Communicating the Facebook Political Ad Library:

Data and Design Challenges

Jennifer Stromer-Galley[†]
School of Information Studies
Syracuse University
Syracuse, NY, USA
jstromer@syr.edu

Brian McKernan
School of Information Studies
Syracuse University
Syracuse, NY, USA
bmckerna@syr.edu

Jeff Hemsley
School of Information Studies
Syracuse University
Syracuse, NY, USA
jjhemsle@syr.edu

Ania Korsunska
School of Information Studies
Syracuse University
Syracuse, NY, USA
akorsuns@syr.edu

Sarah Bolden
School of Information Studies
Syracuse University
Syracuse, NY, USA
sebolden@syr.edu

ABSTRACT

Greater transparency is needed in political advertising given that foreign groups meddled in the 2016 election, and allegations surfaced that the Trump campaign ran ads to suppress voter turnout. Facebook currently has a website, the Facebook Ad Library, where journalists can search for any one candidate and see the content of ads they are running, in which states, to whom (in terms of age and gender), for how much money, reaching how many screens, and over how many days. The Facebook Ad Library in its current form is not a highly usable tool for journalists to monitor and understand which campaigns are advertising to targeted publics. Candidates are running dozens to hundreds of ads at a time, thus making it challenging for one journalist or interested citizen to gain a broader perspective on campaign advertising.

As a part of the Illuminating project, our team has been analyzing online political campaign strategy since the 1996 presidential campaign, and we have developed algorithms to accurately classify the content of campaign messages, including what is an attack, a call-to-action, or uncivil content. We have applied these algorithms to social media messages by political campaigns since the 2014 U.S. gubernatorial elections. We aim to expand these models to classify ads. Our goal is to provide an interactive database that will empower journalists to easily and quickly track what the micro-targeted public is being told by presidential candidates and interest groups in their paid ads.

The challenge with creating this interactive website is in determining how best to communicate the visualization of the large volume of data along the different dimensions of potential interest to journalists. Facebook provides several variables including the state a given ad is running in, the volume of impressions in that state, the gender and age demographics of the target, and how much money has been spent on the ad buy. Our algorithms add several additional variables, including the classification of images on several dimensions, such as: the type of ad (attack, advocacy, call-to-action, campaign information, ceremonial); the topic of the ad (on 9 topics including immigration; social issues, economy, the military, and the environment); and the degree of civility of the ad.

Additional challenges surface in order to effectively communicate the data of what campaigns are advertising and to whom. There have been an unprecedented number of candidates running this election cycle that the news media and public opinion polls have deemed viable. Thus, visualizing comparisons of all of the candidates running, and their different ad strategies requires a highly complex visualization that can become overwhelming to even computational journalists. A related challenge is that this project aims to make its data visualizations available to journalists during the presidential primaries. Candidates will drop out as the voting continues across states, leading to challenges around how best to visualize data as candidates leave the race.

Our project applies design science principles [1] to tackle the challenges of communicating the complex data that we are generating so that we can develop effective data visualizations that drive an interactive website. Our project has conducted interviews with journalists covering the 2020 presidential campaign to understand the tools that they are using and what is lacking that could help them effectively report presidential candidates' social media advertising strategies. After the interviews are finished, we will apply our algorithms to the Facebook and Instagram ads that we are collecting through the Facebook API in order to create a beta version of the website, which we will then use to conduct user testing with journalists to understand how to better visualize the data.

We will ascertain whether the design sciences methodology allows us to overcome the design challenges associated with complex datasets that include multiple variables that shift over time.

REFERENCE

[1] Herbert A. Simon. 1996. *The Sciences of the Artificial.* Cambridge, MA, USA: MIT Press.