INSTRUCTOR

Seth Shaw
Office location: Baker University Center 316C
Office hours: Face-to-face, phone, WebEx: Monday through Friday, 10:00 A.M. – 2:00 P.M. EST, as available. Evening appointments are possible by appointment.

PREREQUISITES

ARST 5000 · Principles and Practices

CREDIT

3 weekly contact hours

SCHEDULE AND LOCATION

Lecture: Wednesday, 6:30 – 8:30 P.M. To join the lecture, open http://claytonstate.webex.edu/ in a browser, then navigate to Meeting Center. Click the link for this class, and when prompted, enter your name, Clayton State email, and the password.

Discussions, Assignments: Throughout the week via Georgia View.

FORMAT OF COURSE

Online course including weekly, synchronous lectures, asynchronous discussions, and tech labs.

DESCRIPTION

An introduction to fundamental concepts and application of technology in an archival context, including hardware, operating systems, programming languages, data storage and backup, networking, and security. Overview of a wide range of digital record formats and their distinguishing characteristics, as well as common archival standards for metadata, including MARC format, Encoded Archival Description, and Dublin Core. Investigates the challenges of acquiring, preserving, and providing access to digital records. Students will be able to design and implement a simple database.

COURSE OUTCOMES

Students will be able to
- Describe the hardware components of a computer and their interactions, including virtualization.
- Describe a variety of operating systems, with an emphasis on the Linux operating system.
- Describe a variety of applications and to demonstrate the installation and configuration of software in a Linux environment.
- Describe the basics of a computer network, including hardware and protocols.
- Design and create a simple, relational database.
- Discuss a variety of techniques for troubleshooting implementation and configuration of computer systems and applications.

**Readings**

*No textbooks are required for this course.* We will be making extensive use of online available resources. In the few cases where readings do include a chapter from a published monograph they will be scanned and placed in GAView. An *optional* text you should consider buying is *UNIX and Linux System Administration Handbook*, 4th ed. (Prentice-Hall, 2011), by Nemeth, Evi, Garth Snyder, Trent R. Hein, and Ben Whaley. This is a large and informative book that can be used as a secondary reference supporting the course content.

**Resources**

- Computer Science E-1: Understanding Computers and the Internet. An extension course offered by Harvard professors Dan Armendariz and David J. Malan. “This course is all about understanding: understanding what’s going on inside your computer when you flip on the switch, why tech support has you constantly rebooting your computer, how everything you do on the Internet can be watched by others, and how your computer can become infected with a worm just by being turned on. Designed for students who use computers and the Internet every day but don’t fully understand how it all works, this course fills in the gaps.” [http://computerscience1.tv/](http://computerscience1.tv/)

- Computer Science 50: Introduction to Computer Science I. “A course offered by Harvard professor David J. Malan. “Introduction to the intellectual enterprises of computer science and the art of programming. This course teaches students how to think algorithmically and solve problems efficiently. Topics include abstraction, algorithms, encapsulation, data structures, databases, memory management, security, software development, virtualization, and websites. Languages include C, PHP, and JavaScript plus SQL, CSS, and HTML.” [http://cs50.tv/](http://cs50.tv/)

**Grading**

Online discussion and participation in lectures: 20%
Midterm: 20%
Assignments: (5% each, 40% total)
1. System Diary
2. Encoding
3. Setup a Virtual Computer
4. Shell Scripting
5. Filesystem Survey
6. Network Discovery
7. Data Modeling
8. Create & Query a Database

Final Exam: 20%

**Expectations**

Students are responsible for their own education. Throughout the course, students should assess their progress towards the course objectives and outcomes. At the same time, the course instructor is
responsible to facilitate students’ learning by structuring content, by providing a foundation of information through readings and lectures, by guiding discussion, and by answering students’ questions.

Students should bring curiosity and creativity to the course. They are expected to think critically about the course content – the readings, the lectures, and discussion. Students are encouraged to (respectfully) challenge the ideas presented in the course. Those challenges must be justified based on the literature, empirical evidence, or other authoritative source. When evidence is contradictory, students should develop a synthesis that finds commonalities, identifies differences, and notes how a particular context may influence that synthesis. As such, there is seldom a “right” answer, but well-reasoned and well-informed points of view. Students are given credit for correcting course materials submitted by the instructor.

Students are expected to find, read, and share additional, relevant works and incorporate the ideas into class discussions.

**Computing Requirements and Responsibilities**

Each student enrolled at Clayton State University is required to have ready access throughout the semester to a notebook computer that meets faculty-approved hardware and software requirements for the student's academic program. Also, each student is responsible for monitoring the Clayton State issued email address on a daily basis for official communications from faculty and administrators. Online courses require a high-speed Internet connection and a headphone with a microphone.

This course uses software that requires a minimum of 2GB RAM; more is preferred. Students are responsible for their own backups; an external hard drive is highly recommended.
## Course Schedule

Note: Dates are for Fall 2014. Dates of class meetings, assignment due dates, and holidays will vary depending on the first day of class (often a week earlier) and holidays (Labor Day, Thanksgiving).

<table>
<thead>
<tr>
<th>20 Aug ~ 1. Introductions and Orientation</th>
<th>Assignment 1. System diary Due: 10 Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students and instructor will introduce themselves. Review syllabus and discuss the expectations and objectives for the course. Overview of topics to be covered in the course and how they relate to each other. How to use Georgia View. How to prepare for discussion.</td>
<td></td>
</tr>
</tbody>
</table>

**You should be able to**
- Distinguish technical fluency (knowledge), competency (basic skills), and proficiency (advanced skills).
- List different kinds of technical skills, including programming, application development, system administration, and database design.
- Discuss the skills a digital archivist needs to work with electronic records and information systems.
- Discuss the impact of legacy systems on archives.
- How to read technical manuals.

**Discussion**
- Introduce yourself to the class. In addition to the introduction in ARST 5000, please describe your experience and knowledge of computers. Are you a novice, or do you have significant experience using a wide range of technologies? What kind of computer are you using for this class?

**Readings**

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital materials have different characteristics and affordances than other media. The prevalence of these materials demands that archivists understand the nature of the medium and see parallels between characteristics of traditional and digital record formats.</td>
<td></td>
</tr>
</tbody>
</table>

**You should be able to**
- Discuss similarities and differences between analog and digital characteristics in terms of paleography; carrier, content, and signal.
- List and define techniques for digital encoding, including binary and hex, ASCII and Unicode and how they are layered.
- List common file formats used to store digital information, including flat text; delimited (e.g. CSV); fixed-field length; raster v. vector graphics; XML; containers; & compression.
- Name common proprietary formats and associated software.
Readings

Discussion
- Explore the notion of affordances in different technologies. For example, what are the different affordances of traditional landline phones (consider rotary phones) and mobile phones? Of vinyl LP records, CDs, and MP3s? Consider other technologies that have analog and digital equivalents. In particular, consider aspects of human interaction with these technologies.
  Responses to posts may suggest other affordances of the different technologies. Or, challenge whether an affordance is a benefit.

3 Sep - 3. Digital Information Systems and Hardware

Overview of physical components of digital information systems and how they are integrated as a system.

You should be able to
- List and describe the major components, including CPU, ROM, RAM, I/O, storage, bus size.
- Describe the differences between different hardware platforms based on their size (super to micro) and function within a system (server, workstation, peripheral).
- Describe virtualization, including host and guest operating systems,
- List common storage media.
- Discuss the importance of backups and perform a basic backup.

Readings

Discussion

Use the discussion boards for the next few weeks to share your experiences configuring the virtual hardware and installing the operating system. Where did you get stuck? What did you try, and what was the solution? You may also use the boards to seek advice from other students.

Also, describe any insights about how computers work “under the hood,” is anything different from how you expected things to be? Consider the impact that technology had on 19th century communications and recordkeeping; can you see any parallels in evolving communication technologies, such as email, Twitter, Facebook, blogs?

10 Sep ~ 4. Operating Systems & System Interfaces

Introduce the components and functions of an operating system. Discuss common operating systems and interfaces (shells).

You should be able to
- To define an operating system and describe its function.
- To name major operating systems (Windows, Mac, Unix/Linux).
- To describe the differences between command line interfaces and graphical user interfaces and to discuss the affordances of each.
- To acquire basic familiarity of command line interfaces

Readings
- Thomas, Keir. Chapter 5, Hands on at the Command-Line. Ubuntu Pocket Guide and Reference, p. 49-96. http://www.ubuntupocketguide.com/index_main.html Note: you may skip any portion not relevant to your operating system. You will also want to be familiar with the table of contents so you’ll know when to refer back to this book.

17 Sep ~ 5. System Administration

Common system tasks (and commands). Creating batch commands, capturing output from commands.
You should be able to
- To discuss the role of system administrator
- To describe basic OS-level operations, such as login, exit, shutdown, restart.
- To identify the common important OS directories.
- To describe the nature and use of system configuration and log files.
- To define patches and know how to apply them.
- To install software using Windows Control Panel and Synaptic package manager (or aptitude).

Readings

24 Sep ~ 6. Software
Users typically interact with specific applications to perform practical tasks. Software may be design for an enterprise or individual. Basic software by functionality: word processing, communication (email, IM), web, spreadsheets, databases, workflow and business process, enterprise.

You should be able to
- To name functional categories of software: word processing, spreadsheet, database, enterprise, communication, graphic processing; application, productivity, groupware.
- To distinguish the affordances of different text processors, including WYSIWYG (Word, Open Office), plain text (NotePad++, Textpad)
- To define open source software and name and describe common licensing schemes (e.g., GNU Public License).

Readings

1 Oct ~ 7. Midterm exam
A combination of short and long answer questions. The exam (in Word format) will be available for download from GAView at 6:30 p.m. It must be
submitted in Georgia View by 9:00 p.m. Contact the instructor immediately if you have technical problems.

**8 Oct ~ 8. Programming**

The act of creating software to be run on computers is called programming. There are a few basic concepts that govern most programming languages.

You should be able to
- Describe how source-code is interpreted into byte-code
- To describe basic programming concepts:
  - Variables
  - Conditionals
  - Loops

Readings
- Holowczak, Richard. “Programming Concepts: A Brief Tutorial for New Programmers.” [http://holowczak.com/programming-concepts-tutorial-programmers/](http://holowczak.com/programming-concepts-tutorial-programmers/). *Note: The Introduction through Loops is required; the rest is optional although “Hashing Functions” is recommended for future reference. Don’t be intimidated by tables of language examples; it simply serves to illustrate that programming languages use different syntaxes to do exactly the same thing.*
- Nemeth, Evi, Garth Snyder, Trent R. Hein, and Ben Whaley. “Chapter 2: Scripting and the Shell.” in UNIX and Linux System Administration Handbook, 4th ed. (Prentice-Hall, 2011), pp. 29-48, 73-75. *Note: The whole chapter will be scanned but only the pages listed are required; however, the rest is recommended.*
- Evans, David, “Programming,” in Introduction to Computing, p. 35-52. (Optional)

**Assignment 4. Shell Scripting**

Due: 17 Oct.

**10 Oct ~ Midterm**

Last day to withdraw from any course and receive a passing grade. Withdrawal after this date will receive an automatic ‘F’ unless the student is granted an exemption for hardship.

**15 Oct ~ 9. Storage**

Digital information may be stored in a number of places, including removable media, local storage, central servers, and the cloud. Each has specific affordances and limitations. Digital archivists must consider the appropriate technology to support the specific needs of the repository and the data.

You should be able to
- To list and describe media commonly found and use in archives, including signal encoding (physical, magnetic, optical); media (paper, plastics, metal, laminates); connectivity (removable, local, system, cloud).
- To define backup and describe backup strategies; to define online, nearline, and offline storage.
- To distinguish ‘deleting’ and ‘erasing’ files completely, define slack space.
- To discuss mechanisms to transfer files using sftp, ssh.

**Assignment 5. Filesystem Survey**

Due: 31 Oct.
- To discuss LOCKSS and SRB as a means to store content.

Readings

22 Oct – 10. File Systems

The organization of files in storage.

You should be able to
- To define a file system.
- To describe path.
- To navigate the file system using absolute and relative paths.
- To navigate and manipulate file systems using a CLI or GUI, the Unix and Windows commands pwd, ls (dir), mv (rename), cp (copy), cd, mkdir, rmdir.
- To describe the rules for naming files, including practical file naming conventions. To define filename and filetype (extension).

Readings

29 Oct – 11. Networks

Computers are frequently connected to share data and processing using the public Internet or private intranets. Discussion of means of connecting computers with emphasis on Ethernet and related protocols for exchanging data securely.

You should be able to
- To describe common network activities: storage, communication, remote desktop/ssh, file transfer
- To define common network protocols: TCP/IP, http, ftp, smtp
- To describe the basic purpose components of TCP/IP and DNS, including packet switching for robust, post nuclear network
- To define: IP address, DHCP, static IP, MAC address
- To describe the purpose of key addresses and ports, including localhost, 127.0.0.1, 192.168.x.x; ports 80, 8080, 21, 22
- To describe basic networking in Windows clients (ipconfig, Network and Sharing Center) and Ubuntu (ifconfig)
- To distinguish the functionality of bridged and NAT configuration in VirtualBox
- Security, including firewall ports

Readings
- “Internetworking Basics” (Cisco, 2009).
  http://docwiki.cisco.com/wiki/Internetworking_Basics

5 Nov – 12. Preservation and Security

Ensuring digital information and records can be access for many years poses a number of interesting challenges. Protecting systems from accidental and malicious hazards, including backup, secure passwords, SQL injection, https, ssh. Tools can help identify common formats to help us prepare for preservation activities as those formats become obsolete.

You should be able to
- To describe the risks and hazards of digital preservation, including software and hardware obsolescence, media and signal degradation.
- To discuss the component of a digital preservation system.
- To list key digital preservation standards, including OAIS, TRAC, and PREMIS.
- To describe basic techniques for system security, including passwords, physical and network security.
- To describe best practices for system backup.
- To describe approaches for system recovery.
- To describe hacker attacks, defenses, and methods to research potential attacks (CERT).
- To define hash values as digital thumbprints for security

Readings

12 Nov ~ 13. Relational Data Modeling

Introduction to concepts. Common database software (Access, SQL Server, Oracle, MySQL, PostgreSQL). Spreadsheets are not databases. Data modeling

You should be able to
- To distinguish structured and unstructured data.
- To describe data types (integer, string, date, currency).
- To define relational databases, tables, columns, and rows.
- To identify data elements and determine the appropriate data type.
- To read a simple UML Class diagram.
- To describe normalization.
- To relate tables using primary and foreign keys.

Readings

Assignment 7. Data Modeling
Due: 31 Nov.

19 Nov ~ 14. Using Relational Databases

Implementing the data model in SQL. Demonstration of MS Access and LibreOffice as front ends to MySQL databases.

You should be able to
- Install MySQL, MySQL Workbench.
- To use management tools to create a simple database schema.
- Insert information into the database using the update query and from a file (insert data infile).
- To write simple SQL queries that retrieve select data element (select [columns]), with limits (where), in order (ordered by) using a simple join.

Readings
- “SQL Tutorial” (w3schools, checked 2 Aug 2012).
  http://www.w3schools.com/sql/. Read all of SQL Basic, including Intro, Syntax, Select, Distinct, Where, And & Or, Order By, Insert, Update, and Delete.
- Mariano Casanova. “BASE TUTORIAL: From Newbie to Advocate in a one, two... three!” (2010). Chapters 7 & 9 (p. 71-84 and 111-131). Note: Chapters 8 & 10 are optional but recommended.
  https://wiki.documentfoundation.org/images/0/02/Base_tutorial.pdf

26 Nov ~ Thanksgiving

Holiday, no class

3 Dec ~ 15. Additional Data Modeling Techniques

An introduction to other means of modeling & querying data.

You should be able to
- To describe different types of databases (e.g. Graph/Triple-stores, object dbs, & document dbs)

Readings

10 Dec ~ Final Exam

No late work accepted after this date.
ASSIGNMENTS

Plan to stay after class on Wednesday nights for group discussions to help solve problems.

Assignment 1. System Diary  
  Due: 10 Dec  
  Keep a record of the process to install, configure, and maintain your virtual computer.

Assignment 2. Encoding  
  Due: 05 Sep  
  A short exercise to help understand binary, hex, and decimal encoding of digital information.

Assignment 3. Setup a Virtual Computer  
  Due: 26 Sep  
  Download and install Oracle VirtualBox software on your computer. Download the Xubuntu operating system and create a virtual computer running the Linux operating system.

Assignment 4. Shell Scripting  
  Due: 17 Oct  
  A short exercise to walk through programming concepts (variables, conditionals, and looping).

Assignment 5. Filesystem Survey  
  Due: 31 Oct  
  Explore and document the contents of a directory structure.

Assignment 6. Network Discovery  
  Due: 07 Nov  
  Locate and document information about your home network.

Assignment 7. Data Modeling  
  Due: 31 Nov  
  Design a catalog database, including identifying data elements and typical values; establish the relationships between elements; define a table.

Assignment 8. Create & Query a Database  
  Due: 05 Dec  
  Create a database based on a model, insert data, and create SQL queries to find records in your database.