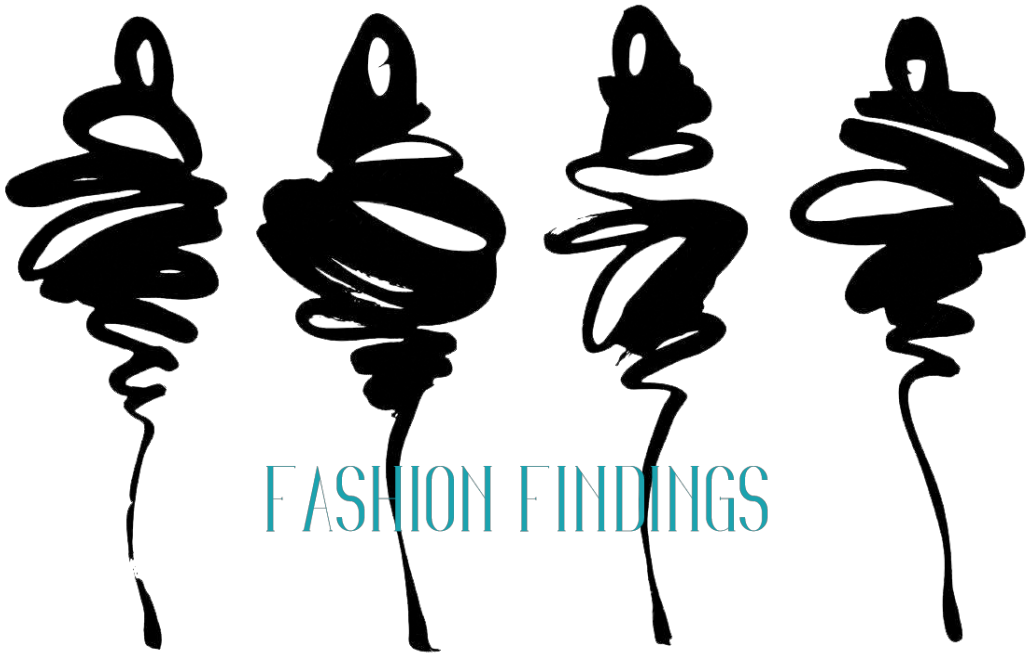
2018



NEH GRANT: FASHION FINDINGS

INTRODUCTION TO DIGITAL HUMANITIES

EMILY SONG

TABLE OF CONTENTS

List of Participants

Emily Song, University of Richmond
Matthew Jockers, Washington State University
David Blei, Princeton University
Google and Google Trends

Narrative

Enhancing the Humanities

Fashion Findings is a project seeking to explore Google search patterns over time as they relate to the fashion industry. This proposed project pursing a Level II Grant employs digital humanities methodologies including text analysis and topic modeling to uncover unique patterns related to clothing and fashion. Fashion Findings strives to add value to an industry that is often misconceived as superficial. There is a certain disdain towards this world seemingly steeped in material objects – an assumption that expensive clothes and the models who wear them can offer nothing beyond their physical attributes. For many, it is difficult to see past the materialistic culture that the fashion industry enables. Projects such as Fashion Findings can uncover the value behind this business. Diana Crane and Laura Bovine's study of the sociology of clothing states, "Fashion can be conceptualized as an example of a broader phenomenon, the creation and attribution of symbolic values to material culture" (Bovine and Crane, Science Direct, 2006). Studying the trends that drive the fashion industry can reveal invaluable information about the cultural and societal values behind the clothes themselves.

Fashion Findings uses data and results from Google searches linked to the industry. By tracking top searched for trends and key words related to fashion, this project aims to uncover patterns that reveal information about ideals in society. Cross examining this information with time period and place can lead to a broader understanding about what shapes the industry and why. This project seeks to answer questions that increase intellectual information in the humanities including: How do fashion related searches reflect social and cultural changes in society? How have these searches changed over time? What else occurred over these periods to potentially lead to these changes? Overall, the purpose of this project is to use techniques from the digital humanities to study cultural developments through the lens of fashion. Consequently, these results and findings have the potential to shift the perspectives of those who view it by revealing the value of a seemingly superficial world.

The essence of this technology driven society formed an environment in which literary scholars can study big data in the form of digital libraries. Data is increasingly accessible to those who desire to use it. Google search results and Google Trends have a plethora of information. Examining search patterns over time through text analysis and topic modeling creates a way to understand what this information means. However, using Google is limiting in itself. Google was created in 1998, meaning that this project is constricted in terms of time. The project will use search results from 2000 to 2018, in order to give it a basis of solid information. Because of this technological limit, the project does not analyze data from earlier time periods that could offer great insights into the way the fashion industry changed as certain movements evolved. An interesting point of study for the future or for another project would be to look at how words

such as “corset” have been used over time, and when their use declined in tandem with movements and calls for women’s rights. The search driven nature of Fashion Findings limits the project’s ability to examine such connections. However, this is an interesting point to potentially examine in the future as the project grows and expands.

This project requires various digital humanities methods including topic modeling and text analysis in order to concretely examine an abundance of search results. In 2008, computers revolutionized scientific research, and later literary research, in their ability to process large-scale data sets. As a result, new methodologies, such as topic modeling and text analysis, were developed. Using these methods allows literary scholars to gather evidence in order to develop potential hypotheses. However, this process is not identical to the scientific way of thinking. While both scientific and literary studies aim to reach the same goal of uncovering new information and intellectual knowledge, the difference lies in experimentation and arguably interpretation and opinion. (Jockers, University of Illinois, 2013).

According to David Blei, “Topic modeling provides a suite of algorithms to discover hidden thematic structure in large collections of texts. The results of topic modeling algorithms can be used to summarize, visualize, explore and theorize about a corpus” (Blei, JDH, 2012). From there, topic modeling uncovers a set of patterns to create topics relating the overarching themes found throughout the text. This method represents the text through categories to enhance understanding and enable interpretation. Topics are based on co-occurring words and the ways in which they are used together. Once topics are chosen, scholars can decide their weight based on overall importance and corresponding words or titles to represent the topics as a whole (Blei, JDH, 2012). However, topic modeling acts as a starting place for discovery and information. While the algorithm reveals patterns, the scholar must determine what they mean. Interpretation is a necessary element of uncovering evidence. “Topic modeling is not necessarily useful as evidence, but it makes an excellent tool for discovery” (Brett, JDH, 2013).

Text Analysis is a form of topic modeling. This method finds words that frequently appear together and gives insights into a piece of writing as a result. This word co-occurrence and clustering transforms free data into structured data, essentially creating “a continual process of coming to know by manipulating representation” (Weingart and Meeks, JDH, 2012). This process can lead to the discovery of unique and unexpected patterns that can be used to create an interface. In many ways, topic modeling is a preliminary step when it comes to research and intellectual growth, rather than developing a concrete conclusion. Topic modeling is one single element in the much larger process of learning.

Unlike scientific studies, literary studies do not have the advantage of exhaustive experimentation. Conclusions in this world cannot be traditionally tested, absolutely confirmed, or fully repeated. While Fashion Findings can identity trends based on the number of times something was searched for and pair those results with historical events and time periods, there is no way to concretely conclude that the two patterns are causal. Instead these processes and methodologies of studying text lead to understandings, which are invaluable in their own way. Interpretations can open up a realm of possibilities to other scholars and allow them to see text in a way that they did not previously consider (Jockers, University of Illinois, 2013). Results from Fashion Findings will give others a new perspective of the fashion industry and the way it relates to society at large.

However, interpretation is limiting. The nature of interpretation is driven by observation, and observation is inherently flawed. Observation is unavoidably bias because it is seen through the filter of an individual's own subjective subconscious. Fashion Findings has the potential to draw connections between trends in society and trends in the industry. Researchers involved in this project want these relationships to exist, introducing an element of bias. While these biases exist, big data and its accessibility changed the way the literary scholars can study and draw conclusions. "Instead of conducting controlled experiments on samples and then extrapolating from the specific to the general or from the close to the distant, these massive data sets are allowing for investigations at a scale that reaches or approaches a point of being comprehensive" (Jockers, University of Illinois, 2013). As a result, the focus of Fashion Findings is to uncover trends and offer potential connections, rather than present absolute answers.

Macroanalysis allows for contextualization in a way like never before. It leads to a better understanding of context and, as a result, gives researchers a better grasp of the text itself. This allows for knowledge of historical placement, growth and development over time, and unexpected patterns (Jockers, University of Illinois, 2013). Using this information for Fashion Findings allows for the discovery of cultural and societal implications. However, note that computers and the technology they harness for this kind of processing do not interpret or draw conclusions. This misconception can lead researchers astray when performing this kind of work. Instead, macroanalysis and the methodologies it uses uncover patterns. These trends can then be analyzed in further studies to develop new theories. The connections and relationships uncovered by Fashion Findings is a starting place for viewers to explore much larger conclusions. This project gives users the information they need to reimagine the fashion world by adding value and humanity to an industry that is often seen as lacking.

Environmental Scan

Fashion Findings offers new insights into the world of clothing and all that it entails. There are, however, existing projects in the digital humanities realm similar to this project. The Rhythm of Food was created by Truth & Beauty and Google News Lab. This project similarly uses Google search results to examine the top searched for foods and food related trends. From veganism, to apricot, to molecular gastronomy, this work tracks and analyzes search patterns throughout the years to learn more about the field of food. These trends can be viewed monthly or yearly or by keyword, all of which contribute to the rhythm of food around the world. These results vary by season and region and reveal peaks and valleys during certain times and special occasions. Users can also choose their own words and results to explore. Similar to Fashion Findings, the goal of The Rhythm of Food is to uncover hidden search patterns in order to better understand the various factors that play a role in the seasonality of different foods (The Rhythm of Food, 2018). One of the most apparent strengths of The Rhythm of Food is the platform on which it was built. The results of this project are displayed on an incredibly interactive and aesthetically pleasing website. The website itself engages viewers in the sense that they want to explore the page regardless of their initial interest in the project itself. This draws in users and creates a space within which they can unintentionally gain new knowledge. The Rhythm of Food is full of attention-grabbing graphics and displays data through charts and maps in a way that allows those examining the project to directly interact with its findings. Fashion Findings aims to visually attract viewers in the same manner with similar Google driven data so that individuals may better learn about all that the industry has to offer. The platform on which the proposed project is delivered is necessary to the success of the project itself. It must capture the beauty of the fashion

industry to please those who work in it and simultaneously give those uninterested or opposed something to be curious about.

Robots Reading Vogue, a project by Yale University Library, is similar to Fashion Findings in two ways. Overall, both projects aim to reveal the value in the fashion industry. Robots Reading Vogue collected data from 2,700 issues of vogue and captured and shared that data through various digital humanities methods. Their website contains multiple forms of visualization including histograms, topic modeling, n-gram searches, and more. This extensive project similarly offers new ways to see the value in the fashion industry by revealing how Vogue has changed over time. Their n-gram search feature tracks words such as pretty, sexy, lovely, and beautiful, and displays how their uses have increased or decreased over time. Comparing terms such as these allows viewers to draw connections between words and their meanings. It reveals the way beauty is discussed, and how those discussions evolved. Users can also search for specific words based on what they want to know more about (Robots Reading Vogue). Fashion Findings seeks to capture data on a topic similar to that of Robots Reading Vogue, but through Google search results and trends rather than a print-based source. The goal of the two projects is also comparable in terms of the information about the industry they intend to reveal. Aside from the source of data, the main difference between Robots Reading Vogue and Fashion Findings is in their delivery. As stated above, an essential element of Fashion Findings is the platform on which it is created. While Robots Reading Vogue is interactive, it fails to captivate audiences in the way that The Rhythm of Food does so beautifully. As a result, Fashion Findings is essentially a combination of the two projects – one that merges topic modeling and text analysis to explore the Fashion Industry in a way that is fascinating for all who see it.

History of the Project

The idea for Fashion Findings was derived from a much larger study of the digital humanities field as a whole. This understanding and previous work with methods of topic modeling and text analysis is necessary to the success of the project. Previous research includes an in-depth study of topic modeling and text analysis as described in the earlier section, Enhancing the Humanities. Project participants, Matthew Jockers and David Blei, and their work, as further outlined in the Biography section of this proposal, also bring innumerable benefits to this project. Their publications “Macroanalysis: Digital Methods and Literary History” and “Topic Modeling and Digital Humanities” as well as David Blei’s discovery of latent Dirichlet allocation (LDA), reveal their vast amount of knowledge on the subject. Fashion Findings will also require exceptional web designers in order to convey the project’s research fully. A great degree of thought went into the planning distribution of this project, as further discussed in the Final Project and Dissemination section of this proposal. Communicating the results of this project is necessary to its success; as a result, experienced web designers are necessary. Additionally, the extensive data from Google makes Fashion Findings viable and gives it the potential to grow and develop in the future as additional resources become available. As individuals inevitably search, this project will continually evolve, creating a cycle of intellect and development. This progression will require additional means in the future, which will be determined based on the required level and area of expertise relevant to the project’s increasing needs.

Work Plan

Fashion Findings will be completed in a series of tasks. The foundation of this project revolves around collecting data and search results from Google Trends. This data will be gathered and analyzed through topic modeling and text analysis over a period of six months by David Blei with the help of a Google representative. Following collection, David Blei and Matthew Jockers will focus on the topic modeling and text analysis aspect of the project. Blei's experience with and understanding of these methods will allow him to successfully oversee and understand this process. During this work, the team will simultaneously build a skeleton for the Fashion Findings website so that the data can be inputted in an efficient matter.

In addition to working on topic modeling and text analysis Matthew Jockers will be responsible for the macroanalysis of the data. His previous work with macroanalysis will allow him to find potential connections and interpretations of the data. Because he is more removed from the fashion aspect of this project, he will be able to do this with limited bias. Additionally, his overall understanding of these methods equips him with a vast understanding of these methods. Jockers will focus of the outcomes and implications of this data. As a result, he will be able to offer insights into the best way to present the data so that it is most beneficial to viewers.

As project director, Emily Song will oversee all elements of this plan and be directly involved throughout each stage. She will work alongside Blei during the collection of data, as this in an imperative step of the process. She will also work on topic modeling and text analysis in order to see the results of the project. During the interpretation stages of the project, Song will work with Jockers in order to offer insights into the fashion industry and its various trends. However, Jockers will ultimately be in charge of this element of the process to remove bias that comes from Song's opinions and experience. In terms of distribution, Song will play a major part in the development of the Fashion Findings website to ensure the platform fits the goals and values of the project. Overall, she will be involved in every stage of the work plan, but in varying degrees.

The success of this project will depend on the public's reaction, beginning with individuals in the fashion industry. Before its official launch, there will be a trial period during which only certain individuals will have access to the website. People with access will include select members of the digital humanities realm and key players in the fashion industry. There will also be a focus group during which these individuals can offer feedback, both positive and negative. These individuals will be selected during project development based on current trends in the digital humanities and fashion worlds and expressed interest. From a technical perspective, views and use of the Fashion Findings website will also be tracked. Specifically, project team members will be able to see how many times the website is views, where traffic comes from, and what elements of the site are clicked on most frequently. This data will allow for future improvements based on what users are most interested in on the platform.

Final Project and Dissemination

Fashion Findings will be disseminated to the public in the form of a website (see Appendix B Figure 2). This website with be both interactive and aesthetically pleasing in order to increase ease of use and viewing. The overall presentation of Fashion Findings is an essential element of the project itself. The website must communicate data and findings effectively in order to reveal the value of this project and of the fashion industry. The graphics on this site must be both

interactive and easy to understand in order to open it up to a broader audience. In terms of visuals, the design elements of Fashion Findings are equally as important as the data it is built on. This website strives to capture the beauty of the fashion industry, both in its physical attributes and in the meaning it has the potential to uncover. As a result, the Fashion Findings platform must communicate these values simultaneously.

Fashion Findings will be accessible to everyone. Trends in the fashion industry stem from all areas of the world. Because Fashion Findings analyzes data and search results internationally, the results of the project will also be available worldwide. This will also increase the span of the project's impact by sharing the values and meaning of the fashion industry to a larger population. The website will also cater to individuals with disabilities, specifically those with impaired sight. The platform will feature audio recordings that explain the project in tandem with its visual representations and textual analyses. Additionally, the website will be available and fully interactive on all supported browsers.

In order to attract press and support, Fashion Findings will host a launch conference to reveal the full project, website, and results. Respected members from both the digital humanities realm and the fashion industry will be invited to the conference. Depending on the number who plan to attend, the launch may also be open to the general public. At the conference, project members Emily Song, Matthew Jockers, and David Blei will present the project foundation, findings, and overall goals. The launch will be recorded and later uploaded to the Fashion Findings website so that users can watch an in-depth explanation of the process behind the project itself in addition to what it means on a broader scale. Overall, this launch will further communicate the value of Fashion Findings and help it achieve its overarching goal of uncovering the significance of the fashion industry in terms of culture and society through the digital humanities.

Biography

Project Director: Emily Song
As project director, Emily Song has experience in both the digital humanities realm and the fashion industry. Experience in the digital humanities includes five months of study in the field and work on projects involving topic modeling, text analysis, networking, and spatial analysis. In terms of fashion, Song worked in retail for over three years as a sales associate and marketing manager. She can offer unique insights into the industry as a result of directly interacting with current trends. Additionally, the creation of her jewelry company, Emily Song Designs, gives her an understanding of how the business side of fashion operates. As owner and creator of this company, she also has experience with web design through the creation of her company website. These unique skills combined allow her to successfully direct and oversee the operations of the project through a unique lens that directly combines the fashion industry and the digital humanities – essentially the essence of Fashion Findings.

Project Member: Matthew Jockers
Matthew Jockers offers great insight into this project as a result of his direct work in the digital humanities. In his piece "Macroanalysis: Digital Methods and Literary History," Jockers discusses digital humanities methods in depth including topic modeling and text analysis. Much of the basis of this proposal is based on Jockers' studies. His intellect and understanding of this field and its challenges allow him to make a significant contribution to Fashion Findings.

Project Member: David Blei
David Blei has a great deal of experience in the digital humanities. As one of the developers of the LDA topic modeling method, Blei has a technical expertise that can place this project above others of its kind. His piece "Topic Modeling and Digital Humanities" reveals his in-depth level of understanding of this field and beyond. Blei's abilities and knowledge related to the digital humanities realm will fuel Fashion Findings from its early stages to its growth and beyond.

Project Contributor: Google
Google and Google Trends are essential elements of Fashion Findings. With the help of Google's team, Fashion Findings will be able to create a basis of data for the project itself. Direct Google members will also be able to offer insights into the structure of their data and the way it is processed overall, allowing the Fashion Findings team to decipher the best way to disseminate it to the general public.

Appendices

Appendix A: Works Cited

Banister, Emma N., and Margaret K. Hogg. "Negative Symbolic Consumption and Consumers' Drive for Self-Esteem." *European Journal of Marketing*38, no. 7 (2004): 850-68. doi:10.1108/03090560410539285.

Blei, David. "Topic Modeling and Digital Humanities." *Journal of Digital Humanities*2, no. 1 (2012). Accessed December 11, 2018. http://journalofdigitalhumanities.org/2-1/topic-modeling-and-digital-humanities-by-david-m-blei/.

Brett, Meghan. "Topic Modeling: A Basic Introduction." *Journal of Digital Humanities*2, no. 1 (2012). Accessed December 11, 2018. http://journalofdigitalhumanities.org/2-1/topic-modeling-a-basic-introduction-by-megan-r-brett/.

Craik, Jennifer. *Face of Fashion: Cultural Studies in Fashion*. London and New York: Routledge, 2015.

Crane, Diana, and Laura Bovone. "Approaches to Material Culture: The Sociology of Fashion and Clothing." *Poetics*34, no. 6 (2006): 319-33. doi:10.1016/j.poetic.2006.10.002.

Jockers, Matthew Lee. *Macroanalysis: Digital Methods and Literary History*. Urbana: University of Illinois Press, 2013. Accessed December 11, 2018. https://muse.jhu.edu/book/21978.

King, Lindsay, and Peter Leonard. Robots Reading Vogue. Accessed December 11, 2018. http://dh.library.yale.edu/projects/vogue/.

The Rhythm of Food. Accessed December 11, 2018. http://rhythm-of-food.net/.

Underwood, Ted. "A Genealogy of Distant Reading." *Digital Humanities Quarterly*11, no. 2 (2017). http://www.digitalhumanities.org/dhq/vol/11/2/000317/000317.html.

Weingart, Scott, and Elijah Meeks. "The Digital Humanities Contribution to Topic Modeling." *Journal of Digital Humanities*2, no. 1 (2012). http://journalofdigitalhumanities.org/2-1/dh-contribution-to-topic-modeling/.

Appendix B: Figures

Figure 1: Fashion Findings Logo

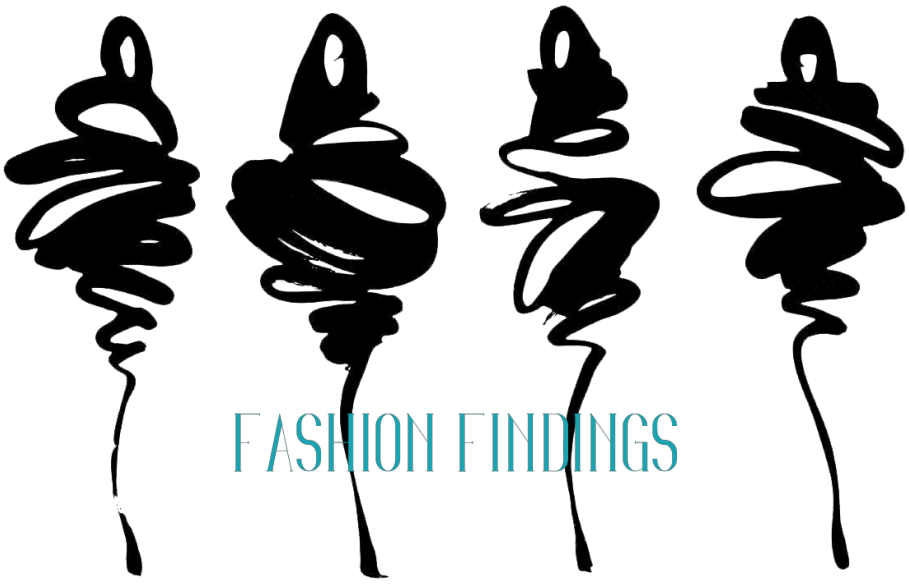


Figure 2: Fashion Findings Website Plan