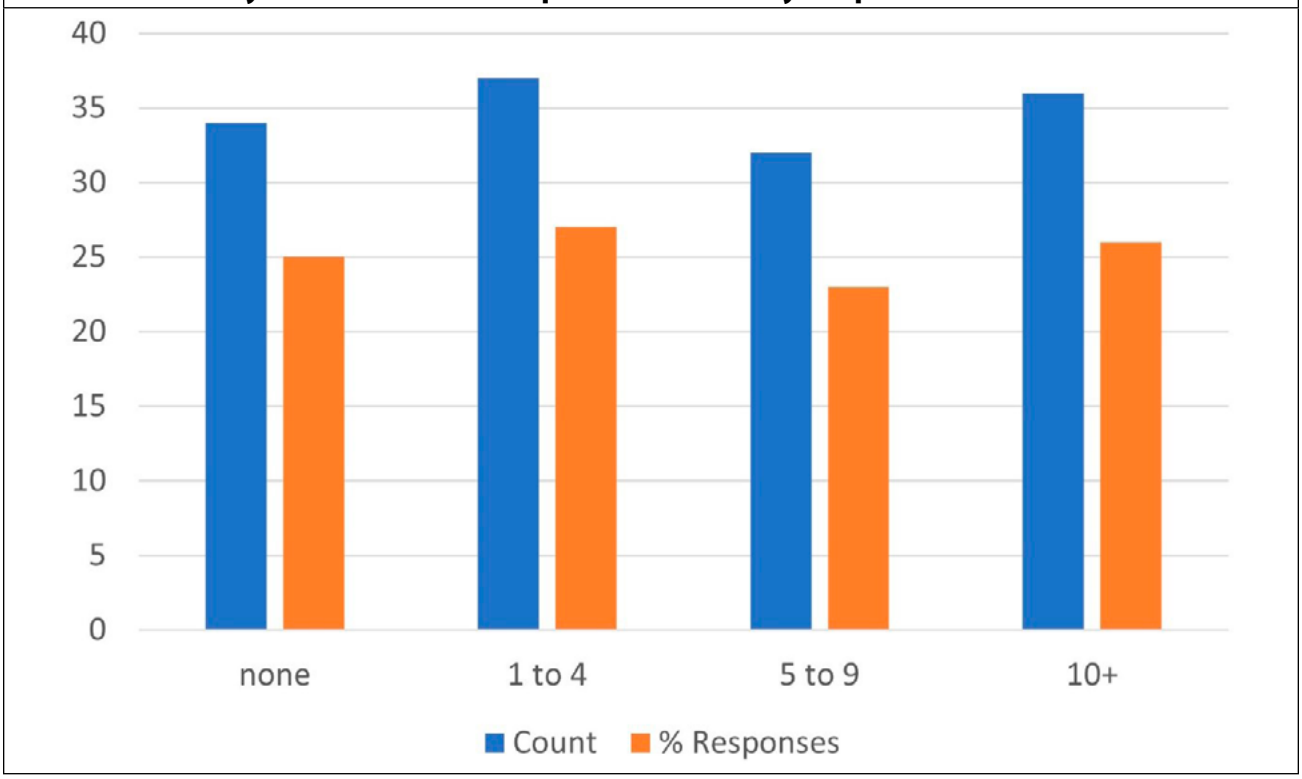
Total Years as a Librarian

To analyze based on years of experience and other factors, respondents were asked how many years they had worked as a librarian. Of the 157 responses, eight (five percent) answered zero to one years; 25 (15.9 percent) answered two to five years; 35 (22.3 percent) answered six to ten years; 20 (12.7 percent) answered 11 to 14 years; and 69 (43.9 percent) answered 15+ years. Spearman's correlation coefficient analyses were done to look for potential relationships between years of experience as a librarian, and either confidence in SR support or the number of SRs completed in the last five years, but no correlations were found.

Systematic Review Requests from The Past 5 Years

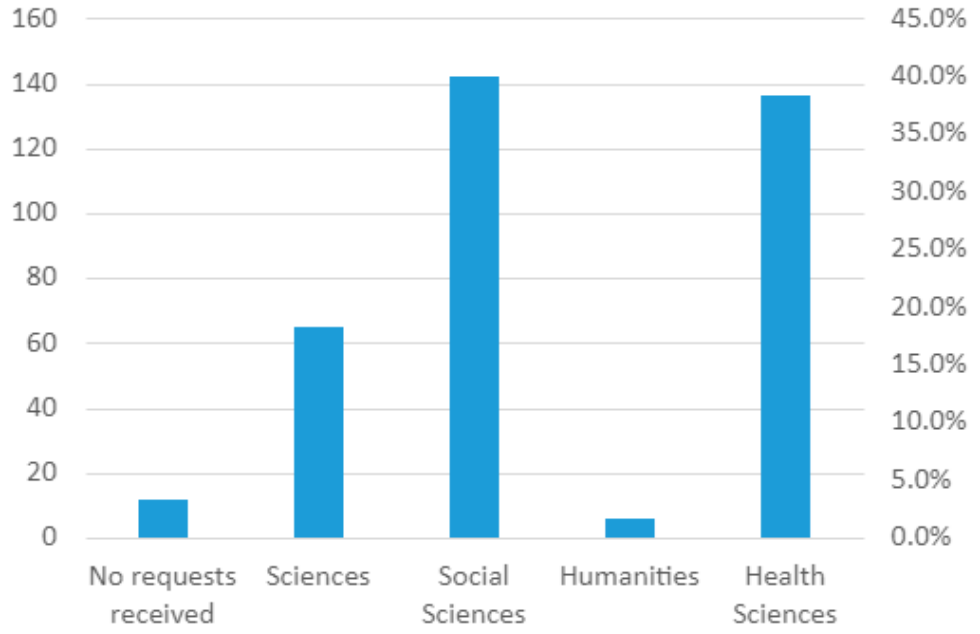
Respondents were asked whether they had been asked by a faculty member, researcher, or student to participate in a systematic review in the past 5 years. Of the 149 responses received for this question, 104 (70 percent) indicated yes; 38 (26 percent) said no, and seven (five percent) were unsure. When asked how many systematic reviews they had supported during the last five years, out of 139 respondents, 34 (25 percent) had not supported any SRs; 37 (27 percent) had supported one to four; 32 (23 percent) supported five to nine; and 26 (25.9 percent) had supported ten or more (see figure 2 below). Statistical analysis between the number of SRs performed in the last five years and years of experience as a librarian, using Spearman's rank correlation, found no significant correlation between the two variables. Our analysis also did not find a difference in the number of reviews completed based on whether respondents were given time away from regular duties to complete reviews. For example, 29.7 percent of those who were not given time away from regular duties to work on SRs completed five to seven SRs in the last five years, compared to 34.5 percent of those who were given time away from their regular duties. When asked if the frequency of SR support requests has changed over the past five years, just over half of respondents (56.6 percent, n=81) indicated that they had experienced a change, 31.5 percent of respondents reported no change, and 11.9 percent were unsure if there had been a change in the frequency of requests. Of the 81 respondents who reported a change, one reported a decrease in the number of requests, while the remaining 80 reported an increase in the number of requests received. This translates to 55.9 percent of respondents indicating an increase in the frequency of requests overall.

FIGURE 2
Distribution of Systematic Review Requests Received by Respondents in the Past Five Years



Most respondents (93, or 67.9 percent) received requests for SR support from only one of the broad disciplinary areas. Close to one third of respondents (44, or 32.1 percent) indicated that they received requests from a combination of two or three disciplinary areas. Over the past five years, Social Science and Health Science disciplines were the main source of requests for SR support, making up 39.3 percent and 37.7 percent of total requests respectively (see figure 3). Respondents received 65 requests (18 percent) from Science disciplines, and six requests (1.7 percent) from Humanities disciplines.

FIGURE 3
Responses to the Question "In The Past Five Years Which Disciplines Have Requested Systematic Review Support From You?"



Note: Answers have been classified into broad disciplinary areas

Types of Systematic Review Support

Respondents were asked to indicate the type of SR support they had provided, with the ability to select as many support types as applicable. As Table 2 shows, the areas with greater number of responses were those that are typically seen as an area of expertise for librarians, such as search development, search execution, and search translation. Consultation, which was not explicitly defined, had the highest number of responses.

TABLE 2
Specific Types of SR Support Provided by Respondents

Type of Support Provided	Count	Percentage (%)
Consultation	116	83.5
Search strategy development	113	81.3
Search strategy execution	85	61.2
Search strategy translation	75	54

TABLE 2 Specific Types of SR Support Provided by Respondents		
Type of Support Provided	**Count**	**Percentage (%)**
Research question development	61	43.9
Protocol development	62	44.6
Deduplication	61	43.9
Write up of methods for publication	54	38.8
PRISMA or equivalent	45	32.4
Search update	42	30.2
Peer review of search strategy (e.g., PRESS)	25	18
Stage 1 screening	23	16.5
Project management	21	15.1
N/A	20	14.4
Stage 2 screening	13	9.4
Write up of other parts of the study for publication	10	7.2

Acknowledgement

Respondents were asked to indicate the types of acknowledgement they had received for their support of SRs; respondents could select as many answers as were applicable. A thank you via email/in-person/phone had the most responses (70; 50.7 percent); followed by: co-authorship of a paper or presentation (55; 39.3 percent); no acknowledgement (42; 30 percent); mention of working with a librarian in a paper or presentation (29; 30 percent); mention of working with a librarian by name in a paper or presentation (29; 20.7 percent); mention of working with a librarian by name in the formal acknowledgement of a paper or presentation (26; 18.6 percent); co-investigator/collaborator/investigator status on a grant application (seven; five percent).

TABLE 3 Specific Types of Acknowledgement Received for SR Support		
Type of Acknowledgment	**Count**	**Percentage (%)**
Thank you via email/in-person/phone, etc.	71	50.7
Co-authorship of a written paper or presentation	55	39.3
No acknowledgment	42	30
Mention of working with a librarian in a written paper/grant/presentation	29	20.7
Mention of working with a librarian BY NAME within the text of a written paper/grant/presentation	29	20.7
N/A	27	19.3
Mention of working with a librarian by name within formal Acknowledgement section of a written paper/grant/presentation	26	18.6
Other	9	6.4
Co-investigator/collaborator/investigator status on a grant application	7	5

Administration Communication, Administration Attitude, and Librarian Workload

Most respondents discussed participation in SR projects with their library administrators or otherwise received communication from administration on the topic (100; 68 percent). Of those who had discussed SR participation with their administration, 80 (54.4 percent) said administration was supportive of participation; 34 (23.1 percent) believed their administration to be neutral on participation; 5 (3.4 percent) perceived their administration as discouraging of participation; 21 (14.3 percent) were unsure of administration attitude; and seven (four percent) marked not applicable.

Statistical analysis shows that administrative attitudes toward SRs (discouraging, neutral, or supportive) did not appear to correlate with the actual numbers of reviews the respondents completed over the last five years. Of those who had had discussions with administration, 27.1 percent of respondents had supported five to nine SRs; 11.5 percent supported ten to 14 SRs; and 13.5 percent supported 20 or more SRs. The majority of respondents (52.6 percent) with no discussion with their administration did not support any systematic reviews.

The survey included two questions related to workload adjustments. It asked respondents whether they were or would be given time away from regular duties, and/or granted a reduction in their existing workload to accommodate the work involved in participating in SRs. A small number indicated that they were, or would be, granted protected time (29 or 19.9 percent) and/or a reduction in workload (nine or 6.1 percent). The majority indicated that they would not receive time away from regular duties (77 or 52.7 percent) and/or reduction in workload (98 or 66.7 percent) to participate in SRs. The remaining respondents were unsure if these accommodations were possible (24 or 16.4 percent for time away from regular duties; 22 or 15 percent for reduction in workload) or responded with "not applicable."

Interest in Supporting Systematic Reviews

When asked to rank their interest in participating in SRs on a 100 point scale, the largest number of respondents indicated a high level of interest (see figure 4 below). Forty-three (28 percent) respondents indicated an interest level ranging from 91 to 100; 24 (16 percent) participants entered a ranking in the 81 to 90 range; and 25 (16 percent) entered a ranking in the 71 to 80 range. Thirty-six participants (23 percent) indicated interest in the middle range (41 to 70), and 26 participants (17 percent) registered their interest in the low range (zero to 40).

Statistical analysis indicated that respondents' interest in working on SRs was not related to administration attitude towards support of SRs, whether that attitude is perceived as discouraging, neutral, supportive, or if the respondent was unsure of administration attitude. However, the data suggests that interest in supporting SRs is highest among those who are now, or would be, given time away from regular duties (average of 88 on a 100 point scale) or have their workload adjusted to accommodate SR support (average of 91). Those who are not, or would not be, given time away from regular duties showed an average interest of 68, and those who do not have other aspects of their workload decreased showed an average interest of 70. Respondents who indicated they were unsure if they are or would be eligible for time away from regular duties indicated an average interest score of 56, and those who were unsure if they are eligible to receive a reduction in other aspects of their workload indicated an average interest score of 66.

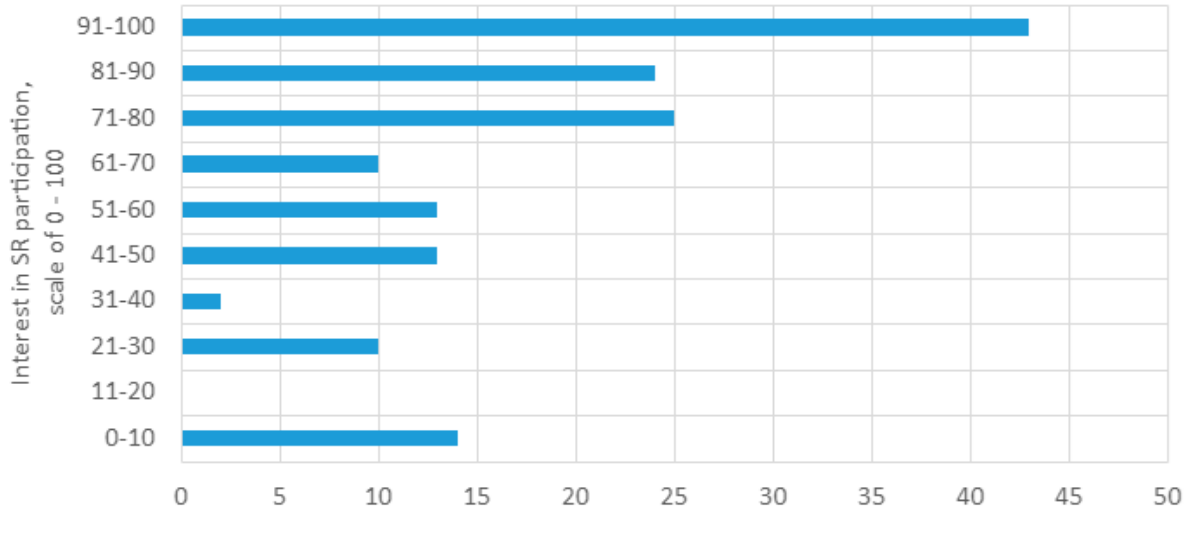
Additional Comments

Respondents received the option to provide additional comments following the closed-ended questions. A common theme from the respondents' comments was the need to create new positions for SR support, such as a SR librarian, or evidence synthesis librarian, to meet increasing demand for SR support coming from areas outside of the health sciences. Unfortunately, this increase in demand is not always matched with the commensurate support needed. One comment stated:"The requests are repeated and overwhelming […] We've talked about how to support this need but honestly, without a reduction in workload, it's just unsupportable." One reason noted for an increase in SR requests was the disruption of in-person research caused by the COVID-19 pandemic. A different respondent noted that, even though researchers are increasingly turning to SRs, faculty do not necessarily understand what SR research involves, writing: "They don't really know what it [systematic reviews] means but it sounds good and could get their work published so they want it – whatever it is." Lack of support for librarians participating in SRs, whether through training or workload reduction in other

TABLE 4
Supports Provided for SR Work With Respondents' Interest in Providing SR Support

Supports provided for participation in SR work	Mean interest reported (1-100)	Standard Deviation
Time away from regular duties		
Yes (n=29)	88	16
No (n=77)	68	28
Unsure (n=24)	56	33
Adjustment in workload		
Yes (n=9)	91	16
No (n=98)	70	29
Unsure (n=22)	66	29

FIGURE 4
Number Of Respondents Reporting Interest In Participating In Systematic Reviews On A Scale Of Zero To 100, Where A Higher Score Indicates A Higher Level Of Interest



areas, was also emphasized in the comments. While the ability to support SRs was seen as an important way to demonstrate value to faculty and the academic community, the need for educating library administrators on this type of research and its importance was also noted.

While most additional comments on SRs were positive, notwithstanding noted areas requiring improvements, some respondents did not see a need for this kind of support in their liaison areas, stating that their faculty do this type of work themselves. For example, one respondent stated: "The other, more newfangled aspects of systematic reviews that you mention here I'd not heard of and do not seem to apply yet to doing comprehensive literature reviews in my collection areas." For some librarians, there was simply a lack of interest in this type of work, stating that "If poorly implemented, [SRs] represents one of the lesser inspiring and also quite robotic activities in which librarians might engage."

Discussion

Librarian support of SRs in the SHSS disciplines can be viewed positively in many ways, including acknowledgement of librarian expertise, demonstrating value in library services, and relationship building between research teams and librarians. Our data indicate that SHSS librarians are facing an increasing number of requests for SR support, and many are rising to the challenge regardless of the number of years of experience as a librarian they may have. This supports what the growth in literature over the last ten years,[15] development of robust training programs,[16] and anecdotal evidence have indicated. Broadly, disciplines within the social sciences are more likely to request librarian support than the sciences or humanities. This aligns with a strong history of SR in psychology, business, and education.[17]

Consultation was the most frequently reported form of SR support requested. Consultation, whether provided over email, video, or in-person, is a chance for a librarian to provide guidance and answer any questions that a research team may have. Consultation as the most common form of SR support aligns with a thank you via email/in-person/phone as the most common form of acknowledgement (50.7 percent). The percentage of librarians who receive co-authorship of a paper or presentation (39.3 percent) is high and similar to numbers reported by health science librarians,[18] although it does not reach the numbers reported by health science library administrators.[19] As noted by Ross-White[20] authorship on SRs is an important indicator that researchers value librarians' work and contributions to the team. Consults alone do not typically lead to authorship, but can be formally acknowledged in a paper; they are often the first step to a librarian taking on a more substantive role on a SR team, such as search strategy development, translation, and execution, which were the next most frequently requested support that respondents reported. These types of requests are unsurprising, as they draw upon fundamental areas of librarian expertise.

One of the most striking themes to emerge from our data is the high level of interest in supporting SRs among SHSS librarians. Library administrators should recognize that there is both demand for SR support in SHSS and an appetite by librarians to take on this work. Administrators who wish to encourage SR services should note that librarians who reported knowing that they had organizational supports in place reported higher levels of interest in participating in SR work, while those who were unsure of the organizational supports available to them expressed a similar level of interest to those who said they had no organizational supports. Providing resources for SHSS librarians involved in SR research is important to prevent burnout[21] and to ensure the long-term viability of offering this type of service to SHSS

researchers and faculty. As such, consideration of SR services and organizational support for SHSS librarians should be a part of library planning.

Librarians and administrators must exercise caution regarding scalability and capacity to provide SR services. Providing high levels of SR support requires extensive training and education by the librarian and a significant time investment. One study found that the average time a librarian spent on a single SR is 26.9 hours (median 18.5 hours).[22] Adding support for multiple SRs to an already full workload can lead to burnout; one study found that SR-related burnout is a problem for health sciences librarians, but those who are assigned to spend more than 80 percent of their time on SRs had lower personal burnout scores.[23] This suggests that employing dedicated SR librarians or providing time away from regular duties for SR work can help reduce burnout. SHSS librarians who are already involved in SRs or are interested in it for the future, must pay attention to how they manage their workload. SHSS administrators or department heads planning on offering or expanding SR services in the SHSS disciplines should carefully examine existing workflow and capacity to ensure there is adequate staffing, tools, and resources to ensure the long-term feasibility of that type of service. Researchers in SHSS disciplines where SR are a new research methodology will likely need even higher levels of librarian support and training to successfully complete the process.

Limitations and Future Research

The findings of this research are limited by the small sample size and use of convenience sampling. Due to restrictions from our research ethics board, we were unable to directly contact ARL or CARL libraries or librarians to increase the sample size and instead had to rely on the use of listservs to recruit respondents. Based on the actual number of ARL or CARL academic libraries that support institutions or researchers in the SHSS disciplines, we believe there is a larger number of librarians doing this work than is indicated in this study. The original design of this study only included librarians working at ARL or CARL institutions to narrow a suspected large potential number of respondents; future research should include librarians working at any academic institution who support SRs in the SHSS disciplines. Another possible limitation is that SHSS librarians who have not been asked to support SRs in their liaison areas, or who serve disciplines where this methodology is rarely used, may not have considered this survey on SRs relevant, which may have biased the sample. However, the fact that 26 percent of respondents had not yet been asked to participate in a SR demonstrates that at least some librarians who had not completed a SR filled out the survey, providing a small amount of information about the views of this group.

Another limitation of this study relates to data analysis. Most of this analysis is descriptive, as more sophisticated analysis was not possible with the data available. While this prevents us from drawing firm conclusions about the data that can be generalized more broadly, it does provide a snapshot of SR support in an area that has been under-examined to date.

A final limitation is the process used to categorize a librarian's liaison area(s) into one or more broad categories (e.g., a librarian who listed psychology as their liaison area was categorized into "Social Sciences"). This is an imperfect process and is open to interpretation, particularly in multidisciplinary areas (e.g., biomedical computing) or those that may be classified differently at different institutions (e.g., kinesiology). However, broad categorization was necessary to maintain participant anonymity and facilitate data analysis. Fortunately, liaison areas that could potentially have been assigned to multiple disciplines were in the minority

(see OSF for the categorization scheme). Future surveys could, instead of asking respondents to list their liaison areas in free-text format, provide standardized liaison areas or disciplines and ask respondents to select those they feel most closely aligns with their liaison areas.

This study is among the earliest investigations into SHSS librarians' involvement in SRs. There are many potential avenues of research and inquiry into library support of SRs in the SHSS disciplines. Examples could include subsequent benchmarking studies that document growth in demand for library support and librarian participation, development of discipline-specific SR standards, and documentation of time spent supporting SHSS SRs.

Conclusion

This study provides evidence that demand for SR support from librarians serving the sciences, humanities, and social sciences at ARL and CARL institutions has increased, and that many SHSS librarians are rising to the challenge of providing that support. However, SR support is a labor-intensive endeavor that is not sustainable as an "off the side of the desk" activity. Library administrators should consider both the benefits and challenges of providing such service and should take steps to adequately provide for the development of library SR expertise and ensure the scalability of library services.

Supplemental Material

All supplemental material, including survey instrument, listservs contacted, and anonymized data are available on the Open Science Framework at https://osf.io/mqxf2.

Acknowledgements

This study was funded by the Canadian Association of Research Libraries' (CARL) *Research in Librarianship* grant, and we are extremely grateful for their support. We are also indebted to our colleagues at the University of Manitoba for testing early forms of the survey. We acknowledge the help of Loring Chuchmach from the George & Fay Yee Centre for Healthcare Innovation (CHI) for statistical consulting services.

Contributorship

M.L., C.J.N, & J.W. contributed to the design and implementation of the research, the analysis of the results, and the writing of the manuscript.

Notes

1. Anthea Sutton, Mark Clowes, Louise Preston, and Andrew Booth, "Meeting the review family: exploring review types and associated information retrieval requirements." *Health Information & Libraries Journal* 36 (2019): 202-222. https://doi.org/10.1111/hir.12276. David Gough, James Thomas, and Sandy Oliver, *An Introduction to Systematic Reviews* (London: Sage Publications, 2nd Ed., 2017).

2. Laura Sheble "Macro-Level Diffusion of a Methodological Knowledge Innovation: Research Synthesis Methods, 1972–2011." *Journal of the Association for Information Science and Technology* 68, no. 12 (2017): 2693-2708.

3. Thomas L. Mead. "Librarian Participation in Meta-Analysis Projects." *Bulletin of the Medical Library Association* 83, no. 4 (1995): 461-64; Jack T. Smith, "Meta-Analysis: The Librarian as a Member of an Interdisciplinary Research Team." *Library Trends* 45, no. 2 (1996): 265-79.

4. Institute of Medicine (US) Committee on Standards for Systematic Reviews of Comparative Effectiveness Research. *Finding What Works in Health Care: Standards for Systematic Reviews*. Eds. J. Eden, L. Levit, and A. Berg, (Washington (DC): National Academies Press, 2011). Lefebvre, Carol, Julie Glanville, Simon Briscoe, Anne Littlewood, Chris Marshall, Maria-Inti Metzendorf, Anna Noel-Storr et al. "Searching for and selecting studies." *Cochrane Handbook for Systematic Reviews of Interventions* (2019): 67-107.

5. Jonathan B. Koffel, "Use of Recommended Search Strategies in Systematic Reviews and the Impact of Librarian Involvement: A Cross-Sectional Survey of Recent Authors." *PLoS ONE* 10, no. 5 (2015): 1-13; Deborah Meert, Nazi Torabi, and John Costella. "Impact of Librarians on Reporting of the Literature Searching Component of Pediatric Systematic Reviews." *Journal of the Medical Library Association* 104, no. 4 (2017): 267-77; Melissa L. Rethlefsen, Ann M. Farrell, Leah C. Osterhaus Trzasko, and Tara J. Brigham. "Librarian Co-Authors Correlated with Higher Quality Reported Search Strategies in General Internal Medicine Systematic Reviews." *Journal of Clinical Epidemiology* 68, no. 6 (2015): 617-26.

6. Angela J Spencer and Jonathan D. Eldredge. "Roles for Librarians in Systematic Reviews: A Scoping Review." *Journal of the Medical Library Association* 106, no. 1 (2018): 45-56.

7. Falk Hoffmann, Katharina Allers, Tanja Rombey, Jasmin Helbach, Amrei Hoffmann, Tim Mathes, and Dawid Pieper. "Nearly 80 Systematic Reviews Were Published Each Day: Observational Study on Trends in Epidemiology and Reporting over the Years 2000-2019." *Journal of Clinical Epidemiology* 138 (2021):1-11.

8. Michelle R. Demetres, Drew N. Wright, and Antonio P. DeRosa. "Burnout among Medical and Health Sciences Information Professionals Who Support Systematic Reviews: An Exploratory Study." *Journal of the Medical Library Association* 108, no. 1 (2020): 89-97; Genevieve C. Gore and Julie Jones. "Systematic Reviews and Librarians: A Primer for Managers." *Partnership: The Canadian Journal of Library and Information Practice and Research* 10, no. 1 (2015).

9. Allison Campbell, Brian Taylor, Jessica Bates, and Una O'Connor-Bones. "Developing and Applying a Protocol for a Systematic Review in the Social Sciences." *New Review of Academic Librarianship* 24, no. 1 (2018): 1-22; Karen Chapman, "Characteristics of Systematic Reviews in the Social Sciences." *The Journal of Academic Librarianship* 47, no. 5 (2021): 102396; Megan Kocher and Amy Riegelman. "Systematic Reviews and Evidence Synthesis: Resources Beyond the Health Sciences." *College and Research Libraries News* 79, no. 5 (2018): 248-52; Ashlynn Kogut, Diana Ramirez, and Margaret J. Foster. "Systematic Review Training Model for Education Librarians: A Case Study." *New Review of Academic Librarianship* (2020): 1-22; Diana Papaioannou, Anthea Sutton, Christopher Carroll, Andrew Booth, and Ruth Wong. "Literature Searching for Social Science Systematic Reviews: Consideration of a Range of Search Techniques." *Health Information and Libraries Journal* 27, no. 2 (2010): 114-22; Zahra Premji, "Exploring the Involvement of Librarians in Systematic Reviews in Business" (Ontario Library Association Super Conference, January 2020), https://www.olasuperconference.ca/wp-content/uploads/2020/01/Premji_OLA2020_BusinessLibrarianInvolvement_StandardPoster.pdf [accessed 21 April 2022]; Zahra Premji, Ryan Splenda, and Sarah Young. "An Exploration of Business Librarian Participation in Knowledge Synthesis Reviews." *College & Research Libraries;* Vol 83, No 2 (2022): 314; Amy Riegelman and Megan Kocher. "A Model for Developing and Implementing a Systematic Review Service for Disciplines Outside of the Health Sciences Amy." *Reference & User Services Quarterly* 58, no. 1 (2018): 22-27; Nedelina Tchangalova, Eileen G. Harrington, Stephanie Ritchie, and Sarah Over. "Working across Disciplines and Library Units to Develop a Suite of Systematic Review Services for Researchers." *Collaborative Librarianship* 11, no. 4 (2019): 267-81.

10. Premji, "Exploring the Involvement,"11. Maribeth Slebodnik, Kevin Pardon, and Janice Hermer. "Investigating Systematic Reviews Outside Health Sciences." *ALA Annual STS Poster Session 2018*. (2018), https://repository.arizona.edu/handle/10150/628518 [accessed 21 April 2022].

12. Koffel, "Use of Recommended," 1-13.

13. Kogut, Ramirez, and Foster. "Systematic Review Training," 1-22.

14. "U15 - Group of Canadian Research Universities." U15, accessed April 29, 2022, https://u15.ca/.

15. Smith " The Librarian as a Member," 265-79. Diana Papaioannou, Anthea Sutton, Christopher Carroll, Andrew Booth, and Ruth Wong. "Literature Searching for Social Science Systematic Reviews: Consideration of a Range of Search Techniques." *Health Information and Libraries Journal* 27, no. 2 (2010): 114-22; Premji, Splenda, and Young. "Business Librarian Participation," 314; Slebodnik, Pardon, Hermer. "Systematic Reviews Outside Health Sciences." *ALA Annual STS Poster Session* (2019). Jennifer Evans and Pauline Benefield. "Systematic Reviews of Educational Research: Does the Medical Model Fit" *British Educational Research Journal*, 27, no. 5 (2001): 37-41; Melanie Nind. "Conducting Systematic Review in Education: A Reflexive Narrative." *London Review of Education* 4, no. 2 (2006): 183-95; Mark Petticrew and Helen Roberts. *Systematic Reviews in the Social Sciences*. (Oxford, UK: Blackwell Publishing Ltd, 2006); Robin A Paynter, "Evidence-Based Research in the Applied Social Sciences." *Reference Services Review* 37, no. 4 (2009): 435-50; Georgina Guilera, Maite Barrios, and Juana Gómez-Benito. "Meta-Analysis in Psychology: A Bibliometric Study." *Scientometrics* 94, no. 3 (2013): 943-54; Jacqueline Davis, Kerrie Mengersen, Sarah Bennett, and Lorraine Mazerolle. "Viewing Systematic Reviews and Meta-Analysis in Social Research through Different Lenses." *SpringerPlus* 3, no. 1 (2014): 1-9.

16. Kogut, Ramirez, and Foster. "Systematic Review," 1-22.

17. Margaret Foster, T. Derek Halling, and Catherine Pepper. "Systematic Reviews Training for Librarians." *Journal of EAHIL* 14, no. 1 (2018): 4-8. 17. Sheble "Macro-Level Diffusion," 2693-2708.

18. Amanda Ross-White. "An Environmental Scan of Librarian Involvement in Systematic Reviews at Queen's University: 2020 Update." *Journal of the Canadian Health Libraries Association* 42, no. 2 (2021): 110-117.

19. Robin Desmeules, Sandy Campbell, and Marlene Dorgan. "Acknowledging Librarians' Contributions to Systematic Review Searching." *Journal of the Canadian Health Libraries Association* 37, no. 2 (2016).

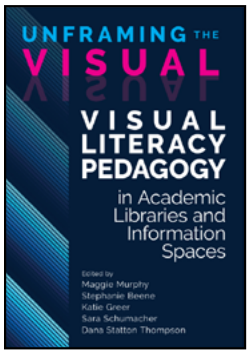
20. Ross-White. " Environmental Scan," 110-117.

21. Demetres, Wright, and DeRosa. "Burnout among Medical," 89-97.

22. Krystal Bullers, Allison M Howard, Ardis Hanson, William D Kearns, John J Orriola, Randall L Polo, and Kristen A Sakmar. "It Takes Longer Than You Think: Librarian Time Spent on Systematic Review Tasks." *Journal of the Medical Library Association:* 106, no. 2 (2018): 198.

23. Demetres, Wright, and DeRosa. "Burnout among Medical," 89-97.

Unframing the Visual: Visual Literacy Pedagogy in Academic Libraries and Information Spaces, Maggie Murphy, Stephanie Beene, Katie Greer, Sara Schumacher, and Dana Statton Thompson (eds.), ACRL, 2024. 452p. Softcover, $108.00. 9780838939918



Visual information is everywhere. Not only that, but visuals are multidisciplinary, making it imperative to be able to analyze, examine, modify, read, and question them both as a part of everyday life, as well as in higher education. While for years librarians have been steeped in the concept of teaching information literacy, *Unframing the Visual* encourages us to consider the importance of *visual* literacy education to students, faculty, and the overall campus community.

This extensive anthology was inspired by the 2022 Association of College and Research Libraries (ACRL) Framework for Visual Literacy in Higher Education: Companion Document to the Framework for Information Literacy for Higher Education, also known as "the VL Framework." That title is a mouthful, and the size of this book is likewise expansive; it includes an abundant collection of research and experience related to visual literacy pedagogy.

The book's five editors were members of the task force that authored the VL Framework. Chapter authors are just as diverse as the visual literacy tools they describe, coming from a variety of library types, locations, and backgrounds. Many different areas of librarianship are represented, including instruction, museums, special collections, DEIA, and more, making it an ideal addition to any academic library's shelf.

The book is divided into four sections that echo the VL Framework's themes: "Participating in a Changing Visual Information Landscape," "Perceiving Visuals as Communicating Information," "Practicing Visual Discernment and Criticality," and "Pursuing Social Justice through Visual Practice." Each section includes an introduction written by an editor and VL Framework author, plus six peer-reviewed chapters that relate to the theme. Multiple chapters make the connection between visual literacy and information literacy, noting that one does not exclude the other. As stated in Chapter 4, "to be information literate is to be visually literate" (p. 54).

Chapter topics include case studies to enhance understanding of the what, how, and why of using visual literacy in the academic library profession. While the book is lengthy at over 450 pages, it includes valuable takeaways and, of course, appealing visual examples throughout the text. All chapters include copious references and bibliographies for further reading, allowing serious researchers to delve deeper. Some chapters also include supplementary materials—such as lesson plans, worksheets, classroom discussion questions, and survey instruments—that will prove helpful to those wanting to recreate these authors' successes.

As an outreach librarian who coordinates my academic library's social media accounts, I found the first three chapters, which focus on "remix media" and online trends, especially relevant to my mission to reach and engage my users online. In Chapter 3, the writers take this concept a step further by encouraging libraries to use social media to not only reach their audiences, but also to educate them. Visual copyright, plagiarism, appropriation, and attribu-

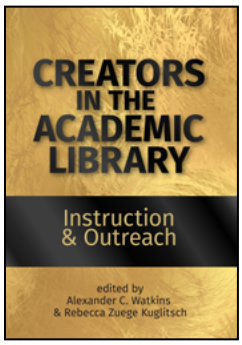
tion are also mentioned throughout the book, which are topics most all library marketing professionals struggled with.

Part 2 of the book focuses on literary instruction and inclusivity, and includes topics such as data visualization and visual rhetoric. Authors assert that visual literacy can be used to teach students about academic honesty and integrity, skills that will help them throughout their college career and into the professional environment. Subject librarians will be happy to know that a wide variety of disciplines are used as examples throughout the book's numerous case studies, including the humanities, social sciences, health sciences, and interdisciplinary studies.

The book wisely extends the definition of "visuals" to include more than just images; discussions on a variety graphic formats—including infographics, comics, memes, diagrams, videos, and more—augment the book. One of the more eye-opening chapters was Chapter 11, "Collaborative Approaches to Teaching and Building Visual Literacies," written by librarians from UCLA who used visual modalities to "frame library instruction and create instructional objects" to better engage learners. This chapter will be especially helpful for librarians who may not have the skills, nor time, to create new visual resources, and who may instead wish to "reframe" existing resources through collaboration.

The importance of evaluating visuals is a prominent theme, especially in Part 3, which is dedicated to the VL Framework theme "learners practice visual discernment and criticality" (p. 151). Just as librarians used to champion the CRAAP test to evaluate text resources, several chapters in this section promote the need for incorporating critical visual literacy into library instruction. Visuals are not without their issues, however, and several chapters focus on accessibility concerns. In Chapter 19, "What We Aren't Seeing: Exclusionary Practices in Visual Media," authors Smith and Malinowski point to the need for visual media inclusivity via critical design, as well as the role of information professionals in addressing exclusionary practices. "Learning and unlearning are necessary," they state, "and we as a profession should continue to position ourselves to evolve accordingly" (p. 339).

Unframing the Visual contains a vast amount of information and numerous case studies on the importance of libraries teaching, using, learning, and evaluating visual literacy. While a casual reader may balk at the book's sheer size, any librarian—particularly those who work in user engagement or instruction—will be sure to find a chapter that resonates with them. If nothing else, librarians will have their eyes opened to the sheer number of visuals that surround us daily, and, hopefully, will reconsider how academic libraries can better utilize visual to inform and connect with learners. — *Maria Atilano, Student Engagement Librarian, University of North Florida*



Creators in the Academic Library: Instruction and Outreach. Alexander C. Watkins, & Rebecca Zuege Luglitsch (eds.). ACRL, 2023. 312p. Softcover, $72.00. 9780838939703

Creators in the Academic Library: Instruction and Outreach, edited by Alexander C. Watkins and Rebecca Zuege Kuglitsch, is an expansive survey of research, instruction, and engagement collaborations between librarians, departmental faculty, and students in higher education institutions. Using numerous case studies, the editors compile a work featuring the librarian advancement of curriculum in service to students completing creative

deliverables in degree programs. Spanning environments from creative writing courses, to engineering projects, to theater research in a community college, the author contributions represent a variety of methods, resources, and spaces giving insight for fellow library employees focusing on similar goals.

Using a similar outline and layout through the entire book, each chapter presents thorough methodologies covering the courses selected, projects completed, and student populations engaged. Multiple chapters adhere to the ACRL information literacy framework, explaining how either the entire foundation or individual frames apply to student learning outcomes and instruction processes. As a result, the work acts effectively as a teach-the-teacher resource. The edited volume presents a library as three realms: a space, a resource, and a service. Chapters highlighting the approach include: "The Web is Your Canvas," by Carmen Cole, in which library space is used for the "Code for Her" program, providing female students a calm, supportive, and nonjudgmental physical space, and "Library Instruction That Sticks," by Tess Colwell and Jessica Quagliaroli, where a group of Yale architecture graduate students—comprised of nontraditional, multigenerational, and international students—engage in multiple library instructional sessions.

A particular strength of the book is the variety in how contributors approach their topics. A book on "creators" in the academic library can easily fall into traditional parameters of makerspaces and traditional artistic curriculum; however, these case studies highlight the creativity of the librarians themselves, emphasizing the method they used to select which courses to engage with in the first place. "Drawing from Life," by Lane Glisson, uses a constructivist pedagogy connecting students with their new content combined with prior knowledge experiences, while "Contemporary Research Methods for Creatives," by Kristina Keogh and Nicole Caron, highlights an embedded librarianship approach as both an outreach and an instructional tactic. "Library as Portal," by Carla-Mae Crookendale and Andrea Kohashi, examine special collections, not solely as a resource, but as a means for generating inspiration, highlighting the role of serendipity in the creative research process. STEM topics including computer program coding and patent research, showcase the creative process in research within fields too often misconstrued as purely mechanical and formulaic.

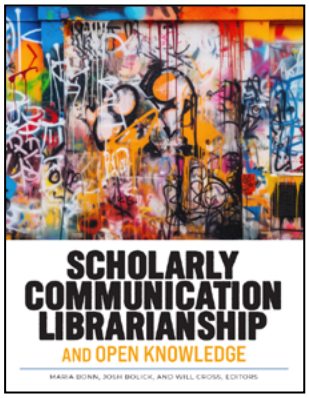
Another underlying theme of *Creators in the Academic Library* is formalizing creative instruction and engagement methodologies. "A Librarian's Guide to Helping Creators Understand and Use Patent Information," by Rachel Knapp, focuses on teaching students the application of resources in the creation of design patents, with "Drawing from Life" moving outside of the university setting and into community college librarianship. Theater students study materials connected with the context of the productions' thematic timeframes. The attention on a variety of instruction approaches is a distinct hallmark of the book.

The book also considers a wide variety of undergraduate and graduate student populations. The instruction styles and assessment practices vary from chapter to chapter and show a design with a customized framework approach for addressing student needs. "Creative Research and Digital Visual Literacy," by Giana Ricci, teaches students to make the most of both popular search engines and research databases in unison, while also addressing copyright factors with creative commons licensing. "Outside of the Digital Dark Room," by Maggie Murphy and Kelley O'Brien, discusses engaging students in a practice of reading comprehension and information synthesis during the creation of art projects. Additionally, "Embracing Messiness: Inspiring and Creating Improvisers," by Liv Valmestad, addresses the

research process as a recursive methodology—a means of discovering new pathways while conducting background research—as well as using strategic searching to move from a form of scaffolding toward improvisation in research and writing practices.

Even though the chapters rely heavily on case studies, the conceptual approaches employed by each of the contributing authors offer insight into theoretical frameworks, understanding of the value of information, and ways of approaching the construction of authority. The organization of the work is clear and thematic, allowing the reader to either select a specific chapter or systematically work through the whole text with related examples flowing easily into each other. Ideal readers of the work are librarians in a research, instruction, and/or outreach and engagement roles at a large higher education institution. Understanding student learning outcomes, information literacy frameworks, instruction and assessment practices, and syllabus formatted curriculum are key to finding the examples in the book relatable. The work can also serve as a model to subject specific faculty who may wish to partner with the librarians at their institutions in ways professors and adjuncts have not yet considered. The work is a compilation of reliable narratives, clearly laid out methodologies, and well selected examples serving as a next step for future higher education library-based research and instruction collaborations. — *Andrew Beman-Cavallaro, Assistant Librarian, University of South Florida*

Scholarly Communication Librarianship and Open Knowledge, Maria Bonn, Josh Bolick, and Will Cross (eds.) ACRL, 2023, 9780838939901 Softcover, 528p., $150



Scholarly Communication Librarianship and Open Knowledge, edited by Maria Bonn, Josh Bolick, and Will Cross, is an openly licensed textbook, packed with foundational knowledge, as well as practical lessons in open data, open access, and open education; it also includes short case studies from those working in the field. Readers hear from nearly 80 scholarly communications (scholcomm) practitioners, offering a range of perspectives on everything from coordinating an OER program to supporting reproducible research through the curation of open data.

The book, intended for MLIS students and those new to scholcomm, addresses several well-known curricular gaps in the profession's training—namely, in the areas of copyright, open data, open education, and scholarly publishing. *Scholarly Communication Librarianship* provides both the contextual information, and the practical skills and resources necessary for a deep, multifaceted understanding of scholcomm and all it entails.

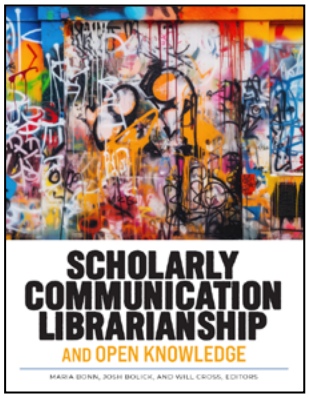
As an open textbook, *Scholarly Communication Librarianship* follows OER best practices in that it is written to be reused. The content is modular and easily adaptable for other audiences and contexts. It makes good use of previously published articles, book excerpts, conversations, charts, and interviews, pulling in relevant content from a variety of open sources. In addition, it is published under a CC-BY Creative Commons license for maximum reusability. The book provides an excellent example for other OER authors on how new and existing content can be interwoven to create an effective curricular resource.

Scholarly Communication Librarianship is divided into three parts. Part I describes the interplay between formal and informal scholarly communications systems and outlines the contexts—social, economic, technological, and legal—that shape scholcomm. The editors author this first section, and their chosen frames are apt, focusing on the tensions faculty ex-

research process as a recursive methodology—a means of discovering new pathways while conducting background research—as well as using strategic searching to move from a form of scaffolding toward improvisation in research and writing practices.

Even though the chapters rely heavily on case studies, the conceptual approaches employed by each of the contributing authors offer insight into theoretical frameworks, understanding of the value of information, and ways of approaching the construction of authority. The organization of the work is clear and thematic, allowing the reader to either select a specific chapter or systematically work through the whole text with related examples flowing easily into each other. Ideal readers of the work are librarians in a research, instruction, and/or outreach and engagement roles at a large higher education institution. Understanding student learning outcomes, information literacy frameworks, instruction and assessment practices, and syllabus formatted curriculum are key to finding the examples in the book relatable. The work can also serve as a model to subject specific faculty who may wish to partner with the librarians at their institutions in ways professors and adjuncts have not yet considered. The work is a compilation of reliable narratives, clearly laid out methodologies, and well selected examples serving as a next step for future higher education library-based research and instruction collaborations. — *Andrew Beman-Cavallaro, Assistant Librarian, University of South Florida*

Scholarly Communication Librarianship and Open Knowledge, Maria Bonn, Josh Bolick, and Will Cross (eds.) ACRL, 2023, 9780838939901 Softcover, 528p., $150



Scholarly Communication Librarianship and Open Knowledge, edited by Maria Bonn, Josh Bolick, and Will Cross, is an openly licensed textbook, packed with foundational knowledge, as well as practical lessons in open data, open access, and open education; it also includes short case studies from those working in the field. Readers hear from nearly 80 scholarly communications (scholcomm) practitioners, offering a range of perspectives on everything from coordinating an OER program to supporting reproducible research through the curation of open data.

The book, intended for MLIS students and those new to scholcomm, addresses several well-known curricular gaps in the profession's training—namely, in the areas of copyright, open data, open education, and scholarly publishing. *Scholarly Communication Librarianship* provides both the contextual information, and the practical skills and resources necessary for a deep, multifaceted understanding of scholcomm and all it entails.

As an open textbook, *Scholarly Communication Librarianship* follows OER best practices in that it is written to be reused. The content is modular and easily adaptable for other audiences and contexts. It makes good use of previously published articles, book excerpts, conversations, charts, and interviews, pulling in relevant content from a variety of open sources. In addition, it is published under a CC-BY Creative Commons license for maximum reusability. The book provides an excellent example for other OER authors on how new and existing content can be interwoven to create an effective curricular resource.

Scholarly Communication Librarianship is divided into three parts. Part I describes the interplay between formal and informal scholarly communications systems and outlines the contexts—social, economic, technological, and legal—that shape scholcomm. The editors author this first section, and their chosen frames are apt, focusing on the tensions faculty ex-

perience in navigating systems for which they have not been trained, the economic realities of shrinking library budgets, the monopolistic practices of commercial academic publishers, and the many biases built into academic publishing systems.

In Chapter 1.4, "The Social Context," the authors examine the colonial legacy that continues to privilege English speaking researchers in Western countries, even as research itself has become a thoroughly networked and global endeavor. The chapter also introduces the concept of open access (OA) as a partial solution to publisher control of information through paywalls. While the book teases out and explores the many facets of OA throughout, it introduces the concept early as a foundational concept.

Part II, "Scholarly Communication and Open Culture," explores the concept of open in greater depth, not just as it pertains to open access journal publishing and OER, but also open data, open pedagogy, and open science. Here, the editors invite a variety of scholcomm professionals to introduce complex concepts, share theory, practices, and resources in their own voices. Many authors pull in articles, quotes, graphics, and essays from others, creating a collage of voices and perspectives.

For example, in Amy Buckland's chapter 2.1, "Open Access," the author incorporates several relevant readings to further illustrate concepts. In service of debunking the myth, "everything should be open," Buckland includes an essay by Tara Robertson about the privacy implications of digitizing a lesbian-created porn zine from the 1980s-early 2000s. Later, the author includes a short piece by Meredith Jacob covering open licensing and how Creative Commons licenses facilitate open access. This technique demonstrates by example how open allows us to avoid recreation of the wheel and can provide learners with a multitude of entry points into a given concept or topic.

Part III of *Scholarly Communication Librarianship* shifts from theory and concept to practical example. It again features voices from the field, this time in the form of short pieces divided into the categories of "perspectives," "intersections," and "case studies." Here, readers will again find a great deal of variety, both in terms of authorial voice, but also job and experience type. Those who find themselves in a similar role or situation to the ones described in this section could use this part of the book as a reference source. Alternatively, readers could also read this section straight through to gain a better sense of the breadth of projects that scholcomm practitioners find themselves taking on.

The pieces focusing on the intersections between scholcomm and other areas of academia and librarianship were the most interesting and potentially useful. Indeed, scholcomm librarianship does involve intersection with many areas, and the focus on university press publishing, collection development, and public libraries is particularly important as budgets diminish and areas of expertise shift.

As a textbook, *Scholarly Communication Librarianship* succeeds in providing learners with a strong foundational understanding of the many intersecting domains within scholarly communications. Each chapter includes discussion questions, extensive bibliographies, and suggested readings for students who wish to dig deeper. The authors have written about complex topics in a conversational and engaging manner, avoiding the terminology soup that can sometimes plague writing about scholcomm. — *Mahrya Burnett, Scholarly Communications Librarian, University of Iowa*