120th ASEE Annual Conference & Exposition
FRANKLY, WE DO GIVE A D*MN
June 23-26, 2013

Paper ID #6956

An E-Learning Approach to Data Information Literacy Education

Mr. Jon N. Jeffryes, University of Minnesota, Twin Cities

Jon Jeffryes is an engineering librarian supporting researchers and students in the areas of Biomedical, Civil, and Mechanical Engineering. He is currently the chair of the scholarly communication committee for the Engineering Libraries Division.

Lisa Johnston, University of Minnesota

Lisa Johnston is the Research Services Librarian and co-director of the University Digital Conservancy, the University of Minnesota's institutional repository. She coordinates the Libraries' efforts around digital scholarship and research data management, access, and archiving. She has a master's of Library Science and bachelor's of Science in Astrophysics, both from Indiana University.

An E-Learning Approach to Data Information Literacy Education

Abstract
This paper presents the University of Minnesota Libraries' contributions to the ongoing Data Information Literacy project, an IMLS-funded project to educate the next generation of e-scientists through developing a library-run curriculum. Our project team at the University of Minnesota interviewed graduate students in Civil Engineering to determine their needs in data management and curation. We found many skill areas that were missing or needed support in the graduate program and proposed a list of learning outcomes that might be addressed through library training efforts. This paper will focus on our resulting approach to providing Data Information Literacy (DIL) instruction utilizing Google tools (Sites, Drive, YouTube) to present a self-paced, interactive online course. The paper also provides an examination of our assessment methodology and discusses our findings after a pilot launch with Civil Engineering graduate students in Fall 2012.

Introduction
Since the announcement from the National Science Foundation that all grant funding proposals submitted after January 2011 must include a data management plan (DMP), academic libraries have experimented with ways of providing support and education for researchers at their institutions. At the University of Minnesota, we saw strong demand for guidance on writing and complying with this requirement. This high level of interest required a cross-disciplinary approach to teaching data management skills[1]. Although successful, this instructional approach did not facilitate in-depth, domain-specific skill building. Starting in October 2011, the University of Minnesota, along with partners at Purdue University, the University of Oregon and Cornell University collaborated on the Data Information Literacy (DIL)[2] project to learn more about the domain-specific needs of graduate students in the area of data management education. Funded by a grant from the Institute of Museum and Library Services, each research team conducted in-depth interviews with multiple graduate students and their research advisor from five different academic science disciplines. The data management skills and needs identified across the five disciplines in the DIL project illustrated a general lack of data management training with a particular gap in skills and knowledge around long-term preservation and access of research data.[3] Our team looked at the data management needs of a research group in the structural engineering area of the civil engineering department. The results of our interviews with the civil engineers found that graduate students had the most needs in the mechanics of managing and transferring data from one student to the next and that they were particularly lacking in skills such as data documentation, access and ownership considerations, and digital preservation techniques.[4] Based on our findings, in Fall 2012 the University of Minnesota Libraries launched an instructional response to address the data management skills absent from the curriculum. This paper will outline the e-learning approach we took in planning and delivering our "Data Management Course," a seven-module online course[5] we created using

Google Sites, ScreenFlow[6] and YouTube[7]. Finally, we will discuss our plans for assessment and implementation.

Literature Review

The benefits of e-learning can be found enumerated in the literature reviews and discussion of many more specific studies.[8 9] The United States Department of Education was more tempered in their meta-analysis of the literature finding "Students in online conditions performed modestly better, on average, than those learning the same material through traditional face-to-face instruction."[10] Gikandi, Morrow, and Davis's review of formative assessment in online learning, citing the influence of Oosterhof, Conrad, & Ely, specifically posits online learning benefits students by providing "many additional opportunities to dynamically interact with and assess learners."[9] Gruca nicely outlines the benefits of libraries' adopting e-learning platforms to deliver instruction.[11] Most resonant with our experience was her assertion that "E-courses are equally accessible for full-time and remote students and may be a step towards inclusion for disabled students."[11] We wanted our instruction to be as accessible as possible to graduate students that we anticipated would be carrying a full course load as well as a time-intensive research schedule. Although Gruca never explicitly uses the phrase, many of the benefits of e-learning that she lists support the scalability of instruction inherent in an e-learning platform (Gruca states that e-learning, "saves teachers' and students' time" and "[o]nce published, an e-course may be improved and used many times.").[11] The ability to scale out our work would be integral to ensuring future expansion at the University of Minnesota, where librarians support tens of thousands of students.

Building the Course

Conceptualization and creation of the course took place over the summer of 2012 by the authors and with the assistance of a library science graduate student, Kevin Cunningham. After synthesizing the findings from our graduate student interviews[4] we decided on several learning outcomes that would meet the most pressing student needs (Table 1). The learning outcomes guided our course content creation. The DIL project emphasized outcomes that were measurable and could be assessed.

Once the learning outcomes were in place, we moved into the course design phase of our project. We met face-to-face with our faculty partner to vet our intended learning outcomes and strategize how to connect students to our course content. Knowing that the graduate-level curriculum was already quite full, we knew that our approach would need to be a voluntary, extra-curricular program for students. An online, e-learning format was clearly a good fit. In addition, modularized video lessons would be easy to download and watch on any device - a functionality that matched the busy graduate student lifestyle.

Along with presenting practical skills in a virtual environment, the course needed a real-world application in which the students might demonstrate or test their newly acquired skills. Building on our earlier success offering data management training to researchers, we chose to use a data management plan (DMP) template as the framing device for course content delivery and evaluation. Writing a DMP, now required in all National Science Foundation grant applications, has become a skill needed by many academic researchers. As a result, each of the course modules map to a corresponding section of a data management plan template, where the student will apply what they have learned in the course. The resulting seven course modules became:

1. Introduction to Data Management
2. Data To be Managed
3. Organization and Documentation Methods
4. Data Access and Ownership
5. Data Sharing and Reuse
6. Data Preservation Techniques
7. Completing your DMP

Table 1: Descriptions and learning outcomes of the 7 modules in the University of Minnesota Data Management Course.

#	Course Module	Brief Description	Learning Outcomes (Students will...)
1	Introduction to Data Management	In this section we introduce the concept of data management using an example from the academic discipline.	1. Describe the benefits of data management in order to explicitly understand the benefits of participating in the course 2. Articulate what they will get out of this program in order to reinforce the learning outcomes of the curriculum.
2	Data to be Managed	This module will help you define what information will be managed, document the data collection process, and create a plan to store, backup and securely house these data.	1. Create a data inventory for their research project (data, project files, documentation, etc.) in order to not overlook any aspects of their DMP. 2. Write a backup and storage plan in order to avoid potential loss of data.
3	Organization and Documentation Methods	This module will help you plan for how to organize your data, track versions, create metadata and document data	1. Plan an organizational structure for their data using a file naming system and directory structure that is well-documented and interoperable with

		collection for reuse.	other data sets in order to decrease versioning issues and data duplication. 2. Articulate a plan to collect and share the supplementary data points of their research in order to assist other researchers in making sense of their data. 3. Fill out a metadata schema example for their data in order to model ideal metadata practices.
4	Data Access and Ownership	In this section we will illustrate some of the intellectual property and access concerns that researchers face when sharing their data with others.	1. Name the stakeholders of their data in order to understand the potential intellectual property and ownership concerns with releasing their data to a broader audience. 2. Report potential access concerns with their data in order to plan for the appropriate access controls. 3. Identify potential access controls in order to secure their data prior to release.
5	Data Sharing and Reuse	This section will describe the benefits of data sharing and potential for reuse a well as introduce students to the concept of data publishing and citation.	1. Name the audience for whom the data will be shared in order to customize the documentation and format for potential reuse. 2. Explain an approach they will use to share the data in order to instill best practices for their future data sharing. 3. Cite their data in a properly structured format in accordance with emerging standards in order to prepare them to ethically reuse data in the future.
6	Preservation Techniques	This module will introduce the preservation and curation	1. Explain the lifespan of potential use for their data in order to

		techniques used by information professionals who manage digital information for long-term access.	recognize the long-term value of their data. 2. Identify the relevant preservation-friendly file format for their research data in order to ensure long-term access to their digital information.
7	Complete Your DMP	This final module will help you create a plan on how to complete and implement your final data management plan within your lab, research group, or future project.	1. Map out an implementation plan in order to poise them for immediate use the previous modules content. 2. Identify the components of a data management plan in order to repeat the process with future research activities.

At the outset of our course design we decided that "utilize pre-existing content" would serve as a guiding principle for creating our online instructional modules. With that philosophy in mind our first step was to find content openly available for reuse, including video, images, and e-learning tools that covered any of our data management topics. Our library science practicum student interested in the topic of data management helped us conduct a review of relevant online content. He discovered many sources labeled for reuse including professional library-generated tutorials, such as MANTRA,[12] along with many informal YouTube videos and cartoons. Short YouTube videos proved well-suited to our needs. After receiving permission of the authors, we embedded several of these throughout our modules.

In addition to the resources that we found online we reused content from the data management workshops that the University of Minnesota Libraries had offered as in-person sessions for several years.[1] We customized the content from these workshops to focus on the particular needs of structural engineering graduate students.

To create the online modules we began a process of writing scripts that cover the content for each of the seven topics. The scripts were written to incorporate a logical flow of information and set up the student to respond to each learning outcome. Next, we built a slide deck in Microsoft PowerPoint and then captured a screencast of the presentation with voice-over using ScreenFlow, an Apple-based video recording software. ScreenFlow was chosen because it allowed us to capture and edit the existing YouTube videos that we embedded in the modules' PowerPoint presentations. ScreenFlow also presented a relatively easy-to-learn editing interface over alternative software such as Apple iMovie or Adobe Captivate. After creating the videos we uploaded them onto a YouTube channel to allow us to link or embed them into content

platforms. YouTube also facilitated closed captioning of the videos, making them more accessible to a variety of learners.

Finally, all of the video content was organized on a Google Site that we created as the course home page at http://z.umn.edu/datamgmt. The Google Site allowed us to create separate web pages for each module. Each of those sub-pages include:

- Textual description of the module's learning outcomes
- Instruction video (embedded from YouTube)
- Assignment
- Links to additional resources (if applicable)
- Cartoon illustration of a relevant data management concept

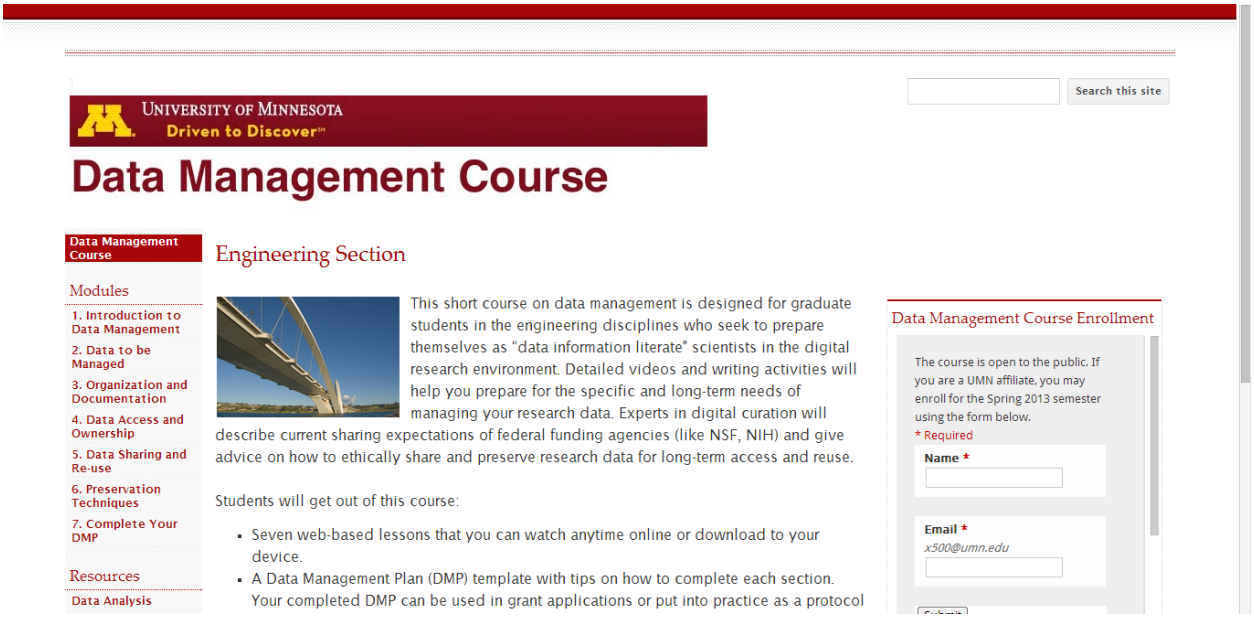

Figure 1: Screenshot of the Data Management Course: Engineering Section

The course site is open to the public and any user may enroll and progress through the course at their own pace. Google Sites was chosen over other campus e-learning tools, such as Moodle, due to their ease of creation, accessibility to users outside the University of Minnesota, and potential for one-click "cloning" if the library adapts the course for future semesters or disciplinary sections beyond civil engineering.

Beta testing of the e-course, held just before the semester began, revealed several minor errors and inconsistencies with the video modules and web site. The beta testers were primarily librarians at the University of Minnesota and members of the Data Information Literacy grant project. ScreenFlow allowed for quick video edits and insertions while the written-out scripts proved easy to edit and re-record.

Assessment Plan

To measure the success of our instructional intervention we decided to use a two-pronged assessment plan including both formative and summative assessment techniques. Throughout the course students would be asked to take the information covered in the individual modules and apply them directly to their own research project through the creation of a data management plan. As mentioned above, upon enrolling in the course the instructors created a unique copy of a data management plan template and "shared" it with the student using Google Drive. Sharing on Google Drive allows the student to customize the template, and allows the instructors access to review and comment. We used the completion of the data management template as a formative assessment throughout the course. Oosterhof, Conrad and Ely describe formative assessment as "those [assessments] that occur *during* learning" analogous to "what a mentor does continuously when working with an apprentice."[13] The different modules strategically mirrored the data management template. This design made it easy for students to take the content learned in the module and immediately utilize a real-world application for it, exercising their new knowledge on their own research data. Since the document was shared with the two instructors via Google Drive we were able to check in on the students' understanding periodically throughout the course and provide feedback via the "Comment" feature. We chose this form of assessment because it allowed us to gauge student understanding in an organic way that would seem relevant to the students, rather than quiz-style assessment that we feared would be viewed as busy work.

The second prong of our assessment plan is to measure the long-term impact of the course via an online survey that we will send out several months after a student completes the online course. For this first cohort the follow-up assessment is scheduled to go out May 2013. The assessment survey will ask the student whether completing the course had any impact on their behavior with regard to how they manage research data. We are curious to find out if, in our student population, completing the data management plan will translate into more intentional and stronger data management practice; e.g., whether the student's knowledge about data management transitions from the declarative to procedural knowledge.[13] This form of assessment also shows us whether the student successfully moves through the "hierarchical order of the different classes of objectives" found in Bloom's taxonomy, from Knowledge, through Comprehension, Application and Analysis up to Synthesis.[14] As Bransford, Brown and Cocking state in the executive summary of their committee's report on the science of learning "It is essential for a learner to develop a sense of *when* what has been learned can be used--the conditions of application."[15]

Implementing the Course

Implementing an extra-curricular course for busy graduate students was a concern. We strategized with our faculty partner and with his facilitation we were able to integrate a face-to-

face introduction to our online program in a weekly seminar required for all civil engineering graduate students on the "Structural Engineering" track. In conversation with the seminar organizer we set a launch date for early September 2012.

At the end of the first week of the fall semester the authors created a short, in-person introduction to the course for the Civil Engineering Structures Seminar (around twenty students). The introduction included demonstrating the content for module one which was largely focused on the topic of "Why Data Management Is Important". At the end of the session we asked the students to complete a "one-minute-paper" to describe how they believed a data management plan would benefit their research. In addition to the one-minute paper activity, we included a checkbox asking students if they would like to enroll in the course. Ten students responded yes (50% of the participants).

After the introduction session, progress through the course was entirely driven by the student. The instructors sent out email reminders three times throughout the semester --once, at the semester's midpoint, and again a week before the course deadline (the last Friday of classes for the fall semester) and finally on the day of the deadline of the course. In addition to these reminders, periodically throughout the semester, the instructors would visit the shared DMP templates in Google Drive to view the draft data management plans to provide private feedback to the student on their work. We did not see any progress on the templates until late in the semester.

By the end of the semester only one student had successfully completed the Data Management Plan template and a second student completed the course over the University's winter break. We heard back from several students asking for extensions or permission to push their enrollment into the next semester. The reasons for postponing included too heavy workloads and lack of an actual data set to which to apply the principles covered in the videos.

Discussion

Our pilot semester proved to be a valuable learning experience in the presentation of this e-course. We have taken some key lessons that we will apply in future iterations of the course.

Connection to Actual Data Sets

We attempted to make this course as applicable to graduate students' experiences as possible, connecting our content to the actual work students were doing in their labs. As a result we learned that students have to have relevant research data to make the course useful. Many students were interested in taking the course, but found it difficult to complete, as they were not far enough along in their graduate program to have started collecting data for their research project.

Creation of Generic Simulation

Another approach to consider for students interested in the topic of data management, but who currently have no active research projects, would be for the instructors to create a generic simulation to which students could apply the principles addressed in the video modules. Some ideas we have discussed would be having students use the data management plan for personal files (photos, drafts of papers, etc.) or have students critique an existing data management plan.

Ensuring Completion

Although we were initially pleased with the unexpectedly high percentage of students that enrolled in the course (around 50% of Structures Seminar) the completion rate was very low. In our first iteration of this course we used the promise of a certificate of completion as an incentive (on the advisement of our faculty partner) but still only two of our original ten students completed the course (although we hope those that deferred will finish by May 2013). Instead of using the graduate seminar as an entry point to students we may try to direct engagement through PIs and lab advisors who could leverage more clout in seeing that students completed the work. They would have more incentive to review the data management plan as it pertains to data created in their lab and for their use after the students graduate. Finally, the total run time of all seven-course modules is less than an hour. We may ask the students to complete the course in a much tighter time frame or provide an option for an in-person workshop-style session.

Final Thoughts

We hope to take what we've learned and apply it as we scale-out our instruction to reach other graduate student researchers across campus. As we created the course we kept an eye on scalability and built the course so it would be relatively easy to remove the discipline-specific content and replace it with that of other research areas. The research services librarian hopes to use this initial template as a starting point in collaboration with other subject librarians and in her consultations with faculty and research groups across campus.

The course has also provided a framework for other librarians either hoping to learn more about data management themselves or attempting to build instruction objects for their own institutions. Through the promotion of the Data Information Literacy website and social media presence and presentations at other conferences we have been in correspondence with other librarians interested in examining our course.

On our campus we have seen a hunger for guidance on these issues from faculty and researchers and currently the Libraries are the only campus entity offering consultations on these services. This area seems a natural extension of classic library services, including information

classification and organization as well as information literacy instruction, to help the library maintain its relevancy in the evolving information landscape.

[1] Johnston, L., Lafferty, M., & Petsan, B. (2012). Training researchers on data management: A scalable, cross-disciplinary approach. *Journal of eScience Librarianship*, *1*(2). 79-87. http://dx.doi.org/10.7191/jeslib.2012.1012
[2] *Data Information Literacy*. www.datainformationliteracy.org.
[3] Carlson, J., Johnston, L., Westra, B., and Nichols, M. (in press) "Developing an Approach for Data Management Education: A Report From the Data Information Literacy Project." 8th International Digital Curation Conference, Amsterdam, The Netherlands, January 18, 2013.
[4] Johnston, L and Jeffryes, J. (in press) "Data Management Skills Needed by Structural Engineering Students: A Case Study at the University of Minnesota" Journal of Professional Issues in Engineering Education and Practice. In press.
[5] Johnston, L. and Jeffryes, J. (2012). *Data Management Course*. https://sites.google.com/a/umn.edu/data-management-course_structures/
[6] Telestream. *ScreenFlow 4*. http://www.telestream.net/screenflow/overview.htm.
[7] Johnston, L. *Data Management Course (2012) Playlist* http://www.youtube.com/playlist?list=PLTPfFZ_tWjErY6cTsThrkqk6bnYAtE87U
[8] Safar, A. H. (2012). The students' perspectives of online training at Kuwait University. *College Student Journal*, *46*(2), 436-458.
[9] Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education*, *57*(4), 2333–2351.
[10] U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. (2010). *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies*. Washington, D.C. http://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf
[11] Gruca, A. N. (2010). E-Learning in academic libraries. *New Review of Information Networking*, *15*(1), 16–28.
[12] *MANTRA Research Data Management Training*. http://datalib.edina.ac.uk/mantra.

[13] Oosterhof, A., Conrad R., and Ely D. P. (2008). *Assessing Learners Online*. Upper Saddle River, N.J.: Pearson/Merrill Prentice Hall.
[14] Bloom, B.S. (Ed.). (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook 1: Cognitive Domain*. New York: Longman, Green and Co.
[15] Bransford J., Brown A. L., Cocking R. R. (Eds.). (1999). *How People Learn Brain, Mind, Experience, and School*. Washington, D.C.: National Academy Press.