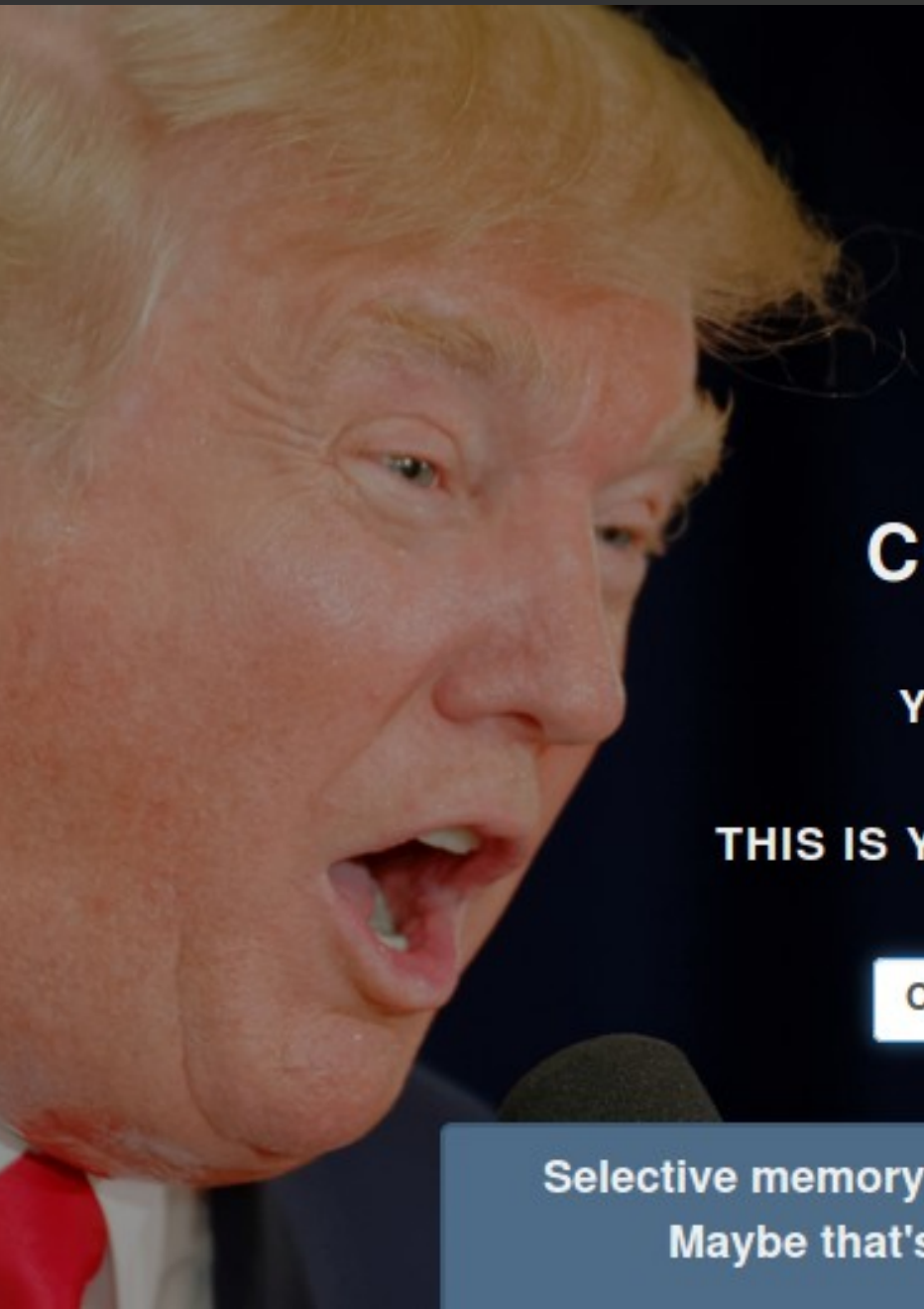


# Trump Bot

---

Speaks just like the real Trump!

# Web App



**CONGRATULATIONS!**

**YOU'VE REACHED TRUMP BOT!**

**THIS IS YOUR CHANCE TO ASK ME ANYTHING.**

Obama

?

Selective memory - Barack obama says that he forgets the recession  
Maybe that's why he is a Celebrity Apprentice Champion!

# @surrealTrumpBot



**Tengen** @towardstengen · Nov 25

@surrealTrumpBot do you like China?



**Trump Bot** @surrealTrumpBot · Nov 25

.@towardstengen And if China isn't treating us fairly, we have Donald Trump - the American dream.



**Susie Sun** @susiexsun · 2 Dec 2015

@surrealTrumpBot When will you admit that you're wearing a wig?



**Trump Bot** @surrealTrumpBot · 2 Dec 2015

.@susiexsun You're right.



**Tyler Patton** @tygpatton · 1h

@surrealTrumpBot What do you think of Ben Carson?



**Trump Bot** @surrealTrumpBot · 1h

.@tygpatton You didn't think I know people like that.

# How does it work?

---

# How does it work?

---

## Bi-gram Markov Model

hair

?

My hair may not be reliable.

$\left( \text{hair}, \text{may} \right) \left[ \text{not} \dots \right]$

$\left( \text{may}, \text{not} \right) \left[ \dots \text{be} \right]$

$\left( \text{not}, \text{be} \right) \left[ \text{reliable} \dots \right]$

$\left( \text{be}, \text{reliable} \right) \left[ \dots \right]$

hair may not be reliable.

# The Model

---

- Stores probabilities of all possible words following a bi-gram in the corpus.
- For the English language, this would be over 100,000 probabilities per bi-gram.
- Fortunately, Trump's vocabulary is way smaller than that!

# Trump's Vocabulary

---

~250,000 words

~14,000 unique words

Collected from:

- Tweets
- Speeches
- Debates
- Essays



# Cleaning Tweets

---

- Tweet completion: Longer tweets end in "(cont) some.link.to.twitlonger"
- Discarded re-tweets
- Removed hyperlinks
- Handle acronyms/titles, such as "U.S." or "Jr."
- Added spaces between words and punctuation

# Dealing with @usernames



A large crowd of people is gathered around Barack Obama, who is wearing a white shirt and has his back to the camera, reaching out to shake hands. In the background, a woman holds a baby, and many people are holding up smartphones to take photos or videos.



A circular profile picture of Barack Obama smiling, wearing a dark suit and tie, with an American flag in the background.

**Barack Obama** ✓  
@BarackObama

TWEETS	FOLLOWING	FOLLOWERS	LIKES	LISTS
14.6K	638K	69.3M	10	3

Tweets   Tweets & replies   Photos & videos

 **Barack Obama** @BarackObama · 59m




A green square profile picture with a white oval shape in the center, resembling a stylized 'O' or an egg.

**cat**  
@gigglemitz

TWEETS	FOLLOWING	FOLLOWERS	LIKES
1,853	82	75	47

Tweets   Tweets & replies

 cat Retweeted

# Dealing with @usernames

---

- Keep only “important” people and organizations
- Replace the @usernames by the name as listed on Twitter.
- Tweets to “unimportant” people of a personal nature – discard.

# Model Summary

---

- Input word or sentence
  - Select a *seed* from sentence
  - Search dictionary
  - If key found, generate sentence
  - Output sentence

What if key not found?

# What about words outside Vocabulary?

---

## Word vectors!

- 400,000 pre-trained word vectors (GloVe) from Stanford NLP project
- For each word, find ten most similar words (cosine-similarity)
- Use one of the words as seed

# What about words outside Vocabulary?

## SUBMARINE

submarines

warship

ship

torpedo

destroyer

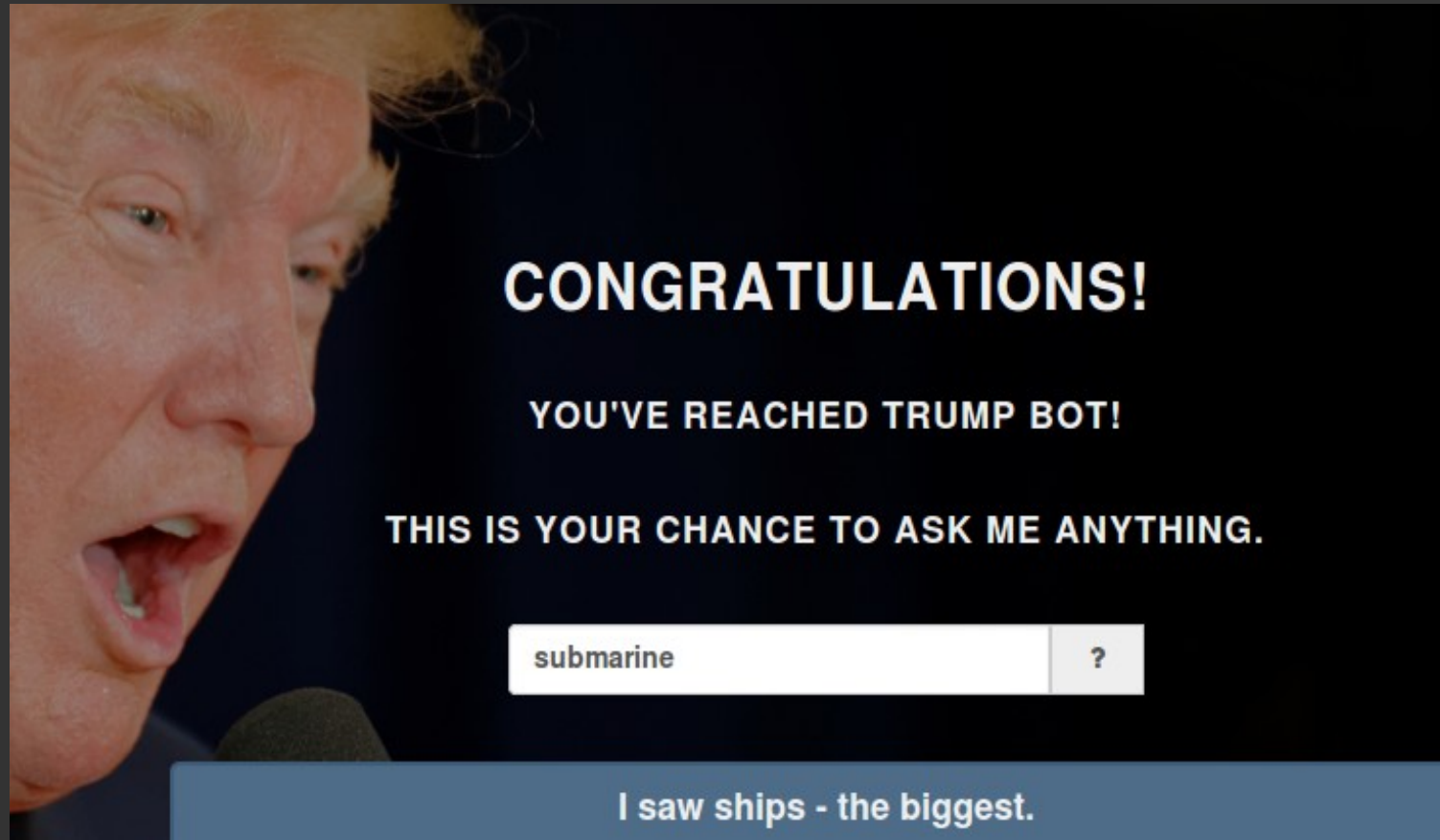
destroyers

vessel

ships

u-boat

navy



# Canned responses



**CONGRATULATIONS!**

**YOU'VE REACHED TRUMP BOT!**

**THIS IS YOUR CHANCE TO ASK ME ANYTHING.**

python

?

**You're fired!**



**CONGRATULATIONS!**

**YOU'VE REACHED TRUMP BOT!**

**THIS IS YOUR CHANCE TO ASK ME ANYTHING.**

adsfhklhl

?

**You need to learn how to type. You're fired!**

---

We are what we speak ...



# Trump's Favorite Words

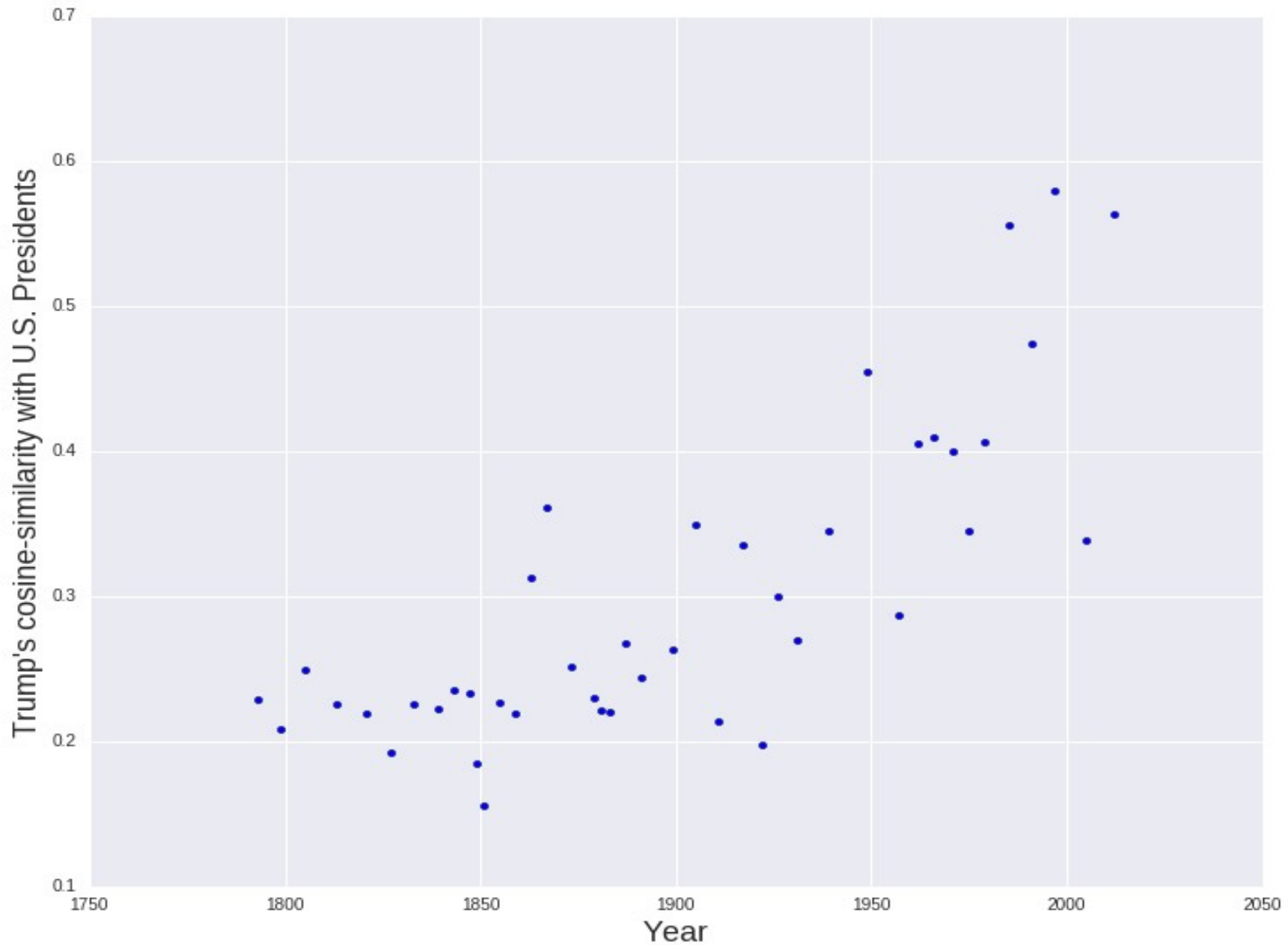


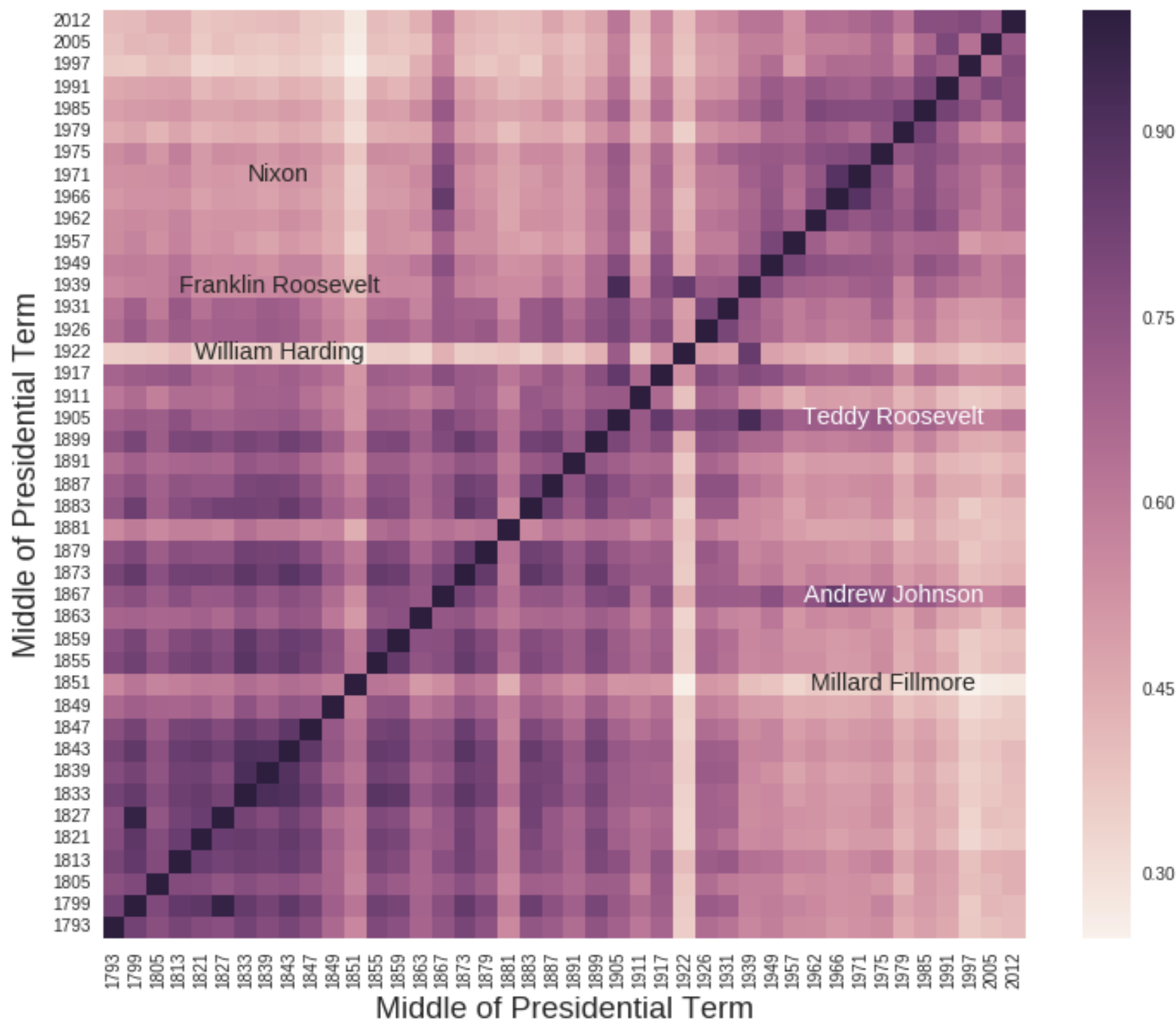
[illegible]

---

How similar or different is he really?

# Trump vs Presidents





# Further Exploration

---

- Look for semantic similarity
  - document vectors
  - sentiment analysis
- For sentence generation
  - RNNs
  - Hidden Markov Models

# Aditi Raval

---

[www.twitter.com/surrealTrumpBot](https://www.twitter.com/surrealTrumpBot)

[www.whatwouldtrumpsay.elasticbeanstalk.com](http://www.whatwouldtrumpsay.elasticbeanstalk.com)

[www.github.com/araval/trump\\_bot](https://www.github.com/araval/trump_bot)