Data Information Literacy Case Study Directory

Volume 4 *University of Oregon*
Number 1 *Ecology*

Article 1

2015

Ecology/ Graduate Students/ Westra & Walton/ University of Oregon/ 2012

Brian Westra
University of Oregon, bwestra@uoregon.edu

Dean Walton
University of Oregon, dpwalton@uoregon.edu

Recommended Citation

Westra, Brian and Walton, Dean (2015) "Ecology/ Graduate Students/ Westra & Walton/ University of Oregon/ 2012," *Data Information Literacy Case Study Directory*: Vol. 4: No. 1, Article 1.
http://dx.doi.org/10.5703/1288284315480

TEACHING ECOLOGY DATA INFORMATION LITERACY SKILLS TO GRADUATE STUDENTS: A Discussion-Based Approach

Brian Westra, University of Oregon

Dean Walton, University of Oregon

INTRODUCTION

At the University of Oregon, our Data Information Literacy (DIL) team worked with a vegetation ecology research group that was in the final year of a 4-year grant-funded project. The purpose of the project was to study climate change impacts on Pacific Northwest prairie ecosystems. The librarian team consisted of the science data services librarian and the subject specialist for biology, environmental science, and geology. We partnered with a professor in the Department of Landscape Architecture within the School of Architecture and Allied Arts and a co-principal investigator (co-PI) on a climate change impacts (CCI) study. All other members of the team, including the lead investigator for the Department of Energy grant, were in the Institute of Ecology and Evolution within the Department of Biology. The CCI research group composition changed as students completed projects, but at the out-set of our work, it consisted of two faculty, two postdoctoral research associates, three graduate students, and one research assistant who had completed an undergraduate degree in ecology.

The CCI team investigated the impacts of increased temperature and precipitation on vegetation ecology in prairie ecosystems. The research used three localities, each with plots where temperature and precipitation were artificially increased above ambient levels, and un-manipulated control plots for comparison. Team members researched a variety of factors, such as growth and reproduction of specific plant populations, transpiration rates, and soil characteristics, with individual projects within this larger context.

LITERATURE AND ENVIRONMENTAL SCAN OF ECOLOGICAL DATA MANAGEMENT BEST PRACTICES

To better understand the data management culture of practice within ecology, as well as cur- rent theory and guidance, we examined the literature on research data management (RDM) practices in biology, ecology, and aligned environmental fields, additional generic best practices, and resources.

The literature revealed a robust set of articles on RDM in established ecological and science journals. The ecology and environmental sciences publications were useful not only be- cause of their applicability to the team's needs, but also because sharing such resources from journals in their research domain might lend greater credibility to instructional efforts with the team. Data management, sharing practices, and related topics have been presented in articles, reviews, and columns in journals such as the Bulletin of the Ecological Society of America (Borer, Seabloom, Jones, & Schildhauer, 2009; Fegraus, Andelman, Jones, & Schildhauer, 2005), Trends in Ecology & Evolution (Madin, Bowers, Schildhauer, & Jones, 2008; Michener & Jones, 2012), PloS ONE (Tenopir et al., 2011; Wieczorek et al., 2012), Global Change Biology (Wolkovich, Regetz, & O'Connor, 2012), and Ecological Informatics (Enke et al., 2012; Madin et al., 2007; Michener, 2006; Michener, Porter, Servilla, & Vanderbilt, 2011; Veen, van Reenen, Sluiter, van Loon, & Bouten, 2012).

These articles make the case for good data management practices and outline specific steps that researchers can take to curate their data. One of the most informative and practical articles was Borer et al. (2009), which we shared with the team as a pre-instruction session reading. The authors provided a list of basic data management steps that could be taken with ecology data, such as

- using scripts to record statistical analyses;
- storing and sharing data in nonproprietary formats;
- archiving original raw data;
- using descriptive file naming;
- creating optimal spreadsheet structure and database schema;
- recording full taxonomic names;
- standardizing date and time formats;
- recording metadata early and frequently.

More recent articles take a similar approach, such as advocating for the publication of bio- diversity data (Costello, Michener, Gahegan, Zhang, & Bourne, 2013), and highlighting steps that will make it easier for others to re- use the data one might publish (White et al.,2013).

Data practices in research teams are often not standardized (Borgman, Wallis, & Enyedy, 2007) and vary from one person to another even within research teams under a common faculty member (Akmon, Zimmerman, Daniels, & Hedstrom, 2011).

Science and engineering faculty interviewed at Purdue University and the University of Illinois at Urbana-Champaign wanted graduate students to better understand and implement good metadata practices (Carlson, Fosmire, Miller, & Sapp Nelson, 2011). Metadata standards and usage have been discussed in a number of articles aligned with the CCI team's ecology focus (Fegraus et al., 2005; Jones, Schildhauer, Reichman, & Bowers, 2006; Kunze et al., 2011; Madin et al., 2007, 2008; Michener, 2006; Michener, Brunt, Helly, Kirchner, & Stafford, 1997).

However, some scientists have been reluctant to provide metadata due to the time it would take to create and record it, concerns about misuse of data, and loss of intellectual property rights (Schmidt-Kloiber et al., 2012). Concerns about data ownership may have more to do with "scientific revenue" (Janßen et al., 2011) than intellectual property that would generate income, particularly since these are fields with less potential for monetization of re- search discoveries through technology transfer. Some posit that a consensus-driven agreement on data ownership is needed to further scientific collaboration and avoid conflict (Fraser et al., 2013). In an attempt to facilitate continuing individual control over data sharing, some proposed an "account-based approach to data property rights management" (Janßen et al., 2011, p. 617). A study of the Center for Embedded Networked Sensing (CENS) noted that data sharing transactions can resemble bartering for goods transactions with other trusted colleagues (Wallis, Rolando, & Borgman, 2013).

There are, however, a growing number of influential proponents for open access to research data (Dryad, 2014; National Evolutionary Synthesis Center, n.d.). Funding agency requirements to share research data (Holdren, 2013) will likely accelerate the transition to practices and services in support of open data. Dryad provides a leading example of a data repository, with Creative Commons Zero (CC0) licensing for all submitted data. This is integrated with the publication review process for a growing number of ecology journals (Dryad, 2014).

INTERVIEWS AND RESULTS

We conducted interviews with several members of the CCI team using the DIL interview protocol (available for download at http://dx.doi.org/10.5703/1288284315510). Our interviews were with the collaborating professor, a postdoctoral fellow, the research assistant, and two graduate students (one completing a master's degree, the other working on a doctorate).

Participants in the interviews provided descriptions of the data life cycles of their re- search, though data sharing processes and project close-out practices were less clear because they did not yet have experience in those areas.

The team primarily collected and created tabular data, such as manually recorded field observation data that were later transcribed into spreadsheets, and data downloaded from field devices and sensors. At least one graduate student was conducting laboratory analyses of soil samples, but those tests did not commence until a few months later. They compiled tabular data using Excel and usually imported them into statistical programs for analysis (typically SPSS, though PC-ORD and R were also noted). They graphed results for review, analysis, and presentation or publication using pro- grams such as SigmaPlot and GIMP.

Interviewees were aware of the types (including format) and numbers of data files (computer files or data sheets) collected and created in their work at almost all stages of the data life cycle. Interviewees were less aware of the typical size of any given data file, but were also confident that the size and numbers were small compared to the storage space available on a typical laptop computer.

Interviewees were generally comfortable using their data collection and analysis tools, though some were in the process of learning tools such as SigmaPlot. The type of statistical analysis tools varied based on personal preference and previous experience. Data conversions were typically between Excel and .csv file formats. In limited instances, there were re- projections of spatial data sets.

Most group members were familiar with the concept of metadata, if not the actual term. The types of annotations and other descriptive information associated with data collection varied slightly between individuals for their own unique project data. However, all individuals who collected data in the field used data sheets and field notebooks to annotate data collection issues. They backed up field notes by transcribing them from the field notebook to a lab book that did not leave the lab. The degree of detail in these records varied based on descriptions by the interviewees. Team members held differing views on how readily another person could reproduce their research or reuse the data if relying solely on the notebooks and metadata.

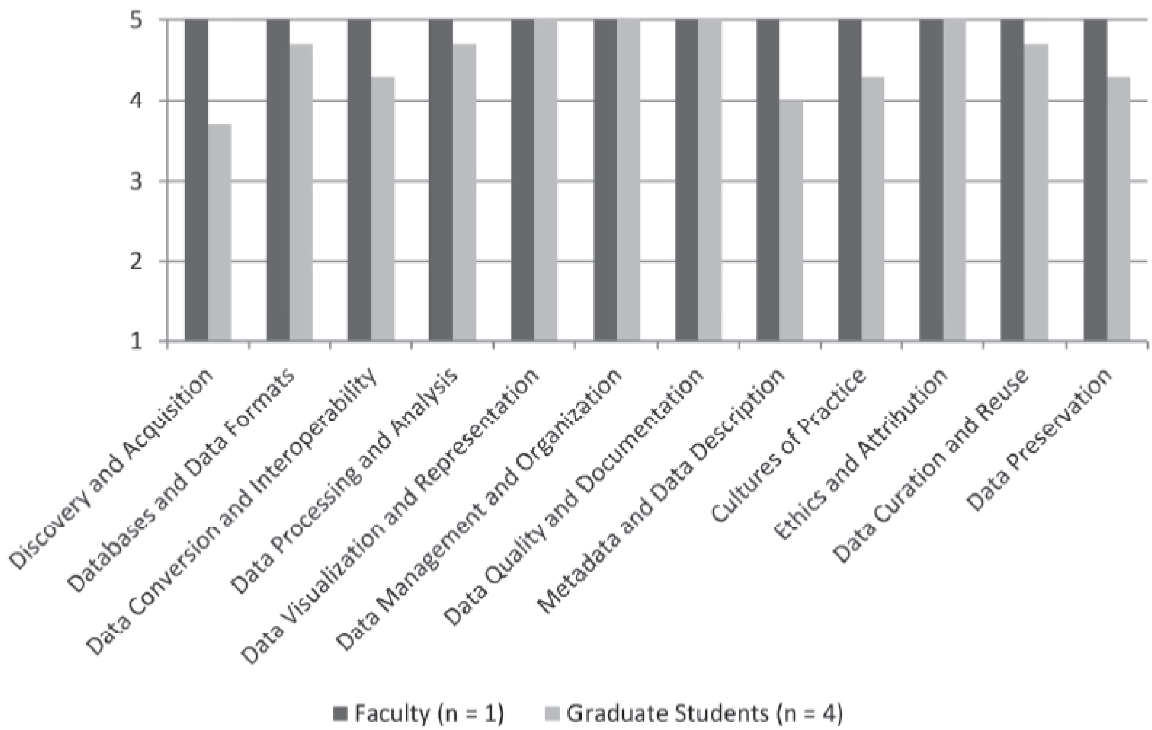

Figure 8.1 Data information literacy competencies as rated by the University of Oregon faculty and graduate students. Ratings based on a 5-point Likert scale: 5 = essential; 4 = very important; 3 = important; 2 = somewhat important; 1 = not important.

There was a lack of consistency across the team in file management practices, from file naming and version control, to storage and backup. All interviewees assumed that they would leave a copy of their data with the faculty, but interestingly, faculty and students both assumed that lab notebooks were the property of the students. Interviewees expressed interest in establishing protocols for handing off work product to the PIs as they completed their respective research projects. Interview responses indicated that the participants were motivated to improve their practices, even as the grant approached its closeout date.

The team members used multiple storage locations, including external hard drives, personal laptops, home computers, and a shared computer in the team's research offices. All team members backed up their data; however, backup intervals differed from person to person.

Because few, if any, had used external data for their own research, and none had published data, their knowledge of practices and resources in these areas was limited. However, all expressed a willingness to share their data and felt that their data could provide a base- line for other studies on the effect of climate change on plant ecosystems. For this reason they believed that their data would be important for many years. Restrictions that they might impose on data sharing were primarily related to proper acknowledgment of the source. They were aware that some journals required the submission of associated data sets with a manuscript, but they did not know how the data would be annotated, preserved, or shared. Most interviewees reported that they had not received training in dealing with intellectual property and data ethics issues and had a limited understanding of privacy, confidentiality issues, and the university's policies on research.

Educational Needs and Priorities

The faculty member who participated in the interview indicated that all 12 of the data literacy competences were important to the research project. He felt that skills in each of the competencies were needed to do proper research and that both he and the students would benefit from training in these areas (see Figure 8.1).

The rest of the team agreed, at least conceptually, about the importance of these data skills. However, in comparison to the professor, the other team members were not as familiar with each of the concepts. Their ratings of the importance of the competencies ranged from "important" to "essential," with the exception of one "I don't know" because of unfamiliarity with metadata concepts. The team reported that self- teaching (or trial and error), peer-to-peer, and student-to-mentor (whether faculty or postdoc) consultations were the common practice for ad- dressing RDM questions as they arose.

A DISCUSSION-BASED APPROACH TO TEACHING DATA INFORMATION LITERACY SKILLS

We scheduled our instruction for the group to be completed during the fall quarter of 2012, which was also the final quarter of their 4-year grant. Seasonal and weather-dependent field data collection events could not be delayed; the potential data to be collected would be irreproducible. With these pressures on the faculty and the rest of the research team, it was reason- able to expect that our access to the team for instruction would be limited.

We negotiated with the two faculty members to schedule a 1.5-hour session in place of a regular team meeting in October. The session incorporated lecture, group exercises, and discussion. Providing training for a small team of research scientists enabled us to design and present the instruction in an informal, conversational setting.

After reviewing the interviews and the results of our literature review, we developed a data management training session on the following:

- Metadata as it relates to documenting, sharing, finding, and understanding data
- File naming
- Data structure and recording methods
- Data repositories and shared data
- Commonly accepted lab notebook policies
- Data ownership and preservation

We believed it would be unrealistic to expect the team to implement many new practices with only a few months left in the project. However, these topics and resources might be applied when handing off data to the faculty and when publishing research results and the skills would applicable to future projects. The topics and respective learning outcomes that we generated for our DIL program are displayed in Table 8.1.

To develop a foundational link to cultures of practice, we provided two assigned readings from the research domain prior to the instruction session and then integrated them into the discussions. A third reading was included to highlight typical policies and best practices for research notebooks. The readings were

- “Some Simple Guidelines for Effective Data Management” from the Bulletin of the Ecological Society of America (Borer et al., 2009);
- a Global Change Biology article on the need for open science and good data management for advancing global change re- search (Wolkovich, Regetz, & O’Connor, 2012);
- an online chapter on lab notebook policies and practices (Thomson, n.d.).

The research team had some turnover be- tween our interviews and the instruction session. Six people attended the training: two faculty, two postdocs, and two graduate students. Only two of this group had participated in the interviews: our faculty partner and one graduate student.

Instructional Components

We created a session outline which included links to examples presented in the class, additional resources, and references (see Appendix A to this chapter).

We anticipated that the readings we as- signed before the team meeting would pro- vide shared understanding and starting points for some of the discussion. The instruction session was a combination of lecture with slides, online resources, hands-on activities, and discussion. Some of the presentation slides were taken from education modules by the DataONE project.

The instruction session began with why data management is important, the risks of poor data practices, and the value of sharing data to the researcher, scientific community, sponsor, and the public.

To direct a discussion of the chapter about lab notebook policies and practices, we asked: (1) what policies or guidelines were new to you? and (2) is there anything you might change or do differently in light of the guide- lines? Here the discussion turned to concerns about the applicability of the notebook practices and policy materials to field research note taking. We highlighted roles and responsibilities for data and notebook stewardship, indicating that these typically are not the property of graduate students, but remain with the PI as a representative of the institution when projects are completed.

TABLE 8.1 - Learning Outcomes for the University of Oregon Training Session

Topics	Learning Outcomes
File formats and conversions	Is aware of and accounts for interoperability issues throughout the data life cycle: considers impacts that proprietary file formats, identifiers, and data access can have on linked data/Semantic Web, and so forth Knows how and why to convert files from one format to another and does so consistently
Publishing data	Knows where to find relevant data repositories and how to evaluate and select where to deposit data, and where to get data Publishing data with Nature, other journals, Dryad?
Preservation and archiving	Knows what data preservation is, why it is important, and what it costs; employs some evaluative criteria in choosing what to preserve and for how long Records metadata in the repository so others can find, understand, use, and properly cite the data set Knows how to properly package and hand off the data to the PI at the close of his or her participation in a project
Data citation	Correctly cites data from external sources Knows what a unique identifier is, and its utility for data citation Knows how to publish/share data/identifiers Understands usage permissions issues, and permissions management tools and restrictions such as creative commons, copyright, and data commons

Next we looked at file management, reviewing common file naming conventions outlined on the University of Oregon data management website, followed by data backup considerations and file conversions and transformations. We discussed data structures and used a short exercise to test whether they could identify errors in a spreadsheet. This exercise was based on materials from the DataONE project.

Several members of the group reported in the interviews that they did not use relational databases for data and were not confident with these concepts. To demonstrate some basic structures of relational databases, we created a hands-on exercise using "flat files" (which were titled sheets of paper) that could be organized into relationships of one-to-one, one-to-many, and many-to-one. The participants arranged the files in a manner that represented data similar to what they might collect and that showed the relationships of the files.

We reviewed Dryad and DataONE Mercury as two examples of ecological data repositories. Navigating to and examining data sets in these two resources provided a concrete introduction to data repositories, metadata standards, data set registration, unique identifiers and DOIs, and linking between data and publications. The data sets provided a foundation for a discussion about publishing data and access and use permissions.

Finally we highlighted the most commonly noted parts of a data citation from the literature, and then opened the rest of the session to questions and discussion about topics of interest to the team.

Assessment

We based our assessment of the DIL program on discussions in the training session, information gathered in two post-training surveys, and conversations and e-mail correspondence with the faculty and other team members. (The training feedback survey questions are in Appendix B to this chapter.) We collected the initial feedback via a Google form linked from the instructional materials. Five of the six attendees filled out the form, while two responded to a more detailed Qualtrics survey that we distributed later. The two faculty were also asked for more information several months later. This section summarizes the collected comments and suggestions and our own observations.

The results of our assessment indicated that we had raised awareness of data management issues and positively impacted the team. Some team members reported that the initial interviews prompted them to think more deeply about how they managed their research data. One researcher reported that since the instructional session the team became more cognizant of data management issues and began to embrace new practices. In particular, the team was more conscientious about providing detailed descriptive information (metadata) in notebooks and electronic records, and the lead faculty member for the project requested that data sets be shared with him in non-proprietary formats to ensure long-term access. Team members reported paying closer attention to data storage, preservation, and sharing issues. More specifically, team members said they planned to

- "do a better job of planning for data management at the onset of a project";
- "explore my options for online backups of my data";
- "save long-term data in a .csv format and provide metadata for that file."

One of the faculty reported that the training had "brought me up to date with growing expectations for sharing of data . . . gave me deeper impetus to apply sound meta practices so that future users could understand how and why data was developed and processed the way it was." The sessions "changed the degree to which we systematically apply protocols for data management across all aspects of the project. They also gave us useful insight into the resources available for data curation."

The team valued guidance that was either very closely aligned with the team's data acquisition practices or easily translated into their workflow and publication processes. Several respondents said they appreciated the open discussion on specific needs and questions that occurred at the end of the session. Several said they would have rather spent more time in interactive work with an immediate application to their current research and data management tasks, and less time on overview and basic instruction.

The article by Borer and colleagues (2009) that provided data management guidelines was particularly well received and provided a useful introduction to a number of practices that were at the heart of the session. The article by Wolkovich, Regetz, and O'Connor (2012) was not mentioned as often in the assessment, but it provided a strong case for data sharing in the multidisciplinary field of global change research, the very topic of the CCI project. Though not its primary focus, the article included a useful table listing some of the actions and skills needed for data and code sharing, as well as supporting website links. We included the chapter by Thompson on lab notebooks in our DIL Program as it had been used by a faculty member in the Department of Human Physiology to introduce good notebook practices to new graduate students. However, the chapter elicited several surprisingly strong negative comments from other participants. One of the faculty and at least one postdoc in the CCI group believed it had no application to their research workflow. Admittedly, the guidelines were established for a research laboratory setting more typical of biochemistry than ecology, but we had believed readers could interpret and apply the recordkeeping guidelines to other forms of research documentation.

DISCUSSION

One of the strengths of the DIL model is that the structured interviews provide librarians with a detailed understanding of the RDM practices, skills and priorities of a particular person or team. That information and the literature translate to targeted instructional interventions. Training can be tailored to the specific needs of the research group, though the amount of content will be determined by the length and number of sessions that can be accommodated by the research team's schedules and faculty prerogatives.

The interview process can open new lines of communication and opportunities to provide RDM services to research faculty, graduate students, postdocs, and research assistants. The interviews and associated conversations raise awareness of library services for research scientists. For the librarians, these experiences can provide insight into the needs of graduate students, and enable librarians to expand their understanding of the research domains they serve.

The instruction session included conceptual information for the competencies and examples of applied RDM principles. The CCI group clearly favored context-based applied learning and application exercises for their instruction. We incorporated some lecture and slides to provide context for some of the DIL competencies. In retrospect, the Borer article was well received and might have sufficed since it grounded the topics in an ecology research ethos. The lecture was not as productive nor well received in this small group set- ting. In the future we plan to put much more emphasis on localized use cases, applied practices, and open discussion.

Developing specific and relevant DIL pro- grams can be time consuming, but it will result in a more engaged group that can adopt new skills toward implementation of better RDM practices. To be effective DIL programs have to respond to the needs of researchers within the environment they inhabit. Researchers are un- der pressure, particularly when time-sensitive field work is on the line. They also want more efficient workflows so they can increase their productivity. This is reflected in a desire to have more immediate application outcomes, through both streamlined and timely instruction and demonstrable improvements in RDM practices. Librarians can gain support for training by connecting learning outcomes to potentially lower risk of data loss, higher research impact, more collaborations, more competitive funding proposals, and more efficient data organization and search and discovery.

There are several considerations in applying the DIL model to smaller research teams. Even with small groups consisting of PIs, re- search associates and postdocs, and graduate students, there may be a high degree of variability in skills across the team, and individuals may be engaged in highly differentiated projects of their own with unique workflows and data management concerns. This will need to be addressed in planning the instruction, and probably acknowledged at the outset of any training. Highly stratified skill sets might be accommodated by distributing this expertise across groups if the team is large enough. In our case the climate change project provided a unifying theme and data sources, and there was some uniformity due to shared project management and logistics, as well as common research methods and workflows across the group.

Should we work with another group that relies on field data collection, we will focus instruction on field notes and documentation methods, and fill in any gaps about policy application, rather than providing laboratory notebook guidance. Clearly several members of the team were looking for materials specific to the form and content of documentation they were using in the field.

In most of the data librarian’s discussions with researchers about RDM, faculty typically preferred that we speak directly with the graduate students and postdocs who were conducting research. Faculty were reluctant to unilaterally impose RDM practices on the team. However, faculty buy-in is critical, and a professor can exert a lot of influence on the DIL process, whether through the degree of librarian access to the students, or via the values and attitudes they impart to the team regarding data sharing and funding agency requirements. This should be kept in mind as librarians select faculty partners and research teams for the significant investment that the DIL model requires. Similarly, creating and nurturing a good working relationship with the team is important and can lead to other collaborations and support opportunities after the initial instruction has been provided.

There are other considerations to be made in selecting groups to participate in implementing the DIL model. The academic calendar and grant cycle must be considered when thinking about optimal timing for scheduling interviews and instruction events. These factors may unduly compress the window of opportunity for interactions with the students. The number of master's students and PhD candidates who are on the team and at what stage they are in their program may influence the type and timing of instruction you can implement.

The educational experiences of the team members may sometimes lead to unforeseen ideas. We were working with a relatively small research group and chose to expand our investigation of the team's practices by including a postdoc and a research assistant in the interviews. The research assistant, who had not yet started a graduate program, received what we considered to be excellent training in re-cording metadata as an undergraduate student. She had worked at a field station previously, where students are required to document field work with metadata and pass reviews of their field notes before they could begin their own projects. Data sets from the students' field projects were deposited for public access. This type of experiential learning, integrated directly with and reinforced by reviews of ongoing re- search practice, is a model that we plan to explore further.

The DIL project may ultimately highlight skills that should be integrated into the curriculum for all STEM students. Within the CCI team a few specific components of DIL are addressed to varying degrees. For instance, our faculty partner in this project remarked that training in information presentation and graphics is a required aspect of the curriculum for students in his department (landscape architecture). In contrast, typical biology students learned data visualization on their own or tangentially through exposure to graphing in foundational statistics courses.

CONCLUSIONS

The DIL model was a very useful tool in developing DIL training for graduate students. The process provides a useful categorization of RDM skills through which research faculty can articulate areas of concern and priori- ties for skill development for themselves and their graduate students. Structured interviews of the students enabled us to identify the data management skills and perspectives of graduate students conducting research on vegetation ecology, and to prepare, present, and assess an instructional session with the team.

Research teams do not always have time for long-term instructional interventions, particularly when grant deadlines are looming. In these situations, shorter, discussion-based sessions focused on specific local DIL issues can yield a measurable positive impact on graduate student RDM skills and attitudes.

It would be risky to assume that the needs and learning outcomes from this particular team were the same as those from other ecology research teams. Taken with care, however, the literature and lessons we learned about RDM practices and DIL instruction through working with this team provided us with a good foundation for working with other graduate students who conduct field research in the biological sciences.

Our results also informed the model by showing that a 1.5-hour training session can be an effective way of supporting and developing graduate student DIL competencies. However, there are caveats to the method. A short window for instruction significantly limits the number of topics and degree of detail to be covered. Various aspects of the training may gain more support if they are previewed or negotiated with the faculty partner(s). There are many factors that will affect uptake, but active, context-based learning activities and discussions carry the potential to help graduate students understand these skills and integrate them into their research practices.

Finally, positive and supportive interactions with graduate students can set the stage for further instructional efforts and other RDM services by librarians.

NOTE

This case study is available online at http://dx.doi.org.10.5703/1288284315480.

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. http://dx.doi.org/10.1007/s10502-011-9151-4

Borer, E. T., Seabloom, E. W., Jones, M. B., & Schildhauer, M. (2009). Some simple guidelines for effective data management. Bulletin of the Ecological Society of America, 90(2): 205–214. http://dx.doi.org/10.1890/0012-9623-90.2.205

Borgman, C. L., Wallis, J. C., & Enyedy, N. (2007). Little science confronts the data deluge: habitat ecology, embedded sensor networks, and digital libraries. International Journal on Digital Libraries, 7(1–2), 17–30. http://dx.doi.org/10.1007/s00799-007-0022-9

Carlson, J., Fosmire, M., Miller, C. C., & Sapp Nelson, M. (2011). Determining data information literacy needs: A study of students and re- search faculty. portal: Libraries & the Academy, 11(2), 629–657. http://dx.doi.org/10.1353/pla.2011.0022

Costello, M. J., Michener, W. K., Gahegan, M., Zhang, Z.-Q., & Bourne, P. E. (2013). Biodiversity data should be published, cited, and peer reviewed. Trends in Ecology & Evolution, 28(8), 454–461. http://dx.doi.org/10.1016/j.tree.2013.05.002

DataONE (n.d.). DataONE education module: Data entry and manipulation. http://www.data one.org/education-modules

Dryad. (n.d.). DRYAD. http://www.datadryad.org/repo/

Dryad. (2014). Joint data archiving policy (JDAP). Retrieved from http://datadryad.org/pages/jdap

Enke, N., Thessen, A., Bach, K., Bendix, J., Seeger, B., & Gemeinholzer, B. (2012). The user's view on biodiversity data sharing—Investigating facts of acceptance and requirements to realize a sustainable use of research data. Ecological Informatics, 11, 25–33. http://dx.doi.org/10.1016/j.ecoinf.2012.03.004

Fegraus, E. H., Andelman, S., Jones, M. B., & Schildhauer, M. (2005). Maximizing the value of ecological data with structured metadata: An introduction to ecological metadata language (EML) and principles for metadata creation. Bulletin of the Ecological Society of America, 86(3): 158–168. http://dx.doi.org/10.1890/0012-9623(2005)86[158:MTVOED]2.0.CO;2

Fraser, L. H., Henry, H. A., Carlyle, C. N., White, S. R., Beierkuhnlein, C., Cahill, J. F., Jr., . . . Turkington, R. (2013). Coordinated distributed experiments: An emerging tool for testing global hypotheses in ecology and environmental science. Frontiers in Ecology and the Environment, 11(3), 147–155. http://dx.doi.org/10.1890/110279

Holdren, J. P. (2013). Increasing access to the results of federally funded scientific research (Executive Office of the President, Office of Science and Technology Policy, memorandum for the heads of executive departments and agencies). Retrieved from http://www.whitehouse.gov/sites/default/files/micro sites/ostp/ostp_public_access_memo_2013.pdf

Janßen, T., Schmidt, M., Dressler, S., Hahn, K., Hien, M., Konaté, S., . . . Zizka, G. (2011). Ad- dressing data property rights concerns and pro- viding incentives for collaborative data pooling: The West African vegetation database approach. Journal of Vegetation Science, 22(4), 614–620. http://dx.doi.org/10.1111/j.1654-1103.2011.01271.x

Jones, M. B., Schildhauer, M. P., Reichman, O.J., & Bowers, S. (2006). The new bioinformatics: Integrating ecological data from the gene to the biosphere. Annual Review of Ecology Evolution and Systematics, 37(1), 519–544. http://dx.doi.org/10.1146/annurev.ecolsys.37.091305.110031

Kunze, J. A., Cruse, P., Hu, R., Abrams, S., Hastings, K., Mitchell, C., & Schiff, L. R. (2011). Practices, trends, and recommendations in technical appendix usage for selected data-intensive disciplines [Report]. Retrieved from eScholarship University of California website: http://escholarship.org/uc/item/9jw4964t#page-1

Madin, J., Bowers, S., Schildhauer, M., Krivov, S., Pennington, D., & Villa, F. (2007). An ontology for describing and synthesizing ecological observation data. Ecological Informatics, 2(3), 279–296. http://dx.doi.org/10.1016/j.ecoinf.2007.05.004

Madin, J. S., Bowers, S., Schildhauer, M. P., & Jones, M.B. (2008). Advancing ecological research with ontologies. Trends in Ecology & Evolution, 23(3), 159–168. http://dx.doi.org/10.1016/j.tree.2007.11.007

Michener, W. K., & Jones, M. B. (2012). Ecoinformatics: Supporting ecology as a data-intensive science. Trends in Ecology & Evolution, 27(2), 85–93. http://dx.doi.org/10.1016/j.tree.2011.11.016

Michener, W. K., Porter, J., Servilla, M., & Vanderbilt, K. (2011). Long term ecological research and information management. Ecological Informatics, 6(1), 13–24. http://dx.doi.org/10.1016/j.ecoinf.2010.11.005

Michener, W. K., Brunt, J. W., Helly, J. J., Kirchner, T. B., & Stafford, S. G. (1997). Non-geospatial metadata for the ecological sciences. Eco- logical Applications, 7(1), 330–342. http://dx.doi.org/10.1890/1051-0761(1997)007[0330:NM FTES]2.0.CO;2

Michener, W. K. (2006). Meta-information concepts for ecological data management. Ecological Informatics, 1(1), 3–7. http://dx.doi.org/10.1016/j.ecoinf.2005.08.004

National Evolutionary Synthesis Center (n.d.). NESCent data, software and publication policy. Retrieved from http://www.nescent.org/public_documents/Informatics_Policy/Data_and_Software_Policy.pdf

Schmidt-Kloiber, A., Moe, S. J., Dudley, B., Strack- bein, J., & Vogl, R. (2012). The WISER meta- database: The key to more than 100 ecological datasets from European rivers, lakes and coastal waters. Hydrobiologia, 704(1), 29–38. http:// dx.doi.org/10.1007/s10750-012-1295-6

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A.U., Wu, L., Read, E., . . . Frame, M. (2011). Data sharing by scientists: Practices and perceptions. PLoS ONE, 6(6), e21101. http://dx.doi.org/10.1371/journal.pone.0021101

Thomson, J. A. (2007). How to start–and keep–a laboratory notebook: Policy and practical guide- lines. In A. Krattiger, R. T. Mahoney, & L. Nelson (Eds.), Intellectual property management in health and agricultural innovation: A handbook of best practices (Chapter 8.2). Oxford, UK: MIHR. Retrieved from http://www.iphandbook.org/handbook/ch08/p02/

Veen, L. E., van Reenen, G. B. A., Sluiter, F. P., van Loon, E. E., & Bouten, W. (2012). A semantically integrated, user-friendly data model for species observation data. Ecological Informatics, 8, 1–9. http://dx.doi.org/10.1016/j.ecoinf.2011.11.002

Wallis, J. C., Rolando, E., & Borgman, C. L. (2013). If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology. PLoS ONE, 8(7), e67332. http://dx.doi.org/10.1371/journal.pone.0067332

White, E. P., Baldridge, E., Brym, Z. T., Locey, K. J., Mcglinn, D. J., & Supp, S. R. (2013). Nine simple ways to make it easier to (re)use your data. PeerJ PrePrints, 1, e7v2. http://dx.doi.org/10.7287/peerj.preprints.7v2

Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., . . . Vieglais, D. (2012). Darwin Core: An evolving community- developed biodiversity data standard. PloS ONE, 7(1): e29715. http://dx.doi.org/10.1371/journal.pone.0029715

Wolkovich, E. M., Regetz, J., & O'Connor, M. I. (2012). Advances in global change research re- quire open science by individual researchers. Global Change Biology, 18(7), 2102–2110. http://dx.doi.org/10.1111/j.1365-2486.2012.02693.x