NEH Grant: Analyzing the News Sources Current Politicians Follow on Twitter

Claudia Vasquez

List of Participants:

Claudia Vasquez
- project director
- first-year student at the University of Richmond

Dr. Lauren Tilton
- project advisor
- professor of digital humanities at the University of Richmond

History of the Project:

This project began as a very simple networking assignment. As so much of our news today comes from social media, I was interested in investigating how politicians gather their news on twitter- which news sources they follow and which politicians follow the same sources. I began by making lists of current politicians around the world, and proceeded to search the names on twitter. If the person did not have an account, I crossed off their name and moved down the list. If they did have an account, I would scroll through the users they followed and search for news sources. If I found more than one news source on their following list, I wrote the sources down by their name. It was a tedious process, as it was all done manually. Next I imputed all of the information on to a google sheets page. The columns were labeled name and year. I created a network to analyze the data. The results were not what I was expecting, but I attributed them to the errors of the experiment. I believe that if the project is improved and expanded, I will find more sufficient results.

Enhancing the Humanities:

Using the methods of networking, web scraping, and GIS mapping, my level II grant proposal will investigate the correlation between news sources and current politicians around the world. Networks are “collections of points joined together in pairs by lines” (Newman, 2010). In my project the points will consist of the names of politicians and the news sources they follow. The lines will show which news sources are followed by each politician. Web scraping is “a technique employed to extract large amounts of data from websites whereby the data is extracted and saved to a local file in your computer or to a database in a table (spreadsheet) format,” (Webharvy). This process will allow me to analyze the accounts followed by each politician in a more efficient manner. I will be able to easily search for news titles, as opposed to scrolling through their following list, which is both inefficient and error-prone. GIS (geographic information system) mapping is “a system designed to capture, store, manipulate, analyze, manage, and present all types of geographical data,” (University of Wisconsin). This technique is a new element I wanted to add to my project. By including a map that will display where each politician and news source is located, it will be easier to visually make connections between and draw conclusions about the sources and politicians. I want to determine if there is a correlation between location and political view, as well as location and followed news sources. For instance, do politicians from the United States follow less foreign news sources than politicians from other countries? Perhaps politicians who favor policies that expand immigration rights follow more news accounts from different countries besides their own. Additionally, I want to know whether or not news sources could be influencing politicians’ decisions. I anticipate that politicians from European countries will follow the most diverse news sources. I am also anticipating a direct correlation between political views and political leaning of news sources. Specifically, I believe that most conservative politicians will follow similarly conservative-based

news sources. I believe this experiment will be enlightening to people on the motives and behaviors of today's current politicians. Though this project is mostly intended for those old enough to vote, it can also benefit future voters. It is designed to give individuals an understanding of where politicians are acquiring their news. Further, it should lead them to question the motivations and habits of our current world leaders. Are they attempting to diversify themselves by following a variety of news sources, with difference views from their own and from foreign countries? Does the news that they receive contain bias from a particular political standpoint? This project could also be informative and helpful for anyone studying politics or psychology. The accounts that politicians follow could reveal trends of politics or underlying psychiatrics. The project also brings new light to the field of Digital Humanities. Kathleen Fitzpatrick describes the current state of the field as resting "in that creative tension between those who've been in the field for a long time and those who are coming to it today, between disciplinarity and interdisciplinarity, between making and interpreting, between the field's history and its future," (Fitzpatrick, 2012). The pairing of myself, a new digital humanist, and Dr. Tilton, an experienced person in the field, will surely bring a sense to the current state of digital humanities.

Environmental Scan:

Though there are no current projects that I was able to find that directly resemble my experiment, there are several that use similar techniques or are investigating similar data. The project "Mapping the Media Landscape in Old Regime France" by Elizabeth Bond explores letters that were published in the newspapers in France from 1770-1788 (Bond). In the eighteenth century, French citizens would write to their local newspapers about their opinions on topics introduced in various periodicals. Their citations of other newspapers demonstrated how and where the French were getting their news. My project is very similar in its goal of locating and identifying sources of news for politicians. Additionally, both projects use networking as a method. Whereas my project shows the connection of politicians and the news sources they follow, Bond's project shows which newspapers had citations of other newspapers. Both networks depict the spreading and sharing of information. Another similar project called "Linked Jazz" is a research project attempting to discover the relationships between musicians and their communities. The project involved creating tools- "a transcript analyzer, a name mapping and curator tool, and a crowdsourcing tool"- in order to analyze the transcribed interviews and decipher the names, questions, and answers mentioned (Linked Jazz, 2011). My project similarly is identifying connections through a technology-based method. Though, the two project vary greatly in their goals and predicted outcomes, both use networking as a method to display the data. The Linked Jazz Network is a beautiful interactive map that includes information on and a picture of each musician, similar to my intent for each politician I include. Many of the musicians also have a linked recording that users can click on to listen to a sample of their music. I love this feature because it allows people to experience the project in a more personal way. This feature would not be applicable for my particular project, but it is a great source of inspiration. Another networking project called "Movie Galaxies" explores the interactions

between movie characters. The project currently includes networks for 775 movies but is still continuing to expand (Kirk, 2013).

Work Plan:

The project will mostly be conducted individually. I will begin with expanding on the project I have currently, as opposed to beginning from scratch. The number of politicians will be expanded, and there will be more information included on each individual. Additionally, I would like to find politicians from a larger variety of countries. In my first project, the politicians came from the United States, Canada, Australia, Argentina, Brazil, France, Portugal, and Norway. Though this data may seem diverse, it consisted largely of white males. I would like my politicians to be more diverse, including politicians of various races and genders. I also want to look further into each politician's background, birthplace, current occupation, implemented policies, and political party. Some of this information may seem superfluous but it may add meaning to the dataset and "even reveal shifts of emphasis of which we were entirely unaware" (Goldstone and Underwood, JDH, 2012). Each of these topics will be added as column labels in my google spreadsheet. I also will include a picture of each politician next to their information on the network. This information will be obtained manually through my own research and resources. Gathering this data may take several weeks, as I will be gathering it on my own time. Additionally, I will need to ensure that my information is factual, so I will have to review a variety of sources and check their credibility. I hope to only take three weeks for this particular part of the project. After I have the information on each politician, I will work on using web scraping to gather my twitter data. This area of the project will require assistance from Dr. Tilton. It is also where most of the funding will go, as we will need to purchase the software for web scraping. I, having less experience with coding and softwares, will rely on Dr. Tilton to assist me in the process of downloading their data from twitter onto excel. Essentially we will download a software that will allow us to copy the URL- the website code- of a twitter account and download the data from the page. From there we will be able to specify the data we want to search, in this case: followers. Once specified, we will be able to search for new sources directly from the profile's following list (urlprofiler). I will search each politicians page for as many news sources as I can find. Afterwards I will input this information onto the google spreadsheet to create my network. After downloading the spreadsheet as a cvs file on my computer, I can transfer it into R studio. This will create my network for me. After this step, I will move on to creating my GIS map. This map will help my audience to understand the significance of location in the project by "visually detecting spatial patterns that remain hidden in texts and tables," (Bodenhamer). For this process I will use the same spreadsheet to input into the website ArcGIS.com. This will display my politicians and news sources as points on the map. I will then create a heat map from the same website that will show where the majority of the politicians and news sources from my data are located. By seeing how the politicians and news sources are clustered on a world map, it may give my project an added meaning.

Final Product and Dissemination:

Considering that this project is focused on the spread of information through Twitter, I think it is only fitting that the project results are shared largely through social media. However, the project is aimed at a large audience, and needs to be available to everyone. Therefore, the project results will be disseminated through twitter, a website, and presentations. First, I will make a twitter account to advertise the project. As a tweet can only contain a maximum of 140 characters, the project details will be cut into short snippets of tweets, with the project as a whole linked in the biography section of the account. The tweets will be easy to comprehend, free of complicated terms and language. They will not only explain the project and its results, but will provide information on the presentations- where they will be located, how to sign up to attend, and what times they will begin. Though many individuals, from grandparents to middle schoolers, have twitter accounts, social media is not always preferred by or accessible to everyone. The project will also be published on a website. This website will include the GIS map, the network, and a written report of the project. It will have a listening feature, so that anyone who has visual impairments is able to access the information. There will be a feature to listen to the written report out loud. Additionally, there will be a video describing the project's procedures, goals, outcomes, and methods. All aspects of the project will also be written, so anyone who has hearing impairments will have no issue accessing this data. The presentations will be conducted by myself in schools, libraries, and any other available venue. The presentations will be available to anyone who wishes to learn about the project. There will be signup sheets available on the twitter account, but they will only be used to estimate the number of people attending. Before any presentation that is not in a school, flyers will be hung up in the area to inform the public of the upcoming event. The presentations will begin with a background of the project- an explanation of why I began this experiment- and will follow with an address to the audience- why does this pertain to you? I want to encourage voters and future voters to look into the habits of current politicians. This project specifically addresses politicians' relationships to news sources, but I want to encourage the audience to go further and investigate who they are voting for. After my appeal to the population, I will continue to explain my project results through a slideshow and tour of the website. I will explain the network and GIS map and what information can be gathered from those resources. After period of time for questions and answers, I will end the presentation by asking my audience to consider how digital humanities can be used in their lives- what questions do they have that could be answered using similar methods to the network and GIS map used in my own project.

Biographies:

Claudia Vasquez is a first year student at the University of Richmond. Born in Winter Park, Florida, she graduated from Winter Park High School in 2018. Claudia has taken an introduction to digital humanities course for the past semester, and is well versed in the fields of networking, GIS mapping, and text analysis. She is resourceful,responsible, and analytical. These traits will

help her to direct this project and make insightful observations and conclusions. Though a new digital humanist, Claudia will have the support of her mentor Dr. Tilton who will help her to reach her full potential in the field.

Dr. Tilton is a professor of Digital Humanities at the University of Richmond. Born and raised in Louisiana, Dr. Tilton has a fun and optimistic personality that make her the perfect mentor for this project. She is excellent at programming and very knowledgeable about technology and the field of Digital Humanities. Knowing the the ropes of the field of Digital Humanities will allow her to successfully guide and aid Claudia in the process of this experiment. Her patience and experience will provide the necessary support to ensure that this project reaches its full potential and is as error-free as possible.

Work Cited:

Dempsey, Caitlin. "Research Guides: Mapping and Geographic Information Systems (GIS): What Is GIS?" University of Wisconsin-Madison Libraries. 1999. Accessed December 11, 2018. https://researchguides.library.wisc.edu/GIS.

Bond, Elizabeth. "Mapping the Media Landscape in Old Regime France." Mapping the Media Landscape in Old Regime France – Current Research in Digital History. 2018. Accessed December 10, 2018. http://crdh.rrchnm.org/essays/v01-11-mapping-the-media-landscape-in-old-regime-france/.

SysNucleus. "What Is Web Scraping ?" Web Scraping Explained. Accessed December 12, 2018. https://www.webharvy.com/articles/what-is-web-scraping.html.

Newman, Mark. "Networks: An Introduction." *A Short Introduction to Networks and Why We Study Them*, September 2010. Accessed December 11, 2018. doi:10.1093/acprof:oso/9780199206650.001.0001.

Pattuelli, Cristina. *Linked Jazz*. Pratt Institute School. 2011. Accessed December 09, 2018. https://linkedjazz.org.

Hallaway, Patrick. "How to Scrape Data From Twitter Profile Pages." URL Profiler. January 08, 2016. Accessed December 10, 2018. https://urlprofiler.com/blog/scrape-twitter/.

Kirk, Andy. "'Movie Galaxies' Social Graphs in Movies." Visualising Data. August 29, 2014. Accessed December 10, 2018. http://www.visualisingdata.com/2013/07/movie-galaxies-social-graphs-in-movies/.

Bodenhamer, David, “Beyond GIS: Geospatial Technologies and the Future of History.” p. 1-13. History and GIS: Epistomologies, Considerations and Reflections. Springer, 2013. Accessed December 10, 2018.

Fitzpatrick, Kathleen. “The Humanities, Done Digitally.” Debates in the Digital Humanities. 2012. Accessed December 9, 2018.

Goldstone, Andrew and Ted Underwood. “What Can Topic Models of PMLA Teach Us About the History of Literary Scholarship?” Journal of Digital Humanities. Vol 2. No. 1 Winter 2012. Accessed December 12, 2018.