

JORDAN PRUETT

Ph.D Candidate

✉ jpruett@uchicago.edu

☎ 5756448505

📍 Tacoma, WA

🌐 jordanpruett.github.io

EDUCATION

Certificate

Machine Learning

University of Washington

📅 Oct 2021 - June 2022

Ph.D (abd)

Digital Humanities and English
Literature

University of Chicago

📅 2016 - current

Bachelor of Arts

English

University of Utah

SKILLS

Data scripting

Python

SQL

R (1 semester)

DS/ML packages

pandas

scikit-learn

spaCy, flair, matplotlib

RELEVANT COURSES

Statistics

Data Analysis in R

Introduction to Machine Learning

Natural Language Processing

WORK EXPERIENCE

Data Management Intern

Post45 Data Collective

📅 June 2021 - September 2021

📍 Remote

- Implemented Python scripts for extracting text from hundreds of screenplays as a preparatory step for computational analysis
- Automated data cleaning and entity resolution tasks to link and prepare 5 datasets for publication

Research Assistant

Textual Optics Lab and the ARTFL Project

📅 Feb 2020 - current

📍 University of Chicago

- Designed a scikit-learn model that could correctly classify fiction genres with 85-90% accuracy for best-represented classes
- Used spaCy NER functionality as a preprocessing step to prevent overfitting, increasing recall for underrepresented classes by 6-7%
- Designed Python scripts to extract, encode, and save text data from over 60,000 scans of books from the British library

Instructor

University of Chicago

📅 Aug 2018 - March 2021

- Taught courses on computational text analysis, literature, and games

Game Design Intern

University of Chicago

📅 June 2017 - August 2017

- Designed puzzles for an alternative reality game for undergraduate orientation
- A profile of the game appeared in Wired magazine

PROJECTS

Cluster Analysis of 3 Million Book Reviews

- Used hierarchical clustering to discover which periodicals tend to review the same books in a dataset of over 3 million book reviews
- Implemented a custom LSTM sequence tagger in Flair for parsing records from semi-structured data
- Presented findings at the 2021 meeting of the Association for Computation in the Humanities (ACH)
- <https://github.com/jordanpruett/book-review-index>