

Joseph Jegier

May 2023

Library Research Award Essay

In the Summer of 2022, I began an ambitious senior thesis: a case study considering the concept of discernment as it manifests at the University of Notre Dame. The resources available at the Hesburgh Libraries, particularly in the Navari Family Center for Digital Scholarship, helped me bring this research project to fruition. I am especially grateful for the Digital Initiatives Librarian, Eric Lease Morgan, who provided essential data analysis together with helpful guidance along the way.

I first met Eric Leas Morgan in as a first-year student, at a workshop on data mining and natural language processing in the Navari Family CDS. In this workshop, Eric described how to conduct “data science with words,” showing how students like me as well as faculty and staff could use the Distant Reader to analyze large sets of texts. The Distant Reader is a powerful tool, created by Eric, that can analyze texts in several ways including topic modelling, a machine learning process that identifies latent themes, concordance, a search function that finds words in their contexts, and word embedding, a function that computes the distances between words in a lexicon. When I first encountered Eric and his Distant Reader, I could have never foreseen using him and his expertise again, three years later, for my senior thesis.

This workshop I attended in 2019 was a fantastic introduction to the Distant Reader and to the world of text mining and natural language processing in general. Eric walked us through the process of using the Distant Reader to analyze texts, using the Odyssey and the Iliad as an example. He passionately presented the various functions of the Distant Reader. He answered questions with simplicity and clarity. Most importantly, Eric made us feel

capable of conducting the analysis ourselves. When I left the workshop, I did not feel like a beginner, but as a researcher ready to use this tool and apply it to my academic career.

I first used the Distant Reader to write an essay for a class, analyzing political speeches in post-war Germany. Three years later, I had the idea of using the Distant Reader again, but this time for my senior thesis. As I was studying discernment in the context of Notre Dame, I was already pursuing traditional methods of data collection such as interviews and surveys, but I wanted to analyze a set of set of student reflections from the Moreau First Year Experience (MYFE) to see how students conceptualized discernment in their writings. I knew that this data would be a creative and informative addition to the research.

I wrote an email to Eric in January of 2023, explaining my idea. Before even getting to campus, he helped me download the software and begin the process. During the semester, Eric and I met roughly once a week in the Navari Center to work on the research. We started by articulating a research question – “In a set of first-year student reflections, what latent themes arise, and what is the context surrounding discernment?” After articulating the research question, we converted and anonymized the collected the texts. The resulting corpus contained over 1,400 files with a total of roughly 2 million words. Eric and I then used several processes to analyze this corpus of student writings. We counted and tabulated the features (pronouns, unigrams, bigrams, etc.). We searched and sorted by the word ‘discernment’ and related words. We tabulated bigrams (two-word phrases) that contained a form of the word ‘discernment,’ resulting in the word cloud seen below.

Figure 1

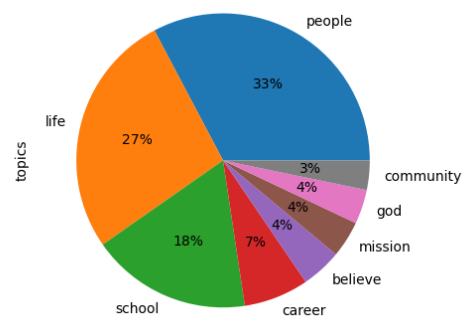
Bigram Word Cloud



We also used topic modelling, a machine learning process, to reveal the most common subjects of first-year student reflections. A pie chart of these themes can be seen to the right. Finally, we used a technique called word embedding to create a table and plot of semantic distances. These functions not only helped me address my research questions, but the resulting graphics also served as useful figures in my final thesis and presentation. The analysis revealed that students related discernment primarily to their careers but also to their lives, more generally speaking. Moreover, the most salient aspects of first-year student experiences seem to be relationships with others (interestingly, “mom” was the most commonly mentioned person related to discernment). This result lends credibility to the claim I make in my thesis, that meaningful relationships are central to effective discernment.

Figure 2

Topic Model Pie Chart



In addition to the services provided by Eric Lease Morgan in the Navari Family Center for Digital Scholarship, several other aspects of the Hesburgh Library played a key role in the successful completion of my senior thesis. First, I used databases such as EBSCO to find relevant literature, and through the library I also had access to digital and print media. Next, I visited the physical spaces in the library for reading and writing. Finally, I used the study rooms as a quiet place to conduct student interviews.

Throughout my entire academic career at Notre Dame, the Hesburgh Library has been monumental. Outside of my residence hall, I have spent more time in the library than in any

other place. While writing my senior thesis, I have come to appreciate the library's staff and services even more. The library's resources, collections, spaces (particularly the Navari Family Center for Digital Scholarship), and the staff (especially Eric Lease Morgan), make the Hesburgh Library a truly remarkable place. I am grateful for all that the Hesburgh Library has done to enhance my intellectual development, and I am particularly appreciative of the role it has played in creating an informative senior thesis.