2018 ASEE Annual Conference & Exposition
June 24–27, 2018
Salt Palace Convention Center
Salt Lake City, UT

Paper ID #21190

Implementing a Graduate Class in Research Data Management for Science and Engineering Students

Dr. Joseph H. Holles, University of Wyoming

Associate Professor, Department of Chemical Engineering

Mr. Larry Schmidt, University of Wyoming

Larry Schmidt is an associate librarian at the University of Wyoming and is the current Head of the Brinkerhoff Geology Library. He holds BS degree's in Chemistry and Biology, MS Degree in environmental engineering from Montana State University and received an MLS from Emporia State University in 2002. His interests lie in providing undergraduate and graduate students with information, data and science literacy skills that will allow them to succeed in a global economy.

Implementing a Graduate Class in Research Data Management for Science/Engineering Students

Introduction:

Research data management (RDM) is an integral part of engineering and science graduate student life, both during graduate school and in their future occupations. Federal agencies, including NSF[1], NIH[2], and USGS[3], are now requiring the submission of a Data Management Plan (DMP) when submitting proposals for funding. Carlson et al. further advocate for RDM by stating "… it is not simply enough to teach students about handling data, they must know, and practice, how to develop and manage their own data with an eye toward the next scientist down the road."[4] Thus, while a RDM requirement may be forced on scientists and engineers from the outside, the growth of our profession also offers a reason for education in this area. Therefore, from a top-down approach, RDM is a required part of many federal funding opportunities. From the bottom-up, RDM leads to effective and efficient research progress.

In a study of current RDM practice, Carlson et al. note that in today's university research laboratories, "graduate students are often expected to carry out most or all of the data management tasks for their own research."[5] While literature studies have shown that faculty understand the need for RDM education for their students[6], the same faculty also acknowledge that graduate students were not prepared to manage data effectively[4] but that they as faculty could not provide adequate guidance or instruction and that they would benefit from experts "helping us to do it right." Carlson's work also points out multiple faculty perceived shortcomings of RDM: self-directed student learning in the laboratory through trial and error, absence of formal policies governing data in the lab, and lack of formal training in data management.[5]

RDM education for graduate students has taken a variety of approaches. These approaches range on the intensity and commitment scale from no-credit seminars and workshops to for-credit stand-alone courses. Information science programs have used the stand-alone course approach,[7, 8] while the seminar/workshop approach is commonly offered through libraries.[9-11] Workshops and seminars offer the advantage of smaller time commitments, broad overview classes, or focused applications, but are often not for credit and consequently suffer retention issues.[9, 10] Another drawback associated with library delivered seminars is that librarians have difficulty providing strong in-class examples since they often do not have extensive experience in basic research or knowledge of discipline specific practices. For the stand-alone course, the advantage is in-depth material coverage at the expense of a larger time commitment. Carlson et al. also observe that it is difficult to attract students to courses that reside outside of their discipline.[5] For-credit courses focused specifically on RDM for graduate students have been taught by librarians and offered through the graduate school[12] or taught by a combination of librarians and faculty and offered by specific research-focused departments.[13, 14]

The goal of this manuscript is to describe a graduate course in RDM for science and engineering students. The course was designed to expose students to both high-level RDM topics and practical, hands-on experience. As part of the course, students were provided the RDM knowledge advocated for by National Academy of Sciences and NSF. This course was co-taught by a research active faculty member and a librarian in an effort to deliver broad knowledge on RDM standards and tools from the expertise of the librarian while allowing research focused examples and experience from the faculty perspective. This manuscript describes the course, course materials, lecture topics, assignments and projects and assessment tools for the course. Comparison with similar approaches and courses in the literature along with lessons learned are also provided. An earlier version of this manuscript appeared in Chemical Engineering Education as "A Graduate Class in Research Data Management".[15]

Methods:

A three credit graduate course, Research Data Management, was developed and taught for the first time during the Fall 2017 semester. The course was team taught by a reference librarian and a research active faculty member. A number of guest speakers from across campus were also incorporated to provide lectures on more specific topics. The course met three times per week for 14 weeks (a total of 42 class periods) plus a final exam meeting. A typical class schedule is in Table 1.

The specific goals of the course were: 1) Expose the students to broad concepts and best practices of RDM, 2) Bring in outside experts to demonstrate specific areas of RDM, and 3) Provide a focused application of RDM to active research projects. Topics discussed under broad concepts and best practices included: Data and Data Lifecycles, Describing Your Data, and Planning Your Research Topics. Lectures from guest experts were focused on specific applications of RDM. For example, these lectures included metadata and RDM tools available within the university. Focused application of RDM to research projects included DMPtool, and development of a data curation profile (DCP) to use for development of a DMP for an ongoing research project. Student work for the course is outlined in Table 2.

"Data Management for Researchers: Organize, maintain and share your data for research success" by Kristin Briney[16] was used as the text for the course. Individual chapters from Briney formed the foundation for 9 lectures in the course. Additional material for the course was developed from the work of Whitmire[12] and Krier and Strasser's *Data Management for Libraries: A LITA Guide*.[17]

The course was team taught by two main instructors, a reference librarian and a research active faculty member from chemical engineering. The reference librarian brought information and expertise on university resources for data management such as data repositories and assistance preparing DMPs. The research active faculty contributed expertise on laboratory data collection, management, storage, preservation, and discipline specific standards. Additional experts from across campus delivered guest lectures on metadata, developing an ORCID profile[18], a university data archive and management, RDM in the humanities, RDM for human subjects, data for re-use, a data manager for an interdisciplinary project, and the PI on the same multi-university data intensive project.

Students were surveyed both pre- and post-course to develop the initial course offering and modify future offerings of the course. Faculty who volunteered to participate in the final project were also surveyed to determine if they found the process effective and to suggest potential modifications and improvements. Pre- and post-course assessment was performed to gauge the students' knowledge about eight specific areas of RDM and their current laboratory RDM practices. All 10 students completed both pre- and post-course assessment. To assess specific areas of RDM, the students self-rated their knowledge levels in the eight areas using a Likert type scale from poor (1) to excellent (5). Each of the eight topic areas was represented by one question. To assess current laboratory practices, responses were yes/no/don't know. The average student response for each question was determined both pre- and post-course. The preand post-course assessment variances for each question were found to be equal using an f-test at $\alpha=0.20$. Since the variances were determined to be equal, hypothesis testing using a t-test at $\alpha=0.05$ demonstrated that the post-course mean (average) exceeded the pre-course mean for each of the eight questions.

Average normalized gain $\langle g \rangle$ for each assessment question was determined to quantify how large each effect was.[19] This has been used to represent a rough measure of the effectiveness of a course in promoting conceptual understanding. It can also be described as the amount students learned divided by the amount they could have learned.

Course Description:

For this course, the textbook was "Data Management for Researchers: Organize, maintain and share your data for research success" by Kristin Briney.[16] This book was selected because it is written from the perspective of helping the researcher accomplish RDM using standard techniques and best practices. The Briney text contains a mixture of practical information (e.g., improving data analysis and documentation) together with high level topics (e.g., planning for data management and data lifecycles). The course topics and individual lectures were developed from the Briney text along with similar material from a RDM course taught at Oregon State University.[12] A similar book by Corti et al.[20] may also provide useful information, however it was more focused on United Kingdom based researchers and examples.

Individual lecture schedule for the initial offering of the course is shown in Table 1. The individual lecture topics were divided into three areas: 1) broad concepts and best practices of research data management, 2) outside experts to demonstrate specific areas of RDM, and 3) a focused application of RDM to active research projects. The Briney text provided the material for the lectures on broad concepts (with the chapter noted in parenthesis in Table 1). Approximately 10 lectures were based on the text. Additional lectures which completed the broad concepts part of the course included RDM sharing mandates, DMPtool[21], and reference managers. Guest experts provided eight lectures on specific applications of RDM and included: RDM tools available through the university, metadata, and the PI for a multi-university data intensive project. These lectures are noted with "Guest" in the topic title in Table 1.

Table 1: Typical Class Schedule

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

Table 1: Typical class schedule for Research Data Management course including week, class, and topic. Numbers in parenthesis indicate chapter of the Briney test used as basis for the lecture topic.

The remainder of the course focused on RDM application to active research projects. Early in the course, DMPtool[21] was used individually by the students to develop a DMP for their research project. DMPtool is an open source online tool available through the university which allows researchers to create a short, 2-page “funding DMP” that is required by federal agencies as part of the grant application process. DMPtool is agency specific in the information entered which allows the students to gain field specific knowledge.

The second focused application of RDM used a Data Curation Profile (DCP) to create data management plans for research projects. A DCP is a tool designed to cover all areas of RDM and to allow data management specialists to work with researchers to develop specific data management plans. The class used the Data Curation Profiles Toolkit from Purdue[22-25] to develop a DCP for their subsequent use. The class developed DCP was then used by the students as part of the Final Project (Table 2) to interview faculty members to obtain the information for the subsequent "project DMP."

Course assignments and objectives are shown in Table 2. The student work can be divided into four categories: 1) Individual assignments reinforcing topics from the class, 2) Student's reflection on guest speakers focused on applications to their RDM, 3) Final project developing a DMP for an ongoing research project, and 4) Student's reflection on their RDM practices.

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 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 2: Assignments and objectives for the Research Data Management course.

Eight guest speakers delivered lectures throughout the course on specialized topics in RDM. The guest speakers were invited to provide the class with applications and specific examples of RDM that were outside the instructors areas. Guest speakers were identified based on personal knowledge from either of the two instructors. Following each guest speaker, a student reflection was completed by each student. The objective of these reflections was for the students to individually consider the talk itself but also to consider applications to the student's current RDM practices. Three of the guest speakers discussed the same project from three different perspectives: 1) PI for the large, multi-university, data intensive project, 2) use and management of the university's data curation repository by an expert, and 3) day-to-day management by an IT expert. Two topics from Briney were also covered or reinforced by the guest speakers (data reuse and managing sensitive data).

For the final project, the class used the DCP to develop a project DMP for four research active faculty members across the campus. The class was divided into two or three member groups and each group worked with a different faculty member on campus to investigate their RDM protocols and to develop a DMP for that researcher. These volunteer researchers were selected by the instructors from a variety of disciplines (civil engineering, chemical engineering, physics, and chemistry). The student team then prepared for the interview and interviewed the researcher. Following the interview, the student team developed a DMP for the researcher based on the DCP which included suggested best practices. Finally, each group then prepared and delivered a presentation on the interview and DMP that they developed. In this way, the observations and experiences across the broad spectrum of researchers was shared with the entire class. Each student also completed a self-reflection on these presentations to again consider how each project could be applied to their research.

The last assignment was an individual reflection of the current RDM for each student's thesis or dissertation project. The goals was for the student to step back and take a high-level review of their current RDM practices in light of the course topics, speakers and DMP project and to consider and suggest possible changes and revisions.

Results:

Student background knowledge in eight RDM topics was assessed prior to the course and results are shown in Figure 1. As part of the pre-course assessment, students were also asked about their research funding, research topics, laboratory RDM practices, RDM needs, and knowledge they would like to obtain from the course. This was in order to possibly modify the course in advance to address student needs. As a result of the pre-course survey, two additional topics were added to the course. At the end of the course, the student's RDM knowledge was again assessed for the same eight topics. These results are also in Figure 1.

Figure 1: Assessment Results

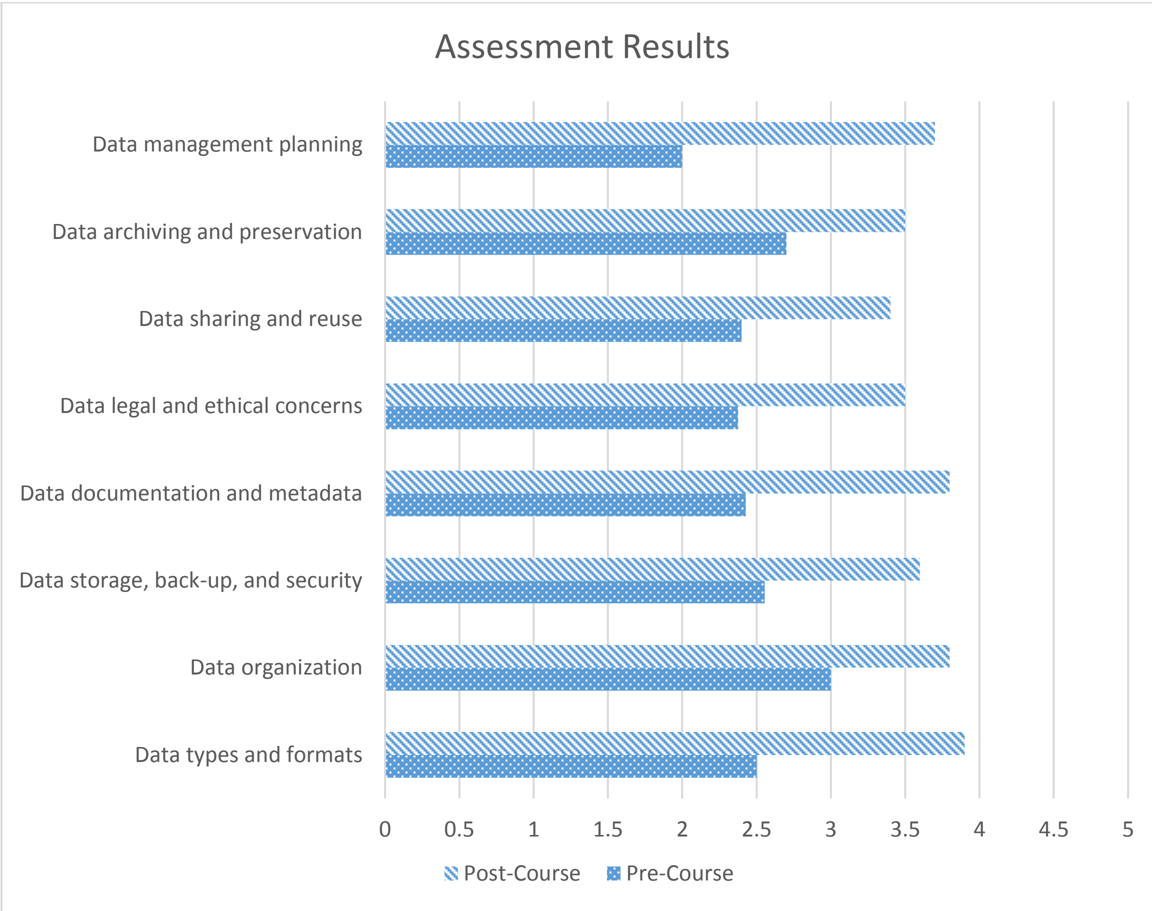

Figure 1: Pre- and post-course assessment results for eight specific areas of Research Data Management. Results are self-reported using a Likert-type scale from poor (1) to excellent (5). Pre- and post-course assessment variances for each questions were equal using an f-test at α=0.20. Subsequent hypothesis testing using an f-test at α=0.05 demonstrated that the post-course mean exceeded the pre-course mean for each question.

Assessment results demonstrated that student-rated knowledge increased as a result of the course for each of the eight RDM topics. The average increase in the score from the pre- to the post-assessment for all eight topics was 1.15 points. Data management and planning (1.7) and data types and formats (1.4) had the largest increase while data organization (0.8) and data archiving and preservation (0.8) had the smallest increase. Comparing pre- and post-course means using hypothesis testing demonstrated that the means were different for all eight topics indicating student reported knowledge growth in all eight topics.

Additional assessment demonstrated that students improved their knowledge of laboratory RDM protocols and their ability to write a DMP. For the three questions focused on these topics, 12 of 30 pre-course assessment responses were “don’t know” compared to only four “don’t know” from post-course assessment. Finally, the students who could write a DMP for their research increased from two to nine while those with protocols for managing their research data increased from four to eight.

When examined quantitatively using average normalized gain <g>[19], the average <g> across all 8 topics was 0.45 with individual <g> values ranging from 0.35 for archiving and presentation to 0.57 for management and planning. Average normalized gain is used to quantitatively represent a measure of the effectiveness of a course in promoting conceptual understanding. According to Hake, this would be a "medium-g" course for values 0.7 > "<g>" >0.3.[19] Average normalized gain for all eight topic areas are reported in Table 3.

Table 3: Average Normalized Gain

Topic	Avg. Normalized Gain
Management and Planning	0.57
Archiving and Preservation	0.35
Sharing and Reuse	0.38
Legal and Ethical Concerns	0.43
Documentation and Metadata	0.53
Storage, Backup, and Security	0.43
Organization	0.40
Types and Formats	0.56

Table 3: Average normalized gain for each of the eight assessment topic areas.

All materials associated with the course are available at: https://doi.org/10.15786/M28D50. This includes: lecture notes, online resources, assignments, project, assessment tool, and some of the guest speaker presentations.

Discussion:

A variety of different approaches to RDM education have been tried over the last few years. This variety of approaches indicates that there is an effort to broaden and improve the education on this topic. The education approaches have ranged from low intensity to high intensity. For example, at the University of Washington, Muilenberg et al. developed a low intensity seven-module workshop taught by librarians that met weekly for one hour.[9] These weekly topics were developed from the New England Collaborative Data Management Curriculum (NECDMC).[26] A similar approach was used at UMass Amherst[11] and the University of Minnesota.[10] In contrast, in a high intensity approach, the Information Studies (IS) department at UCLA delivers a four quarter-credit 11-week class on Data Management and Practice taught by a department faculty member.[7] While focused on the department needs, the IS course is available to students outside of the IS program. As expected, each approach offers positive results and specific drawbacks. The library workshop approach suffered from low retention while the full-term IS course is more broadly focused on data management in general without a specific RDM focus. However, both courses are "outside" the department of the students seeking the education.

The "inside" the department and more discipline specific approach has also been used by offering a RDM course through specific departments. For example, the Natural Resources program at Cornell offered "Managing Data to Facilitate Your Research"[13] and the Climate and Space Sciences and Engineering Department at the University of Michigan offered "Data Management."[14] Both courses were offered for credit. The Cornell course met for six sessions for 1 credit while the Michigan course was 14 weekly offerings for 2 credits. In contrast, the Cornell course was co-taught by a data librarian and a faculty member in the department, while the Michigan course was taught by a team of three librarians. However, in both cases, offering the course in conjunction with and through the department provided the opportunity for subject specific examples. Cornell concluded that it was desirable to expand the class to introduce more exercises and hands-on learning.[13] Michigan concluded that future offerings of the course would benefit from having a department professor co-teaching the course along with the librarian. [14]

"Research Data Management" offered in partnership between the library and the graduate school at Oregon State University is the course most parallel to this one.[12] The Oregon State course was taught by a librarian (with a Ph.D. in a scientific field). This course met twice each week for 50 minutes for 11 weeks (one quarter) and was a 2-credit course. The two major projects for the course were a DMP for a final project and a DCP as a midterm assignment. For our course, we reversed the order of these two assignments. The author also recommends that a better approach may be to have the students create sections of the DMP as weekly assignments throughout the course. Additionally, since the sole instructor was a librarian, the student feedback indicated a desire for more "real world" context and practical applications instead of theoretical discussion.[12]

The Thielen et al. course also had numerous similarities with the course developed in this work. For example, they also used the Briney text, used frequent student reflections, employed multiple guest speakers (five), and focused on a class structure with the students helped to educate each other.[14] Consistent with this paper, Thielen et al. suggested that instead of librarians, a co-taught model with a departmental faculty member and a librarian would improve the ability for the course to put theory into practice. The DMP assignment in the Thielen work resulted in a large range of DMPs: from 3 to 26 pages. The authors note that they will revise the assignment in the future to be more explicit. Our experience from this course indicates that this might be a result of no distinction between the short, funding proposal DMP and the longer, laboratory DMP. A distinction between these two different DMPs and suggestions on how to teach each is included in a separate presentation.[27]

From the variety of previous approaches to teaching RDM at the graduate level, two themes emerge: 1) a broad enough course to cover both theory and practice, and 2) a demonstrated need for a combined librarian/faculty teaching approach. As noted above, in this work the course was co-taught between a research active engineering faculty member and a librarian. The course was designed in this way in order to bring both sets of skills into the classroom. The faculty member provided experience and examples of how data is currently collected and shared along with discipline specific examples of publication requirements and available archives. To complement this, the librarian led the topics on available campus resources for RDM such as archives, DMPtool along with experience using the Purdue DCP. Both instructors also had different networks of colleagues across campus from which we were able to identify potential guest speakers. Comments and lessons learned from Whitmire[12], Wright and Andrews[13], and Thielen et al.[14] all suggested this combine teaching approach to maximize the results of RDM classes.

Although the initial offering of this course was offered under the Chemical Engineering (CHE) title, it is not a chemical engineering specific course. All of the topics apply generally to most research graduate degrees in engineering and the hard sciences. However, due to the student enrollment, many of the examples were chemical engineering focused in this initial offering. The library seminars are commonly offered as seminars without formal course numbers[9-11], the department courses under the department's name[13, 14], or under a broader graduate (GRAD) nomenclature.[12] Similar to the Whitmire GRAD nomenclature, Muilenberg et al. recommend offering the course under a name with broad reach such as the graduate school in order to increase enrollment.[9] In light of this suggestion, we will be incorporating this course as an engineering science (ES), library science (LIB), and graduate course (GRAD) in order to reach broad audiences. This broad audience can then increase the classroom learning experience as the student can become exposed to standards and practices in other disciplines.

Multiple approaches in the class worked well and should be emphasized. Requiring a student reflection for each guest speaker and for the final student presentations was a strong learning tool. Through this assignment, each student was encouraged to focus on the messages from the guest speakers and the student developed DMPs and consider how to apply this to their individual laboratory RDM. The approach of using multiple guest speakers worked well and was well-received by the students. As anticipated, the guest speakers were able to complement the main course instructors by providing in-depth knowledge on RDM topics where the instructors did not have the depth of experience or knowledge. Thielen et al. also made similar use of guest speakers for topics similar to our use of guest speakers and reported positive results.[14] In contrast, we did not provide enough focused guidance to two of the guest speakers. As a result, the presentation was not as complete and thorough as envisioned. Experience will allow us to provide future speakers with more direct guidance.

The final project where the student groups used the DCP developed in class to work with other campus researchers and develop a DMP for that faculty member's research project had many positive outcomes. The initial work for this project was the class time spent revising the Purdue DCP. This revision forced the students to delve deeply into the RDM process and consider RDM not only from the researcher's perspective, but also the perspective of archivists and researchers who re-use data and librarians attempting to rapidly learn the research project and assist with the project RDM. This class time spent on the revision was effective as having the students individually revising the DCP would likely required too much work. Focused selection of campus faculty to participate in the final project was also successful. Working with the students to explain their current research and RDM issues was a time intensive commitment by these faculty. The alternative of requiring each student to complete the project with their own advisor was also considered. While this approach may also work, the selected faculty approach was chosen due to their demonstrated interest in participation. Broad student learning of RDM issues was also increased by including the diverse, campus-wide selection of participating faculty. For the future, we will also suggest to these faculty that they focus their participation on one research project in their lab instead of considering all of their laboratory research. While not suggested in advance, the faculty seemed to innately sense this refinement. Finally, consider selecting young faculty for participation in this project and using this project as an opportunity to help them develop their research group by writing a DMP. Faculty feedback indicates that the revised DCP prepared by the class covered all appropriate aspects of RDM and that participation in the project resulted in useful and effective observations and suggestions from the students through the prepared DMP.

RDM ethics were not covered in the initial offering of the course but will be incorporated into future offerings. RDM ethics topics will include, data manipulation and data sharing ethics.

While considered during course development, the topic was not included due to oversight. Post-course assessment also indicated student desire to incorporate RDM ethics. In addition, we will acknowledge Thielen et al. and incorporate appropriate readings on current topics in RDM.[14] The major anticipated change to the course will focus on the use of DMPtool. For the initial offering, we had the students use DMPtool and their current funding source to generate an appropriate DMP for their research consistent with the requirements of their funding agency. The instructors then graded these DMPs. Since the students are funded from multiple sources, the learning could be increased if the students shared their DMPs with the class for additional comments and suggestions. Therefore, we will incorporate an additional assignment to have the students review their classmates DMPs followed by revisions/updates based on the class comments.

The goal of the full semester, three-credit course developed here was to combine theory and practice for the students. This subject depth and time commitment may not be optimal for all cases. For example, focused RDM subtopics may be incorporated into the departments seminar series[28] or introductory graduate mentoring class[29-32]. A smaller, 1 credit course with high level content may be appropriate at the department level with follow-on courses coving more specific and specialized RDM topics. Offering this course in a focused, modular manner over a 1-month winter term between semesters might also appeal to students. These modules could be delivered either in the traditional instructor/student classroom or online.

Conclusion:

A research data management course for graduate students in science and engineering has been developed and delivered. This course covers both high-level topics and practical experience in research laboratory data management. As a result of the course, students were prepared to develop proposal DMPs using DMPtool as required by their funding agency. The students also obtained practical, hands-on knowledge developing a laboratory DMP for an ongoing research project. The team instructional approach (two main instructors with a series of guest speakers) provided instruction on broad and focused topics. Student reflection throughout the course provided multiple opportunities for revision and improvement of their laboratory RDM practices. Finally, student knowledge growth was demonstrated in eight areas of RDM.

Acknowledgements:

We gratefully acknowledge the faculty members and guest speakers who contributed to this course.

References:

1. National Science Foundation, *Proposal and Award Policies and Procedures Guide Part I - Grant Proposal Guide*, National Science Foundation, Editor 2013.
2. National Institutes of Health, *NIH Data Sharing Policy*. 2017 April 17, 2007 [cited 2017 March 16, 2017]; Available from: https://grants.nih.gov/grants/policy/data_sharing/.
3. United States Geoligical Survey, *USGS Data Management: Data Management Plans*. 2017 March 7, 2017 [cited 2017 March 16, 2017]; Available from: https://www2.usgs.gov/datamanagement/plan/dmplans.php.

4. Carlson, J., M. Frosmire, C.C. Miller, and M. Sapp Nelson, *Determining Data Information Literacy Needs: A Study of Students and Research Faculty.* Portal: Libraries and the Academy, 2011. **11**(2): p. 629-657.
5. Carlson, J., L. Johnston, B. Westra, and M. Nichols, *Developing an Approach for Data Management Education: A Report from the Data Information Literacy Project.* The International Journal of Digital Curation, 2013. **8**(1): p. 204-207.
6. 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, 2014. **14**(2): p. 139-164.
7. Borgman, C.L., *Syllabus for Data Management and Practice, Part I, Winter 2015*, 2015, University of California, Los Angeles: Selected Works, bepress.
8. Qin, J. and J. D'Ignazio. *The Science Data Literacy Project: Educator Resources*. [cited 2017 March 16, 2017]; Available from: http://sdl.syr.edu/?page_id=15.
9. Muilenburg, J., M. Lebow, and J. Rich, *Lessons Learned from a Research Data Managment Pilot Course at an Academic Library.* Journal of eScience Librarianship, 2014. **3**(1): p. e1058.
10. Johnston, L. and J. Jeffryes, *Civil Engineering/ Graduate Students/ Johnston & Jeffreys/ University of Minnesota/ 2012.* Data Information Literacy Case Study Directory, 2015. **3**(1): p. Article 1.
11. Adamick, J., R.C. Reznik-Zellen, and M. Sheridan, *Data Management Training for Graduate Students at a Large Research University.* Journal of eScience Librarianship, 2013. **1**(3): p. e1022.
12. Whitmire, A.L., *Implementing a Graduate-Level Research Data Management Course: Approach, Outcomes, and Lessons Learned.* Journal of Librarianship and Scholarly Communication, 2015. **3**(2): p. eP1246.
13. 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.
14. 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, 2017. **86**(Spring).
15. Schmidt, L.O. and J.H. Holles, *A Graduate Class in Research Data Management.* Chemical Engineering Education, 2018. **52**(1): p. 52-59.
16. Briney, K., *Data Managment for Researchers: Organize, maintain and share your data for research success*. Research Skills Series. 2015, Exeter: Pelagic Publishing, UK.
17. Krier, L. and C.A. Strasser, *Data Management for Libraries: A LITA Guide*. 2014: American Library Association.
18. ORCID Inc., *ORCID: Connecting Research and Researchers*. 2018; Available from: https://orcid.org/.
19. Hake, R.R., *Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanical test data for introductory physics courses.* American Journal of Physics, 1998. **66**(1): p. 64-74.
20. Corti, L., V. Van den Eynden, L. Bishop, and M. Woolard, *Managing and sharing research data : a guide to good practice*. 2014: SAGE Publications.
21. University of California Curation Center, *Data Management Plan Tool*. 2014; Available from: https://dmptool.org/.

22. Carlson, J., *The Data Curation Profiles Toolkit: User Guide*, in *Data Curation Profiles Toolkit*, 2010, available from: http://dx.doi.org/10.5703/1288284315650.
23. Carlson, J., *The Data Curation Profiles Toolkit: Interviewer's Manual*, in *Data Curation Profiles Toolkit*, 2010, available from: http://dx.doi.org/10.5703/1288284315651.
24. Carlson, J., *The Data Curation Profiles Toolkit: Interview Worksheet*, in *Data Curation Profiles Toolkit*, 2010, available from: http://dx.doi.org/10.5703/1288284315652.
25. Carlson, J., *The Data Curation Profiles Toolkit: The Profile Template*, in *Data Curation Profiles Toolkit*, 2010, available from: http://dx.doi.org/10.5703/1288284315653.
26. Russo Martin, E. *New England Collaborative Data Management Curriculum*. 2017 [cited 2017 March]; Available from: http://library.umassmed.edu/necdmc/index.
27. Holles, J.H. and L.O. Schmidt. *Graduate Research Data Management Course Content: Teaching the Data Management Plan (DMP)*. in *2018 ASEE Annual Conference & Exposition*. 2018. Salt Lake City, UT.
28. Madihally, S., *Reviving graduate seminar series through non-technical presentations.* Chemical Engineering Education, 2011. **45**(4): p. 231.
29. Burrows, V. and S. Beaudoin, *A Graduate Course in Research Methods.* Chemical Engineering Education, 2001. **35**(4): p. 236.
30. Ollis, D., *The Research Proposition.* Chemical Engineering Education, 1995. **29**(4): p. 222.
31. Ollis, D., *Catalyzing the Student-to-Researcher Transition: Research Initiation and Professional Development for New Graduate Students.* Chemical Engineering Education, 2016. **50**(4): p. 221-229.
32. Holles, J.H., *A Graduate Course in Theory and Methods of Research.* Chemical Engineering Education, 2007. **41**(4): p. 226-232.