

## **History 9877a**

### **Digital Research Methods**

Fall Term  
Wednesdays, 6:30-9:30 pm, Lawson Hall 2270C

William J. Turkel, Associate Professor  
Department of History  
Office: Lawson Hall 2267  
E-mail: wturkel@uwo.ca

Course website:  
*<http://williamjturkel.net/teaching/history-9877a-digital-research-methods-fall-2014/>*

GitHub repository:  
*<https://github.com/williamjturkel/Digital-Research-Methods>*

#### **Short Course Description**

Historical research now crucially involves the acquisition and use of digital sources. In History 9877A, students will learn to find, harvest, manage, excerpt, cluster and analyze digital materials throughout the research process, from initial exploratory forays through the production of an electronic article or monograph which is ready to submit for publication.

#### **Three Scenarios**

Why should you take a course in digital research methods?

1. You've just returned from a whirlwind trip to the archives. On your laptop you have about nine thousand digital photographs of various documents. You could spend the next few years going through the pictures one at a time and typing notes into a word processor. Or you could write a small script to convert each image into readable text and drop the whole batch into a custom search engine. In less than an hour you could be searching for words and phrases anywhere in your primary sources.

2. You discover that the Internet Archive has a collection of eight hundred online texts that are directly related to your research. You could look through the list of titles in your web browser and click on the links one at a time, scanning each to see if it is relevant. Even if you cut-and-paste notes from the sources to a word processor, it will still take you at least a few months to go through the collection. Or

you could write a small script to download all of the sources to your own machine and run a clustering program on them. This sorts the texts into folders of closely related documents, then subdivides those by topic. In less than an hour, you would be able to visualize the contents of the whole collection and focus in on the topics that are of immediate interest to you.

3. You've been working with the written corpus of a historically significant figure. You have the books and essays that he or she wrote, their diary entries and their correspondence with a large number of other individuals. How do you make sense of a lifetime of writing? Can you chart important changes in someone's conceptual world? Spot the emergence of new ideas in the discourse of a community? Map the ever-changing social relations between a network of correspondents?

In this course you will learn to apply techniques that are currently used by fewer than one percent of working historians. Computation won't magically do your research for you, but it will make you *much* more efficient. You can focus on close reading, interpretation and writing, and use machines to help you find, summarize, organize and visualize sources.

### **Prerequisites and Workload**

There are no prerequisites for the course other than a willingness to learn new things and the perseverance to keep working when you're confused or when you realize that you could spend a lifetime learning about the topics and technologies that we will cover in class, and still not master them all. Students will come into the course with very different levels of experience and expertise. Some, probably most, will be familiar only with the rudiments of computer and internet use. A few may already be skilled programmers.

This course also requires that you spend at least a little bit of time each day (say 20-30 minutes) practicing your new skills. It's a lot like learning a new language, learning to play a musical instrument or going to the gym. With daily practice, you will soon find ways to do your research and coursework faster and more efficiently. If you can't commit to regular practice, however, you should probably not take this course.

### **Course Materials**

There is one required text. It is available in the Western Bookstore, or online from sites like Amazon and Indigo.

Shotts, William E., Jr. *The Linux Command Line: A Complete Introduction*. No Starch Press, 2012.

If you have a laptop (Windows, Mac or Linux) please bring it to every class. You should also have a USB Flash Drive with at least 16 Gb of free space.

We will be working with a wide variety of freely-available open source software packages. More information about these will be provided in class.

## **Evaluation**

The techniques that you learn in this class build cumulatively week-by-week. In addition to regular practice, it is essential that you attend every meeting of the class and do the readings carefully.

You will be graded on your participation in class (20%) and on your reflective blogging (80%; see below).

There will be no midterm or final examinations, and no final paper.

## **Blogging**

Every student in the class will have an academic blog and will be required to make weekly posts to it. These entries do not have to be long (300-500 words per week is ample). The use of blogging is to encourage you to engage in 'reflective practice,' that is, to force you to think about your learning and research as you are doing it. It also provides me with feedback for how the course is going. You can use each week's blog entry to talk about what you learned, things that were clear or not, things you would like to know how to do, and so on.

Before the first class you should go to either WordPress or Blogger (not both) and create an account and a blog. If possible, create the blog under your own name; if not, choose something professional sounding. Post an introductory message about yourself and then send me the URL of your blog so that I can add you to the course blogroll for History 9877A.

Wordpress: <http://wordpress.com>

Blogger: <http://www.blogger.com>

## **Preliminary Schedule (Weekly Topics)**

This is a preliminary list of topics. Consult the course website for up-to-date information.

1. Basic text analysis and word frequencies
2. Regular expression pattern-matching and permuted term indexes
3. Batch downloading sources and building a simple search engine
4. Named entity recognition
5. Optical character recognition
6. Working with PDFs
7. Information trapping and text editing

8. Bibliographic work
9. A more advanced search engine
10. Clustering
11. Working with structured data
12. Visualization
13. Scripting and workflow

### **Statement on Academic Offences**

Scholastic offences are taken seriously and students are directed to read the appropriate policy, specifically, the definition of what constitutes a Scholastic Offence, at the following Web site:

[http://www.uwo.ca/univsec/handbook/appeals/scholastic\\_discipline\\_grad.pdf](http://www.uwo.ca/univsec/handbook/appeals/scholastic_discipline_grad.pdf)

### **Health and Wellness**

As part of a successful graduate student experience at Western, we encourage students to make their health and wellness a priority. Western provides several on campus health-related services to help you achieve optimum health and engage in healthy living while pursuing your graduate degree. For example, to support physical activity, all students, as part of their registration, receive membership in Western's Campus Recreation Centre. Numerous cultural events are offered throughout the year. Please check out the Faculty of Music web page

<http://www.music.uwo.ca/>

and our own McIntosh Gallery

<http://www.mcintoshgallery.ca/>

Information regarding health- and wellness-related services available to students may be found at

<http://www.health.uwo.ca/>

Students seeking help regarding mental health concerns are advised to speak to someone they feel comfortable confiding in, such as their faculty supervisor, their program director (graduate chair), or other relevant administrators in their unit. Campus mental health resources may be found at

[http://www.health.uwo.ca/mental\\_health/resources.html](http://www.health.uwo.ca/mental_health/resources.html)

To help you learn more about mental health, Western has developed an interactive mental health learning module, found here:

[http://www.health.uwo.ca/mental\\_health/module.html](http://www.health.uwo.ca/mental_health/module.html)

This module is 30 minutes in length and provides participants with a basic understanding of mental health issues and of available campus and community resources. Topics include stress, anxiety, depression, suicide and eating disorders. After successful completion of the module, participants receive a certificate confirming their participation.