Teaching Tools, Applications, and Infrastructure for Digital Curation Through the Use of a Virtual Lab

Joseph Helsing
Department of Computer Science
University of North Texas
Joseph.Helsing@unt.edu

Paulette Lewis
College of Information
University of North Texas
paulette.lewis@unt.edu

Edward Warga
College of Information
University of North Texas
edward.warga@unt.edu

Abstract

With an increased demand for digital curation, data management, archiving, and preservation the library and information science community has begun offering new education and training in these fields. A trend in this education is the use of online virtual labs to offer students hands-on experience with the tools and methodologies inherent to these fields. At the University of North Texas, we are also developing a four course education program in digital curation and data management including a course called *Tools, Applications, and Infrastructure for Digital Curation* that employs a virtual lab component. This paper discusses the first offering of the course; several challenges we faced; and the new strategies developed to address these issues and improve the course and virtual lab for the next offering.

Keywords: virtual lab, digital curation, elearning, online course

Introduction and Project Goals

Libraries have new opportunities to meet the needs for data management, sharing, archiving, and digital curation through new roles and services (Walters and Skinner, 2011). The library and information science community is responding by offering training and education programs focused on digital curation, data management, archiving, and preservation (Fulton, Botticelli, and Bradley, 2011; Harvey, and Bastian, 2012; Yakel, Conway, Hedstrom, and Wallace, 2011). A common trait among these academic programs is the development of virtual labs in which students can gain hands-on experience with the technology infrastructure and tools of the curation environment while working through problems and scenarios concerning digital content management. To work in an information technology rich environment, these professionals need practical knowledge and skills with the tools and processes used in digital curation. Such technology competencies are described by Lee (2009) in the Matrix of Digital Curation Knowledge and Competencies. The virtual lab component of these academic programs seeks to support students as they develop these competencies.

The University of North Texas received a grant from the Institute of Museum and Library Services to create a four course, competency-based curriculum in digital curation and data management (Moen, et al., 2012). The *Information: Curate, Archive, Manage, Preserve* (iCAMP) *Project* is a three-year curriculum development project to meet the need for digital curation professionals in libraries and other information centers. During the first year of the iCAMP Project, two courses were designed and implemented: *Digital Curation and Data Management Fundamentals* and *Tools, Applications, and Infrastructure for Digital Curation*. These courses are intended to give students an introduction to the concepts and technologies used in digital curation and data management.

As the second of four courses, *Tools, Applications, and Infrastructure for Digital Curation* offers an introduction to tools and processes used to conduct digital curation activities. The iCAMP Virtual Lab was developed to provide an environment for students to perform activities and gain experience with the tools, applications, and processes of digital curation. The Virtual Lab is a network of technologies including Linux environments pre-loaded with tools; VirtualBox files loaded with applications such as

---

Acknowledgements: This paper presents work from the iCAMP project, funded by a generous grant from the U.S. Federal Institute of Museum and Library Services, Librarians for the 21st Century Program, RE-05-11-0073-11, with additional support from the University of North Texas College of Information and UNT Libraries.

Helsing, J., Lewis, P., & Warga, E. (2013). Teaching tools, applications, and infrastructure for digital curation through the use of a virtual lab. *iConference 2013 Proceedings* (pp. 781-784). doi:10.9776/13387

DSpace, EPrints, Omeka, and Archivematica; and a shared Fedora repository using a Drupal Islandora front end.

We first implemented the course in Summer 2012, and collected feedback from students using surveys and a focus group session. The data collected inform our course revisions during Fall 2012. Course revisions focus on improving the Virtual Lab technology infrastructure, the learning materials, and the activities of the course, with the goal of improving student performance, learning, and satisfaction. The revised course will be implemented and evaluated again in Spring 2013.

Virtual Lab Infrastructure and Tools

To enable students' hands-on experience with the technologies for digital curation in an online course format, we built a robust cyberinfrastructure environment. This included two primary areas for knowledge development: basic Linux and server knowledge; and digital curation tools and systems knowledge. We used two different systems to cover these areas of knowledge: VirtualBox and an Ubuntu Server.

The first system implemented used the VirtualBox by Oracle. We assumed that some students would be novices in the use of the console in the Linux environment, and we wanted to provide them a safe, destructible system. VirtualBox allowed us to make pre-setup Ubuntu installations that could be redownloaded if someone accidentally destroyed their home directory or an assignment. Students first downloaded the emulator, and then an Ubuntu VirtualBox preloaded with tools and files. Assignments required students to interact with the Linux environment and learn how to use the basic commands and operations to complete certain tasks.

Once students had gained sufficient skill with Linux, they started working on a local server installed and configured for the course. This server was similar to their VirtualBox but had a Fedora repository attached to a Drupal site via Islandora. On the server, we provided tools such as BagIt, ImageMagik, and a Fedora repository. Students performed activities informed by the OAIS reference model and the Digital Curation Centre Lifecycle Model. They engaged in these activities through guided exercises, using the tools in a manner similar to a professional digital curator. Students also performed activities with other digital curation software platforms including Archivematica, EPrints, and DSpace available to students as VirtualBox files.

Since all students were working on the same server for a portion of the course activities, it was very important to provide security and privacy for each student's files. Each student was chrooted into their home directory and only allowed sym links to necessary folders. This, along with strict permission settings, allowed students to work on the same server, while preventing cheating and potential destruction of other students' property.

Technology Challenges, Student Evaluation, and Improvements

Students expressed overall satisfaction and appreciation for the hands-on experience with the tools and technologies presented in the course and through the Virtual Lab. However, student evaluations and feedback pointed out three major areas that presented problems which we are now addressing in the revision of the course and technologies. First, some students realized that they lacked the technical skills necessary for full and ready participation in the course. Second, students reported that it took too long installing and configuring the technology used in the course. Third, students expressed the view that there was a lack of resources and materials available to help solve technology issues in the Virtual Lab, thereby causing delays in learning.

Lack of Technical Skills

Students expressed they lacked overall technical skills necessary for engagement in an information technology rich environment. Some students were unfamiliar with file transfer protocols and related software applications. Additionally, students had little to no fundamental knowledge of server administration and command line interfaces. Students also indicated the need for a Linux primer before beginning the course. To address this issue, we are creating a scaffolding course in Linux to provide the students with basic command line skills. We will also outline and publish the prerequisite skills and knowledge required for success in the course.

Time Installing, Configuring VirtualBox

The use of VirtualBox and appliances served the purpose of providing a safe environment for student exploration. However, this technology proved ineffective due to our inability to ensure a controlled laboratory environment for student practice. The variability among operating systems, memory, RAM and other software in the students' computing environment proved to be a consistent challenge. Students expressed the need for more extensive assistance with the initial technology setup on their individual computers. Not only was it time consuming to download the VirtualBox and several appliances, students also reported that they spent too much time configuring hardware and software for the course. We provided installation tutorials for these tools, but some students encountered technical issues that we did not anticipate. For example, the automatic configuration in the Omeka appliance failed and students were not sure how to manually configure the appliance. Students also encountered problems using SSH/SFTP to transfer files into and out of the Ubuntu virtual machine. Students experienced numerous connection errors because the VirtualBox was unable to accept outside network connections. A considerable amount of time was spent testing various network configuration schemes to determine the correct settings needed.

Since several of our students had problems with using VirtualBox, we decided to forego the use of VirtualBox. Students will have access to an Ubuntu server to learn the basics about Unix environments. Additionally, instead of providing software like Archivematica and EPrints as a VirtualBox, these programs will be hosted on the server and accessed through a web interface. This will hopefully remove additional complications added by using a server emulator and give us greater abilities to diagnose and solve problems.

Lack of Resources and Materials

Students reported that we did not provide appropriate resources and materials to help them solve technology issues in the Virtual Lab. Students communicated they had to troubleshoot various technology errors and software conflicts throughout the course. This hampered the students' abilities to complete assignments and projects and also caused delays in their learning. Students expressed a need for dedicated technical support personnel to help troubleshoot problems throughout the course. Our revised course will provide a knowledge base and helpdesk to support students with technology problems. We will work to identify common technology issues, gather more publicly available online tutorials, and create better documentation for existing infrastructure.

Conclusion

The iCAMP Virtual Lab is our first step toward building successful infrastructure to support this course. Our goals are to have students learn digital curation tools and methods in an environment modeled on the OAIS reference model and the DCC Digital Curation Lifecycle Model and develop the practical technology competencies required in the field of digital curation and data management. By listening to student feedback and responding with revisions to the course, the iCAMP team will refine the technology infrastructure and learning materials to provide an improved student learning experience in future implementations of the *Tools, Applications, and Infrastructure for Digital Curation* course.

References

Fulton, B., Botticelli, P., and Bradley, J. (2011). Digin: A hands-on Approach to a digital curation curriculum for professional development. *Journal of Education for Library and Information Science*, *52*(2). 95-109

Harvey, R., and Bastian, J. A. (2012). Out of the classroom and into the laboratory: teaching digital curation virtually and experientially. *IFLA Journal, 38*(25)

Lee, C. (2009). Matrix of digital curation knowledge and competencies, Version 13. Retrieved from http://ils.unc.edu/digccurr/digccurr-matrix.html

Moen, W, Kim, J., Warga, E., Wakefield, J., and Halbert, M. (2012). iCAMP: building digital information curation curriculum. Poster presentation. 2012 iConference.

Walters, T., and Skinner, K. (2011). New roles for new times: digital curation for preservation. Report Prepared for the Association of Research Libraries. Washington DC.

Yakel, E., Conway, P., Hedstrom, M., & Wallace, D. (2011). Digital curation for digital natives. *Journal of Education for Library and Information Science, 55*(1), 23-31.