A GRADUATE CLASS IN RESEARCH DATA MANAGEMENT

Lawrence O. Schmidt and Joseph H. Holles
University of Wyoming • Laramie, WY 82071

Effective practices in research data management (RDM) are integral to research for two reasons: (1) RDM leads to thoughtful and thorough use of the data for sharing and publication, and (2) RDM has become a required part of federal funding for research. As a result, practical knowledge and training in RDM is critical for researchers.

The National Academy of Sciences (NAS) has advocated for: (1) stewardship of research data to ensure research integrity and data accessibility; (2) the development of standards and policies regarding the dissemination and management of data; and (3) data management training for all researchers.[1] Carlson, et al. follow this up with "… it is not simply enough to teach students about handling incoming data, they must know, and practice, how to develop and manage their own data with an eye toward the next scientist down the road."[2] Federal agencies (*e.g.*, NSF,[3] NIH,[4] and USGS[5]) are also requiring the submission of a Data Management Plan (DMP) when submitting proposals for funding. Thus, it is clear that a systematic and thorough education on RDM is appropriate and necessary for a graduate curriculum. Thielan, et al. also present a case for how a RDM course can meet multiple outcomes associated with ABET accreditation.[6]

Literature studies have shown that faculty understand the need for RDM education for their students (and themselves) and would benefit from experts "helping us do it right the first time."[7] Similar work also found that faculty reported graduate students were not prepared to manage data effectively, but also acknowledged they (faculty) could not provide adequate guidance or instruction.[2] Carlson, et al. note that "graduate students are often expected to carry out most or all of the data management tasks for their own research."[8] This work also pointed to several faculty-perceived RDM

Joseph H. Holles *is an associate professor of chemical engineering at the University of Wyoming. He received a B.S. in chemical engineering in 1990 from Iowa State University and a M.E. and Ph.D. from the University of Virginia in 1998 and 2000. His research area is nanoscale materials design and synthesis for catalytic applications with an emphasis on structure/property relationships and in-situ characterization.*

Lawrence O. Schmidt *is an associate librarian at the University of Wyoming. He holds B.S. degrees in chemistry and biology, an M.S. degree in environmental engineering from Montana State University, and received an MLS from Emporia State University in 2002. His research interests are related to providing undergraduate and graduate students with information, data, and science literacy skills that will allow them to succeed in a global economy.*

shortcomings: lack of formal training in data management, absence of formal policies governing data in the lab, and self-directed learning through trial and error.[8] The National Science Foundation suggests "full engagement of students is vitally important since they are in a special position to inspire future students with the excitement and understanding of cyberinfrastructure-enabled scientific inquiry and learning."[9] Faculty studies concluded that researchers felt that some form of data information literacy was needed for their students.[2]

Graduate student education on RDM has generally taken one of two approaches: stand-alone courses or seminars/workshops. The stand-alone course has been used by information science programs,[10,11] while the workshop/seminar approach is commonly offered through libraries.[12] The stand-alone course offers the advantage of in-depth material coverage while the workshops require less time. However, workshops and seminars are often not for credit and can suffer retention issues.[12] Two drawbacks to library-delivered seminars/workshops include the need for continuing advocacy to offer the sessions and that most librarians do not have a lot of basic research experience to provide strong in-class examples. Conversely, for stand-alone courses, Carlson, et al. observe that it is difficult to attract students to courses that reside outside of their discipline.[8] More recently, for-credit courses specifically focused on research data management have been offered through the graduate school and taught by librarians[13] or offered by specific research-focused departments and taught by a combination of librarians and faculty.[14]

The goal of this paper is to describe a graduate course in RDM taught in a specific discipline, without necessarily being discipline-specific. This course is designed to provide the in-depth RDM knowledge that the NAS and NSF encourage for graduate students and that faculty acknowledge they cannot provide. The course herein was co-taught by a librarian and a faculty member with an active research program. This was in order to deliver broad knowledge on RDM tools and standards from the expertise of the librarian while also allowing the faculty member to provide research-focused examples and experience. Another benefit provided by this arrangement was the added value of different perspectives on research management processes and the research lifecycle. This paper describes the course, course materials, lecture topics, assignments and projects, lessons learned, and assessment tools.

Methods: Research Data Management was taught as a 3-credit graduate course during the Fall semester of 2016. The course met for 42 class periods of 50 minutes each. For its initial offering the course was taught under the ChE 5150, Special Topics, course number. Ten graduate students (nine chemical and one civil engineering; three in the first year of graduate study, one in the second year, and six in the third year) took the initial course. This course was offered as an elective course and had no pre-requisites.

The major goals of the course are: (1) expose the students to broad concepts and best practices of research data management; (2) bring in outside experts to demonstrate specific areas of RDM; and (3) provide a focused application of RDM to active research projects. These goals were then separated into individual learning objectives as reflected in the assessment, below.

One recommended textbook was used for the course: *Data Management for Researchers* by Kristen Briney.[15] Additional resources for the course included DMPtool[16] and the Data Curation Profile.[17] Krier and Strasser's *Data Management for Libraries: A LITA Guide* was also used to develop several lectures.[18]

Pre-course and post-course assessment was performed to determine the students' knowledge about their current laboratory RDM practices and eight specific areas of RDM. For assessment of current laboratory practices, the student responses were yes/no/don't know. For assessment of specific areas of RDM, the students were asked to rate their level of knowledge in each of the eight areas on a scale from poor (1) to excellent (5). The average student response for each question was determined for both the pre- and post-assessment. The pre- and post-assessment variance for each question was determined to be equal using an f-test at α=0.20. Once the variances were shown to be equal, hypothesis testing using a t-test at α=0.05 demonstrated that the post-assessment mean (average) exceeded the pre-assessment mean for each question. Average normalized gain <g> for each assessment question was determined to quantify how large each effect was.[36]

Pre- and post-course feedback from the students was obtained to develop the initial course and refine future offerings of the course. Feedback was also obtained from the faculty who volunteered to participate in the final project to determine if they found the process effective.

Results: One recommended textbook was selected for the course: *Data Management for Researchers: Organize, Maintain and Share your Data for Research Success* by Briney.[15] The Briney text covers all of the relevant areas of RDM from the perspective of the academic researcher. It includes a mixture of high-level topics (*e.g.*, data lifecycle, planning for data management) together with more practical information (*e.g.*, documentation, improving data analysis) for a student researcher to apply to their own RDM. The course topics were developed from the Briney text together with material from a similar RDM course at Oregon State University.[13] *Managing and Sharing Research Data* by Corti, et al. may also provide useful information.[19] However, the Corti book was more focused on researchers in the United Kingdom and a number of the specific examples were too focused on the U.K. There are also a number of other RDM-focused textbooks but they tend to be written from the perspective of the research librarian or information technology specialist. Since these texts are not student-researcher focused, we did not select them for the course.

The class schedule for the initial course offering is shown in Table 1. The individual course topics were divided into three general areas: (1) broad components of RDM; (2) examples of specific concepts in RDM as demonstrated by guest experts; and (3) focused application of RDM to ongoing research projects. For the broad concepts in RDM, the Briney text provided the course materials. These topics included, for example, Data and Data Lifecycles, Describing Your Data, and Planning Your Research Topic. Approximately 10 lectures were based on the text and these lectures are noted in the table with the specific chapter from Briney in parenthesis. An additional five lectures completed the broad concepts portion of the course and included RDM sharing mandates, DMP Tool,[16] and reference managers. Eight lectures were provided by guest experts and were focused on specific applications of RDM. These topics included, for example, metadata, data management on an interdisciplinary project, RDM tools available within the university, and the primary investigator for a multi-university data intensive research project. All of these lectures include "guest" in the topic title in Table 1.

The remainder of the lectures were focused on applications of RDM to an ongoing research project. First, DMPtool[16] was used by the students to develop a DMP for their research project. This application is somewhat high-level (overarching with minimal details) but funding-source specific. DMPtool is an open-source online tool to assist researchers in creating data management plans as required by specific funding agencies as part of the proposal application process. The second application involved the development of a Data Curation Profile (DCP) and the application of the DCP to an active research project. The purpose of the DCP is to "provide a foundational base of information about a particular data set that may be curated…" and is intended "to address the needs of an individual researcher or research group with regards to the 'primary' data generated or used for a particular project."[17] For this course, the class developed a DCP modeled on the Data Curation Profiles Toolkit from Purdue University.[17, 20-22] The objective of the DCP assignment was to develop a DCP from the researcher's perspective. Once developed, the DCP was applied to an ongoing research

TABLE 1
Typical Class Schedule

Week	Class	Topic	Week	Class	Topic
1	1 2 3	Introduction/Syllabus What is Research Data? RDM and Sharing Mandates	9	24 25 26	DCP Draft/Revision DCP Draft/Revision Guest – PI on Multi-University Data Intensive Project
2	 4 5	Holiday Overview of Data & Lifecycles (1 & 2) Planning Your Research Project (3)	10	27 28 29	Guest –RDM in the Humanities DCP Draft/Revision DCP Draft/Revision
3	6 7 8	Organization, File Naming, and Structure (5) Lab notebooks & Readme files (4) DMP Tool	11	30 31 32	Long Term Storage & Preservation (9) Guest – RDM for Human Subjects (7) Data Sharing (10)
4	9 10 11	Resources at Univ., National, and International Tools Support Reference Managers	12	33 34 35	Student Projects Student Projects Student Projects
5	12 13 14	Citation Management Data Curation Profile (DCP) Setting up for Interview	13	36	Student Projects/Help Session Thanksgiving Thanksgiving
6	15 16 17	Class Canceled Guest - Metadata Guest – Orchid Profile	14	37 38 39	Data Sharing and Governance (10 & 11) Guest –Data for Re-use Improving Data Analysis (6)
7	18 19 20	DCP Preview DCP Exercise Guest – Data Management on Interdisciplinary Project	15	40 41 42	Student Presentation Student Presentation Post-Assessment & Student Feedback
8	21 22 23	Guest - Canceled Guest – Univ. Data Archive and Management DCP Draft/Revision	16		Final Exam Week

Numbers in parenthesis indicate Chapter of Briney.

problem as the Final Project (Table 2). Further details on the project are discussed below.

Assignments for the course and the objectives for each assignment are shown in Table 2. The student work for the course falls under four areas: (1) individual assignments to reinforce topics from class; (2) student reflections on guest speakers with potential applications to the student's RDM; (3) the final project focused on generation of a DCP; and (4) a student reflection on their RDM practices.

A total of eight guest speakers presented lectures for the course on specialized areas of RDM. The goal of inviting the guest speakers was to provide the students specific examples or applications of RDM that were outside the course instructor's areas of expertise. For each guest speaker, each student completed a student-reflection assignment. The learning objective of these assignments was for each student to consider possible RDM issues associated with the talk and also to consider possible applications and changes to the student's research RDM resulting from the talk. Two of the guest speakers covered or reinforced topics from the Briney text (managing sensitive data and data reuse). Three of the guest speaker topics reinforced each other and examined the same project from three different angles. These three topics examined a large, multi-university, data-intensive research project. The three perspectives were: day-to-day data management by an information technology expert, the use of the university's data curation archive and management of the repository by an expert, and the primary investigator for the project.

The objective of the final project was a focused application of RDM to an active research project through the use of DCP. A DCP contains two types of information about a data set: information about the data set itself and information about the researcher's need for the data.[17] The DCP provides a means to guide the management and curation of the data through its lifecycle.[17] A DCP can also contribute towards the development of a DMP required as part of a funding proposal submission. As noted above, several class periods were spent reviewing, and for use in this project, revising the DCP Toolkit developed by Purdue University.[17, 20-22] The approach used in this class was to simplify the DCP where possible and to focus the DCP on the researcher. Using the Purdue DCP provided a base case for discussion as the students considered the collections storage, reuse, and sharing of data. For this project, the class was divided into four teams (of two or three students) each of which worked with one of four research-active faculty members from across the campus. The research faculty were volunteers and chosen broadly from across campus (physics, chemistry, civil engineering, and chemical engineering). Two of the faculty were chosen due to known specific consideration about their research data while two were chosen based on a desire to participate in developing RDM standards for their research laboratories. Each student team then prepared for and interviewed the researcher, and developed a DCP with RDM best practices for the faculty member/research project. To share their experiences and observations, each student team then prepared and delivered a group presentation on the results of their DCP development for the class. Following each presentation, each student in the class completed a student reflection on how lessons from that work could be applied to their research.

TABLE 2
Assignments

Topic		Objectives
Perceptions of Data		Holistic examination of data; define student knowledge base
Data Lifecycle		Overview of all aspects of data management
DMP Tool		Experience with DMPs and creating DMPs
DCP Module Refinement		Critical examination of RDM details
Guest Speaker Reflections		Potential application of speaker's experience to student RDM
Final Project		Application of the DCP developed by the students to campus research faculty
A	Planning Document	Establish roles and tasks; examination of DCP as applied to researcher; practice session
B	Interview Session	Interview of researcher to gain knowledge for development of DCP
C	Combined Document	Synthesis of individual material from interview into one document for refinement into DCP
D	Post Interview Reflection	Examination of positives/negatives of interview and DCP template
E	Data Curation Profile	Suggested RDM best practices for the researcher
F	Presentation Outline	Distillation of DCP experience into presentable format
G	Presentation	Sharing of experience/knowledge with the broader class; Presentation skills
H	Student Presentation Reflections	Potential application of results/observation from other groups to student's RDM
Student Data Reflection		Self-examination of the student's RDM and potential changes/additions to data management from taking the course

TABLE 3 **Assessment Results**			
	Assessment Results		
	Pre-course	Post-course	Avg. Normalized Gain
Management and Planning	2.00	3.70	0.57
Archiving and Preservation	2.70	3.50	0.35
Sharing and Reuse	2.40	3.40	0.38
Legal and Ethical Concerns	2.38	3.50	0.43
Documentation and Metadata	2.43	3.80	0.53
Storage, Backup, and Security	2.56	3.60	0.43
Organization	3.00	3.80	0.40
Types and Formats	2.50	3.90	0.56

The final assignment of the course was an individual student reflection of the current RDM for their thesis/dissertation work. The intent of this work was for the student to take a holistic examination of their current RDM approaches and to contemplate and suggest potential changes and additions to their DCP based on the course topics, guest speakers, and DCP project.

Prior to the course, the students' background knowledge in eight areas of RDM was assessed. The pre-assessment results of this assessment are shown in Table 3. Additional information was also obtained to determine other background about the student's research funding, research topics, laboratory RDM practices, and RDM needs. The pre-course questionnaire also asked the students to provide specific areas of knowledge that they would like to obtain from the course. The goal of this question was to determine information to potentially modify the course to better meet the student needs. As a result, discussion on how to create a file-naming system and ways to share data among multiple institutions were added to the course. Following completion of the course, the student's RDM knowledge in the same eight areas was again assessed. Post-assessment results are also shown in Table 3. The post-course questionnaire also asked if the covered topics delivered what the students were interested in, what was missed, and what was found to be most useful.

Pre- and post-course assessment showed that student-rated knowledge for each of the eight RDM topics increased as a result of the course by an average of 1.15 points. The largest increases were for data management planning (1.7) and data types and formats (1.4), while the smallest increases were for data organization and data archiving and preservation (0.8 each). Hypothesis testing comparing pre- and post-course assessment averages demonstrated that the means were different. Thus, the students reported knowledge growth in all eight areas. Student knowledge growth can also be shown since six responses on the pre-course assessment were "don't know" while there were none on the post-course assessment. When examined quantitatively using average normalized gain <g>,[36] the average <g> across all eight topics was 0.45 with individual <g> values ranging from 0.35 to 0.57. According to Hake, this would be a "medium-g" course for value 0.7 > "<g>" > 0.3.[36]

Additional pre- and post-course assessment showed that the students improved their knowledge of laboratory protocols and their confidence in their ability to write a DMP. Three questions were focused on these topics: (1) Does your laboratory have a DMP?; (2) Could you write a DMP for your research?; and (3) Do you have protocols for managing your research data? Pre-course assessment resulted in 12 of 30 responses to these three questions being "Don't Know." Post-course there were a total of only four "Don't Know." In addition, the number of students who could write a DMP for their research increased from two to nine and the number of students with protocols for managing their research data increased from four to eight.

All materials associated with the course are available at: <https://doi.org/10.5072/FK2MG7P42S>. This includes: lecture notes, online resources, assignments, project, assessment tool, and some of the guest speaker presentations.

DISCUSSION

The direct comparisons for this course are limited, however a number of different approaches to educating students on RDM over the last few years demonstrate the desire to improve the education on this topic. Muilenberg, et al. developed a seven-module course at the University of Washington taught by librarians that met weekly for one-hour workshops.[12] The weekly topics were similar to the chapters from the Briney text and developed from the New England Collaborative Data Management Curriculum (NECDMC).[23] One drawback of this approach was the low retention, and the authors state that future plans will be to offer the course for credit through a department with broad campus reach such as the graduate school.[12] In contrast, the Information Studies (IS) Department at UCLA offers a full 11-week class for four quarter-credits on Data Management and Practice.[11] This course is taught by a faculty member in the IS program. It is more broadly focused on data management in general, but does contain some material more focused on RDM and is available to students outside of the IS program.

Cornell University took a more discipline-specific approach by offering a course through the Natural Resources program titled "Managing Data to Facilitate Your Research."[14]

This for-credit course met for six sessions for one credit and was co-taught by a data librarian and a faculty member in the natural resources program. The topics for this course paralleled those of Muilenberg.[12] Offering the course through and in conjunction with a department allowed for subject-specific examples and helped the students better understand what was happening in the RDM process and why it should happen. From this offering, the authors concluded that they would like to expand beyond the six-session course to introduce more exercises and include more hands-on learning such as the creation of a DMP.[14]

The course most closely comparable to this one is "Research Data Management" offered in partnership between the library and the graduate school at Oregon State University.[13] The Oregon State course was a 2-credit, 11-week long (one quarter) course meeting twice per week for 50 minutes. The course was taught by a librarian (with a Ph.D. in a scientific field). The course topics again paralleled those of NECDMC.[23] Lessons learned include the student's desire for more "real-world" context and practical applications of theoretical topics. Major coursework included a DCP as a midterm assignment and DMP for a final project. For our course, this was reversed: the DMP was an individual assignment and the DCP was a group, final project. The DMP and DCP were included because they provided the "most straightforward mechanism for facilitating student self-reflection on how theoretical data management practices were directly related to their research processes."[13] The author also recommends that a better approach may have the students create sections of the DMP as assignments throughout the course.

Another similar course, "Data Management and Stewardship in the Climate and Space Sciences," was offered through the Climate and Space Sciences and Engineering department at the University of Michigan.[6] This course was a 2-credit, 14-week course meeting weekly for 2 hours. It was taught by a team of three librarians based on discussions with the department faculty and graduate students to define their needs and desires. Student coursework included weekly personal reflections, weekly homework assignments to write sections of a DMP, and a final project of a presentation on their topic of choice and compilation of DMP components and instructor feedback into a final DMP.

Similar to this paper, Thielan, et al. used the Briney text, employed multiple (five) guest speakers, used student reflections, and focused on a class structure where the students helped to teach each other. Under lessons learned, the authors note that they would revise the DMP work to be more explicit about their intent.[6] Final DMPs ranged from three to 26 pages. Based on our experience, this might be an opportunity to divide this work into (1) a short executive summary DMP appropriate for a funding agency (grant/funding-agency focused) and (2) a more thorough DCP appropriate for laboratory use to serve as a personal data management plan (researcher-focused). These authors also suggest a course model co-taught between research faculty and a librarian in order to have faculty demonstrate putting theory into practice.[6]

Two recurring themes come out of the previous experience teaching RDM: (1) expansion of the course to cover theory and practice and (2) the need for a combined faculty/librarian teaching approach. As noted above, a librarian and a research-active engineering faculty member co-taught the initial offering of this course. We designed the course this way to bring both skill sets into the classroom. The librarian was more familiar with campus resources for RDM (archives and DMPtool, for example) along with the Purdue DCP, while the faculty member had more experience with how data is currently collected and shared along with more specifics such as discipline-specific archives and publication requirements. We elected to use this approach based on comments from Whitmire (a librarian) that the students expressed a desire for more real-world content,[13] Wright and Andrews (librarian and faculty) that a subject-specific focus with real-life examples helped the student to better understand,[14] and Thielan, et al. who recommended a co-taught approach.

The core of a chemical engineering graduate degree is advanced chemical engineering topics such as thermodynamics, kinetics, and transport. The graduate degree is then completed with research and additional elective coursework as needed for the students to conduct research and prepare for careers with a research focus. These additional elective classes have included courses to mentor students through the research process,[24-28] critical review of the literature,[29, 30] oral communication,[31] teaching,[32] and safety.[33] Research data management fits well into this group for two reasons. First, following best practices in RDM will make the student a more efficient researcher as well as contribute to the success of their research group. Second, use of proper RDM will similarly lead to efficient research in their future jobs. In addition, by virtue of the graduate degree, these students will become leaders in their future companies or institutions. They will be expected to effectively manage the research data from their teams as well as contribute to larger company/institution projects to develop and maintain RDM policies and platforms.

This initial offering of this course was taught under the chemical engineering nomenclature (CHE), however, it is not a chemical engineering specific course. While many chemical engineering examples were used in the course because of the enrollment, all of the topics apply broadly to all research graduate degrees. Similar RDM courses have been taught as library seminars,[12] and under the nomenclature of natural resources (NTRS)[14] or graduate (GRAD)[13] programs. Muilenberg, et al. suggest that offering the course through a department with broad reach such as the graduate school would increase enrollment.[12] For permanent incorporation of this course into the course catalog, we will be cross listing it as chemical engineering (CHE), engineering science

(ES) to appeal to the engineering college students, library science (LIB) due to the use of resources through the library, and as a graduate course (GRAD) to appeal broadly across campus. A broad array of student disciplines in the course can also encourage learning as the students are exposed to RDM approaches and standards from other disciplines. Student feedback from literature offerings of similar courses have suggested that different versions of the course may be preferred for humanities/social science students and another for natural and applied sciences[13] or following up a more general RDM course with a more discipline specific course as required.[14]

While the initial full three-credit version of this course offers the ability to combine theory and practice, the full version offered here may not meet needs in other cases. Where discipline-specific knowledge is desired, a smaller (1 credit) course of high-level issues followed by 2- or 3-credit-specific courses may be beneficial. Multiple versions (1- to 3-credit) of this course offered over a 1-month winter term between semesters might also appeal to students while delivering the material in a more modular approach. Finally, specific topics of RDM important for broader audiences may be incorporated into a department's graduate seminar series[33] or introductory graduate mentoring class.[24-26, 34] These specific topics might include, for example, a workshop series using R—a language and environment for statistical computing and graphics that includes data handling and storage.[35] In order for the course to appeal to a broad array of students, topics with focused application tools were not included.

Several aspects of the class worked well and should be emphasized. The final project to develop a DCP resulted in many positives. The class discussions to develop a DCP from the Purdue model helped all of the students consider the DCP from the research perspective (what data is obtained and how it is stored, used, shared, and preserved). Individual development of a DCP would likely have been too much work, thus, teamwork allowed for the learning associated with the DCP while balancing the work required. Selecting willing faculty to participate in the DCP (as it required a time commitment) likely resulted in more thorough interviews than if the students were simply assigned to interview their advisor. The discipline diversity in the interviewed faculty also helped to expose the students to a wider variety of RDM standards and practices. Finally, although we did not suggest it, the interviewed faculty tended to focus the interview and DCP on one specific research project in their lab. We would suggest that this be part of the guidance in the future as many labs have multiple research projects that may be related but are still somewhat independent. Feedback from interviewed faculty indicated that all agreed that the RDM questions from the DCP were appropriate and that no aspects of RDM were missed.

The use and variety of guest speakers were also well received by the students. The guest speakers provided in-depth knowledge of topics where the instructors did not have deep knowledge. The Michigan course also extensively used guest speakers, including one to focus on the university's online resource for data curation.[6] Requiring a student reflection for each guest speaker and the final presentation was also a positive. This assignment encouraged the student to learn something from each presentation and apply it to their RDM on a continuing basis. Unfortunately, we did not start this assignment until after several guest speakers had already presented. Future offerings will require student reflections for all guest speakers. We also did not give focused-enough advice to two guest speakers on what we'd like them to cover. Having taught the full class for the first time will allow this to be corrected in the future.

For future offerings, we will change the structure of the DMP assignment and how DMPtool is used. Within DMPtool, the user is allowed to select the funding agency to which they are applying and the software then helps the user build a DMP using that agency's requirements. The students then generated DMPs and they were graded by the instructors. However, the learning can be broadened if the students shared their DMPs with other students for comments and suggestions. Therefore, we will include an assignment to review the DMPs from the class and then have the student revise/update their DMP based on this additional exposure. RDM ethics were not covered but will be added in the future. This will include discussions on, for example, data manipulation and data sharing ethics. Post-course assessment also suggested the incorporation of RDM ethics. Finally, we would like to implement the Thielan, et al. idea to incorporate additional readings on current topics in RDM.[6]

CONCLUSIONS

A course in RDM covering both high-level topics and practical experience in laboratory data management has been developed and delivered. Student knowledge growth was demonstrated in eight areas of RDM ranging from data management planning, to documentation and metadata, to archiving and preservation. Students were prepared to develop DMPs as required by funding agencies. The students also obtained experience developing and using a DCP for their research. Finally, the course presented multiple opportunities for student reflection on their current RDM as related to lecture topics, guest lectures, and project results.

ACKNOWLEDGMENTS

We gratefully acknowledge the guest speakers and faculty members who contributed to the DCP projects.

REFERENCES

1. National Academy of Sciences, Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age, *Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age* 2009, Washington, D.C.: The National Academy Press.
2. Carlson, J., M. Frosmire, C.C. Miller, and M. Sapp Nelson, "Determin-

ing Data Information Literacy Needs: A Study of Students and Research Faculty," *Portal: Libraries and the Academy*, **11**(2), 629 (2011)

3. National Science Foundation, *Proposal and Award Policies and Procedures Guide Part I - Grant Proposal Guide*, National Science Foundation, Editor 2013
4. National Institutes of Health, NIH Data Sharing Policy. April 17, 2007 [cited March 16, 2017]; Available from: <https://grants.nih.gov/grants/policy/data_sharing/>
5. United States Geological Survey, U.S.G. USGS Data Management: Data Management Plans. March 7, 2017 [cited 2017 March 16, 2017]; Available from: <https://www2.usgs.gov/datamanagement/plan/dmplans.php>
6. Thielen, J., S.M. Samuel, J. Carlson, and M. Moldwin, "Developing and Teaching a Two-Credit Data Management Course for Graduate Students in Climate and Space Science," *Issues in Science and Technology Librarianship,* **86**(Spring) (2017)
7. McLure, M., A.V. Level, C.L. Crabston, B. Oehlerts, and M. Culbertson, "Data Curation: A Study of Researcher Practices and Needs," *Portal: Libraries and the Academy*, **14**(2),139 (2014)
8. Carlson, J., L. Johnston, B. Westra, and M. Nichols, "Developing an Approach for Data Management Education: A Report from the Data Information Literacy Project," *Int. J. Digital Curation*, **8**(1), 204 (2013)
9. National Science Foundation, Cyberinfrastructure Council, *Cyberinfrastructure Vision for 21st Century Discovery*, 2007, National Science Foundation
10. Qin, J., and J. D'Ignazio, The Science Data Literacy Project: Educator Resources. [cited March 16, 2017]; Available from: <http://sdl.syr.edu/?page_id=15>
11. Borgman, C.L., Syllabus for Data Management and Practice, Part I, Winter 2015, University of California, Los Angeles: Selected Works, bepress
12. Muilenburg, J., M. Lebow, and J. Rich, "Lessons Learned from a Research Data Management Pilot Course at an Academic Library," *J. eScience Librarianship*, **3**(1), e1058 (2014)
13. Whitmire, A.L., "Implementing a Graduate-Level Research Data Management Course: Approach, Outcomes, and Lessons Learned," *J. Librarianship and Scholarly Communication*, **3**(2), eP1246 (2015)
14. Wright, S.J., and C. Andrews, *Natural Resources / Graduate Students* / Wright & Andrews / Cornell University / 2013. Data Information Literacy Case Study Directory, 2015. 2(1): p. Article 1
15. Briney, K., *Data Management for Researchers: Organize, maintain and share your data for research success*. *Research Skills Series*, 2015, Exeter: Pelagic Publishing, UK.
16. University of California Curation Center, Data Management Plan Tool. 2014; Available from: <https://dmptool.org/>
17. Carlson, J., *The Data Curation Profiles Toolkit: User Guide*, in Data Curation Profiles Toolkit 2010 <http://docs.lib.purdue.edu/dcptoolkit/>
18. Krier, L., and C.A. Strasser, *Data Management for Libraries: A LITA Guide*. 2014: American Library Association
19. Corti, L., V. Van den Eynden, L. Bishop, and M. Woolard, *Managing and sharing research data: a guide to good practice*, 2014: SAGE Publications.
20. Carlson, J., *The Data Curation Profiles Toolkit: Interviewer's Manual*, in Data Curation Profiles Toolkit 2010 <http://docs.lib.purdue.edu/dcptoolkit/>
21. Carlson, J., *The Data Curation Profiles Toolkit: Interview Worksheet*, in Data Curation Profiles Toolkit 2010 <http://docs.lib.purdue.edu/dcptoolkit/>
22. Carlson, J., *The Data Curation Profiles Toolkit: The Profile Template*, in Data Curation Profiles Toolkit 2010 <http://docs.lib.purdue.edu/dcptoolkit/>
23. Russo Martin, E., *New England Collaborative Data Management Curriculum*. 2017 [cited 2017 March]; Available from: <http://library.umassmed.edu/necdmc/index>
24. Ollis, D., "The Research Proposition," *Chem. Eng. Ed*., **29**(4), 222 (1995)
25. Burrows, V., and S. Beaudoin, "A Graduate Course in Research Methods," *Chem. Eng. Ed*., **35**(4), 236 (2001)
26. Holles, J.H., "A Graduate Course in Theory and Methods of Research," *Chem. Eng. Ed*., **41**(4), 226 (2007)
27. Aucoin, M.G., and M. Jolicoeur, "Is there room in the Graduate Curriculum to Learn 'How to be a graduate student?' An approach using a graduate-level biochemical engineering course," *Chem. Eng. Ed*., **43**(4), 306 (2009)
28. Ollis, D., "Catalyzing the Student-to-Researcher Transition: Research Initiation and Professional Development for New Graduate Students," *Chem. Eng. Ed*., **50**(4), 221 (2016)
29. Hill, P.J., "Teaching entering graduate students the role of journal articles in research," *Chem. Eng. Ed*., **40**(4), 246 (2006)
30. Minerick, A., "Journal club: A forum to encourage graduate and undergradaute students to critically review the literature," *Chem. Eng. Ed*., **45**(1), 73 (2011)
31. Wilkes, G.L., "The importance of oral communication skills and graduate course to help improve these skills," *Chem. Eng. Ed*., **46**(4), 251 (2012)
32. Wankat, P.C., "Pedagogical training and research in engineering education," *Chem. Eng. Ed*., **42**(4), 203 (2008)
33. Madihally, S., "Reviving graduate seminar series through non-technical presentations," *Chem. Eng. Ed*., **45**(4), 231 (2011)
34. Frugoli, J., A.M. Etgen, and M. Kuhar, "Developing and Communicating Responsible Data Management Policies to Trainees and Colleagues," *Science and Engineering Ethics*, **16**(4), 753 (2010)
35. The R Foundation, *The R Project for Statistical Computing*. [cited 2017 May 30, 2017]; Available from: <https://www.r-project.org/>
36. Hake, R.R., "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanical test data for introductory physics courses," *American J. Physics*, **66**(1), 64 (1998) ❐