Examining landscapes of research data management services in academic libraries is timely and significant for both those libraries on the front line and the libraries that are already ahead. While it provides overall understanding of where the research data management program is at and where it is going, it also provides understanding of current practices and data management recommendations and/or tool adoptions as well as revealing areas of improvement and support. This study examined the research data (management) services in academic libraries in the United States through a content analysis of 185 library websites, with four main areas of focus: service, information, education, and network. The results from the content analysis of these webpages reveals that libraries need to advance and engage more actively to provide services, supply information online, and develop educational services. There is also a wide variation among library data management services and programs according to their web presence.

Introduction
The importance of research data management has been emphasized over the past few decades. Tenopir et al. argue that, as science grows and moves toward more collaborative, data-intensive, and computational research, researchers are faced with various data management needs. Research data management is also mandated for scholarly researchers. Despite this, many researchers are unprepared for or lack sufficient time to handle the requirements of data management. Researchers also have a number of concerns about data management issues, such as data storage, integrity, and backup options. Keil thus argues the researchers will need the help of a team of experts.

As a response to researchers’ need and request for help with data management, academic libraries have been actively involved in research data services: that is, “services that address the full data lifecycle, including the data management plan, digital
curation (selection, preservation, maintenance, and archiving), and metadata creation and conversion.” Tenopir et al. argue that libraries may be ideal centers for research data services; thus, libraries’ involvement throughout the data lifecycle is vital to knowledge creation, which is one of the core missions of academic libraries. Furthermore, the Association of College and Research Libraries’ Research Planning and Review Committee has identified data curation as one of the top trends for libraries in 2012 with collaboration with research communities, which demonstrates the importance of libraries’ involvement in the data lifecycle and in interactions with researchers.

Although libraries are aware of their potential role in and impact on research data management and curation and have started providing (or planning) data management services, not all libraries are in the same phase due to the different perceptions and needs related to data management at the institutional level, which vary with institutional capacity (and/or boundaries) and policies. Previous research reports that libraries face many challenges in data management program development (such as funding for personnel and equipment and lack of broader institutional support). Tools and recommendations have been developed by leading libraries to overcome these challenges, but other areas of support are needed, such as professional training and collaboration with other institutions to develop more skills in identifying appropriate materials.

This study examined the research data (management) services in academic libraries in the United States through a content analysis of 185 library websites. Examining the current landscape of research data management services is timely and significant for both libraries that are currently planning to improve their data management services and those libraries that are already ahead in providing services. This research provides an overall understanding of where research data management programs are and where they are going; it also facilitates the understanding of current practices and data management recommendations and/or tool adoptions and reveals areas that need improvement and support.

**Literature Review**

Many researchers attest to the role of libraries in data management. Flore et al. argue that data management involves a number of stakeholders, both within and outside the university community. The academic library is especially important because it occupies “the unique position, as both a facility with staff who have expertise in many of the issues surrounding research data management and a campus-wide service with relationships among these many stakeholder groups, that favorably positions it to lead the RDM effort.” Libraries have a long history of working with and training researchers and of curating and preserving information, which makes them key players in research data management. Reinhalter and Walter emphasize the role of libraries in helping researchers deal with federal agencies’ new requirement for data management planning; they state, “Navigating this brave new world is simply an extension of the work that has been going on for decades.”

However, the path to providing data management services has been difficult for many academic libraries. One major concern is the new skills that librarians must have to provide data management services. During a study of librarians who provide data management services in some capacity, Tenopir et al. found that a number of librarians felt they did not have the skills needed, they were not given enough time by their institutions to provide data management services, and their libraries do not consider data management services a priority. Similar findings are reported by Antell et al. from a study of science librarians at Association of Research Libraries (ARL) institutions: Less than a quarter of survey participants (23.2% of 175 librarians) thought they had the skills necessary to provide data management services. Kennan et al. found that
librarians’ concerns and need for new skills and training were similar across different countries, including Australia, New Zealand, the United Kingdom, and Ireland. Financial constraints can be another barrier to providing data management services. Erway and Rinehart point out that research data management services are often more expensive and require more personnel than simply running an institutional repository, and the “recognition that [data management service] requires sustainable funding is slower in coming.”

Despite these difficulties and barriers, some libraries—such as those at Purdue University, the Massachusetts Institute of Technology (MIT), Cornell University, and the University of California—have been successful in providing research data management services, and they have shared their experiences and knowledge. With these initial efforts, many libraries have begun providing (or planning) data management services, and a number of case studies from a variety of institutions show that these services are being implemented across the board. Xia and Wang’s study also demonstrates the growth in data management services in academic libraries. They found that the term “data management” became much more common in librarian job descriptions after 2008.

With the growth in data management services in libraries, several studies have attempted to systematically understand the practices and services relevant to research data at academic libraries. Tenopir et al. conducted a survey study with library directors and academic librarians; Antell et al. also conducted a survey study with ARL science librarians. Both Briney et al. and Steinhart focused specifically on data policy and tried to understand how data policy is implemented and how it impacts data services at institutions. Other studies have tried to understand data management services, practices, and needs in an international context. These studies provide the baseline understanding of the current status of the field and highlight important implications for strategic data management planning in libraries, as well as in library information science (LIS) education for future professionals.

Although these previous studies greatly contribute to the understanding of the current data landscape in libraries, they also have some limitations. First, most studies thus far have relied on self-reporting data through either interviews or surveys, which represents a methodological limitation. In addition, many studies have reported findings at a high level rather than providing a detailed and thorough look at the services that have been offered. This study fills these gaps and contributes to the existing literature by employing a thorough content analysis of library websites. Previously, Si et al. conducted a content analysis of 50 international academic libraries’ data services through content analysis and reported six types of services (including introduction to research data, data management planning guide, data curation and storage, data management training, data management reference, and resource recommendation). The present study employed a larger sample (185 websites) in the United States (whereas only 24 American websites were analyzed in the study conducted by Si et al.), and analyzed data using a predeveloped protocol with broader coverage of services based on standard/best practices for data services not only to understand the current services provided but also to identify gaps in services. Because the study conducted by Si et al. was completed in 2012, this study will also present a comparison in terms of how data services have grown since then, although direct comparison is not possible owing to the different study design and methods of analysis.

**Methods**

A major method used in this study is content analysis. Content analysis is a recognized “versatile tool” for social science researchers, and it has been widely used for objective,
systematic, and quantitative examination of content that researchers are interested in. Content analysis is appropriate for this study because it is unobtrusive and context sensitive, and it allows for examination of artifacts or texts that is not subject to the influence of interests. This study examines library websites because these are often the easiest vehicles for researchers to learn about the data services and information that libraries offer.

Sample
The project team identified academic libraries in the United States offering any service relevant to research data management in September 2015. Different methods were employed to develop an inclusive study sample. First, two authors conducted an online search using the following search terms: “data management library,” “research data library,” “data curation library,” “data planning library,” and “data services library.” The authors also checked the member list of the ARL and the Carnegie list of libraries labeled as having “high” and “very high” research activity to identify any possible additions to our initial online search. These searches resulted in an initial list of 193 institutions with data management services websites.

Protocol Development
Before analyzing the data, the project team developed a protocol from a practice coding session that addresses four main areas of focus: service, information, education, and network, with a total of 40 subcategories. Although this study considered the libraries’ webpages themselves (as well as content) as a form of service to users or visitors of the webpages, this study categorized content displayed on the webpages into different types based on the purpose of the content.

This study defines service as active library engagement with intended users (researchers) to help them and provide necessary information. This study differentiates service and information based on the “activeness” of the services, and we define information as the content provided on the libraries’ webpages. For instance, when libraries provided consultations for data management planning or noted that they provide “services” to help with data deposit, this study considered this an active form of service and thus analyzed it within the service category. On the other hand, when libraries only provided descriptions of data management through their websites, offering information about what it is and how researchers can do it, this study considered these passive services and coded them under the information category. We also considered wording on the webpages, as several libraries specifically noted that they provide certain types of services. Education refers to the libraries educational efforts: that is, whether the libraries offer any educational services to faculty, staff, and students at their institutions. This study captured classes and workshops not only specifically about research data management but also those about other topics relevant to research data. Finally, network refers to how libraries link to internal or external resources to provide either a service or additional information.

The subcategories of service, education, and network were developed from practice coding, but we used predeveloped subcategories for information to capture details concerning not only the information provided but also that missed by the libraries. This study adopted the fields from Research Data Management published by the National Information Standards Organization (NISO), which covers the basics of data management for researchers and was slightly modified in this study context.

Data Collection and Analysis
The identified websites were coded from October to December 2015. During this period, several websites were eliminated because they were either no longer available or
they contained minimal information, such as a few sentences about data management. The final list included 185 websites. Two authors tested interrater reliability before proceeding and reached an 88.3 percent agreement rate. Coding for most of the fields consisted of determining whether each item was present, marked as “yes” or “no,” with no judgment about the quantity or quality of services offered in any respect. For instance, a website that provided a paragraph of information on data management was counted the same as a website that dedicated an entire page to data curation.

The raw data from 185 library websites were reassessed for completeness and exported to SPSS23, a software for statistical analysis, for further descriptive analysis. Text data were analyzed using Excel.

Findings
Basics of the Websites
The libraries’ websites for data management varied in nature, coverage of content, and manner of displaying information. The term “data management” was the most commonly used one when describing the research data management webpages. The majority (65%) used the term “data management” in their titles, such as “data management,” “data management services,” or “manage your data.” Some data management webpages (17%) were part of a bigger website of data services at the libraries. Although they might still use the term “data management,” the webpages existed as a subsection of data services, which included not only data management services but also other types of data-related services, such as methodology consultation. The term “data curation” was used only by a small number of libraries (3%), and other websites (15%) about data management were part of “scholarly communication services,” “research services,” or “grant writing.”

The coverage of basic information about data management provided on the webpages—which is perhaps more useful for first-time visitors or researchers who are

<table>
<thead>
<tr>
<th>Types of Name (n = 185)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data management 65%</td>
</tr>
<tr>
<td>Data service 17%</td>
</tr>
<tr>
<td>Other 15%</td>
</tr>
<tr>
<td>Data curation 3%</td>
</tr>
</tbody>
</table>

FIGURE 1
unfamiliar with data management—varied among libraries. Surprisingly, not many libraries defined what data are and what data management is, and only a small percentage of libraries clearly explained what these terms meant in general and how the libraries understood them (29% and 20%, respectively). The aims of the data management services or purpose of the webpages were relatively well explained (50%), yet a low percentage of libraries defined the intended audience of the services and the webpages (31%) and even fewer defined the scope of the services (7%).

**FIGURE 2**
Basic Information Provided (n = 185)

<table>
<thead>
<tr>
<th>Service</th>
<th>Definition of data</th>
<th>Definition of data management</th>
<th>Aim</th>
<th>Audience</th>
<th>Scope</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0%</td>
<td>10%</td>
<td>20%</td>
<td>30%</td>
<td>40%</td>
</tr>
</tbody>
</table>

**Service**

The libraries’ webpages provided different types of services for their intended users, but the extent of such offerings also varied. Most libraries provided basic services, such as “help/ask librarians” (79.5%), on their data management webpages through online chat, web inquiry forms, or e-mail. However, this service was usually embedded on the library webpages; it is therefore hard to argue that this service is devoted to research data management.

Among other services, data deposit services were well placed by more than half of libraries (60%). These libraries tried to encourage researchers to deposit their data to the repositories at their institutions for preservation. The libraries provided instructions for procedures about data deposit, introduced deposit-related services, connected researchers to the staff at the repositories, and led the researchers to institutional repositories’ websites. Data management planning was the second-most offered service at libraries (41.1%) when general help and “ask a librarian” services were excluded. Data consultation was a more specific service than a help service or “ask a librarian” service, and it was offered by 38.4 percent of libraries. The consultation services offered usually broadly covered the lifecycles of data, from search, use/reuse, data management and preservation, and data sharing. More than a quarter of libraries (27.6%) also offered a service for data publishing and sharing. Although this percentage is not high, data sharing and data deposit are closely related practices, and both reflect the libraries’ efforts toward the promotion of data sharing. Finally, although it is not solely related to data management, 15.7 percent of libraries provided a service related to methodology, and these were primarily libraries that had a data service unit.
The libraries’ webpages covered various topics about research data management. This study addressed four main areas of information that are considered necessary for researchers: 1) background information; 2) data documentation; 3) data administration; and 4) data use and reuse. A relatively large number of libraries provided a basic description about research data management, specifically the importance of data management (64.3%) and how to plan data management (79.5%).

When it comes down to the detailed level (see figures 5, 6, and 7), fewer libraries provided full detail for each topic, and the coverage of topics provided on the webpages varied widely. Despite the significance of documentation in data management, not all libraries provided information about data documentation: 61.1 percent of libraries provided information about metadata standards and practices, and only 43.8 percent of libraries provided information about how to properly organize data files. Very few libraries provided information about software and tools for documenting data (10.8%) or facilitating proper workflow (10.3%).

Regarding data administration, 42.2 percent of the libraries provided some information about data governance (such as copyright issues or ownership), but fewer
libraries (29.7%) addressed the issue of confidentiality in data management, which is important when researchers prepare their data to share. More than half of the libraries (53.5%) provided information about data storage and preservation. Finally, libraries’ data management webpages also provided information relevant to data reuse, primarily regarding how to publish managed data for reuse, with some explanation about benefits (49.2%). Some libraries also addressed procedures for reusing published data, such as how to search existing data (17.8%) and how to properly cite data (37.8%).

### Education

Not all libraries have an educational program or service relevant to research data. A total of 64 libraries in this study (34.5%) provided educational services relevant to data through workshops and classes. Two types of classes and workshops were directly relevant to research data management: topics under data management and those under data sharing and reuse. Among those 64, most libraries (73.4%) offered classes for data management, although half of them were offered on an irregular basis (see figure 9). Classes and workshops about data sharing and reuse were offered by 21.8 percent of libraries, and they usually emphasized the importance of data sharing and reuse within data management.

### FIGURE 6

**Information 3: Data Administration**

(n = 185)

<table>
<thead>
<tr>
<th>Governance</th>
<th>Confidentiality</th>
<th>Storage &amp; preservation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### FIGURE 7

**Information 4: Data Use & Reuse**

(n = 185)

<table>
<thead>
<tr>
<th>Data search</th>
<th>Data citation</th>
<th>Publishing data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### FIGURE 8

**Education: Workshop & Classes**

(n = 64)

<table>
<thead>
<tr>
<th>Data management</th>
<th>Data sharing &amp; reuse</th>
<th>Data methodology</th>
<th>Specific datasets</th>
<th>Not specified</th>
<th>Online instruction</th>
</tr>
</thead>
</table>
Some libraries provided other types of educational services for specific data methodology (21.8%) or for using popular, field-specific datasets (15.6%). Some programs also briefly touched upon topics relevant to data management (such as publishing your data) during data-related classes or workshops. Usually, these libraries had specialties in research data, and most of them already had a class relevant to data management or plans to offer such a class. About half (45.3%) did not specify what types of educational programs the libraries provided, other than referring to “data workshops.” Some libraries (20.3%) also provided instruction through live, online sessions or sharing of recorded sessions online.

Network
When the libraries provided any service or information through their webpages, they also relied on or referred to other internal or external sources. A total of 135 libraries (73%) had internal links on their webpages. Internal links are usually for connecting researchers to other internal resources at their institutions. The most common type of internal link was a link to institutional repositories and resources (73.3%). This is not surprising, given that most libraries encouraged researchers to deposit their data to the institutional repositories. University offices of research were the second most popular internal connection to the libraries’ data management webpages (60.7%). This was primarily due to the role of the office of research, which provides administrative oversight related to research as well as research data. Fewer libraries had linked their webpages to other libraries on campus (14.1%), which were mostly health science libraries (for further information for health data) and data centers (23.7%); data-intensive research centers (13.3%) were also referred to for further information.
Almost all libraries (a total of 180, 97.3%) had external links on their data management webpages. The majority of the external sources were linked to: 1) tools (90.6%) that help researchers manage data; 2) data repositories (80.1%) that provide extensive resources for data management (such as best practices and standards and/or promotion of deposit to any relevant repositories); 3) government regulations and policies (86.7%) (mostly regarding requirements of funding agencies such as National Science Foundation and National Institute of Health); and 4) other libraries’ websites (69.4%) that have more information about their own libraries’ webpages.

The project team conducted additional examination of these external links (particularly for the tools, data repositories, and libraries) to identify the most frequently adopted tools in data management, as well as the institutions that were most frequently referred to by others. DMPTool (https://dmptool.org/) was more commonly recommended by most libraries (154 libraries, 85.6%) for data management planning than any other data management tools, such as Bulk Rename Utility (www.bulkrenameutility.co.uk/) (20 libraries, 11.1%) or ReNamer (http://renamer.com/) (18 libraries, 10%) for file renaming, followed by EZID (http://ezid.cdlib.org/) (16 libraries, 8.9%) for creating long-term identifiers for data.

ICPSR was the most commonly referred data repository by 112 libraries (62.2%) for additional resources for data management. Although the number of references to a particular resource is just one way of identifying key players in research data management, it is obvious that ICPSR provides a number of useful resources. Some of the popular external resources also included data repositories, such as re3data (www.re3data.org/) (85, 47.2%) and Open Access Directory (http://oad.simmons.edu/oadwiki/Main_Page) (61, 33.9%). The libraries introduced these repositories as resources to help researchers find and deposit relevant data.

Other libraries’ data management websites were usually referred to as sources for further information. The top three library data management websites most referred to by the libraries were the MIT Library (64 libraries, 35.6%), the University of Minnesota Library (47 libraries, 26.1%), and the University of California, San Diego Library (37 libraries, 20.6%). Although this study did not evaluate the depth and quality of the content provided on the websites, these libraries’ data management websites were extensive and stood out by providing a variety of information under structured themes, which was probably why they were referred to by many others. Unlike ICPSR, which was the unrivaled repository referred to by the majority of libraries, the counts for the
referred libraries had little variations in the extent to which the fourth through the
tenth most referred-to libraries are recommended, and these included the University of
Virginia Library, the University of Wisconsin Library, the Purdue Library, the Oregon
University Library, the Cornell University Library, and the Stanford University Library.

Discussion
Antell et al. noted that research data management work in libraries is in its emergent
phase based on their survey study conducted in 2012. Libraries have been engag-
ing in various activities and services related to data management since then, and the
fact that a number of libraries have a web presence for data management services is
encouraging. The growth is clear from the most recent study in 2015 that analyzed
data policy, as they found approximately 100 libraries’ data services, and this study
was able to search 185 data-service websites. This also aligns with reports from the
2012 survey study of Tenopir et al., as creating web guides was the most commonly
planned research-data service. However, the results from the content analysis of
these webpages also reveals that libraries need to advance and engage more actively
to provide services, provide information online, and develop educational services.
There is also a wide variation among library data management services and programs
according to their web presence.

As seen in figure 2, some libraries’ data management webpages seem not to provide
a good basic introduction to their services, and they also lack clear explanations of what
data management is and the purpose and intended audience of the data management
webpages and services. Because visitors (researchers) to these webpages might not
always be familiar with the concept of data management or the libraries’ services, it is
important to efficiently inform researchers, particularly if the libraries are interested in
reaching out to researchers and encouraging them to engage with data management.
It is surprising that only a small number of library webpages provided a definition of
data management, considering that the purpose of these webpages is to help research-
ners with data management. Many libraries’ webpages were also unclear about whom
they intended to serve and with what scope of data.

The libraries offered different types of services: an active form of service through
interactions with users either in person or online, a passive form of service by provid-
ing useful information on the webpages, and educational services through workshops
and classes. For the active service, it is worth noting that data deposit was the most
frequently offered service, except for the services of general help and “ask a librarian,”
which is a default, embedded function throughout entire library websites. This reflects
the libraries’ emphasis on promoting and encouraging researchers’ deposit practices
at their institutions, which are perhaps influenced by funding agencies’ data sharing
requirements, and this is why most of the libraries’ data management webpages had a
link to the institutional repositories. It is also interesting that data management plan-
ning was the second-most offered service. However, for less than half of libraries to
provide this service seems not to be enough, considering funding agencies’ emphasis
on data management planning and researchers’ need to meet funders’ requirements.

The coverage of information and topics about data management provided on the
websites varied among libraries. Although many libraries tended to provide some
basic information for researchers—such as what data management is and why it is
important—significantly fewer libraries provided a detailed level of information about
data documentation, administration, and reuse. However, about half of the libraries’
websites in this study provided at least some information about several core areas of
data management, such as metadata, data preservation and storage, and data publi-
cations, which is a good starting point and a promising sign for future development.
Although this study did not evaluate the quality and depth of information, this result suggests the need to improve content and breadth of coverage, at least in the areas suggested by NISO. Providing useful information that is sufficiently broad and in-depth would be a relatively easy and good starting point for libraries planning to initiate data management work.

Educational service for data management is one active form of service offered by libraries. More than one third of libraries were devoted to providing some type of educational service to researchers, and many of them addressed the topic of data management in the context of researchers’ needs, as well as in the context of funding agency requirements for data sharing. It is expected that more libraries will start providing educational services in the future; there were several libraries that stated that they were in the process of planning educational services at the time of data analysis.

The libraries’ data management websites also included various internal and external links to refer the researchers to relevant services and additional information. Further analysis of external sources linked to on the webpages revealed the most frequently adopted tools and the most commonly recommended library or repository websites. DMPTool was the most frequently recommended tool by numerous libraries. It seems that this tool is widely known to libraries and is introduced to researchers through libraries, which indicates its national impact. ICPSR was the top recommended repository due to its extensive resources relevant to data management, which demonstrates its leading role and expertise. Although ICPSR appeared to be the unrivaled repository, most recommended libraries’ websites were more evenly distributed, without a huge gap among them, and it seems that more than just a few libraries have played an influential role in data management.

Conclusion
This study examined library data management webpages to understand the landscape of research data management services. This study does not suggest the need for all libraries to develop the same level of website, as the core services needed in institutions would differ depending on the types and characteristics of institutions (for instance, if they are larger or doctoral-granting institutions, if they receive NSF funding, if they are derived from the data policies at institutional level). Still, libraries’ webpages on data management are instructional sources for researchers, as well as an important way for researchers to learn more about the libraries’ services. How key information is displayed and how the libraries’ services that they intend to provide are introduced are thus important to libraries’ efforts to reach out to potential users and researchers. The results of this study reveal many aspects of library data management webpages that need to be improved, including service development and the thoroughness of information offered.

Although this study contributes to the understanding of the landscape of data management services in libraries, it also has some limitations, which direct some future research. Because this study did not evaluate the quality of the webpages, it did not capture the difference between good or bad services and thorough or cursory information. For instance, a library website with a long page of description about metadata standards for research data was coded in the same way as a library website with a brief paragraph about metadata. There is also a possibility that the webpages did not always correctly reflect what the libraries actually did. Self-reporting survey studies or interview studies would complement the results of this study, because these studies help capture libraries’ actions that may not be reflected on the webpages. Finally, it is critical to study the impact of data services and websites on researchers, the users of data services. The content analysis does not provide information about usefulness
from the users’ perspectives or how the librarians actually work with or without websites. A follow-up study with librarians to learn about their interactions with researchers and listening to users about library services, including websites, is a necessary next step.

Notes
10. Ibid., 88.
11. Ibid., 88–89.


23. Si, Xing, Zhuang, Hua, and Zhou, “Investigation and Analysis of Research Data Services.”


30. Strasser, “Research Data Management.”