



CONSTELLATE

# Teach text & data analysis with Constellate

Prepare your students for a data-driven future



# Who I am



Amy Kirchhoff  
Constellate, Sr. Manager

[Amy.Kirchhoff@ithaka.org](mailto:Amy.Kirchhoff@ithaka.org)

# Who we are

Constellate is part of ITHAKA's portfolio of non-profit services, along with trusted resources like JSTOR and Portico. ITHAKA's services are all aligned around a shared mission to improve access to knowledge for people around the world as affordably and sustainably as possible.



JSTOR



CONSTELLATE



ARTSTOR



PORTICO



ITHAKA S+R

## Why are text & data analysis skills important?



Nearly 97% of businesses said their companies now use data analytics

Job openings for data analysts are expected to grow by 23% between 2021 and 2031

Big data analytics is in the top 37 technologies and initiatives considered strategic to business intelligence

Almost 90% of organizations say success depends on data-driven decisions made by frontline employees

# ITHAKA's journey to Constellate

**We believe that data skills are essential for success in education and employment.**

So, we designed a solution with a range of personal and technological components to help instructors create, sustain and scale text analysis teaching in any discipline.

# Core tenets of Constellate

## Teach

- Train the trainer model
- Cloud based—no local configuration and Constellate scales to classes of any size
- Ready-to-use curriculum
- Create, save, and share your own lessons

## Learn

- Clear learning paths
- On-demand tutorials
- Live classes taught by experts
- Instant visualization of search results - refine ideas and form research questions
- Supportive community of users

## Perform

- Build and download custom data sets from JSTOR, Portico and others for free
- Refine and analyze data with popular, open tools
- Explore your datasets with a variety of visualizations

# Explore Constellate

The screenshot shows the Constellate platform interface. At the top, there's a blue header bar with the Constellate logo and navigation links for "Classes & Tutorials", "Builder", and "Help". Below the header, the main content area has a title "Run your own notebook in Constellate's notebook environment" and a sub-section "Tutorial and FAQ on Importing GitHub repositories to the Constellate Lab". It includes fields for "URL for the GitHub repository" (https://github.com/amykirchoff/text-analysis) and "Notebook to open (optional)" (exploring-word-frequencies.decade.ipynb). There's also a "Branch (optional)" field. Below these, there's a "Constellate Lab URL" section with a "Launch in Constellate" button and a note about adding a badge to the repository's README.md file.

The screenshot shows a Jupyter Notebook tutorial titled "Getting Started with Jupyter Notebooks". The page includes a "Description" section, "Use Case" (For Learners), "Difficulty" (Beginner), "Completion time" (15 minutes), "Knowledge Required" (None), "Knowledge Recommended" (None), "Data Format" (None), and "Libraries Used" (time). Below the text, there's a screenshot of a Jupyter Notebook interface with one cell labeled "1 Example Markdown Cell". At the top right, there are "Open in Constellate Lab" and "View on GitHub" buttons.

The screenshot shows the Constellate dashboard. It features a "Dashboard" section with a message about recent activity, a "Constellate Lab" section with a "Lab session: running" status, an "All datasets" section listing "Foster Care in Secondary 2022", and a "Pinned" section with a "Word Frequencies" notebook. A modal window titled "Analyze your dataset in the Constellate Lab" is open, providing instructions for analyzing the "Foster Care in Secondary Lit 1970-2022" dataset using Python, mentioning that it uses a 1,000 item sample of the dataset. It also notes that larger datasets may require more time to process.

## Teach

Start from your own learning materials

## Learn

Start from Constellate learning materials

## Perform

Start from a Constellate dataset



# Teaching Examples

## Hear from faculty:

“In my Digital History class, I include the *Introduction to Constellate* as part of my modules on text analysis. The Builder function is very handy for showing students the basics of text analysis, like n-grams and other word frequency-based visualization techniques.

**Constellate is a very good entry point for students** to start learning Python and get hands-on experience with more advanced methodologies, like topic modeling, when they don’t necessarily have the skills to code themselves.”

Keyao “Kyle” Pan  
Assistant Professor of Digital History  
Florida International University

# The Student Experience

## In the Library

UNIVERSITY OF TORONTO LIBRARIES

WORKSHOPS

Help With... Find and Borrow Visit Us About Us

University of Toronto / LibCal / Tri-campus workshops

### Constellate: A new platform to learn text analysis [Online](#)

Are you interested in analyzing large bodies of texts, or teaching text analysis? Learn about [Constellate](#), a new tool for text analysis. Constellate provides tools to dataset, using items from JSTOR and other collections, cleaning it, analyzing it and visualizing the results. For advanced users, or those wishing to learn Python offers customizable Python scripts along with hands-on tutorials. This session Constellate, show how to create a dataset and visualizations, and then will touch advanced options.

For more information about Constellate, including instructions on how to access our guide to Constellate.

Also, note that Constellate will be offering three, week long classes:

- Introduction to Python (November 2-5)
- Opening, Writing, and Cleaning Data with Python (November 30-December 1)
- Text Analysis for Humanists (December 14-17)

A link to join the event online will be sent to registrants.

Date: Tuesday, November 30, 2021

Time: 11:00am - 12:00pm

Time Zone: Eastern Time - US & Canada ([change](#))

Campus: St. George (Downtown) Campus

Online: This is an online event. Event URL will be sent via registration email.

Categories: ■ Data & Statistics ■ Digital Tools ■ Programming & Software

Northeastern University Library

Northeastern University / Subject Guides / Departments and Services / Data Management for Research / Workshops and training

### Data Management for Research: Workshops and training

Search this Guide

Managing your data Data management plans

NIH data management & sharing plans as of January 2023

Preparing data for deposit

Workshops and training L2 data access

Upcoming workshop

#### Python and Text Analysis for Absolute Beginners

- Thursday, January 5, 2023
- 1-3 pm Eastern
- Online via Zoom

Pondering Python? Tantalized by text analysis? Wondering how Jupyter notebooks work? Read on!

In this hands-on session, attendees will learn some basic Python while working in Jupyter notebooks, an interactive web tool for running and writing about code. Next, we'll use Python and Jupyter to run a simple text analysis on a custom dataset built with Constellate.

This event is co-sponsored by The NULab for Texts, Maps, and Networks (CSSH) and Research Data Services (Northeastern University Library).

[Learn more & register](#)



## Events Calendar

Campus Life Arts & Culture Academics Athletics Recreation & Wellness Community

Florida International University / Event Details



### Text Analysis with Python for Absolute Beginners

○ Wednesday, November 16  
More dates through November 18, 2022

Virtual Event

November 16th and 18th 2022, 11am-12pm

Come learn basic Python coding skills through JSTOR's [Constellate](#) text analysis platform. We will utilize Jupyter Notebooks (a web-based coding tool) to analyze textual data contained in document collections such as JSTOR, Portico, Chronicling America, Reveal Digital, and more. This hands-on workshop will be divided into two sessions. Part one (Nov. 16) will introduce Constellate's website and dataset builder. Part two (Nov. 18) will introduce basic Python and Jupyter Notebook skills to perform analysis on our previously built Constellate datasets. No previous experience with coding or text analysis is assumed or required!

#### About the presenter:

Molly Castro is the Digital Humanities Librarian at FIU, where she manages the Digital Scholar Studio at the Green Library and provides instruction and consultation on digital humanities and digital scholarship. Her research interests include computational text analysis, cultural analytics, and critical digital literacy.

#### Dial-In Information

Register via Zoom @ [go.fiu.edu/constellate](http://go.fiu.edu/constellate)

#### EVENT TYPE

#### DEPARTMENT

Academics, Workshops & webinars

Digital Scholar Studio, Libraries

#### AUDIENCE

#### ADD TO CALENDAR

# Getting started



# Accommodating fee structure



Classification	Fee
Very Large	\$15,000
Large	\$10,000
Medium	\$7,500
Small	\$5,000
Very Small	\$2,000

20% discount on the first year of participation  
10% discount on the second year



CONSTELLATE

[constellate.org](http://constellate.org)

Amy Kirchhoff  
[amy.kirchhoff@ithaka.org](mailto:amy.kirchhoff@ithaka.org)

