

# Constellate: A New Platform for Text Analysis

## Build your dataset

Search for keywords to begin analyzing and visualizing text data





Eric Clarke et al, Royal College of Surgeons in Ireland. Graphical visualisation of Tweets containing #breastcancer. Image and data from Wellcome

Nick Field, Reference Specialist and  
Kelly Schultz, Data Visualization Librarian







# What is Constellate?

## About

-  [About](#)
-  [Providers to Analyze](#)
-  [Participate and Beta Evaluation Period](#)
-  [What is Constellate?](#)
-  [Participating Institutions](#)
-  [Differences from JSTOR's Data for Research \(DfR\)](#)

[See all About \(6\) →](#)

## Why Learn Text Analysis?



-  [Why Learn Text Analysis? \(for digital humanities\)](#)
-  [What Text Analysis Methods Should I Learn? \(for digital humanities\)](#)
-  [Text Analysis Glossary](#)
-  [What is a Jupyter Notebook?](#)
-  [Why Learn Text Analysis? \(for business and data science\)](#)
-  [What Text Analysis Methods Should I Learn? \(for business and data science\)](#)

[See all Why Learn Text Analysis? \(6\) →](#)

## Data Sources

-  [Dataset Builder File Types](#)
-  [Can I download a dataset I created in the builder?](#)

## For Beta Participants

-  [Build Your Own Class and Jupyter Book - April 2021](#)
-  [Introduction to Text Analytics - May 2021](#)

## About

[About](#)[Providers to Analyze](#)[Participate and Beta Evaluation  
Period](#)[What is Constellate?](#)[Participating Institutions](#)[Differences from JSTOR's Data for  
Research \(DfR\)](#)[Why Learn Text Analysis?](#)[Data Sources](#)[For Beta Participants](#)[Technical Details](#)[Teaching with Constellate](#)

# Providers to Analyze

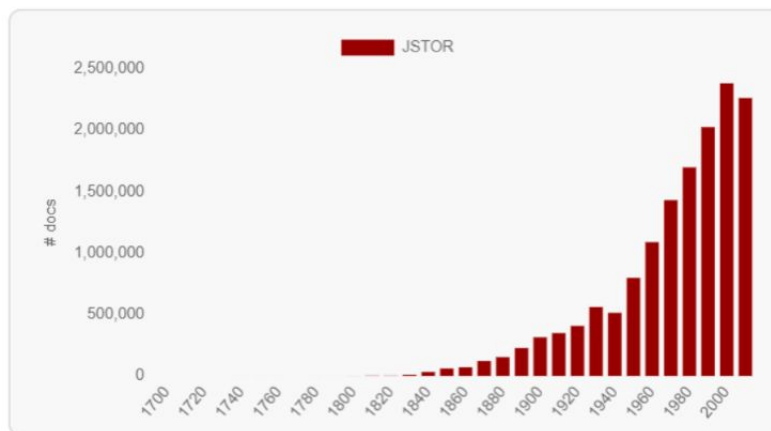
## JSTOR

[JSTOR](#) is a digital library for the intellectually curious. We help everyone discover, share, and connect valuable ideas.

**Content Type:** journal articles, book chapters, research reports

**Size:** over 14 million documents

**Document Publication Date Distribution:**



Distribution of JSTOR Content in Constellate by Year from 1700-Current





# What is Text Analysis?



Photo by dinakimm at Morguefile.com



# Beta Institutional Access



Photo by davidpwhelan at Morguefile.com

# Training Materials



Photo: "Media Studio class - large group" by college.library on flickr



# Preprocessing and Cleaning



Photo: "Woman in Gray Dress Standing on Brown Wooden Floor" by cottonbro on pexels

# Stop Words



Photo: "Traffic Signs Stop Road" by Alexas\_Fotos on pixabay



# Word Frequencies

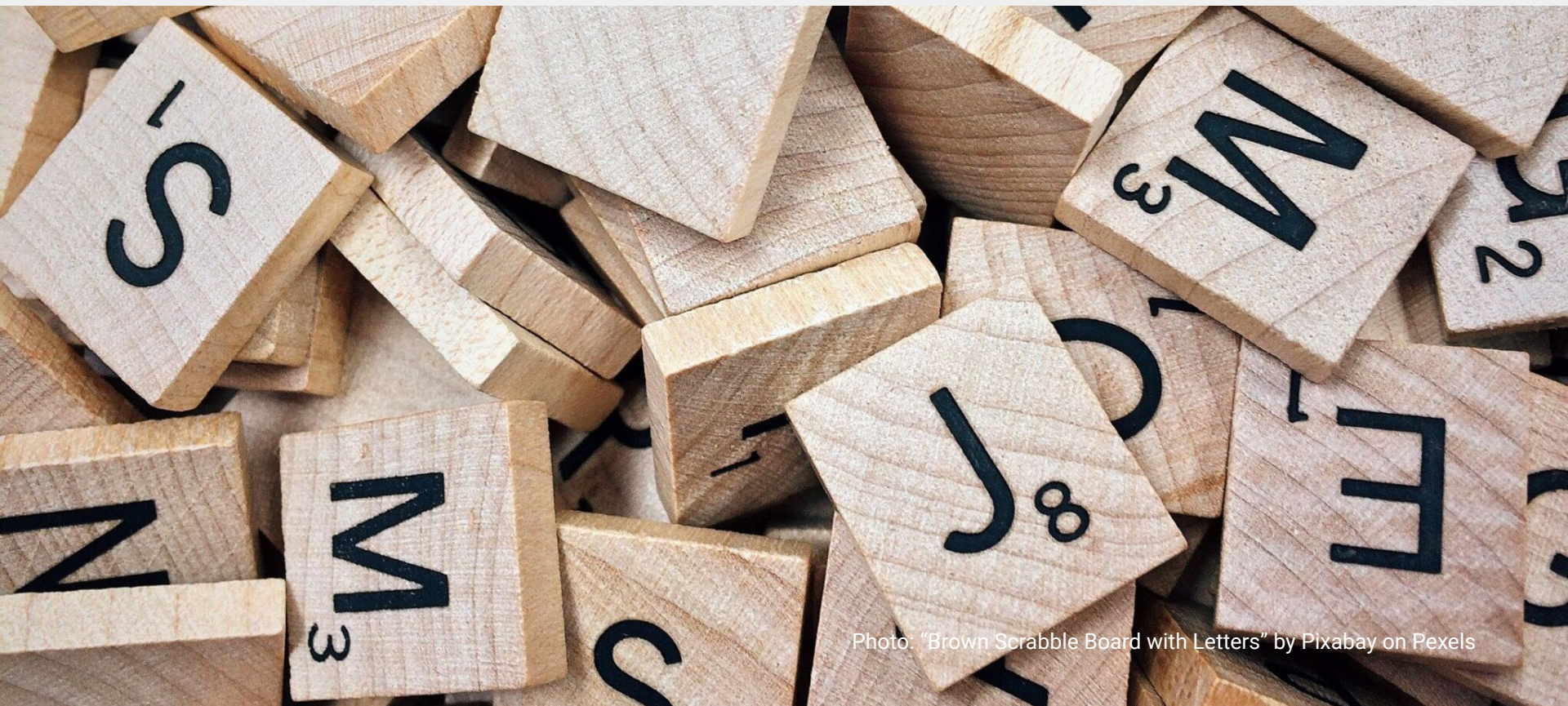


Photo: "Brown Scrabble Board with Letters" by Pixabay on Pexels

# Where to go for more help?

- Constellate Office Hours
  - <https://docs.constellate.org/office-hours-for-beta-participants-2/>
- Constellate Courses
  - [Build Your Own Class and Jupyter Book - July 2021](#)
  - [Introduction to Text Analytics July/August 2021](#)
  - [Recordings](#)
- MDL Help
  - [mdl@library.utoronto.ca](mailto:mdl@library.utoronto.ca)
  - [Help request form](#)



# Feedback Form

- Part of being in Beta
- [Feedback form for all users](#)
- [Feedback form for instructors](#)
- Email Constellate: [tdm@ithaka.org](mailto:tdm@ithaka.org)

# Other Bits and Bytes Workshops

<https://mdl.library.utoronto.ca/support/workshops-training/bits-and-bytes>



# Questions?



Photo: "Question Mark Pile" by qimono on Pixabay