School of Information and Library Science
University of North Carolina, Chapel Hill
INLS 540: Building a Personal Digital Library

Spring 2015
Meeting Time: Friday 9:00 – 11:45, January 9 – May 1, 2015
Location: 303 Manning
Credits: 3
Instructor: Reagan Moore
Office: 22 Manning (Moore)
Phone: 919-962-9548 (Moore)
E-Mail: rwmoore [at] renci [dot] org
Office Hours: 2:00 PM on Friday, or by appointment
Course Web Site:

COURSE DESCRIPTION

This class will prepare students to implement their own personal digital LifeTime Library. The collection that is assembled at SILS will be maintained throughout their academic career, and will be available for migration to their own laptop when they graduate. Topics covered will include: creation of a personal digital library including organization of the material, creation of descriptive metadata, management of the collection, preservation of the collection, and sharing of the collection.

The LifeTiime Library is based on the integrated Rule-Oriented Data System (iRODS). The iRODS software implements policy-based data management. Up to 250 Gigabytes of storage space will be provided for each student. During the course, we will investigate choice of user interface, automation of digital library management functions, enforcement of management policies, and assessment of collection properties. Students will be able to use the LifeTime Library to manage class material, organize multi-media material, archive personal data collections, and share selected material. Collections that have been assembled by prior students range from photograph collections, to music collections, to reference collections of class material, to project collections, to web archives.

Students will be asked to test the digital library, help evaluate the functionality, and generate ideas for improving the system. Students are encouraged to think of themselves as the innovators in the development of digital libraries that will be used by other students.

Students will receive accounts in the LifeTime Library, and will develop policies for managing their personal digital library.

Institutions applying similar policy-based data management systems include:
- UNC-CH SILS LifeTime Library
- UNC-CH Carolina Digital Repository
- French National Library
- Texas Digital Libraries
- UNC-CH Genomics data grid
- NOAA National Climatic Data Center
- NASA Center for Climate Simulations
- iPlant Collaborative
- International Neuroinformatics Coordinating Facility
- Datanet Federation Consortium

The use of the LifeTime Library is governed by the UNC Honor Code. UNC resources may not be used for commercial applications of your personal digital library.

COURSE OBJECTIVES

Upon completion of this course, you should be able to:

- Build a personal digital library
- Select appropriate organization for the material
- Select appropriate descriptive metadata
- Evaluate effectiveness of multiple types of user interfaces
- Build a collection that spans multiple storage resources including your laptop
- Use the personal digital library to build a portfolio
- Use the personal digital library to build a reference collection

PREREQUISITE KNOWLEDGE REQUIRED

Students should have taken INLS 560 (Programming for Information Professionals) or have equivalent knowledge of programming languages. Knowledge of the C programming language or a scripting language is desirable, but it is not required. Drag and drop user interfaces will be available. Students will be expected to write policies for controlling their personal digital library through use of a rule language.

HARDWARE AND SOFTWARE REQUIREMENTS

Students should have access to a laptop (Mac, Windows, or Unix operating system) and will install on their laptops– with help from the instructors – a dropbox style interface, which is available through an open source automated installation system. If you foresee any problems with this laptop requirement, you should let the instructors know as soon as possible. Students will have access to the personal digital library through the computers in the SILS computer lab.

COURSE EXPECTATIONS

- Complete readings BEFORE CLASS each week. Manage your time accordingly.
- Come to class on time.
- Participate in discussions – counts as 25% of your total grade for the course.
- Demonstrate concerted effort to successfully complete all lab exercises, and submit products from the exercises that reflect this effort.
- Practice "respectful and informed ignorance." Will Rogers said, "Everybody is ignorant, only on different subjects." This class will be most effective if everyone feels comfortable asking questions, so respect the questions of others. Bring to class your own informed questions about the week's materials (i.e. be able to convey how you've tried to understand the issues and what still remains unclear to you).
- Final exam. This will be an essay covering formation of a digital library and associated policies.
-

Special Needs: If you feel that you may need an accommodation for a disability or have any other special need, please make an appointment to discuss this with one of us. We will best be able to address special circumstances if we know about them early in the semester.

EVALUATION

- Class participation: 25%
- Completion of in-class lab exercises and homework: 25%
- Presentation of your personal digital library to the class: 25%
- Final exam: 25%

Based on UNC Registrar Policy for graduate-level courses (http://registrar.unc.edu/academic-services/grades/explanation-of-grading-system/), both assignment and semester grades will be H, P, L or F. Few students will obtain an "H," which signifies an exceptionally high level of performance (higher than an "A" in an A-F systems). The following is a more detailed breakdown:

H Superior work: complete command of subject, unusual depth, great creativity or originality

P Satisfactory performance that meets course requirements (expected to be the median grade of all students in the course)

L Unacceptable graduate performance: substandard in significant ways

F Performance that is seriously deficient and unworthy of graduate credit

According to UNC Registrar Policy, undergraduate grades are based on the following definitions:

A Mastery of course content at the highest level of attainment that can reasonably be expected of students at a given stage of development. The A grade states clearly that the students have shown such outstanding promise in the aspect of the discipline under study that he/she may be strongly encouraged to continue.

B Strong performance demonstrating a high level of attainment for a student at a given stage of development. The B grade states that the student has shown solid promise in the aspect of the discipline under study.

C A totally acceptable performance demonstrating an adequate level of attainment for a student at a given stage of development. The C grade states that, while not yet showing unusual promise, the student may continue to study in the discipline with reasonable hope of intellectual development.

D A marginal performance in the required exercises demonstrating a minimal passing level of attainment. A student has given no evidence of prospective growth in the discipline; an accumulation of D grades should be taken to mean that the student would be well advised not to continue in the academic field.

F For whatever reason, an unacceptable performance. The F grade indicates that the student's performance in the required exercises has revealed almost no understanding of the course content. A grade of F should warrant an advisor's questioning whether the student may suitably register for further study in the discipline before remedial work is undertaken.

AB Absent from final examination, but could have passed if exam taken. This is a temporary grade that converts to an F* after the last day of class for the next regular semester unless the student makes up the exam.

FA Failed and absent from exam. The FA grade is given when the undergraduate student did not attend the exam, and could not pass the course regardless of performance on the exam. This would be appropriate for a student that never attended the course or has excessive absences in the course, as well as missing the exam.

IN Work incomplete. This is a temporary grade that converts to F* at the end of eight weeks into the next semester unless the student makes up the incomplete work.

W Withdrew passing. Entered when a student drops after the six-week drop period.

COURSE READINGS
Required Text:
The main text for this course is:

- Rajasekar, Arcot, Michael Wan, Reagan Moore, Wayne Schroeder, Sheau-Yen Chen, Lucas Gilbert, Chien-Yi Hou, Richard Marciano, Paul Tooby, Antoine de Torcy, and Bing Zhu. *integrated Rule-Oriented Data System.*

This book is available as a pdf file through the UNC-CH library and in the LifeTime Library at /lifelibZone/home/rwmoore/LifeTime/Papers/iRODS-primer.pdf. When indicated as one of the required readings for a given class session, we refer to this as the "*iRODS Book.*"

Access to Other Readings:
The texts for this course will be provided as pdf files. Chapters 1-3 of the following book describes the iRODS approach for creating a computer actionable rule. We will use similar rules to control properties of the LifeTime Library.

Ward, Jewel, Michael Wan, Wayne Schroeder, Arcot Rajasekar, Antoine de Torcy, Terrell Russell, Hao Xu, and Reagan Moore. *The integrated Rule-Oriented Data System (iRODS) Micro-service Workbook*, ISBN 978-1-46646-912-9, Amazon.com. (pdf file available through the class Sakai web site and in the LifeTime Library at /lifelibZone/home/rwmoore/LifeTime/Papers/iRODS-micro-services.pdf)

We will use a version that is being edited, /lifelibZone/home/rwmoore/LifeTime/Papers/iRODS-micro-services4.0book.docx.

Additional references will be available from the Sakai web site and from the LifeTime Library at /lifelibZone/home/rwmoore/LifeTime/Papers, and include:
> Actionable Management Policies
> Comments-on-Digital-Libraries
> D3.4Digital_Library_Technology
> Evolving Persistent Archives
> ISO16363MetricsEvidence-100213
> LifeTime-Library-at-UNC-CH
> PolicyDrivenIngestion
> Towards-a-Theory-of_Digital_Preservation

COURSE SCHEDULE AND TOPICS
NOTE: Most weeks of this course will follow a structure of lecture and discussion, followed by a laboratory for digital collection building.

Week 1 (January 9) - Introduction to a Personal Digital Library
Discussion of the characteristics of a digital library, and how a digital library can be used as a personal data collection. We use data grid technology to implement the SILS LifeTime Library. We will examine a framework for building a personal digital library based on ingestion, arrangement, description, management, preservation, and access.

Lab: Please acquire a LifeTime Digital Library account by logging onto http://lifetime-library.ils.unc.edu We will log onto the LifeTime Library and start creating a collection. Bring a laptop computer to class.

Homework: What type of digital library will you want to create? Describe your expectations for the digital library you will be able to build during the semester.

Read:
- LifeTime Library at UNC-CH
- LifeTime Library tutorial - http://lifetime-library.ils.unc.edu/video-tut.html

Week 2 (January 16) – Policy-based data management
Discussion of policy-based data management systems, the basis for the LifeTime Library. We will examine a framework for organizing policies that control a digital library: purpose, properties, policies, procedures, persistent state information, and property verification.

Homework: What type of operations will you need to perform while building your digital library? Are all of these operations available within the LifeTime Library interface?

Read:
- *iRODS Micro-service WorkBook* - Chapter 1 (Introduction)
- *iRODS Primer* - Appendix A (iRODS Shell Commands)
- DL.org Digital Library Reference Model http://bscw.research-infrastructures.eu/pub/bscw.cgi/d222816/D3.2b%20Digital%20Library%20Reference%20Model.pdf
- DL.org Conformance Check List - Policy Oriented Criteria http://www.dlorg.eu/index.php/outcomes/conformance-check-list/policy-oriented-criteria

Week 3 (January 23) – LifeTime Library Purpose

Discussion of the purpose behind the LifeTime Library – archive, portfolio, or reference collection. Discussion of the mechanisms that should be provided by an archive, or digital library, or collaboration environment to satisfy the desired purpose. We will examine a staging policy for managing data ingested into the digital library.

Homework: What purpose do you want your LifeTime Library to serve? Will this purpose change over time?

Read:

- *iRODS Micro-services WorkBook* – Chapter 2 (Rule language)

Week 4 (January 30) – LifeTime Library Interfaces

We will explore the properties of the iDrop interface and implement an initial collection. We will use a web browser interface, and Unix shell command interfaces. We will examine use of an integrity policy for automatically generating checksums on file ingestion.

Homework: What are the minimal capabilities that a digital library interface should provide? What are the extended capabilities that a digital library interface should provide?

Read:

- Evolving Persistent Archives

Week 5 (February 6) – LifeTime Library Properties and Policies

Discussion of the properties that the LifeTime Library should possess. These are assertions that can be made about your collection for completeness, consistency, authoritativeness, and composition.
We will explore rules for automation of data management policies (retention, replication, integrity checking).

Homework: What standard properties should a LifeTime Library possess? What are the minimal policies that should be automatically enforced?

Read:

- Moore, Reagan W. "Building Preservation Environments with Data Grid Technology." *American Archivist* 69, no. 1 (2006): 139-58.

Week 6 (February 13) – LifeTime Library Organization

We will explore how to organize digital material, and assemble an initial collection hierarchy. Approaches range from arrangement in sub-collections, to use of tags, to use of metadata. We will explore rules for arrangement based on file type.

Homework: What organizational structures do you prefer for your digital holdings? Examples include sub-directories, tags, and descriptive metadata.

Read:

- D3.4Digital_Library_Technology
- "Comments on Digital Libraries"

Week 7 (February 20) – LifeTime Library Management

We will explore the policies that are automatically enforced by the LifeTime Library. The policies include enforcement of strict access controls, storage quotas, replication, and metadata backup.

Homework: Would a staging area improve your ability to manage your digital collection?

Read:

- **http://www.proteacher.org/c/388_Classroom_Library_Organization.html. What ideas does this suggest for organization?**

Week 8 (February 27) – LifeTime Library Descriptive Metadata

We will explore addition of descriptive metadata to collections and files. This is normally the hardest task associated with building a digital library. Two methods for automating metadata ingestion will be tried, through parsing of text and import of XML files.

Homework: What descriptive metadata will enable your discovery of relevant material within your collection?

Read:

- **http://www.loc.gov/standards/metadata.html**
- **Compare the difference between descriptive metadata and system metadata**

Week 9 (March 6) – LifeTime Library property verification

We will explore use of computer actionable rules to verify properties of your digital library. Typical assessment criteria include integrity checking, verifying the number of replicas, enforcing a retention period, and checking on time-dependent access controls.

Homework:

- What are the advantages to distributing the LifeTime data collection across multiple storage locations?
- Classify your digital holdings by value, and propose a policy for ensuring the integrity and minimizing risk of data loss.

Read:

- *Actionable Management Policies*

Week 10 (March 13 - NO CLASS (Spring Break, March 9-13)

Homework:

- Consider what standard descriptive metadata should be automatically extracted from your LifeTime Library?
- Propose a set of metadata or tags that you would use to discover material in your digital library.

Week 11 (March 20) – LifeTime Library as an Archive

We will explore the policies (replication, retention, authenticity, integrity) that should be enforced for safely managing an archive. These policies can also be applied to a digital library. We will explore whether the LifeTime Library can serve as both an archive and a digital library.

Homework:

- What are the minimal policies you want enforced on the LifeTime Library to enable use as an archive?

Read:

- http://www.metapress.com.libproxy.lib.unc.edu/content/176p51l2w5278567/fulltext.pdf
- Moore, Reagan. "Towards a Theory of Digital Preservation." *International Journal of Digital Curation* 1, No. 3 (2008). http://www.ijdc.net/index.php/ijdc/article/viewFile/63/42

Week 12 (March 27) – LifeTime Library as a Reference Collection

We will explore the development of a personal reference collection that can be used as either a portfolio for future jobs or a knowledge base. We will consider how the structure of the digital library can be changed over time, how new metadata attributes can be added, and how new policies can be enforced.

Homework: What types of queries should be supported to find reference material?

Read:

- *"ISO 16363 Metrics Evidence-100213"*

Week 13 (April 3) – Holiday

Week 14 (April 10) - LifeTime Library Federation

We will explore how multiple LifeTime libraries can be federated into a larger collection, including access to remote data collections and knowledge bases. We will explore links to social media systems, including the automated retrieval of digital objects and archiving within the LifeTime Library.

Homework:
- What external collections will you want to integrate with your personal Digital Library?
- Identify the published collections that are important to your field of study. Which items would you want to integrate into your personal digital library? Note that the external sources could also include Flickr, Facebook, YouTube.

Read:
- i*RODS Micro-services Workbook* – rules for soft links in module MSO Drivers
- DL.org Cookbook - Policy Domain Interoperability Best practices and Solutions, pp. 68-73 in http://bscw.research-infrastructures.eu/pub/bscw.cgi/d222825/D3.4%20Digital%20Library%20Technology%20and%20Methodology%20Cookbook.pdf

Week 15 (April 17) – Taking the LifeTime Library with you

We will explore how you can integrate the LifeTime Library with your personal computer, and build a life-long digital repository. The LifeTime Library can serve as a life-long archive of your course material. You can also build a second digital library which references data within the LifeTime Library.

We will spend half of the class on presentations of student LifeTime Library collections. Topics of interest include: purpose, ingestion method, arrangement, descriptive metadata, access controls, scale (number of files and attributes), and preferred access methods.

Homework:
- What are the minimal policies you want enforced on the LifeTime Library to enable life-long use?
- Present a collection, organization, descriptive metadata, and control policies for your personal digital library.

Week 16 (April 24) - Synthesis and Conclusions

We will summarize lessons learned during the semester and prioritize the importance of new features for the LifeTime Library. We will spend half of the class on presentations of student LifeTIme Library collections.

Homework:
- We will continue presentations of collection, organization, descriptive metadata, and control policies for your personal digital library.

Week 17 (May 1) - Final

The exam will be an essay question on formation of digital libraries.