

UNIT 2: ORGANIZING AND GATHERING

Project #2: Weeks 5-9

Textual Analysis Data Collection

Lucy A. Johnson

In your syllabus, the description of the **Textual Analysis Data Collection** is as follows:

Now that you've situated yourself within the field of digital technology and culture as a user, the next step is to critically engage with digital tools to produce your own digital project. The focus of the textual analysis data collection is to allow you to organize and archive information that interests you within popular culture. By creating your own textual analysis project it will help you to not only understand how others are using analytics, but also to recognize their prevalence in hosts such as Facebook and Google, learning not only about the literature surrounding text analysis but also about the potential—and the limits—of surveillance analytics. In addition to conducting your textual analysis you will be required to present your findings to the class in a 5-minute presentation synthesizing your findings in juxtaposition with your initial hypothesis.

The Project at Length:

For this project, you will be asked to come up with a popular culture topic that is of interest to you (one in which you engage with personally). Remember, this topic needs to be one that is current. **Once you have a topic in mind you will need to come up with a scaffolded hypothesis concerning the prevalent conversations surrounding your topic and the issues you assume will be addressed within multiple online platforms.** You will select 5 keywords or concepts surrounding this topic that you feel will be displayed prominently within a variety of sources. After you have framed your prediction, you will then begin collecting your material. **It is required that you find at least 10 sources to inform your textual analysis interpretation using Voyant. 5 of these sources must be hand-selected (and justified) by you as to why you can expect to find your keywords within these spaces. You will also need to create an omission list of 8-10 disposable words that might skew your data.** After entering your data into [Voyant](#), you will need to engage with your findings critically, preparing a presentation for the class addressing how your hypothesis either informs or contests your *actual* findings as they are depicted within the textual analysis tool.

In other words, do you know more about your culture, or does your machine?

Detailed Directions:

LOOK/SEE: What are you engaged in within popular culture?

DESCRIBE

1. Establish your ethos: Why do you have some stake in this topic? How are you positioned to engage with this topic within a digital space?
 - a. Define the topic or issue in relation to current society and its digital implications (why would it live within a digital space. In what ways?)
2. Create your prediction about your topic: Based on your *personal* understanding of your topic, create a prediction concerning an overall slant or argument concerning the conversation.
 - a. In other words, what do you think will happen in regards to representation? why?
 - b. With this prediction in mind, form a **HYPOTHESIS** statement utilizing your keywords and prediction (this is where your **five keywords** come to play to help track your prediction).
3. **Justify** each of the 5 keywords in 100 words
4. **Order** your keywords (which word will be the most prevalent, why? Think of the nature of the conversation based on the time in which it is occurring, sources you choose, etc.)
5. Based on your topic, create an **8-10 word omission list**. These are words that do not identify as a particular concept (example: comment, news, election, show) as well as discourse-specific terms that may be popular to your topic, but aren't directly specific to what you're looking for (example: you're researching the immigration ban, but need to filter out Betsy DeVos).

ANALYZE

1. **Collect** 10 web sources that you wish to include within your analysis of your keywords (these sites must hold relevance to your topic, 5 of them need to be hand selected, the others educated guesses)
2. **Archive** these 10 sites in the same document as your hypotheses. (In other words, I want to see the URL's, I also want you to have record of them too)
3. **Justify** why these sites are appropriate to include within the corpus. (Have you had engagement with them?). **Reminder: 5 sources must be spaces you know. 5 can be informed selections that are justified.**

INTREPRET

1. Plug your URL's into Voyant
2. Turn your Stopword on, add your 8-10 omitted words.
3. Record the 5 keywords it shows.
4. Pick **one** of the top 5 keywords, view it within a different visualization tool (bubblelines, scatterplot, context, etc). What does this tool tell you that the word cloud doesn't? Why do you find that interesting?
5. Pick **one** smaller word that surprised you. Why are you surprised this word/concept showed up based on what you know about your topic?