



FACULTY LEARNING COMMUNITY PEW FTLC WINTER 2017

Our Learning Community will focus on text mining, the analysis of quantifiable **features of text corpora**. Examples of text mining include the automated

- discovery of authorship in anonymous and disputed literary texts, through letter and word frequency analysis
- modeling of thematic topics in official records, such as testimony in national parliaments or political party manifestos
- parsing of social media posts to explore correlates of comment sentiment
- characterization of news media framing of international events

CO-FACILITATORS:

Whitt Kilburn, Department of Political Science, and Jerry Scripps, School of Computing and Information Systems. Questions? Feel free to contact both facilitators at kilburnw@gvsu.edu and scripps@j@gvsu.edu.

TEXT AS DATA

Our goals will be to learn about applications of text mining across disciplines, explore freely available software tools for text mining, and evaluate on campus needs to support text mining in our scholarship and teaching. Participants will be given an introductory text on the subject, *Text Mining* by Ignatow and Mihalcea (2016, Sage Publications), as a basis for discussion.

No prior experience with text mining is required or expected, only an interest in the subject and getting to know like minded colleagues around the university. The facilitators will provide participants with sample text datasets for analysis in meetings planned around a particular technique of text mining. Participants with prior experience in text mining will be encouraged to share their expertise with fellow participants.

Some prior familiarity with data analysis software, including the R statistical programming environment, is helpful, though not required. Introductory courses on using R will be held through the Data Inquiry Lab, scheduled at <http://www.gvsu.edu/datainquirylab>.

Learning Community meetings will be held in person on the Allendale campus, four times over the Winter semester: Thursdays from 11am to 1pm on Feb 9, March 2, March 23, and April 13.