

HISTORY 4816A
Digital Research Methods
Fall/Winter 2022-23
4 hours per week, Fall Semester
Asynchronous online course

Instructor: Professor William J. Turkel

Office Hours: online by appointment
Department of History, Office: Lawson Hall 2267
Email: wturkel@uwo.ca

This is a draft syllabus. Please see your course OWL site for the final syllabus.

Course delivery with respect to the COVID-19 pandemic

The intent is for this course to be delivered in-person but should the COVID-19 landscape shift, the course will be delivered on-line synchronously (ie., on Zoom at the times indicated in the timetable) / asynchronously (eg., posted on OWL for students to view at their convenience) [delete as necessary]. The grading scheme will not change. The course will return to an in-person mode of delivery when the University and local health authorities deem it safe to do so.

Course Description:

In this course students will learn how historical content is produced, presented and published online; how to find and evaluate digital primary and secondary sources; and how to use computational techniques to work with digital resources. No previous background in the subject area is required.

Prerequisite(s):

Registration in any academic program at the second-year level or above.

Unless you have either the prerequisites for this course or written special permission from your Dean to enroll in it, you may be removed from this course and it will be deleted from your record. The decision may not be appealed. You will receive no adjustment to your fees in the event that you are dropped from a course for failing to have the necessary prerequisites.

Antirequisite(s):

Digital Humanities 2130A/B, Digital Humanities 3902F/G, the former History 2816F/G.

Course Syllabus:

Research of all kinds now crucially involves the acquisition and use of digital sources, both primary and secondary. In this course, you 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. The digital sources that we work with include texts, images, maps, code, social media, spreadsheets, databases and linked open data. The techniques that we will employ are currently used by relatively few humanists and social scientists—although they have become much more crucial in an age of pandemic lockdowns. 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.

Prior Experience

The only thing that you need to do well in this course is 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.

Workload

This course requires that you spend at least a 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. At first it is going to be hard but be patient with yourself and ask a lot of questions. 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. The techniques that you learn in this class build cumulatively week-by-week, and there are a number of evenly spaced assignments of equal value.

Learning Outcomes

At the conclusion of this course, you will

- Be aware of a wide variety of different kinds of digital sources and strategies for making use of them computationally
- Be familiar with a range of research approaches in digital humanities, computational social science and related disciplines
- Be acquainted with methodological applications of artificial intelligence
- Be more familiar with using computer code as a medium of expression

Course Materials:

To get the most out of this class, you will need a Windows, Mac or Linux laptop or desktop. A tablet or netbook will probably not be sufficient. If you have any concerns, email me.

You will need to purchase a **student desktop license** for Wolfram Research's *Mathematica* software. (Don't let the name scare you, you won't need any particular training in mathematics to do well in the class). A one semester license is \$56 CAD. If you think you might want to use Mathematica for some of your other courses or in the summer, a one-year license is \$88 CAD. You can purchase the software here

<https://www.wolfram.com/mathematica/pricing/students/>

If you are unable or unwilling to purchase the software, please do not take the course.

The required text and screencasts for the course are freely available online:

William J. Turkel. *Digital Research Methods with Mathematica*, 2nd rev ed. (2020).

<https://williamjturkel.net/digital-research-methods-with-mathematica/>

Methods of Evaluation:

Students must submit all course assignments in order to pass the course.

<i>Assignment</i>	<i>Worth</i>	<i>Handed Out</i>	<i>Due Date</i>	<i>Zero Date</i>
01	20%	Week 01	Week 02	Week 03
02	20%	Week 03	Week 04	Week 05
03	20%	Week 05	Week 06	Week 07
04	20%	Week 08	Week 09	Week 10
05	20%	Week 10	Week 11	Week 12

Each assignment has a suggested due date and, a week later, a zero date. If you hand in your assignment before the zero date, you will receive full credit for your work. After the zero date, it will be worth nothing.

If you are unable to meet a course requirement due to illness or other serious circumstances, you must provide valid medical or other supporting documentation to Academic Counseling as soon as possible.

Course Schedule and Readings:

- **Week 01.**
 - **Introduction to *Mathematica*.** Interacting with notebooks
 - **Reading Code.** Word frequency, word clouds and stopwords.
- **Week 02.**
 - **Computable Knowledge.** Entities, tables, timelines and maps.
 - **Text Content.** Mathematica notebooks and expressions, strings and natural language processing.

- **Week 03.**
 - **Data Structures.** Lists, associations and datasets.
 - **Reusing Code.** Defining and developing functions, keyword in context (KWIC).
- **Week 04.**
 - **Networks.** Metadata, matrices and social network analysis.
 - **Indexing and Searching.** Pattern matching, topic classification and term distribution.
- **Week 05.**
 - **Geospatial Analysis.** Geographic information: raster, vector and attribute data.
 - **Images.** Computer vision, face detection, feature extraction and image mining.
- **Week 06.**
 - **Page Images.** Optical character recognition (OCR), figure extraction and classification.
 - **Crawling.** Browser automation, batch downloading, web archives and WARC files.
- **Week 07.**
 - **Linked Open Data.** Resource description framework (RDF), SPARQL queries and endpoints, JSON-LD.
 - **Markup Languages.** Scraping and parsing, XML, really simple syndication (RSS) and text encoding initiative (TEI).
- **Week 08.**
 - **Studying Societies.** Computational social science, search data, social media and social networks.
 - **Extracting Keywords.** Information retrieval, term frequency-inverse document frequency (TF-IDF) and rapid automatic keyword extraction (RAKE).
- **Week 09.**
 - **Word and Document Vectors.** Feature extraction, dimension reduction, word embeddings and global vectors.
 - **Citations.** References, web services, bibliographic linked open data and citation networks.
- **Week 10.**
 - **Natural Language.** Multilingual analysis, computational linguistics and sentiment analysis.
 - **Web Services.** Entity networks, publication search, dashboards, manipulating JSON.
- **Week 11.**
 - **Databases.** Parts, selections and transformations, computations and querying, relations.
 - **Measuring Images.** Photogrammetry, georectification, handwriting and facial 3D reconstruction.
- **Week 12.**
 - **Machine Learning.** Unsupervised clustering, classify, predict and transfer learning.

Additional Statements:

Please review the Department of History Course Must-Knows document, <https://www.history.uwo.ca/undergraduate/Docs/Department%20of%20History%20Course%20Must-Knows.pdf>, for additional information regarding:

- Academic Offences
- Accessibility Options
- Medical Issues

- Plagiarism
- Scholastic Offences
- Copyright
- Health and Wellness

DRAFT