LIS9???

Collecting digital documents

2008–09–22

See the course web site at http://openlib.org/home/krichel/courses/lis653p09s for the latest online version of this file.

Course Description
More and more archives and libraries are trying to build collections of digital documents to support the parent institutions' mission of dissemination. Since the documents have a long-run value, the documents receive curatorial efforts that go beyond simple storage on a web site. Such documents are typically housed in formal repositories. Repository building is a crucial avenue for libraries to increase relevance in the digital age. However, many libraries a crippling shortage of technical expertise that is required to even get started with repository building.
This is hands-on course that develops a digital repositories. Each student will build a server and host the server at home or in the Palmer School. All software used is open-source. Once the server is build students set up and configure the `DSpace`[1]. DSpace is configurable to be suit a variety of repository scenarios. Therefore, its configuration is quite abstract and complicated.

Course objectives
After taking this course students

- will understand basic computer hardware and computer networking issues. This will be tested by quizzes.
- will be able to setup and maintain a Linux server for the storage of digital document. This will be measured by a rubric on the final server.
- will be able to set up a reasonably complete set of functional requirements for a repository. This will be measured by a written submission.
- will be able to translate an institution requirements into a functional system based on DSpace. This will be measured by comparing the written submission with the final server outcome.

Prerequisites
There are no other formal prerequisites for this course. However this course is not suitable for computer neophytes or technophobes.
Students should have an old computer that they can use to run the server on. The computer should be an Intel or AMD processor based PC. Installing on another computer is possible, but would add difficulty. An Intel-based computer needs to have at least a 386 processor. It should have 64 megabytes of RAM, and 1 gigabyte of disk space. Most old computers will do a lot better than than. The computer needs to be dedicated to the course during the run-time of the course but can be put to other usage once the course is over. The instructor will try to collect old computers for those who have difficulty finding a computer.
Students should have a network connection at home. It is best if the network goes via a cable mode connection, that leads to an Ethernet connection. This is the most common scenario. If a student does not have such a connection at home, it is possible to host the computer in the instructor's office, but public service on such computers would only be visible on campus, and the computer would be at the risk on networking policy changes at the CW Post campus.

Instructor

[1] http://dspace.org

Thomas Krichel

Palmer School of Library and Information Science
C.W. Post Campus of Long Island University
720 Northern Boulevard
Brookville, NY 11548–1300
krichel@openlib.org
work phone: +1–(516)299–2843
skype: thomaskrichel
Private contact details may be obtained from the online CV at /home/krichel/cv.html.

Class structure
Classes are held on the CW Post campus on Sundays between 13:00 and 18:00. The instructor promises to be there shortly after 12:00 for extra help and questions.
Most classes will have lengthy presentation by the instructor. Some class time is spent by students working directly with their computers under the supervision of the instructor. However, give the hefty weight of the class material, students are expected to do much of the work on their repositories at home. To support the students in this process, the instructor will be on campus in extra sessions for students who need additional support. Support via skype is available pretty much around the clock, i.e. unless the instructor is riding his bicycle or is asleep.
Class details:

2009–??–??	13:00 to 18:00	Important concepts of open source software, computer hardware and computer networks. Inspection of ha
2009–??–??	13:00 to 18:00	Installing GNU Debian Linux. Transfer servers home.
2009–??–??	13:00 to 18:00	Introduction of console-based work. Configuration of critical server components.
2009–??–??	13:00 to 18:00	Basic overview of DSpace concepts. Discussion of business plans.
2009–??–??	13:00 to 18:00	Installing and configuring DSpace
2009–??–??	13:00 to 18:00	Populating the DSpace installation with contents. Security and housekeeping.

Slides for all classes are downloadable from the course web site. The slides on the course website are drafts until the time that the class is held.

Assessment

Mailing list
There will be a mailing list for the course at https://lists-1.liu.edu/mailman/listinfo/cwp-lis653-krichel. All students are encouraged to subscribe. As a rule, answers to email sent to the instructor are copied to the list. There are exceptions to this rule

- if the question writer requests the answer not to be posted;
- if the question is a purely private matter.

Literature
There is no text for this course. The expertise acquired in the course is very difficult to find in existing literature because it spans a wide area of subjects. Some Internet sites include

- Debian – Debian-Installer
- DebianInstaller - Debian Wiki
-

Formally authored documents include

- Mary R. Barton, "Project Planning Matrix", available at `http://www.dspace.org/images/stories/service-model.pdf`

- Mary R. Barton, "Defining a Service Model", available at http://www.dspace.org/images/stories/project-plan.pdf
- Javier Fernández-Sanguino Peña, "The Debian GNU/Linux and Java FAQ", available at http://www.debian.org/doc/manuals/debian-java-faq/.
- W. Martin Borgert, "Debian GNU/Linux Reference Card. The 101 most important things when using Debian GNU/Linux", available at http://xinocat.com/refcard/refcard-en-lt.pdf
- Stuart Lewis and Chris Yates, "The DSpace Course", available at http://cadair.aber.ac.uk/dspace/handle/2160/615

  The DSpace Foundation, "DSpace 1.5.1 Manual", available at http://dspace.svn.sourceforge.net/viewvc/dspace/branches/dspace-1_5_x/dspace/docs/html/index.html

Assessment

There will be a quiz at the beginning of each session except the first and third session. These will count for 40% of the grade. By the third session, students will hand in a description of basic functionality of the installation that they will built, based on the "stories" in the papers by Mary R. Burton. The repository is most likely fictions, but real collections could also be built. This will count for 20% of the grade. By the fifth session, the instructor will publish a set of requirements that the DSpace installation will have to conform to. These will be assessed on the server by the week following the class. This counts for 20% of the final grade. Finally class participation is a crucial component of the course. This is includes demonstrated ability to build their own server as well as willingness and capability to help struggling comrades with their servers. This will be counted for 20% of the grade.