

Visualizing Common Words and Sentiment Trends Related to Burning Man from 2009-2015

Jennifer Lu and Christina Kao

WHAT IS BURNING MAN?

Burning Man is an annual music and art gathering that takes place in the Black Rock Desert in Nevada. The event name comes from the symbolic, ritual burning of a large wooden effigy ("the Man"). Each year a different theme is proposed for the festival which art installations base their work off of as well as camp themes, and the design of "the Man".

PROBLEM

Existing visualizations of Burning Man are mainly static, and focus less on the underlying cultural trends and the context of what is happening in the real world.

MOTIVATION

By performing NLP analysis on datasets from different sources, we can examine:

- (1) correlation among Burning Man theme, event, art, camp datasets and twitter feeds
- (2) current event context
- (3) cultural trends
- (4) demographic trends
- (5) sentiment

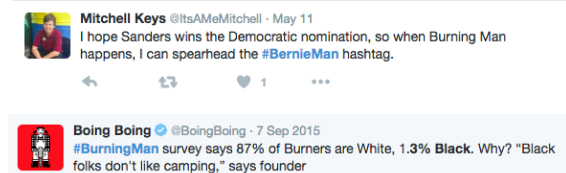


Figure 1. Tweets relating Burning Man with current politics and demographic issues

APPROACH

Data

- Burning Man site about themes, art, camps, and events from 2009-2015
- Twitter posts with the term "burning man" from 2009-2015

Word Count

We eliminated common words like 'the', 'and', 'burning man', 'black rock city'. Then for each source and year, the top 30 most frequent words were ordered

Sentiment Analysis

To do our analysis, we used a Sentiment Analysis, Categorization and Named Entity Extraction Tool called Semantria from Lexalytics. This gives us back sentiment analysis for words among the sections of themes, entities, queries, and categorization.

- Themes: lexically important noun phrases. These represent the buzz words of the document and give a feel of what people are saying.
- Entities: proper nouns like person, place, or product
- Queries: simple search categories
- Categorization: Labels large bodies of text by popular topics
- eg) "I wonder who will win in the California mid-term congressional elections? Is classified with the topic "Politics".

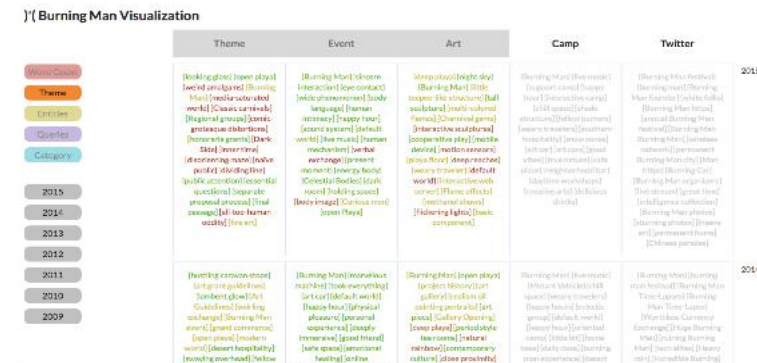


Figure 1. Select Datasets

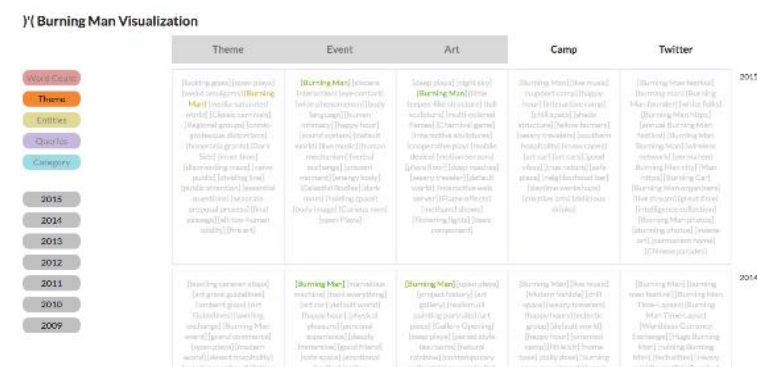


Figure 2. Select Individual Words

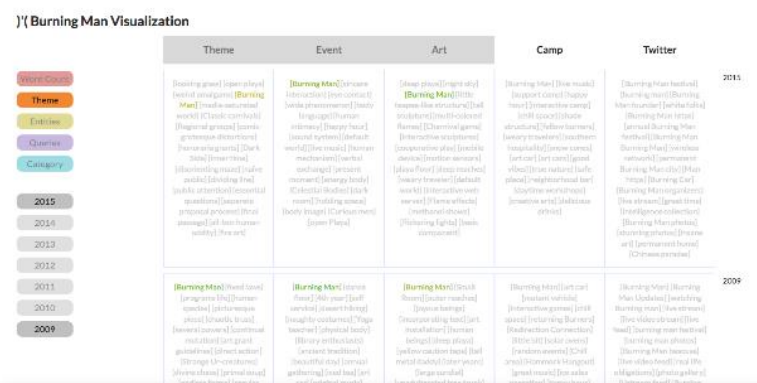


Figure 3. Filter Years

RESULTS

Datasets: BM official datasets, Twitter feed from '09 – '15

NLP Techniques: word count, theme / entity / category extraction, sentiment analysis

Features: filter by year, NLP technique and dataset; top words ordered by frequency; color indicative of sentiment

FUTURE WORK

- **Automate end-to-end workflow:** allow the user to directly upload raw .csv or .json files, perform analysis and generate the interactive visualization on the fly
- **More context:** hover over a word to view the original tweet and word count for more context
- **Diverse datasets:** test on other text based content, and improve the system to be able to process most types of text datasets

ACKNOWLEDGEMENTS

Professor Maneesh Agrawala and the CS448B TA staff

