

Paragraph for submission to Dean for class project showcase

A Core Engine for Social Media Investigations by News Outlets

In the wake of current political trends, tweets and other social media have infiltrated the news cycles to a large degree. Whether it's online, print or even the TV, tweets are oft cited as primary sources. However, the mainstream media is often lacking in their analysis of up to date trends on Twitter, especially those that influence their own brand. For our W205 final project, we created a core engine for news outlets to investigate their own influence within social media, including trend detection, brand protection, and awareness of one's own relationship to the social-media universe.

Using the lens of #fakenews on Twitter as an example, we built a lambda architecture that will stream all the tweets with this hashtag and give live updates on trending words and related hashtags, @mentions, and other metrics from these tweets. Simultaneously we scraped the headlines of 4 mainstream news websites for ad hoc queries and analysis of how these headlines relate to their social media counterparts in our batch layer. Both of these data sets are then fed to machine learning programs in our serving layer that draw out topic-based relationships between the tweets and the hashtags. All analysis is pushed to a Tableau dashboard for graphical consumption.

At scale, this process could be instituted to look at any media outlet's headlines and any searchable word or hashtag on a given social media platform, making a very flexible platform for analysis.