

The International Journal of Digital Curation

Volume 8, Issue 1 | 2013

Developing an Approach for Data Management Education: A Report from the Data Information Literacy Project

Jake Carlson,
Purdue University

Lisa Johnston,
University of Minnesota

Brian Westra,
University of Oregon

Mason Nichols,
Purdue University

Abstract

This paper describes the initial results from the Data Information Literacy (DIL) project designed to identify the educational needs of graduate students across a variety of science disciplines and respond with effective educational interventions to meet those needs. The DIL project consists of five teams in disparate disciplines from four academic institutions in the United States. The project teams include a data librarian, a subject-specialist or information literacy librarian, and a faculty member representing a disciplinary group of students. Interviews with the students and faculty members present a detailed snapshot of graduate student needs in data management education. Following our study, educational programs addressing identified needs will be delivered in the fall of 2012 and spring of 2013. Our findings from the project interviews are analyzed here, with a preview of the training approaches that will be taken by the five teams.





Introduction

Data-driven research methodologies are increasingly prevalent across academic disciplines and in commercial research enterprises. As expectations rise for making data more accessible and reusable, the need for educated personnel in data management, sharing and preservation is critical (Beagrie, 2008). But what knowledge and skills will researchers need to be able to respond to these expectations? Moreover, how and when will such skills be acquired? Graduate school is the time when students learn the cultural norms and practices established within their chosen discipline; however, facility in data management and curation is not typically included in the curricula.

The Data Information Literacy (DIL) project is exploring what data management skills are needed by graduate students as future scientists to fulfill their professional obligations in ways that align with disciplinary cultures and practices. The DIL project is comprised of librarians from four institutions: Purdue University (lead), Cornell University, the University of Minnesota and the University of Oregon. As a librarian-driven project, we seek to apply the relevant elements and approaches of information literacy programs conducted by academic libraries to build capacity and promote the engagement of librarians in data education. Currently, the project is a little more than halfway through its two year life cycle. In this paper, we will report on the results from the first phase of the project, in which we interviewed faculty and students about their current practices in working with data and their perceptions concerning what knowledge and skills graduate students should possess by the time they graduate.

Background

Expanding the scope of information literacy to include data management and curation is a logical extension of information literacy concepts. The Society of College, National and University Libraries' (SCONUL) Seven Pillars information literacy model (SCONUL, 2011), and Vitae's Researcher Development Framework¹ incorporate data management skills into their definitions of information literacy, and support holistic approaches to helping doctoral candidates acquire skills and knowledge in data management. A recent report from the Research Information Network (RIN, 2011) argues that a broader interpretation of information literacy that recognizes research data as information is needed to ensure that students gain the skills they will need to be successful in their careers. The 2012 LIBER working group on eScience selected data as a critical area for involvement by libraries in eScience support and recommended that libraries assist faculty with the integration of data management into the curriculum (Christensen-Dalsgaard et al., 2012).

Graduate students are a natural audience for educational programming on data management and curation issues. In the STEM disciplines, graduate students are often expected to carry out most or all of the data management tasks for their own research, and frequently participate in data activities to support lab/team projects as well (Akmon, Zimmerman, Daniels & Hedstrom, 2011; Westra, 2010). Gabridge (2009)

¹ Vitae Researcher Development Framework: <http://www.vitae.ac.uk/rdf>

←—————→

observed that graduate students compose “a constantly revolving community of students who arrive with... uneven skills in data management.” Another RIN initiative applied the SCONUL Seven Pillars model of information literacy and Vitae’s Research Development Framework towards the development of data management skills in postgraduate courses in the UK. The results from their initiative demonstrated that data management skills were needed in a wide range of disciplines and that core skills as well as discipline-specific training should be embedded into the postgraduate curricula (RIN, [2010](#)).

The approaches taken on data management and curation education for graduate students in the sciences thus far have generally fallen into one of two categories: standalone courses and programs or one-shot workshops. The standalone course approach has been used by several schools of information science, including Syracuse University² (Qin & D’Ignazio, [2010](#)) and the University of Michigan.³ Syracuse designed a course to teach science data literacy, defined as “[understanding], [using] and [managing] science data,” with a particular focus on preparing students for positions in science work or as data management professionals. Michigan developed a research fellowship program centered on building a community of practice around managing, sharing and reusing scientific data. Their curriculum includes a core course on data curation and elective courses from multiple disciplines. Courses are also being launched by research centers, such as the “Data Science” course offered by Tetherless World Constellation at Rensselaer Polytechnic Institute.⁴ The advantage of the standalone approach to teaching data skills is the depth of coverage that can be achieved. However, it may be difficult to attract students to commit their time to a course that resides outside of their discipline.

Becoming more and more prevalent at academic institutions, one-shot workshops represent a second approach to data management and curation education. Many of these workshops focus on helping faculty and graduate students address the recent requirements for data management plans by funding agencies, such as those offered by MIT (Graham, McNeill & Stout, [2011](#)) and the University of Minnesota (Johnston, Lafferty & Petsan, [2012](#)). Other workshops cover data management as one component of a broader training in research ethics or responsible conduct of research, as required by the National Science Foundation and the National Institutes of Health (Coulehan & Wells, [2006](#); National Institutes of Health, [n.d.](#); National Science Foundation, [n.d.](#); Frugoli, Etgen, & Kuhar, [2010](#)). These workshops require less of a time commitment and are likely to reach more people, but they cannot provide much depth on important issues and they may not address important disciplinary considerations.

The DIL project is a part of an evolving third approach towards educating students about the data concepts they will need to be successful in their careers. We seek a balance between a full semester’s commitment and the one-shot, one-size-fits-all model. Our educational programming for graduate students in data management is intended to be aligned with disciplinary needs, as articulated by researchers themselves, and integrated with current research practice. Previous research has stressed the need for curation service providers to understand the nuances and disciplinary practices of the research communities with which they would like to work

² The Science Data Literacy Project: <http://sdl.syr.edu/>

³ Open Data: <http://opendata.si.umich.edu/>

⁴ Data Science 2012: <http://tw.rpi.edu/web/Courses/DataScience/2012>

(Martinez-Uribe & Macdonald, 2009). We believe this to be true of educational services in data management and curation as well. There are several other initiatives actively developing programs using this type of approach. MANTRA has developed online materials designed to be embedded into post-graduate programs in social science, clinical psychology and geoscience.⁵ The University of Massachusetts Medical School and Worcester Polytechnic Institute have developed “Frameworks for a Data Management Curriculum” for teaching research data management to undergraduate and graduate level students in the sciences, health sciences and engineering disciplines (Piorun et al., 2012).

Data Information Literacy

The DIL project builds on earlier research in which competencies in working with research data were identified and categorized (Carlson, Fosmire, Miller & Nelson, 2011). This foundational list of data competencies was generated from two sources: interviews conducted with faculty as a part of the Data Curation Profiles project (Witt, Carlson, Brandt & Cragin, 2009) and from student experiences in a geoinformatics course taught at Purdue University (Miller & Fosmire, 2008). The 12 competencies are listed in Table 1. We are now seeking to test the efficacy and potential application of these data competencies through working with faculty and graduate students directly.

Data Processing and Analysis	Data Curation and Reuse
Data Management and Organization	Data Conversion and Interoperability
Data Preservation	Data Visualization and Representation
Databases and Data Formats	Discovery and Acquisition
Ethics and Attribution	Metadata and Data Description
Data Quality and Documentation	Cultures of Practice

Table 1. The 12 competencies of DIL as identified by Carlson et al. (2011).

Methodology

The DIL project is comprised of five teams including a data librarian, a subject librarian and at least one faculty researcher from a science or engineering discipline recruited for the project. An underlying assumption made by participating librarians is that the success of our efforts will be determined by our ability to align our educational programs with current disciplinary cultures and norms, as well as with local practices and needs. We began by conducting literature reviews in the disciplines of our faculty partners to uncover how the 12 competencies were described and addressed. One important outcome of the literature reviews was recognition of the need to clarify our definitions of the 12 competencies in the subsequent interviews as the faculty and students participating were likely to have different understandings and definitions of the competencies based on their experiences and backgrounds. Therefore, rather than assigning strict definitions, we described each competency by listing several activities and skills that would reflect the nature of the competency.

⁵ Research Data MANTRA Course: <http://datalib.edina.ac.uk/mantra/>

Interviews were conducted in the spring and summer of 2012. Eight of the interviews were with faculty. The other 17 interviews were with current or former graduate students or post docs of the interviewed faculty, or in one case with a lab technician. The interview protocol was based on the structure of the Data Curation Profiles Toolkit developed at Purdue.⁶ The protocol consisted of an interview worksheet, which contained a series of questions for the interviewee to complete in writing during the interview, and an interviewer's manual, which contained follow up questions for the interviewer to ask based on the responses written by the interviewee. Our interview protocol is available for download from the project website.⁷

The interviews were conducted with two objectives in mind. First, we sought to obtain a thorough understanding from faculty and students about the data being generated in each lab and how it was being managed. Second, we asked the participating faculty and students to indicate how important it was for graduate students to become knowledgeable in each of the 12 competencies using a five-point Likert scale, and then to explain their choices. Interviewees were also asked to identify any additional skill sets they saw as important for graduate students to acquire.

Each team compiled their interview data and, armed with a better understanding of how issues with data are conceptualized and expressed within the discipline through local practice, crafted educational interventions to address important DIL competencies. The DIL project is still underway as project teams offer training during fall 2012 and spring 2013. The rest of the paper will focus on what we have learned thus far about DIL, with a particular emphasis on the interviews we conducted with our faculty partners and their graduate students in the project's five case studies.

Results

The results of the five case studies revealed both similarities and differences between faculty and students in how they perceive the importance of the DIL competencies for graduate students.

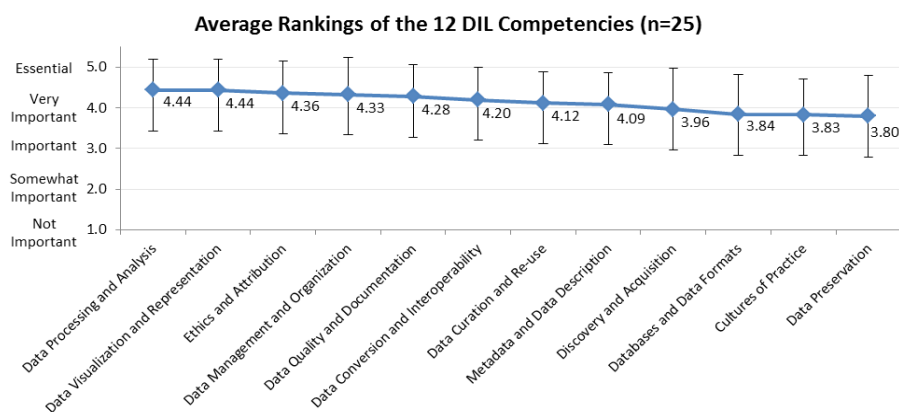


Figure 1. The average ranking of the importance of the 12 DIL competencies as reported by faculty and students.

⁶ Data Curation Profiles Toolkit: <http://datacurationprofiles.org/>

⁷ Data Information Literacy: <http://datainfolit.org>

Due to the small sample size of the project and the use of convenience sampling, the results presented here cannot be generalized outside of this cohort. Nevertheless, we feel that the findings offer a useful starting point for larger investigations into the current environment of the educational needs of graduate students.

The DIL competency rankings show that, on average, participants valued each competency, as all of them were ranked as “Important” or higher. However, there was considerable variance in the responses received, as indicated by the high standard deviations (ranging from .75 to 1.02). In reviewing the results, the competencies that pertain more directly to keeping a research lab operational and in publishing outputs, such as “Data Processing and Analysis,” “Data Visualization and Representation” and “Data Management and Organization” tended to rank higher than competencies that are less central to these activities, such as “Discovery and Acquisition” and “Data Preservation.” Some of the lower ranked competencies, such as “Data Preservation,” were deemed important but difficult to address. In the interviews, many of the faculty stated that they lacked the experience or knowledge to effectively educate students about these competencies. Several of the faculty and students questioned if their field had any “Cultures of Practice” in managing, handling or curating data.

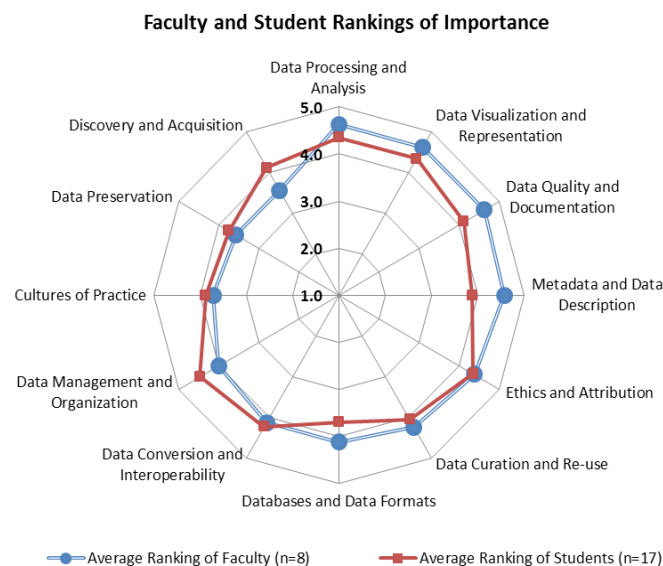


Figure 2. Rankings of importance with faculty response averages compared to student response averages.

As demonstrated in Figure 2, which displays the differences between the average rankings given by the faculty and those given by the students, there were noticeable differences in how the participants viewed some of the competencies. Faculty generally placed a higher value on students developing competencies in actively working with data (“Data Processing and Analysis,” “Data Visualization and Representation”) and in competencies that would sustain the value of the data over time (“Metadata and Data Description,” “Data Quality and Documentation”) than the students did. Students indicated in the interviews that the “Discovery and Acquisition” competency was important to them in learning their field and contextualizing their research. Two of the faculty, both of whom were working with code as their data,

gave “Data Management and Organization” a lower ranking than the other participating faculty. One faculty member believed that, individually, students should know how to manage their own data, but did not necessarily need to know how to develop systems or management plans for larger units. The other found it difficult to respond, not knowing what constituted good management practice and therefore unable to say if it would be worth the investment of his and his student’s time.

Comparison of the Team Interviews


Analyzing the interview transcripts revealed several high-level commonalities across the five case studies. Among the observed commonalities were: the overall lack of formal training in data management, the absence of formal policies governing the data in the lab, self-directed learning through trial and error and a focus on mechanics over concepts.

First, none of the five research groups provided their students with formal training in data management. Instead, faculty reported that they expected that their students had acquired most of these and other competencies prior to joining their lab. As the University of Oregon faculty member noted, “[students may have] picked up [their skills] at on-the-job training, because a lot of them had a former life in a professional field...or [it’s] something they got as an undergraduate.” In contrast, student interviews revealed wide variations in their prior experiences with working with data. Some had a degree of previous experience from work or courses, others had not. Most of the students had attended a responsible conduct of research (research ethics) seminar, but reported that research data practices were not covered in much detail, and that they could not recall what was said about the subject. It should be noted that none of the five case studies involve sensitive data that would require training to deal with human subject or private data.

In lieu of formal training, most graduate students’ data management skills were self-taught through “trial and error,” by reading manuals, asking their peers for help or searching the Internet for information. Of the five labs participating in this project, only one has written policies for the treatment and handling of the data in the lab. Disciplinary norms and processes for data management were predominantly expressed as underlying expectations that tend to be delivered informally and verbally. Some of the students interviewed had inherited data from previous students or others in the lab; this transference process also tended to be informal with a minimal amount of introduction to the data.

Faculty expected their graduate students to be independent learners. For example, one faculty member summed up the skills acquisition process as the “pain and suffering method,” which was described as “[graduate students] try it, they fail, they see what failed, they come back to their advisor and you say, ‘Ah, well maybe you should try X.’ It is not something that we have attempted to teach, certainly.”

When asked how well their students had mastered the DIL competencies, faculty stated that students tended to focus more on the mechanics of working with or analyzing data rather than the theories and assumptions underlying the software or tools they used. For this reason, some of the faculty expressed concern that students’ understanding of these competencies may be somewhat superficial. For instance, one



faculty member stated that students may be able to collect data from a sensor, but they do not necessarily understand the equipment variables that might impact data quality or accuracy, and may be more focused on getting the data than on understanding the steps and settings that created it. Similarly, some faculty were concerned that though students may be able to use tools to work with data, they do not always use them very effectively or efficiently. For example, one faculty member commented “I certainly think that they learn basic visualization tools, but there’s a difference between learning how to draw a histogram and how to draw a histogram that’s informative and easy to read.”


This differentiation between basic project-driven skills and deeper, transferable understanding is also seen in questions about managing and curating data. Most students described idiosyncratic methods of data management and generally overestimated the capacity of their methods to support local collaboration. Only three of the seven faculty interviewed felt that their students provided enough information about their data for the faculty member to understand it, and only one faculty member thought that his students provided enough information for a researcher outside of the lab to understand and use their data. On the other hand, 15 of the 17 students believed that they provided sufficient information for someone outside of the lab to understand and use their data.

Faculty want their students to acquire a richer understanding and appreciation for good data management practices, but there are several barriers that restrain faculty from taking action. First, spending time on data management can be deemed detrimental if it is seen as distracting or delaying the research process. Faced with this pressure, faculty accept that a minimal skill set is sufficient for their students to succeed in school. One faculty member stated, “[Students] can do their work without understanding this. It’s not essential that they have this. It’s best if they do, but they don’t. I guess I could be doing more, but we don’t talk about all of these functions... I’m not sure they all understand why data has to be curated.”

Second, faculty do not necessarily see themselves as having the knowledge or resources to impart these types of skills to their students themselves. One faculty member mentioned requirements by funding agencies for data management plans and journals accepting supplemental data files as positive steps, but researchers in his field were ill-prepared to respond. Most of the faculty stated that there were no best practices in data management to follow in their particular field. Faculty in this study do not believe funding agencies, publishers or scholarly societies in their discipline are providing the guidance or resources to support effective practices in managing, sharing or curating data. In the absence of such support, the data practices in their labs remain more centered on local needs rather than larger perspectives.

Case Studies

Each of the five teams defined learning outcomes and developed targeted pedagogies for teaching and evaluating these learning outcomes based on the particular needs found in their interviews (see Table 2). The five approaches give the DIL project an opportunity to explore educational training in a variety of settings and test multiple approaches to training while remaining grounded in disciplinary and local needs.



Team	Disciplinary Focus	Needs of Project Partner	Educational Approach
Cornell	Natural Resources	Data Sharing Databases & Stewardship	Mini-course (For credit)
Purdue Team #1	Electrical and Computer Engineering	Documenting Code Organizing Code Transfer of Responsibility	Embedded Librarianship
Purdue Team #2	Agricultural and Biological Engineering	Standard Operating Procedures Metadata	Workshops
Minnesota	Civil Engineering	Data Ownership Long-Term Access	Online Course
Oregon	Ecology	Cultures of Practice Data Sharing & Metadata Closing Out a Grant	Readings & Team Meeting

Table 2. A summary comparison of the needs and approaches taken by each project team.

Natural Resources – Cornell University

The DIL project team at Cornell University is working with a research lab in the Natural Resources department. This lab is collecting a variety of different data pertaining to fishing and water quality. Specifically, lab researchers are collecting data examining longitudinal changes in fish abundance, growth and consumption. Some of the data sets contain information collected over a number of years, emphasizing the crucial need for data curation and maintenance over the extended lifespan of the data. Because this longitudinal data cannot simply be reproduced, a more formalized approach to data curation and management would be of great utility to students in the lab. Due to the ongoing nature of data collection, lab researchers use databases extensively. For this reason, acquiring the skills necessary to work with databases and handle data entry is described as essential, otherwise, as the faculty member stated, “it’s [as if] the data set doesn’t exist.” While there are tight controls in place and a formal process followed in the lab for data entry, training in this area is informal, and graduate students typically enter the lab with limited database skills. In fact, across all of the competencies discussed, a lack of formal training for acquiring important skills arose as a common theme with the students noting that most of their skills, such as generating visualizations and ascribing metadata to files, are acquired informally.

Interventions will take place in a classroom setting through offering a spring 2013 semester one-credit course entitled “Managing Data to Facilitate Your Research” taught by the Cornell Team. This course is geared towards graduate students in the Natural Resources department and is meant to serve as a foundational experience in acquiring needed competencies with data. In preparation for teaching this course, the Cornell Team is offering several pilot workshops on general and specific aspects of data management in the fall of 2012. Topics covered in these workshops include an introduction to data management and funding agency requirements, relational databases, and metadata.



Electrical and Computer Engineering (ECE) – Purdue University Team #1

Purdue University Team #1 formed a collaboration with the Engineering Projects in Community Service (EPICS) center. EPICS is a service learning center devoted to providing undergraduate students real world, practical experience through applying their engineering skills to assist local community based organizations. Many of these service projects involve developing and delivering software code as a component of the completed project. Students' work is overseen and guided by graduate teaching assistants (TAs) who provide instruction and serve as a resource to students as they encounter obstacles. Purdue Team #1 is working with the TAs to develop approaches and resources to teach undergraduates data management skills as a part of their education.

The interviews revealed significant concerns about the organization and documentation skills of students. Both faculty and the graduate student TAs stated that students were not effective in describing the work they performed on the project generally and with the code they developed specifically. Students may complete their work satisfactorily, but are not taking the care to make sure that their work is documented or organized in ways that would sustain their code beyond their immediate involvement in the project. This situation presents barriers in transferring the code to new students for its continued development, delivering the code and other project outputs to the community client, and for the center's administration in understanding and evaluating the impact of the center on student learning. The need for documenting data is emphasized to students by the TAs who oversee their work, but the center has not developed specific policies or articulated expectations.

EPICS is a highly structured environment. Students are provided with detailed learning goals, project design specifications and rubrics for evaluating their work. In response, Purdue Team #1 is taking an "embedded librarian" approach for their program by integrating themselves into existing structures to enable close collaborations (Dewey, 2008). This team is developing short skill sessions to deliver to team leaders, crafting a rubric to follow in documenting code and other data, serving as critics in student design reviews and attending student lab sessions to observe and consult on student work.

Agricultural and Biological Engineering (ABE) – Purdue University Team #2

Purdue University Team #2 worked with a research group in Agricultural and Biological Engineering (ABE). While the students working in the lab are covering a wide range of varying topics and data types in their research, the group's central focus is hydrology. An important aspect of the research process for all students is comparing observed data collected in the field to simulation data generated by an array of hydrologic models. Although the faculty researcher has created formal policies on data management practices for his lab, students stated that their adherence to these guidelines was limited at best. One student admitted, "I didn't go through [the policies] very carefully," while another noted that the immense file sizes associated with her data prevent her from following the established guidelines appropriately. Similar patterns arose in discussions concerning the quality of metadata currently being appended to files within the lab. Students reported that they understood the



concept of metadata in the interviews. However, the faculty member noted that his students are not very proficient in incorporating metadata into their files.

These findings suggest that students appear to be aware of the need to manage their data; however, they do not address this need effectively in practice. The second Purdue Team is working with the ABE faculty to develop the means to implement the policies created by the faculty member in a more structured fashion. Their educational program centers on fashioning a checklist to serve as a means of comparing individual practice against the recommended procedures, and to promote a smooth transition of the data from student to faculty upon the student's graduation. In support of propagating the checklist, Purdue Team #2 will be offering three workshops addressing core skills in data management, metadata and data continuity and reuse.


Civil Engineering – University of Minnesota

The University of Minnesota Team is collaborating with a Civil Engineering lab working on researching the structural integrity of bridges within the state of Minnesota. Students collect various types of data – primarily from sensors placed on the bridges themselves – to study factors which may lead to bridges being classified as unsound. The lab works with and receives funding from national and state agencies to conduct its research projects. These project partnerships have a noticeable effect on the treatment and handling of the data. For example, a national engineering data repository has developed processes and standards for sharing and curating the data. The state agency claims ownership over the data and its approval would be required before the data could be shared. Although the work of the lab is influenced by the expectations of its external partners, the lab itself does not have formal policies or procedures in place for documenting, organizing or maintaining data. As a result, individual students approach data storage and management in different ways. The faculty researcher expressed concern over his students' abilities to understand and track issues affecting the quality of the data, the transfer of data from their custody to the custody of the lab upon graduation, and the steps necessary to maintain the value and utility of the data over time.

To respond to these needs, the University of Minnesota Team developed an online e-learning course composed of seven modules with additional reading and links. The course is self-paced, allowing students to complete it outside of their formal coursework and research activity. After completing the course, students will have a written data management plan for creating, documenting, sharing and preserving their data. The University of Minnesota Team is also offering an in-person group session as part of the instruction that all structural engineering graduate students must take as a means to introduce the concepts of the online e-learning course and to promote its use.

Ecology – University of Oregon

The University of Oregon Team is partnering with a professor who is a Co-PI on a grant funded project measuring the impact of climate change on Pacific Northwest prairies. Their grant is in its final year and the research group is currently focused on wrapping up their work. Multiple streams of data were generated in support of this project, including sensor-generated data on temperature and precipitation, data collected from the field on plant growth and survival, and measurements generated in



the lab from soil and plant sample analysis. While the research team shared field equipment manuals and some standard operating procedures via their internal project web site, they did not have similar written data management guidelines. The interviews revealed that students developed their own practices and approaches for working with their data, and an overall lack of awareness of best practices or guides for managing ecological data. Their approaches were not formally documented, but instead promulgated through the experiences team members brought to the project, or through team discussions and other informal methods.

Given the demands of closing out a research project, the Oregon Team recognized that they would not be able to engage in much face-to-face time with the research group. They developed a two-pronged approach by assigning independent readings that were followed up with a discussion-based instruction session during one of the research team's meetings. The readings included an article outlining basic best practices in data management by the Ecological Society of America, an article on the value of data sharing in climate change science, and a lab notebook guide. These resources served as the focal point for the training session. The topics of the session included lab notebooks and note-taking, data backup and storage, file management, data repositories, metadata and links to tools and further information.

Conclusion

The Data Information Literacy (DIL) project introduces a library-led approach towards educating students about the data competencies they will need to be successful in their careers. Librarians have made great strides in developing information literacy programs that align with the specific objectives of a course or academic department. Likewise, the DIL educational programming for graduate students in data management is meant to be linked to disciplinary and local needs, and integrated into current research practice. A significant objective of this project is to increase the capacity of librarians to engage students and faculty in addressing their educational needs in managing and curating their data. We believe that work in identifying and addressing the educational needs of students with data has just begun and that there are many possible avenues for librarians to explore in this area. To support this objective, the next phase of the DIL project will be to produce a model detailing how other librarians can build their own DIL programs based on our collective experiences and knowledge gained from this project.

Acknowledgements

The DIL project is supported by a National Leadership Grant from the Institute of Museum and Library Services (IMLS), LG-07-11-0232-11.

Project personnel are from Purdue University: Marianne Stowell Bracke, Jake Carlson (PI), Michael Fosmire, Mason Nichols (Graduate Assistant), and Megan Sapp-Nelson; Cornell University: Camille Andrews and Sarah Wright; the University of Minnesota: Jon Jeffryes and Lisa Johnston; and the University of Oregon: Dean Walton and Brian Westra.



References

- Akmon, D., Zimmerman, A., Daniels, M. & Hedstrom, M. (2011). The application of archival concepts to a data-intensive environment: Working with scientists to understand data management and preservation needs. *Archival Science*, 11(3-4), 329–348. [doi:10.1007/s10502-011-9151-4](https://doi.org/10.1007/s10502-011-9151-4)
- Beagrie, N. (2008). Digital curation for science, digital libraries, and individuals. *International Journal of Digital Curation*, 1, 3-16. [doi:10.2218/ijdc.v1i1.2](https://doi.org/10.2218/ijdc.v1i1.2)
- Christensen-Dalsgaard, B., van den Berg, M., Grim, R., Horstmann, W., Jansen, D., Pollard, T., Roos, A., et al. (2012). Ten recommendations for libraries to get started with research data management. Retrieved from [http://www.libereurope.eu/sites/default/files/The research data group 2012 v7 final.pdf](http://www.libereurope.eu/sites/default/files/The%20research%20data%20group%202012%20v7%20final.pdf)
- Carlson, J., Fosmire, M., Miller, C.C., & Nelson, M.S. (2011). Determining data information literacy needs: A study of students and research faculty. *portal: Libraries and the Academy*, 11, 629-657. [doi:10.1353/pla.2011.0022](https://doi.org/10.1353/pla.2011.0022)
- Coulehan, M.B., & Wells, J.F. (2006). Guidelines for responsible data management in scientific research. Retrieved from <http://ori.hhs.gov/images/dblock/data.pdf>
- Dewey, B.I. (2004). The embedded librarian: Strategic campus collaborations. *Resource Sharing & Information Networks*, 17(1-2), 5-17. [doi:10.1300/J121v17n01_02](https://doi.org/10.1300/J121v17n01_02)
- Frugoli, J., Etgen, A.M. & Kuhar, M. (2010). Developing and communicating responsible data management policies to trainees and colleagues. *Science and Engineering Ethics*, 16(4), 753–762. [doi:10.1007/s11948-010-9219-1](https://doi.org/10.1007/s11948-010-9219-1)
- Gabridge, T. (2009). The last mile: Liaison roles in curating science and engineering research data. *Research Library Issues*, 265, 15–21. Retrieved from <http://www.arl.org/bm~doc/rli-265-gabridge.pdf>
- Graham, A., McNeill, K., & Stout, A. (2011). *Managing research data 101*. Powerpoint slides presented at IAP 2011. Retrieved from http://libraries.mit.edu/guides/subjects/data-management/Managing_Research_Data_101_IAP_2011.pdf
- Johnston, L., Lafferty, M. & Petsan, B. (2012). Training researchers on data management: A scalable, cross-disciplinary approach. *Journal of eScience Librarianship*, 1(2), 79-87. [doi:10.7191/jeslib.2012.1012](https://doi.org/10.7191/jeslib.2012.1012)
- Martinez-Urbe, L. & Macdonald, S. (2009). User engagement in research data curation. In M. Agosti, J. Borbinha, S. Kapidakis, C. Papatheodorou & G. Tsakonias (Eds.), *Research and advanced technology for digital libraries: 13th European Conference, ECDL 2009* (Lecture Notes in Computer Science, 5714, pp. 309-314). Berlin: Springer. [doi:10.1007/978-3-642-04346-8_30](https://doi.org/10.1007/978-3-642-04346-8_30)

