

# Case Study: Social Media Conversations

## Open Ended Survey Questions

- Don't limit number of questions
- Don't limit number of words in response
- Ask psychographic questions

## Online Reviews

- Categorize with more confidence
- Increase quality of topic names

## Customer Feedback

- Company research forums/platforms
- Help desk, chat, email, logged phone calls

## Social Media Platform Conversations

- Categorize with more confidence
- Increase quality of topic names

# Objectives

**What topics were most tweeted about for the January 6<sup>th</sup> Committee's hearing that took place on Oct 13<sup>th</sup>?**

**How does BERTopic perform against an older algorithm that is used within a lot of products that claim they have AI?**

## Context

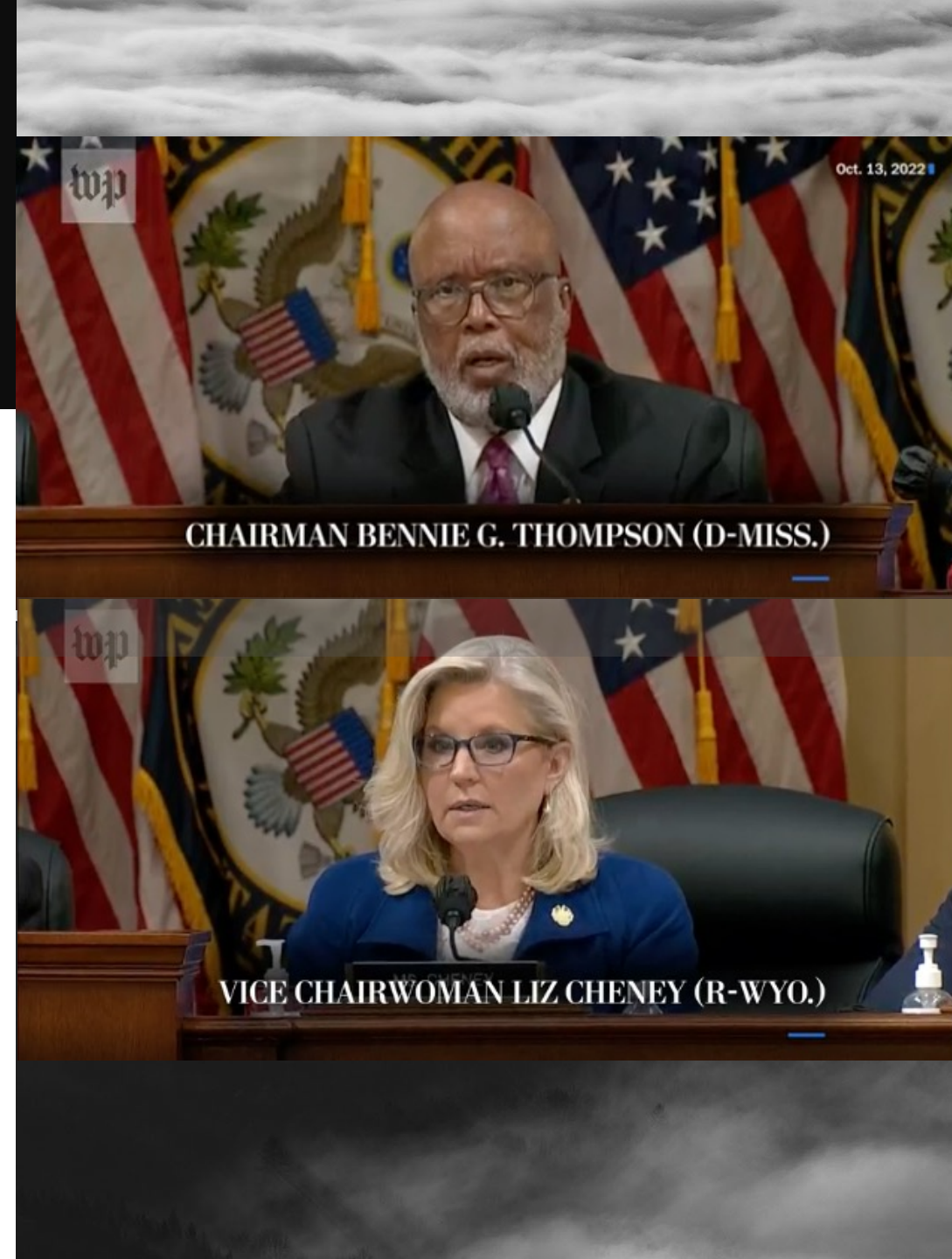
- October 2022 text analysis took place and utilized the prevalent technology of the time - what is now considered an older algorithm.
- September 2023 a new analysis of the text was conducted using BERTopic with GPT4 for topic labeling
- A quantitative testing technique was applied to determine the effectiveness of BERTopic vs LSA
  - Testing asks humans to evaluate the content within the topic categories and contains a scoring method.

# January 6<sup>th</sup> Committee

Hearing on October 13<sup>th</sup> 2022 to present additional video evidence,  
culminating in the vote to subpoena Trump.

What topics were most tweeted about  
during the Oct 13<sup>th</sup> hearing?

34,000 tweets  
including retweets



# Top 4 Topics

Get an understanding before reviewing the underlying content

Tweet Volume	Named Topics
6006	Politicians' Response to Trump's Responsibility
1368	Trump's coup attempt with the help of Giuliani, Meadows, Stone, and others
1176	Testifying Donald Trump under oath
1117	Mark Hamill's views on loving the country without creating violence by overthrowing democracy

# Politicians' Response to Trump's Responsibility

## Representative Tweet

RT @AdamKinzinger:  
Mitch McConnell.  
Kevin McCarthy.  
They both knew Trump was responsible.  
They both called him out.  
They both backed down.

## Observations

Tweet became a chant that was retweeted  
to be the #1 topic

@AdamKinzinger is a political influencer

## How it Looks on Twitter



**Adam Kinzinger #fella**  @AdamKinzinger · Oct 13 · 

Mitch **McConnell**.

Kevin McCarthy.

They both **knew Trump** was **responsible**.

They both **called** him out.

They both backed down.

That's cowardice - not leadership.

[#January6thCommitteeHearings](#)

# Trump's coup attempt with the help of Giuliani, Meadows, Stone, and others

## Representative Tweet

RT @StandForBetter: NEW Video:  
Trump Lost, yet he decided to hatch a coup with the help of Mark Meadows, Roger Stone, Rudy Giuliani, and others

## Observations

@StandForBetter drove 2nd largest volume of tweets

Tweets include video

## How it Looks on Twitter



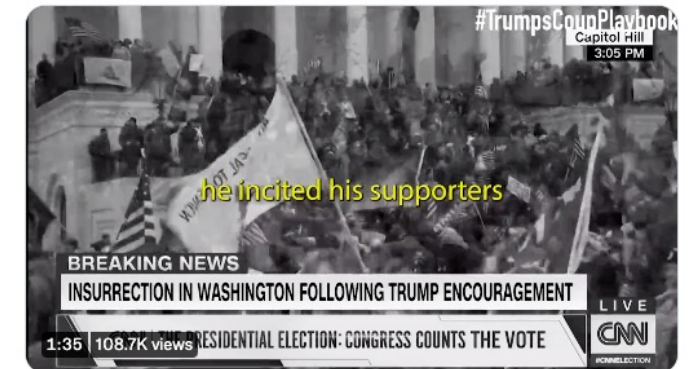
StandForBetter.org @StandForBetter · Oct 13

NEW Video:

Trump Lost, yet he **decided** to **hatch** a coup with the help of Mark Meadows, Roger Stone, **Rudy Giuliani**, and others. Liz Cheney lays it out. I am never voting for a Republican.

Retweet this video: [#TrumpsCoupPlaybook](#)

[#January6thCommitteeHearings](#)  
[#Jan6thHearings](#)



# Testifying Donald Trump Under Oath

## Representative Tweet

RT @AdamKinzinger: We just voted unanimously to subpoena Donald Trump to testify UNDER OATH.  
Democracy demands it.  
History deserves it.

## Observations

@AdamKinzinger created 1st and 3rd  
most tweeted topics

## How it Looks on Twitter





# Mark Hamill's views on loving the country without creating violence by overthrowing democracy

## Representative Tweet

RT @Resist\_MAGA\_GOP: Mark Hamill is correct. You can love this country without creating violence by overthrowing democracy. #January6thComm

## Observations

Topic label creates understanding about the role of Mark Hamill in the topic

Mark Hamill is a political influencer around the October 13<sup>th</sup> Hearing conversation

## How it Looks on Twitter





AI Research:

*How do we know BERTopic works better?*

# AI Research

How do we know BERTopic is correctly categorizing 34,000 pieces of content?

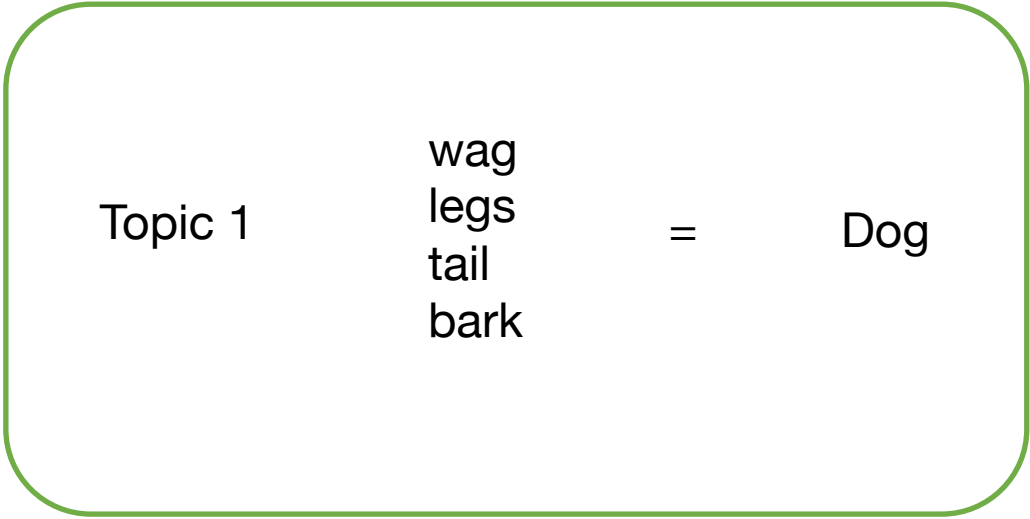
## **How do we measure success?**

- **Usefulness of Topic Names:** Qualitative Observation
- **Accuracy of Content Categorization:** Quantitative Research using human evaluation of content for each topic category

# Accuracy Testing

Does the content within a topic bucket make sense to a human?

Concept Example



Topic 1      wag  
                 legs  
                 tail  
                 bark      =      Dog

**Coherent Topic**



Topic 2      Turtles  
                 Street  
                 Tylenol  
                 Pen      =      ?


**Not a Coherent Topic**

# Accuracy Scores

BERTopic outperforms Old AI

	Old AI	BERTopic
Model Coherence Score (Accuracy)	42%	94%
Key Takeaway	Less Confidence for Business Decisions	High Confidence for Business Decisions

This is important to ensure that the algorithm is correctly categorizing the content from a human perception, and the resulting topic volume is accurate



*Let's look at the usefulness of  
topic names*

# Topic Name Analysis

BERTopic combined with GPT4 are able to tell us specifically what the topic is about

Old AI Topic Name	Old AI ChatGPT Prompt	<b>SightAlchemy</b> Topic Name	Content Representation
Mark, Democracy, Violence, Hamill	Violence & Maga	Mark Hamill's views on loving the country without creating violence by overthrowing democracy	RT @Resist_MAGA_GOP: Mark Hamill is correct. You can love this country without creating violence by overthrowing democracy.
Decided, Hatch, Rudy, Giuliani	Trump & Coup Attempt	Trump's coup attempt with the help of Giuliani, Meadows, Stone and others	RT @StandForBetter: NEW Video: Trump Lost, yet he decided to hatch a coup with the help of Mark Meadows, Roger Stone, Rudy Giuliani, and others

Same dataset used for all methodologies