HISTORY 4816A / 9877A
Digital Research Methods
Fall/Winter 2020-21
4 hours per week, no fixed times scheduled, offered online only

Instructor: Professor William J. Turkel

Office Hours: Mondays through Thursdays, 5-5:45pm or by appointment, meeting online only
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.

This course will take place in a virtual, asynchronous format - on-line with no scheduled meeting time.

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In this course students will learn how historical content is produced, presented and published online; how to find and evaluate digital primary and secondary sources; and how to use computational techniques to work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 should 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).

http://www.wolfram.com/mathematica/pricing/students-individuals.php

You can purchase the license for the semester, for the year or permanently. If you are unwilling or unable to purchase the software, please do not take this course.

The required text for the course is freely available online:

William J. Turkel. *Digital Research Methods with Mathematica*, 2nd ed. (2019).
https://williamjturkel.net/digital-research-methods-with-mathematica/

Methods of Evaluation:

Assignment	Worth	Handed Out	Due Date	Zero Date
01	20%	Sep 17	Sep 24	Oct 01
02	20%	Oct 01	Oct 08	Oct 15
03	20%	Oct 15	Oct 22	Oct 29
04	20%	Oct 29	Nov 12	Nov 19
05	20%	Nov 19	Nov 26	Dec 03

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.

Accommodation for missed assignment deadlines with a Self Reported Absence:

If a student reports an SRA for an assignment (i.e. an essay) the new due date will be 48 hours after the SRA was submitted. For example, if you complete an SRA on March 19 at 3pm, your new due date will be March 21 at 3pm.

Course Schedule and Readings:

- Sep 10. **Reading Code**. Word frequency, word clouds and stopwords.

- Sep 15. **Computable Knowledge**. Entities, tables, timelines and maps.

- Sep 17. **Text Content**. Mathematica notebooks and expressions, strings and natural language processing.

- Sep 22. **Data Structures**. Lists, associations and datasets.

- Sep 24. **Reusing Code**. Defining and developing functions, keyword in context (KWIC).

- Sep 29. **Networks**. Metadata, matrices and social network analysis.

- Oct 01. **Indexing and Searching**. Pattern matching, topic classification and term distribution.

- Oct 06. **Geospatial Analysis**. Geographic information: raster, vector and attribute data.

- Oct 08. **Images**. Computer vision, face detection, feature extraction and image mining.

- Oct 13. **Page Images**. Optical character recognition (OCR), figure extraction and classification.

- Oct 15. **Crawling**. Browser automation, batch downloading, web archives and WARC files.

- Oct 20. **Linked Open Data**. Resource description framework (RDF), SPARQL queries and endpoints, JSON-LD.

- Oct 22. **Markup Languages**. Scraping and parsing, XML, really simple syndication (RSS) and text encoding initiative (TEI).

- Oct 27. **Studying Societies**. Computational social science, search data, social media and social networks.

- Oct 29. **Extracting Keywords**. Information retrieval, term frequency-inverse document frequency (TF-IDF) and rapid automatic keyword extraction (RAKE).

- Nov 02-08. FALL STUDY BREAK

- Nov 10. **Word and Document Vectors**. Feature extraction, dimension reduction, word embeddings and global vectors.

- Nov 12. **Citations**. References, web services, bibliographic linked open data and citation networks.

- Nov 17. **Natural Language**. Multilingual analysis, computational linguistics and sentiment analysis.

- Nov 19. **Web Services**. Entity networks, publication search, dashboards, manipulating JSON.

- Nov 24. **Databases**. Parts, selections and transformations, computations and querying, relations.

- Nov 26. **Measuring Images**. Photogrammetry, georectification, handwriting and facial 3D reconstruction.

- Dec 01. **Machine Learning**. Unsupervised clustering, classify, predict and transfer learning.

Additional Statements

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 website: http://www.uwo.ca/univsec/pdf/academic_policies/appeals/scholastic_discipline_undergrad.pdf.

Accessibility Options:

Please contact the course instructor if you require material in an alternate format or if you require any other arrangements to make this course more accessible to you. You may also wish to contact Student Accessibility Services at 519 661-2111 x 82147 for any specific question regarding an accommodation. Information regarding

accommodation of exams is available on the Registrar's website:
www.registrar.uwo.ca/examinations/accommodated_exams.html

Medical Issues

The University recognizes that a student's ability to meet his/her academic responsibilities may, on occasion, be impaired by medical illness. Please go to:
https://www.uwo.ca/univsec/pdf/academic_policies/appeals/accommodation_illness.pdf
to read about the University's policy on medical accommodation. In the event of illness, you should contact Academic Counselling as soon as possible. The Academic Counsellors will determine, in consultation with the student, whether or not accommodation should be requested. They will subsequently contact the instructors in the relevant courses about the accommodation. Once the instructor has made a decision about whether to grant an accommodation, the student should contact his/her instructors to determine a new due date for tests, assignments, and exams.

Students must see the Academic Counsellor and submit all required documentation in order to be approved for certain accommodation.

Please visit https://www.uwo.ca/univsec/academic_policies/index.html to view all updated academic policies regarding medical accommodations.

Plagiarism:

Students must write their essays and assignments in their own words. Whenever students take an idea, or a passage from another author, they must acknowledge their debt both by using quotation marks where appropriate and by proper referencing such as footnotes or citations. Plagiarism is a major academic offense (see Scholastic Offence Policy in the Western Academic Calendar).

Students are expected to retain all research notes, rough drafts, essay outlines, and other materials used in preparing assignments. In the unlikely event of concerns being raised about the authenticity of any assignment, your instructor may ask you to produce these materials; an inability to do so may weigh heavily against you.

The following rules pertain to the acknowledgements necessary in academic papers.

A. In using another writer's words, you must both place the words in quotation marks and acknowledge that the words are those of another writer.

> You are plagiarizing if you use a sequence of words, a sentence or a paragraph taken from other writers without acknowledging them to be theirs. Acknowledgement is indicated either by (1) mentioning the author and work from which the words are borrowed in the text of your paper; or by (2) placing a footnote number at the end of the quotation in your text, and including a correspondingly numbered footnote at the bottom of the page (or in a separate reference section at the end of your essay). This footnote should indicate author, title of the work, place and date of publication and page number. Method (2) given above is usually preferable for academic essays because it provides the reader with more information about your sources and leaves your text uncluttered with parenthetical and tangential references. In either case words taken from another author must be enclosed in quotation marks or set off from your text by single spacing and indentation in such a way that they cannot be mistaken for your own words. Note that you cannot

avoid indicating quotation simply by changing a word or phrase in a sentence or paragraph which is not your own.

B. In adopting other writer's ideas, you must acknowledge that they are another writers ideas.

You are plagiarizing if you adopt, summarize, or paraphrase other writers' trains of argument, ideas or sequences of ideas without acknowledging their authorship according to the method of acknowledgement given in "A" above. Since the words are your own, they need not be enclosed in quotation marks. Be certain, however, that the words you use are entirely your own; where you must use words or phrases from your source; these should be enclosed in quotation marks, as in "A" above.

Clearly, it is possible for you to formulate arguments or ideas independently of another writer who has expounded the same ideas, and whom you have not read. Where you got your ideas is the important consideration here. Do not be afraid to present an argument or idea without acknowledgement to another writer, if you have arrived at it entirely independently. Acknowledge it if you have derived it from a source outside your own thinking on the subject.

In short, use of acknowledgements and, when necessary, quotation marks is necessary to distinguish clearly between what is yours and what is not. Since the rules have been explained to you, if you fail to make this distinction, your instructor very likely will do so for you, and they will be forced to regard your omission as intentional literary theft. Plagiarism is a serious offence which may result in a student receiving an 'F' in a course or, in extreme cases, in their suspension from the University.

Scholastic 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 website:
www.uwo.ca/univsec/pdf/academic_policies/appeals/scholastic_discipline_undergrad.pdf

Copyright

Lectures and course materials, including power point presentations, outlines, and similar materials, are protected by copyright. You may take notes and make copies of course materials for your own educational use. You may not record lectures, reproduce (or allow others to reproduce), post or distribute lecture notes, wiki material, and other course materials publicly and/or for commercial purposes without my written consent.

Health and Wellness

Students who are in emotional/mental distress should refer to Mental Health@Western http://www.uwo.ca/uwocom/mentalhealth/ for a complete list of options about how to obtain help.

Please contact the course instructor if you require material in an alternate format or if you require any other arrangements to make this course more accessible to you. You may also wish to contact Student Accessibility Services at 661-2111 x 82147 for any specific question regarding an accommodation.

If you have any further questions or concerns please contact, Heidi Van Galen, Administrative Officer, Department of History, 519-661-2111 x84963 or e-mail vangalen@uwo.ca.