

Chapman University

Chapman University Digital Commons

Library Presentations, Posters, and Videos

Leatherby Libraries

3-2-2023

Text and Data Mining Applications for Teaching Music Bibliography

Taylor Greene

Laurie Sampsel

Follow this and additional works at: https://digitalcommons.chapman.edu/library_presentations



Part of the [Data Science Commons](#), and the [Music Education Commons](#)

Text and Data Mining Applications for Teaching Music Bibliography

Comments

Presented at the [2023 Meeting of the Music Library Association and the Theatre Library Association](#).

The slides that can be downloaded above only cover Taylor Greene's portion of the presentation. However, the video link shows the full presentation with both presenters.

Text and Data Mining Applications for Teaching Music Bibliography

Taylor Greene, Chapman University
Laurie Sampsel, University of Colorado Boulder

MLA/TLA Joint Meeting, 2023

Introduction

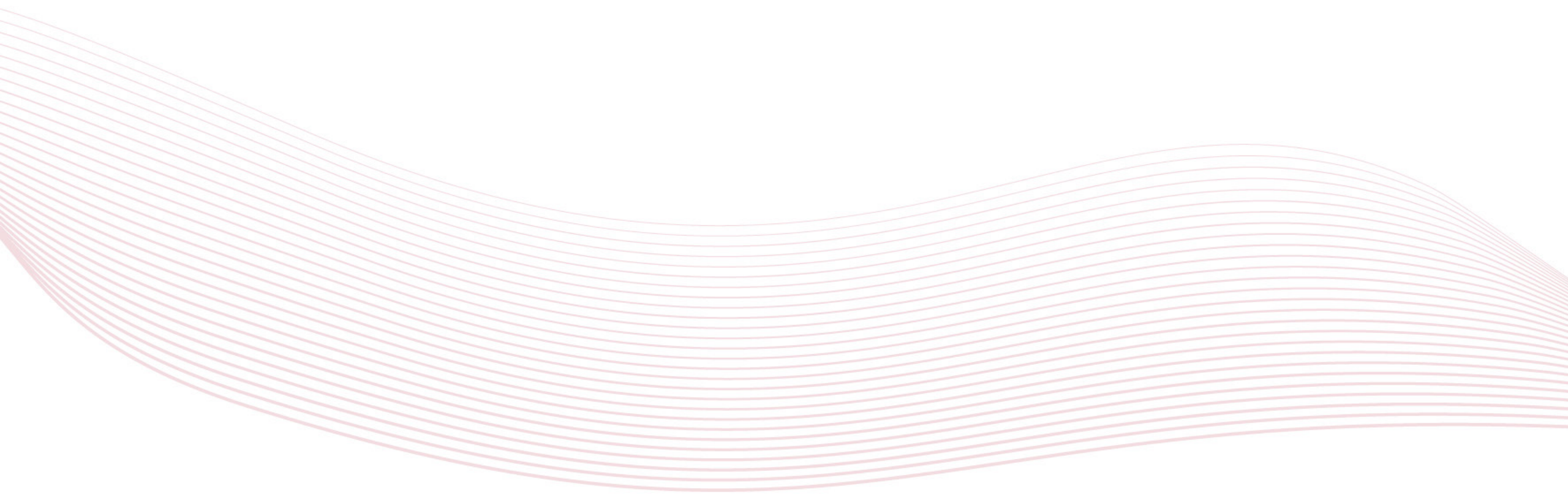
- Two approaches of incorporating text and data mining into music graduate courses:
 1. Using Google Ngram Viewer and JSTOR Text Analyzer in a large graduate music bibliography course at an R1 public institution.
 2. Using ProQuest's TDM Studio in a small Research Methods for Music Performers seminar at a medium-sized R2 institution

What is Text and Data Mining?

- An automated process of analyzing a large amount of unstructured text to uncover hidden patterns and relationships within.

Why TDM?

- Benefits of including TDM in Music Bibliography course:
 - Gaining familiarity with digital humanities as an approach to music research
 - Interdisciplinary research potential
 - Increasing students' ease and engagement with finding research information
 - Discovery of new research pathways for music research subjects



CHAPMAN
UNIVERSITY

Using TDM Studio in Research Methods for Music Performers

Background

- Chapman University
 - R2 status since 2019, new emphasis on increasing research activity
- Leatherby Libraries
 - New focus on supporting data services, data management
- Graduate course: “Research Methods for Music Performers”
 - Small class, all studying Keyboard Collaborative Arts
 - Introduces performers to the many types and applications of research for music
 - Students were completely unfamiliar with TDM and digital humanities

The Resources

- TDM Studio
 - Visualization Dashboard
 - Primarily focus on topic modeling
 - Texts for corpus: ProQuest Dissertations and Theses Global

What is Topic Modeling?

- Topic modeling creates a “bag of words” from the corpus of texts
- The algorithm ingests and analyzes a large amount of text and find themes (topics) based on word relationships.
- Goal = to discover ‘latent’ or hidden topics within a collection of documents

In other words

- Topic modeling allows us to discover meaning, patterns, and new areas to explore in our research
- Without having to read through all of the documents
- And to do this with an extremely large number of documents, even corpora that are too large to ever be read by one person.
- Note:
 - The program has no idea what the “topics” are about that it creates, just that they go together
 - It simply groups words into distinct topics based on how often they occur together
 - It finds when words commonly occur together in distinction to other words.

How is topic modeling useful for the students?

- Discoverability
 - By finding “topics” from the “bag of words” within the texts, topic modeling can point you toward new sources with similar topics
 - Or uncover new topic areas that you might not have thought to explore
- Pattern discernment
 - Topic modeling allows you to see underlying patterns within a corpus of texts
 - Which can give you new insight into:
 - How different ideas are related (consciously or unconsciously) within the literature
 - How coverage of a research topic has changed over time

TDM Studio assignment

- Part 1 - Experiment with Geographic Analysis and Sentiment Analysis
- Part 2 – Topic Modeling analysis using dissertations:
 - A) explore topics, look at patterns
 - B) Select a topic – what is underlying narrative focus, How has research shifted over time?, explore associated dissertations
- Part 3 – Reflection on experience using TDM Studio and ways that text data mining could be useful to performers or to music scholars

Guitar Repertoire

Jan 01, 1953 to
Dec 31, 2022

"guitar repertoire" OR "classical
guitar" OR (guitar AND "art music")
NOT (biology OR mathematics)

6094

ProQuest Dissertations and
Theses



Geographic Analysis



Topic Modeling



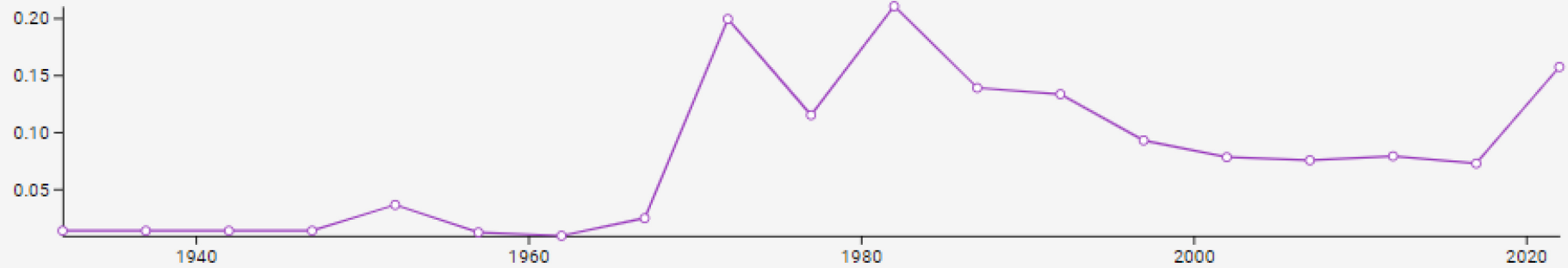
Sentiment Analysis

Number of Topics: 25 topics



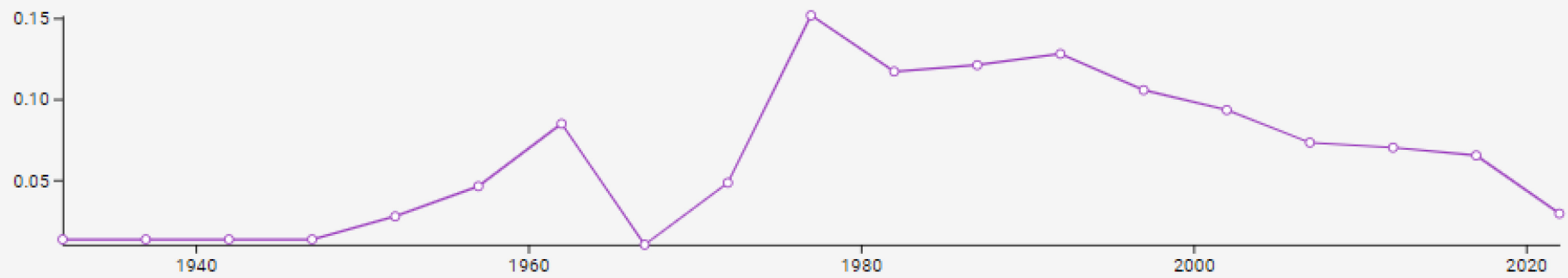
TOPIC 1

students, music, study, school, teachers, education, teaching, participants, data, student...



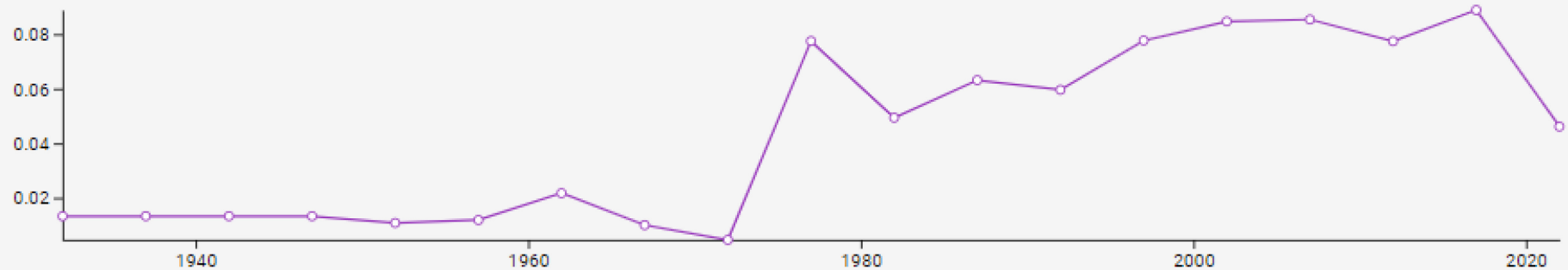
TOPIC 2

works, piano, music, study, songs, style, composer, composers, folk, musical...



TOPIC 3

music, works, composers, jazz, compositions, orchestra, composer, new, work, musical...



The Results

“it was helpful to see how the data mining works on an interactive level rather than reading an article about it. ... If I need to find a specific dissertation [topic], the topic modeling feature would be extremely helpful to filter out irrelevant topics.”

“Overall, this style of research seems to show much promise in the world of musicological academics. ... [Searching] for stronger academic materials conveniently will give us the option to create more convincing arguments having consulted literally everything.”

“In terms of usefulness, this program... ties in other sources to the topic that might not have ever been considered or discovered by the researcher in the first place. All in all, it allows for a more involved and thorough research experience, which will ultimately result in a better final product.”

Recommendations

- Think of ways to introduce performing arts students to TDM
- Work with faculty to develop TDM assignments in classes on research methods, senior seminars, or other research with openness to TDM methods
- Demonstrate potential of TDM and ease of tools like TDM Studio to demystify digital humanities

Next Steps

- Continue to develop and assess students in 2023/2024 year
- Explore Constellate – new TDM tool that analyzes datasets created from content in JSTOR and other open sources.

Thank you!

Taylor Greene

tgreene@chapman.edu

