

Digital Scholarship Fundamentals Workshops

Sponsored by the [University Libraries](#) and the [Department of Library and Information Science](#)

Fall Semester Theme: Text Data Mining

Text data mining is the process of extracting valuable insights and patterns from unstructured textual data using various computational and statistical techniques to analyze and interpret this textual information. This data can include a wide range of sources such as documents, articles, emails, social media posts, reviews, and more.

Starting a Text Data Mining Project

Mon., Sept. 18, 12:00 pm – 1:00 pm

Planning your project is a critical skill in contemporary research. Many students and faculty embark on a project without understanding fully the resources needed and the time commitment involved. This workshop will walk you through the process of a text/data mining project by asking the right questions: what is my research question, how can I locate and acquire texts/data, what tools are relevant for cleaning and analyzing texts/data, and what legal issues may limit my access and use of texts/data?

Using AI to Expand Your Research Toolbox

Fri., Sept. 29, 12:00 pm – 1:00 pm

Tools such as ChatGPT, Elicit, the new Bing, Google Bard and browser extensions can be effective in the research process. How and when to discover, evaluate, and cite resources can be challenging. Come deepen your information literacy skills by understanding the best practices for using these new technologies, and how they complement Google Scholar, SearchBox, and even your favorite library subscription database. This workshop is for anyone in the university community who is curious about the impact of new technologies on traditional research methods. A brief overview of text data mining possibilities will be provided.

Gale Digital Scholar Lab

Fri., Oct. 13, 12:00 pm – 1:00 pm

Learn how to use the lab for locating, analyzing, and visualizing texts. Using Gale's Primary Sources archive, we will demonstrate the workflow process in building, cleaning, and analyzing content. We will explore some of the tools including document clustering, Named Entity Recognition, Ngrams, parts of speech, sentiment analysis, and topic modeling. No previous experience necessary.

Legal and Ethical Issues in Text Data Mining

Mon., Oct. 23, 12:00 pm – 1:00 pm

I am not sure if I can text data mine a particular dataset. How can I determine what my rights are? We will explore best practices in copyright, fair use, licensing agreements and terms of use, privacy and ethical issues, digital rights management, and other issues involving non-consumptive use of text for research. Part of Open Access Week.

HathiTrust for Text Data Mining: Introduction

Mon., Nov. 6, 12:00 pm – 1:00 pm

You may have used the HathiTrust Digital Library for acquiring books and articles. Now use the HathiTrust Research Center for computational analysis! We will provide an overview of the HTRC platform and features by working on such as finding textual data, creating a workset, and performing basic analyses. **Instructors:** Benjamin Cushing, Research and Instruction librarian, and Kevin Gunn, Coordinator of Digital Scholarship

HathiTrust for Text Data Mining: Analytics

Fri., Nov. 17, 12:00 pm – 1:00 pm

Building on the introductory workshop, we will examine extracted features, text analysis algorithms, and data capsules. No coding experience necessary.

Data Visualization Basics

Mon., Dec. 4, 12:00 pm – 1:00 pm

Having performed text data analyses, you must now present your findings visually. Should you use a pie chart (rarely), a scatter plot, or a heat graph? Understand how to present your work in an accurate and ethical manner by joining us for an overview of best practices in data visualization. We will examine some visualization methods and how best to apply them to different kinds of data.

Register through the Events page at libraries.catholic.edu (CUA members only) or by contacting Kevin Gunn (gunn@cua.edu). Unless otherwise indicated, the instructor for each session will be Kevin Gunn, Coordinator of

Digital Scholarship. All workshops will take place on Zoom, recorded, and made available on the Catholic University Libraries' [YouTube Channel](#).



University Libraries

<https://libraries.catholic.edu/services/digital-scholarship/index.html>