



# Understanding Information Weaponization via Data-driven Analysis

# Emiliano De Cristofaro emilianodo.com







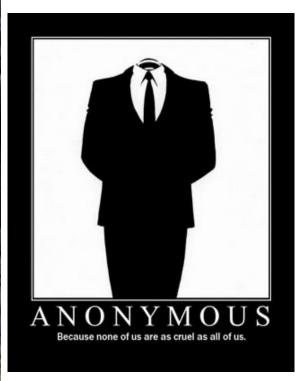


## **Heard of 4chan?**

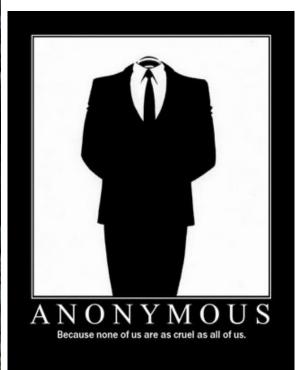
## **Heard of 4chan?**























#### WHERE DID EVERYTHING

## GO SO REICH



## Why Do We Care About 4Chan?

## Why Do We Care About 4Chan?

# How a racist, sexist hate mob forced Leslie Jones off Twitter



Over the past 24 hours, Leslie Jones has been inundated with racist, hateful vitriol on Twitter.

### Why Do We Care About 4Chan?

# How a racist, sexist hate mob forced Leslie Jones off Twitter



Over the past 24 hours, Leslie Jones has been inundated with racist, hateful vitriol on Twitter.

TAYWARD YOUTH 3/24/16 3:10 PM

Who turned Microsoft's chatbot racist? Surprise, it was 4chan and 8chan

By Ethan Chiel

## Why Do We Care About 46han2

# How a racist, sexist hate mob forced Leslie Jones off Twitter

By Kristen V. Brown



Over the past 24 hours, Leslie Jones has been inundated with racist, hateful vitriol on Twitter.

#### **BUSINESS**

French fear Putin and Trump followers are using 4chan to disrupt presidential election

CHRIS O'BRIEN @OBRIEN MAY 5, 2017 3:40 AM

TAYWARD YOUTH

3/24/16 3:10 PM

# Who turned Microsoft's chatbot racist? Surprise, it was 4chan and 8chan

By Ethan Chiel





#### An image-board forum

Organized in boards (70 at the moment)



#### An image-board forum

Organized in boards (70 at the moment)

An "original poster" (OP) creates a new thread by making a post Single image attached



#### An image-board forum

Organized in boards (70 at the moment)

An "original poster" (OP) creates a new thread by making a post Single image attached

#### Other users can reply:

With or without images, possibly add references to previous posts, quote text, etc.



#### An image-boa

Organized in bo

#### An "original p

Single image a

#### Other users c

With or without

#### Boards

#### **Japanese Culture**

Anime & Manga Anime/Cute Anime/Wallpapers

Mecha

Cosplay & EGL

Cute/Male Flash

Transportation Otaku Culture

#### Video Games

Video Games Video Game Generals

Pokémon

Retro Games

#### Interests

Comics & Cartoons

Technology Television & Film

Weapons

Auto

Animals & Nature Traditional Games

Sports

Alternative Sports Science & Math

History & Humanities

International Outdoors Toys **Creative** 

Oekaki Papercraft & Origami

Photography
Food & Cooking
Artwork/Critique

Wallpapers/General Literature

Music Fashion 3DCG

Graphic Design Do-It-Yourself

Worksafe GIF Quests Other

Business & Finance
Travel

Fitness Paranormal

Advice LGBT Pony

Current News

Worksafe Requests Very Important Posts

Misc. (NSFW)

Random ROBOT9001 Politically Incorrect

International/Random
Cams & Meetups

Shit 4chan Says

Adult (NSFW)

Sexy Beautiful Women

filter ▼

Hardcore

Handsome Men

Hentai Ecchi Yuri

Hentai/Alternative

Yaoi Torrents

Iorrents
High Reso

High Resolution Adult GIF Adult Cartoons

Adult Requests

## /pol/ - Politically Incorrect Board

## /pol/ – Politically Incorrect Board





#### Welcome to /pol/ - Politically Incorrect

Anonymous (ID: M9eTFJC) = 01/08/15(Thu)23:40:44 No.40489590 🔊 🗎 [Reply] ▶



This board is for the discussion of news, world events, political issues, and other related topics.

Off-topic and /b/-tier threads will be deleted (and possibly earn you a ban, if you persist). Unless they are quality, well thought out, well written posts, the following are examples of off-topic and/or /b/-tier threads:

- >Red pill me on X. (with no extra content or input of your own)
- >Are X white?
- >Is X degeneracy?
- >How come X girls love Y guys so much?
- >If X is true, then how come Y? Checkmate Z.

The variety of threads allowed here are very flexible and we believe in freedom of speech, but we expect a high level of discourse befitting of the board. Attempts to disrupt the board will not be tolerated, nor will calls to disrupt other boards and sites.

## /pol/ - Politically Incorrect Board



#### Users do not need to register an account to participate

Anonymity is the default (and preferred) behavior

#### Users do not need to register an account to participate

Anonymity is the default (and preferred) behavior

#### "Some" degree of permanence and identifiability is supported

Can enter a name along with their posts (no authentication though)

#### Users do not need to register an account to participate

Anonymity is the default (and preferred) behavior

#### "Some" degree of permanence and identifiability is supported

Can enter a name along with their posts (no authentication though)

#### Threads get "archived" after a while

Actually all posts deleted after a week

1. Not your typical social network (anonymous/ephemeral)

- 1. Not your typical social network (anonymous/ephemeral)
- 2. Their actions not limited to 4chan, need to look at other platforms to understand their impact

- 1. Not your typical social network (anonymous/ephemeral)
- 2. Their actions not limited to 4chan, need to look at other platforms to understand their impact
- 3. Knowing what they're talking about is not easy

- 1. Not your typical social network (anonymous/ephemeral)
- 2. Their actions not limited to 4chan, need to look at other platforms to understand their impact
- 3. Knowing what they're talking about is not easy
- 4. You might get attacked, doxxed, etc.

## **Datasets**

	/pol/	/sp/	/int/	Total
Threads	217K	14.4K	24.9K	256K
Posts	8.3M	1.2M	1.4M	10.9M
June 30 to September 12, 2016				

#### **Datasets**

#### Methodology:

Visit the "catalog"

Snapshot every 5 mins

 /pol/
 /sp/
 /int/
 Total

 Threads
 217K
 14.4K
 24.9K
 256K

 Posts
 8.3M
 1.2M
 1.4M
 10.9M

 June 30 to September 12, 2016

Once a thread is archived, retrieve full/final contents from 4plebs.org

### **Datasets**

#### Methodology:

Visit the "catalog"

Snapshot every 5 mins

 /pol/
 /sp/
 /int/
 Total

 Threads
 217K
 14.4K
 24.9K
 256K

 Posts
 8.3M
 1.2M
 1.4M
 10.9M

 June 30 to September 12, 2016

Once a thread is archived, retrieve full/final contents from 4plebs.org

#### We're still crawling...

## **Hate Speech?**



Crowdsourced dictionary
Manually filtered a bit

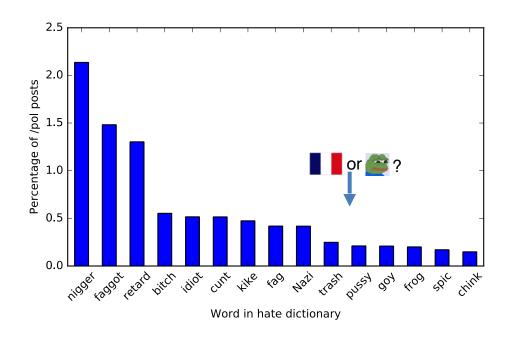
/pol/ by far most hate speech use

/pol/ 12%

/sp/ 7.3%

/int/ 6.3%

Twitter 2.2%



## Raids on other platforms (e.g., YouTube)

Someone posts a YouTube link

Maybe with a prompt like "you know what to do"

Someone posts a YouTube link

Maybe with a prompt like "you know what to do"

Thread is an aggregation point for raiders E.g., "Hah! I called that person a n\*\*\*\*\*\*!"

Someone posts a YouTube link

Maybe with a prompt like "you know what to do"

Thread is an aggregation point for raiders E.g., "Hah! I called that person a n\*\*\*\*\*\*!"

#### If raid is taking place:

Peak in YouTube comments while thread alive? /pol/ thread and YT comments *synchronized?* 

## Raids on other platform



When will the normies learn?



View all 3 replies >



Jack Harper 1 month ago

mfw he accidentally redpills his whole class, they all found the memes funny even the so called "hate speech" he talks about



Jack Harper 1 month ago

I guess I haven't tightened my fedora enough to be as enlightened by my own enlightenment as you sir. Carry on brave knight!



Adem Adiyaman 1 month ago

23:21 "the victory of donald trump and the rise of ultra nazi people" you dense fuck, nice brain washing





Szgerle 1 month ago

Not even just your regular nazi but ULTRA NAZI!



Drakewood 1 month ago

+Potato Jenova A FUCKING LEAF!



A Measurement Study of 4chan's Politically Incorrect Forum and Its Effects on the Web



+ Add to A Share \*\*\* More

11,013 views









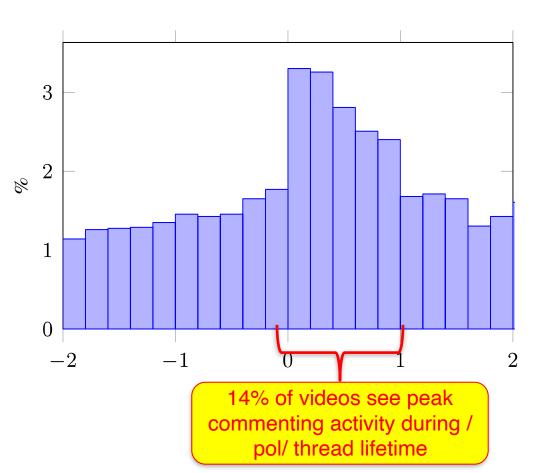
Holy fuck, ((((they)))) are actually intimidated by the powers of Kek.







# **Activity Peaks**

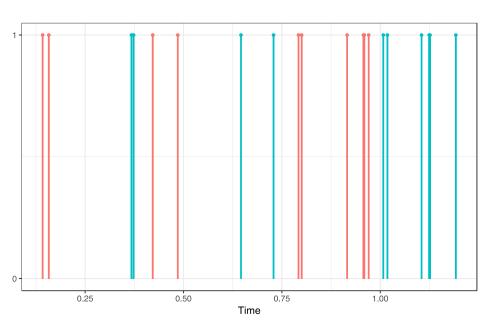


YT videos with peaks during 4chan thread

Determined via PDF of commenting timeseries

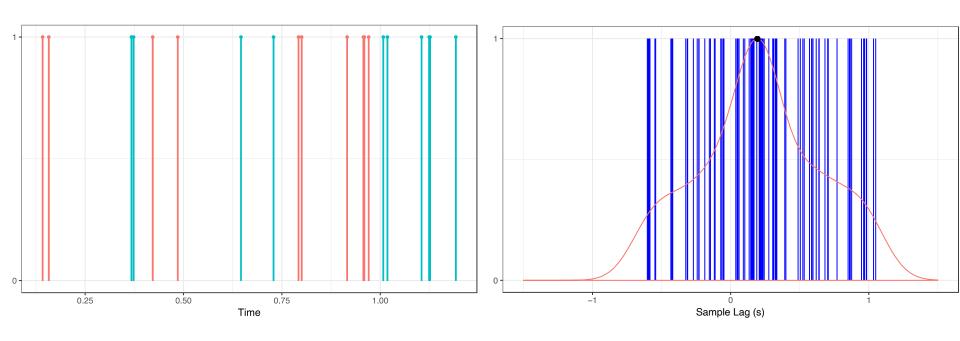
# **Synchronization**

# **Synchronization**



Two series, second randomly shifted from first by 0.2s on avg

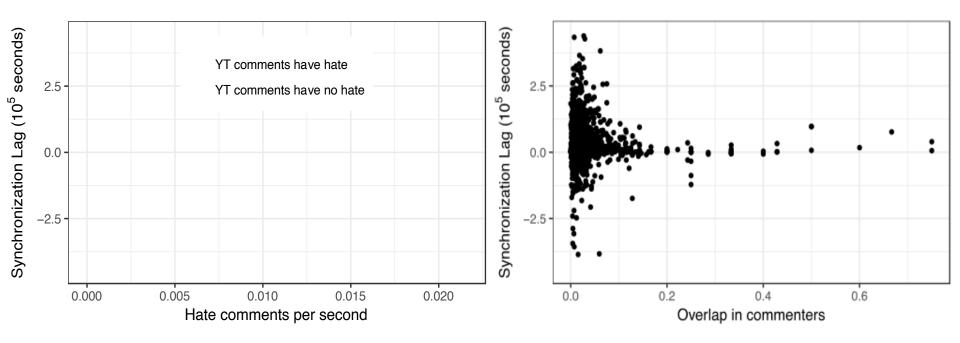
# **Synchronization**



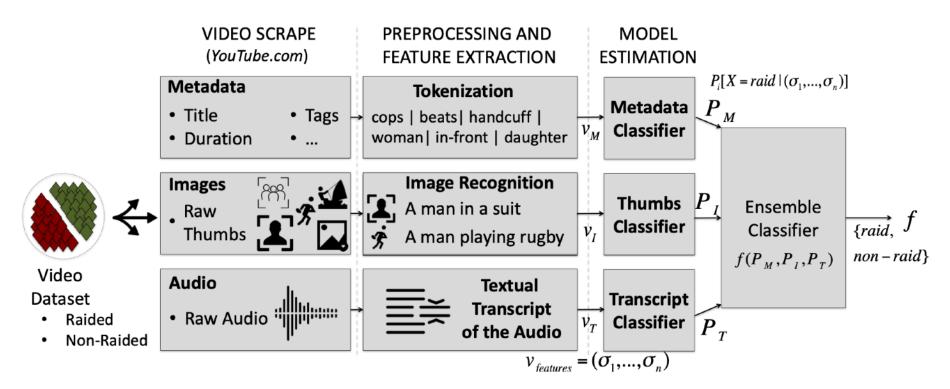
Two series, second randomly shifted from first by 0.2s on avg

Blue lines → per-sample lag
Red area → density of the lags
Peak of density curve = 0.2s

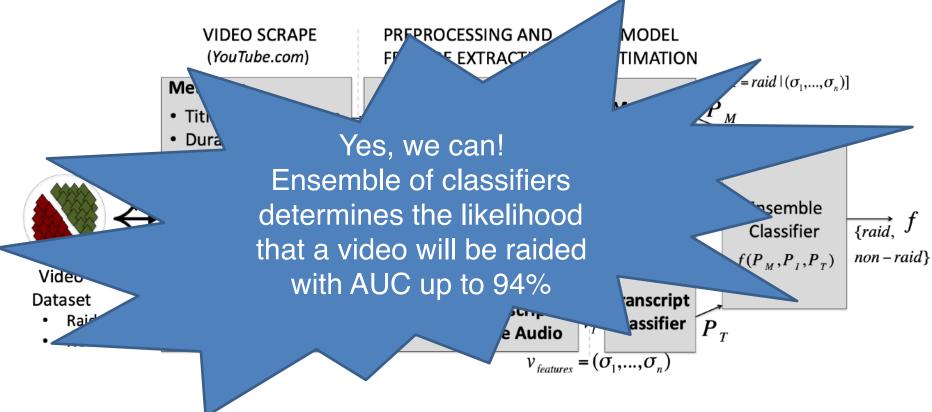
### **Validation**



## In fact... can we predict?



In fact... can we predict?



# What About Cyberbullying?

70.6% of young people say they have seen bullying in their schools
9% of students in grades 6–12 experienced cyberbullying
15% of high school students (grades 9–12) were electronically bullied in the past year

**Cyberbullying:** intentionally aggressive behavior, repeated over time, which involves an imbalance of power

**Cyberbullying:** intentionally aggressive behavior, repeated over time, which involves an imbalance of power

**Cyberaggression**: purposefully saying or doing something to hurt someone

**Cyberbullying:** intentionally aggressive behavior, repeated over time, which involves an imbalance of power

**Cyberaggression**: purposefully saying or doing something to hurt someone

### Can we detect it?

## **Building a Classifier**

Crawling data from Twitter from June to August 2016

**Baseline:** 1M random tweets

Hate-related: 650K tweets based on 309 bully- and hate-related

hashtags

#GamerGate + 308 co-appeared ones

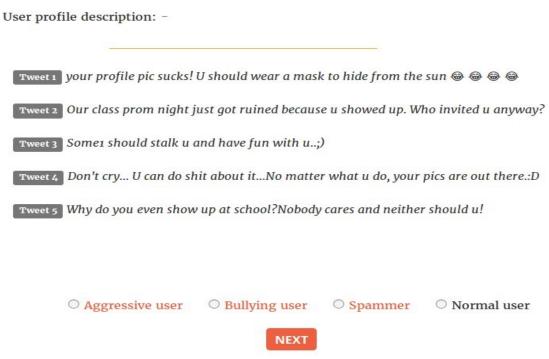
(GamerGate was a coordinated campaign of harassment in the online gaming world, revolving around sentiments of sexism, feminism, and "social justice")

Pre-processing: removing spam, stop words, punctuation marks, etc.

## **Crowdsourced labeling**

#### How do you characterize the following user?

1/10



#### Aggressive

Someone who posts at least one tweet or retweet with negative meaning, with the intent to harm or insult other users (e.g., the original poster of a tweet, a group of users, etc.).

#### Bullying

Someone who posts multiple (at least two) tweets or retweets with negative meaning for the same topic and in a repeated fashion, with the intent to harm or insult other users (e.g., the original poster of a tweet, a minor, a group of users, etc.) who may not be able to easily defend themselves during the postings.

#### Spammer

Someone who posts tweets or retweets of advertising/marketing or other suspicious nature, such as to sell products of adult nature, phishing attempts, etc.

9,484 tweets, 1,307 users (4.5% bullies, 3.4% aggressors, 31.8% spammers, 60.3% normal)

9,484 tweets, 1,307 users (4.5% bullies, 3.4% aggressors, 31.8% spammers, 60.3% normal)

User Features: avg. #posts, account age, #subscribed lists, verified account, posts' interarrival time, default profile image

9,484 tweets, 1,307 users (4.5% bullies, 3.4% aggressors, 31.8% spammers, 60.3% normal)

User Features: avg. #posts, account age, #subscribed lists, verified account, posts' interarrival time, default profile image

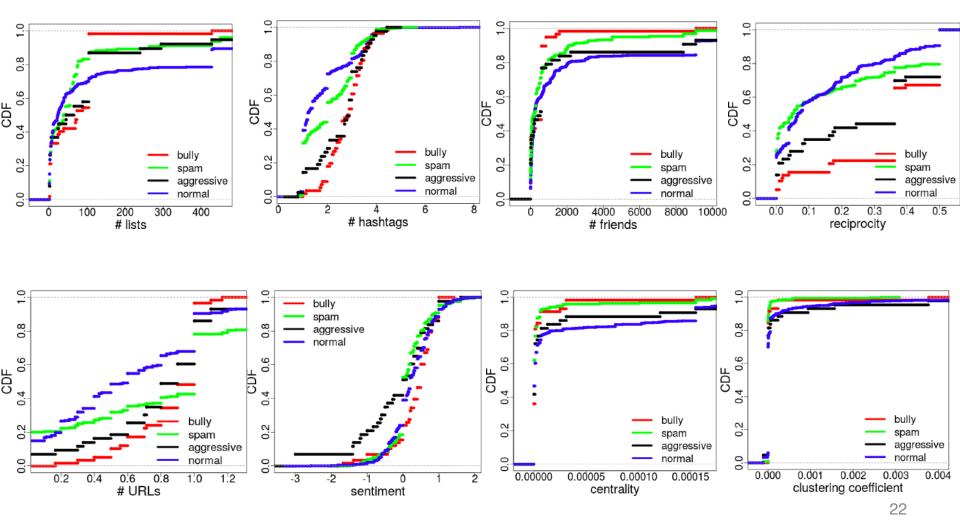
Text Features: #hashtags, #uppercases, #emoticons, #URLs, sentiment, avg. word embedding score, hate and curse scores

9,484 tweets, 1,307 users (4.5% bullies, 3.4% aggressors, 31.8% spammers, 60.3% normal)

User Features: avg. #posts, account age, #subscribed lists, verified account, posts' interarrival time, default profile image

**Text Features:** #hashtags, #uppercases, #emoticons, #URLs, sentiment, avg. word embedding score, hate and curse scores

**Network Features:** popularity (#follower, #friends), reciprocity, avg. power difference with mentioned users, hubs and authority, influence (eigenvector centrality, closeness centrality), communities (clustering coefficient, louvain modularity)

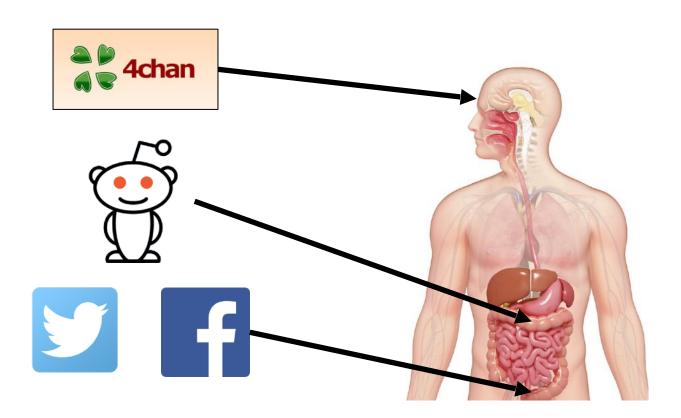


## **Results**

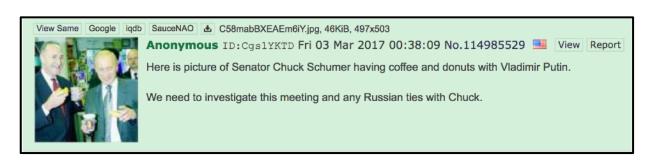
	Prec. Rec. ROC		
bully	0.555	0.609	0.912
(STD)	0.018	0.029	0.009
aggressive	0.304	0.114	0.812
(STD)	0.039	0.012	0.015
normal	0.951	0.976	0.911
(STD)	0.018	0.029	0.009
overall (avg.)	0.899	0.917	0.907
(STD)	0.016	0.019	0.005

	Prec.	Rec.	ROC
bully	1	0.667	0.833
aggressive	0.5	0.4	0.757
normal	0.931	0.971	0.82
overall (avg.)	0.909	0.913	0.817

### How do communities influence each other?











SauceNAO 🕹 C58mabBXEAEm6iY.jpg, 46KiB, 497x503

Anonymous ID:Cgs1YKTD Fri 03 Mar 2017 00:38:09 No.114985529 📕 View

Here is picture of Senator Chuck Schumer having coffee and donuts with Vladimir Putin.

We need to investigate this meeting and any Russian ties with Chuck.



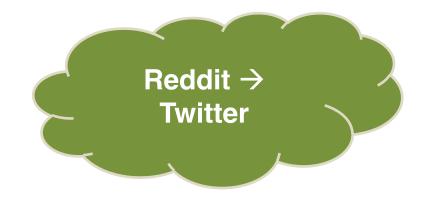




4:21 PM - Jul 2, 2017

□ 141,936 □ 368,969 ○ 605,362







Donald J. Trump 📀 @realDonaldTrump

#### #FraudNewsCNN #FNN

4:21 PM - Jul 2, 2017

141,936 🕽 368,969

○ 605.362



#### Trump's anti-CNN tweet originated from Reddit's largest right-wing extremist forum

Regardless of his intentions, Trump's tweet aligned him with violent hate speech.

Updated by Aja Romano | @ajaromano | Jul 3, 2017, 5:00pm EDT

# The Pizzagate Conspiracy Theory



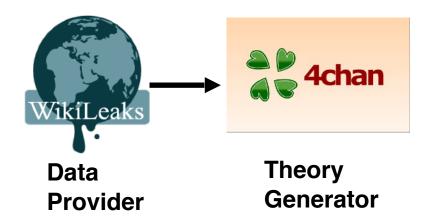
# The Pizzagate Conspiracy Theory





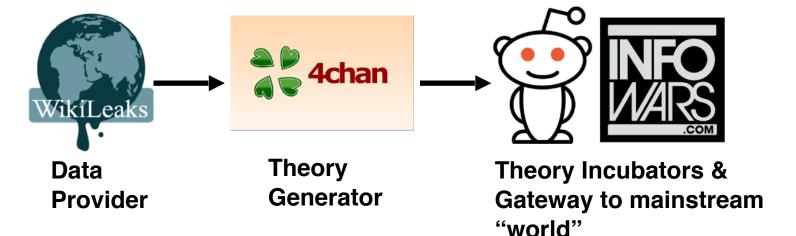
# The Pizzagate Conspiracy Theory





# The Pizzagate Conspiracy Theory



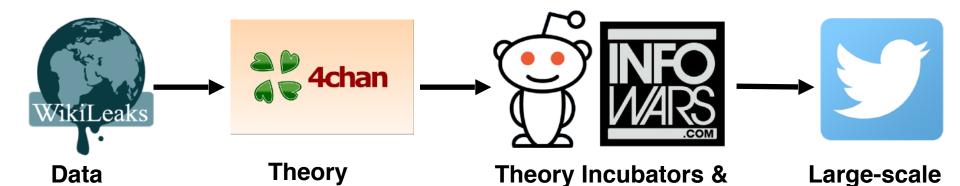


# The Pizzagate Conspiracy Theory

Generator

**Provider** 





"world"

**Gateway to mainstream** 

**Disseminator** 

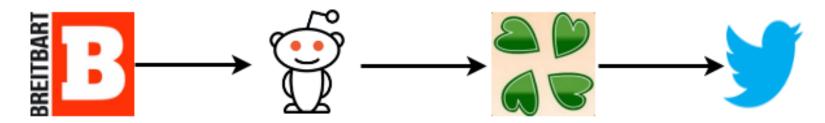
Study the appearance of alternative and mainstream news URLs within the platforms

Study the appearance of alternative and mainstream news URLs within the platforms

Build a sequence of appearance for each URL according to the timestamps

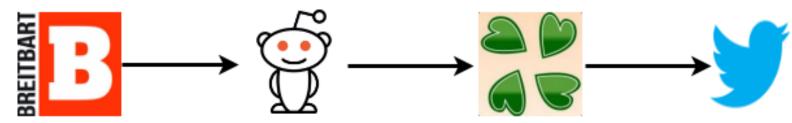
Study the appearance of alternative and mainstream news URLs within the platforms

Build a sequence of appearance for each URL according to the timestamps



Study the appearance of alternative and mainstream news URLs within the platforms

Build a sequence of appearance for each URL according to the timestamps



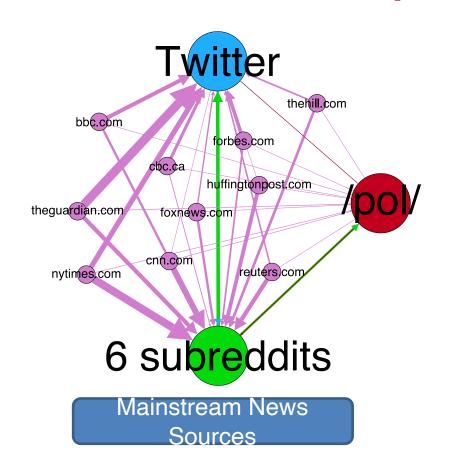
Build a graph with the sequences

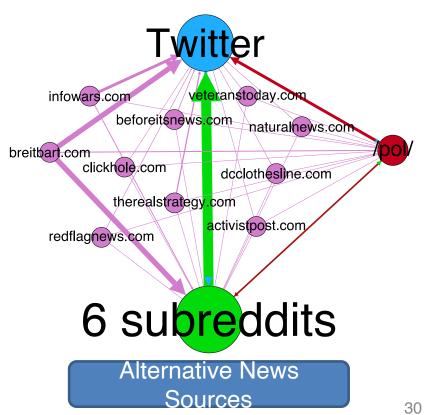
#### **The Data**

99 mainstream and alternative ("fake") news sources

Platform	Posts/Comments	Alternative URLs	Mainstream URLs
Twitter	486K	42K	236K
Reddit (six selected subreddits)	620K	40K	301K
4chan (/pol/)	90K	9K	40K

#### Here's What The Graph Looks Like





#### Assume K processes

Each with a rate of events (i.e., posting of a URL), called the background rate

#### Assume K processes

Each with a rate of events (i.e., posting of a URL), called the background rate

An event can cause **impulse responses** in other processes Increases the rates of other processes for a period of time

#### Assume K processes

Each with a rate of events (i.e., posting of a URL), called the background rate

An event can cause **impulse responses** in other processes Increases the rates of other processes for a period of time

Enables us to be confident about the number of events caused by another event on the source process (weight)

**Reveals causal relationships** 

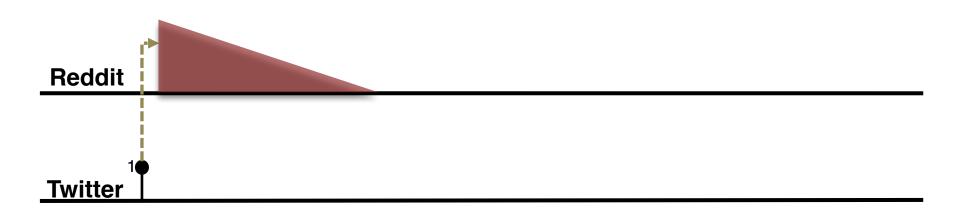
#### Reddit

Reddit

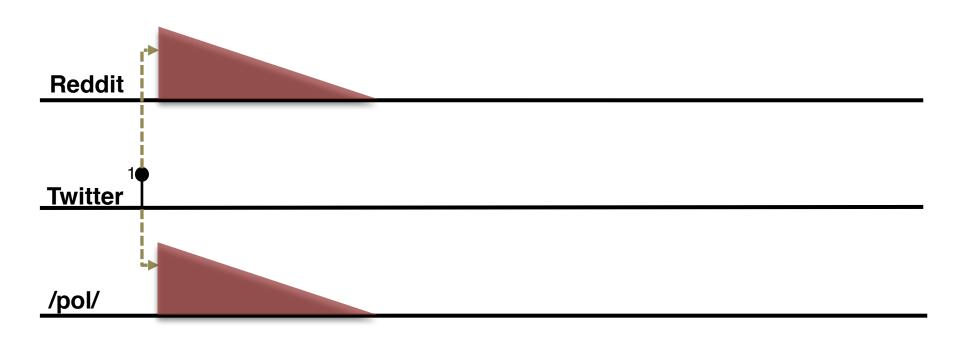
Twitter

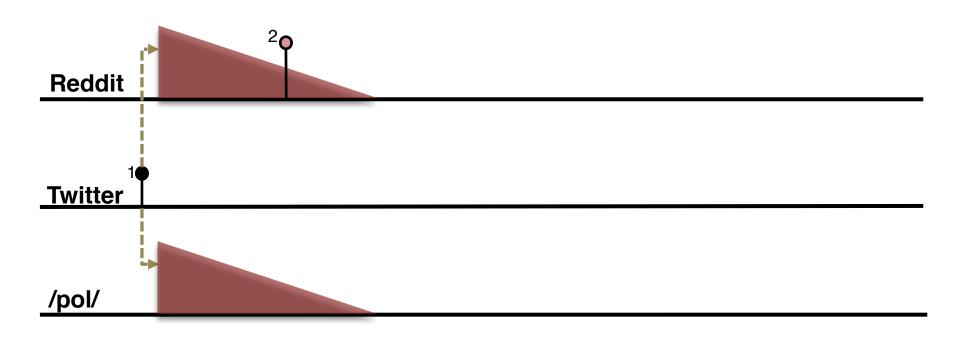
Reddit			
Twitter			
/pol/			

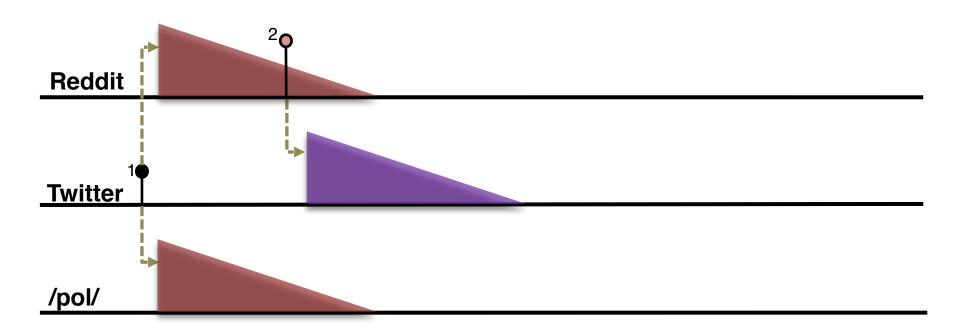
Twitter 1
/pol/

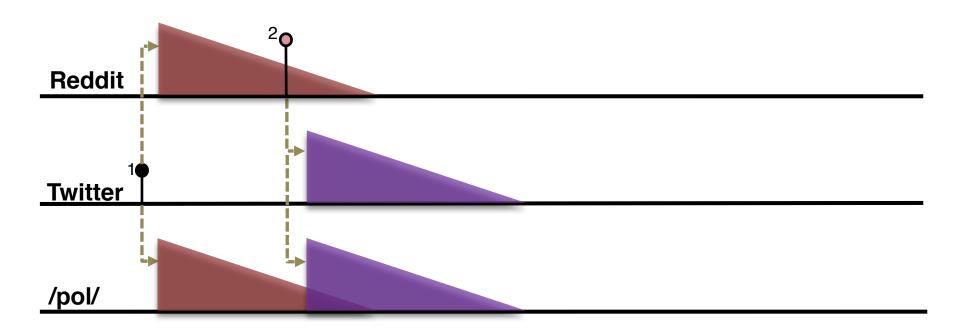


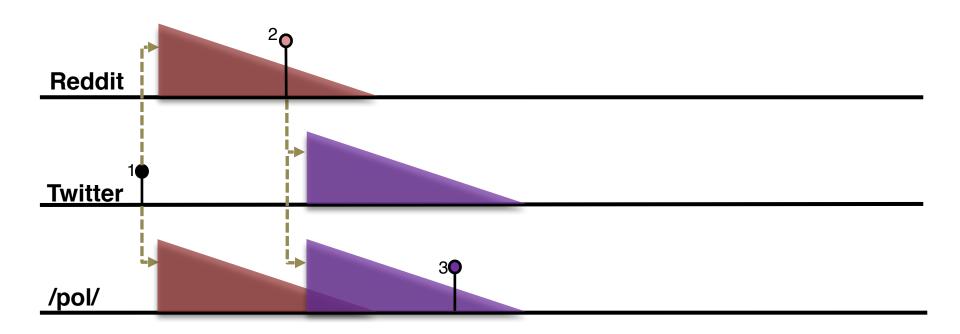
/pol/

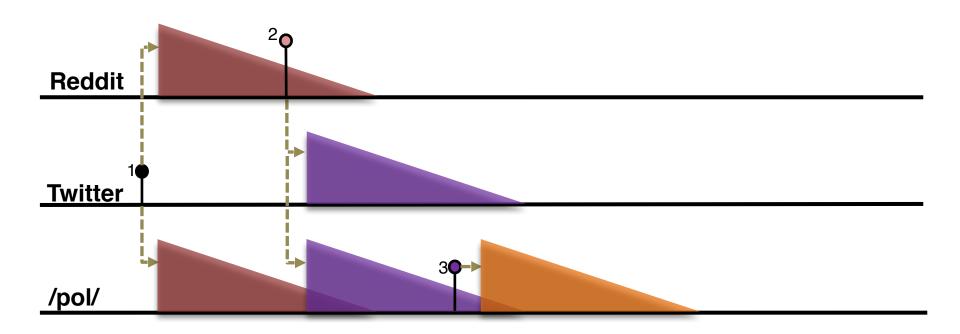


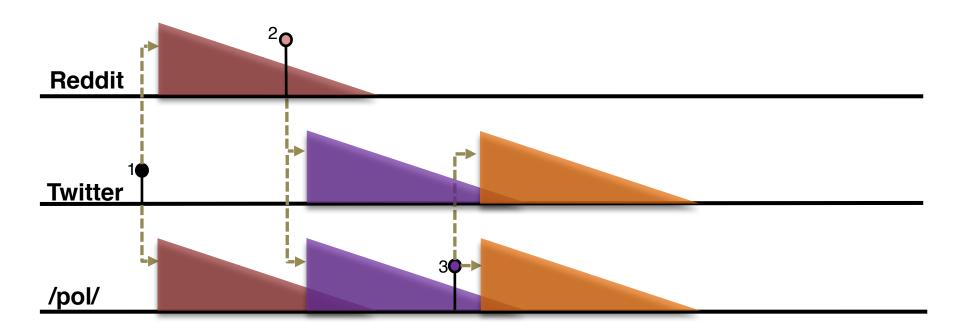


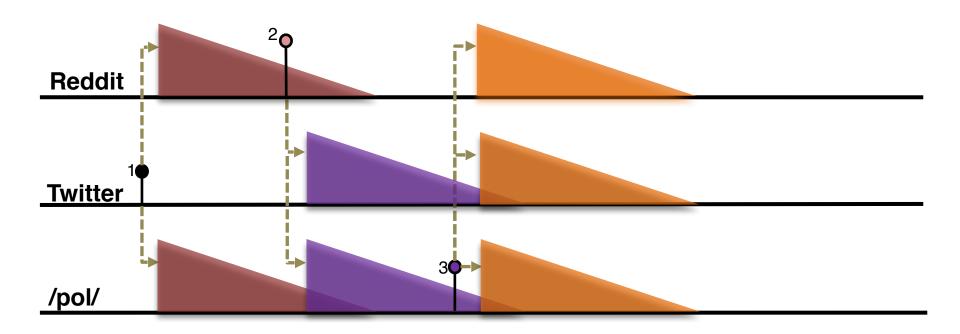


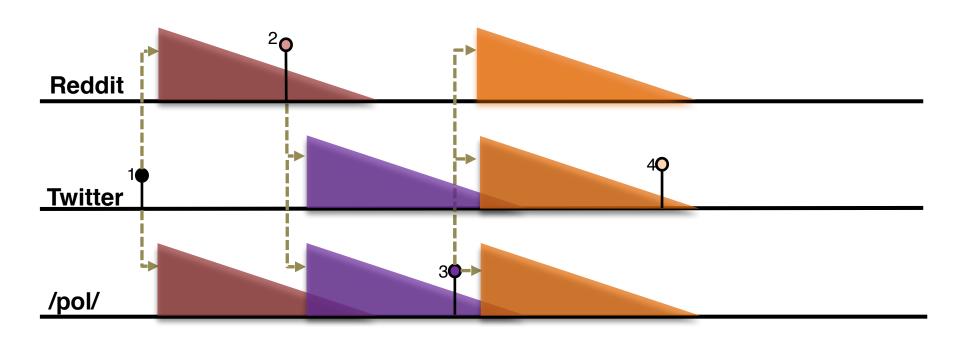


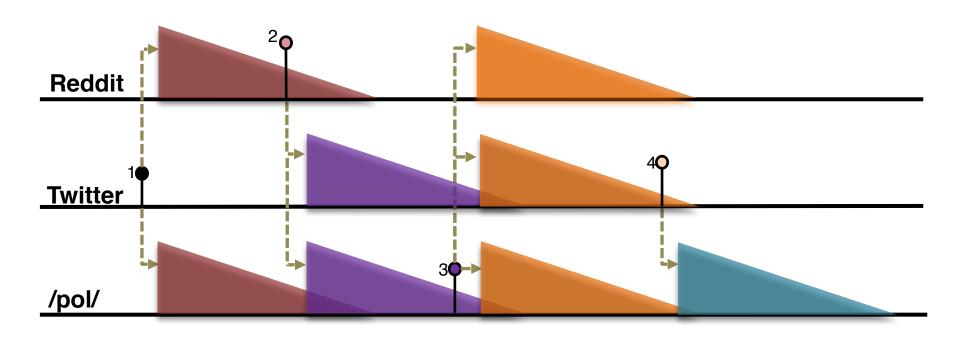


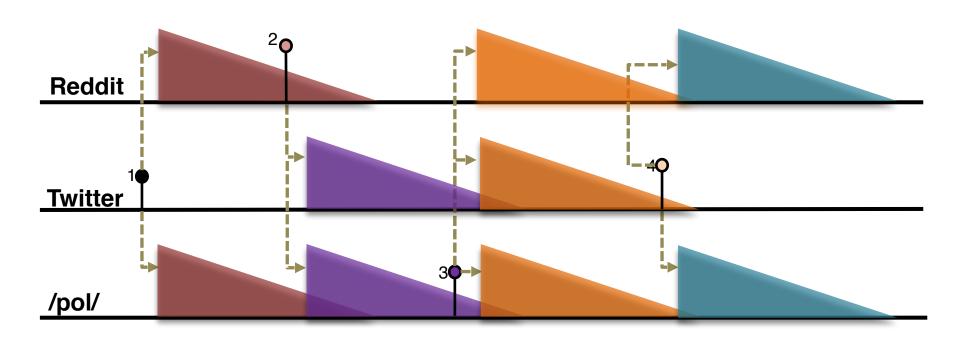


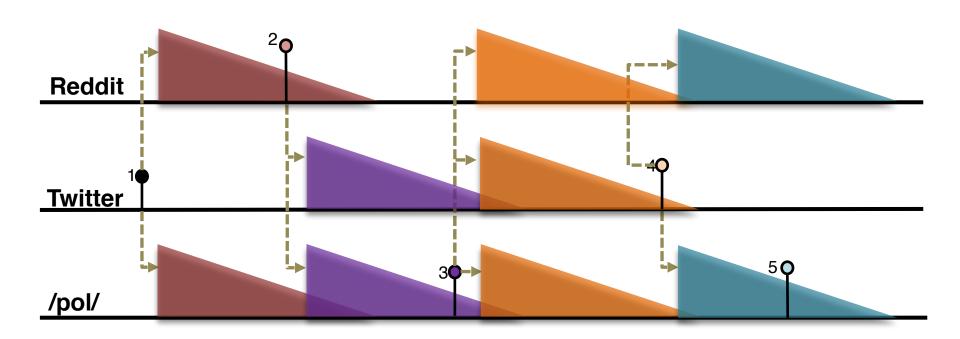


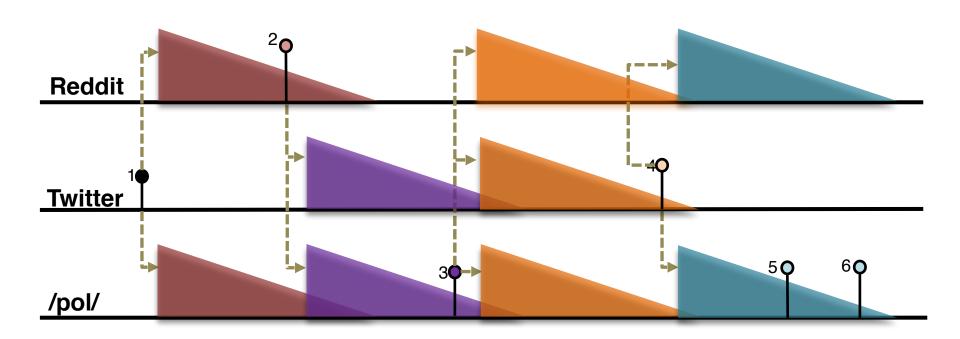


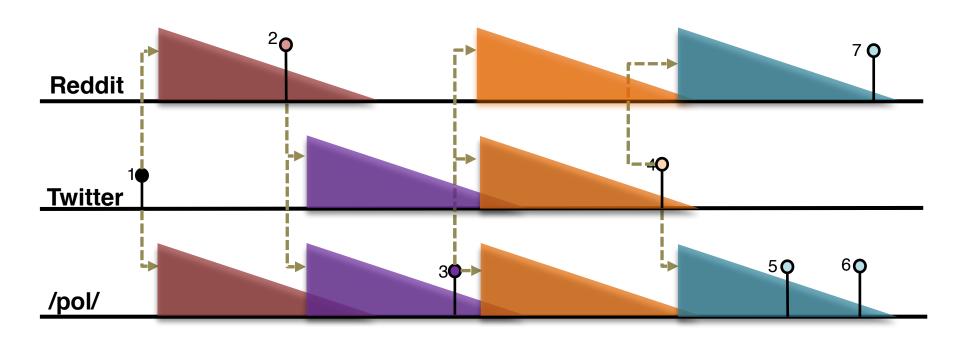












#### For Our Purposes...

#### Hawkes model with 8 processes

One for each platform/community

Distinct model for each URL

#### Fit each model with Gibbs sampling

Calculate the percentage of events **created because** of events happened in each of the other processes

#### **How Communities Influence Each Other?**

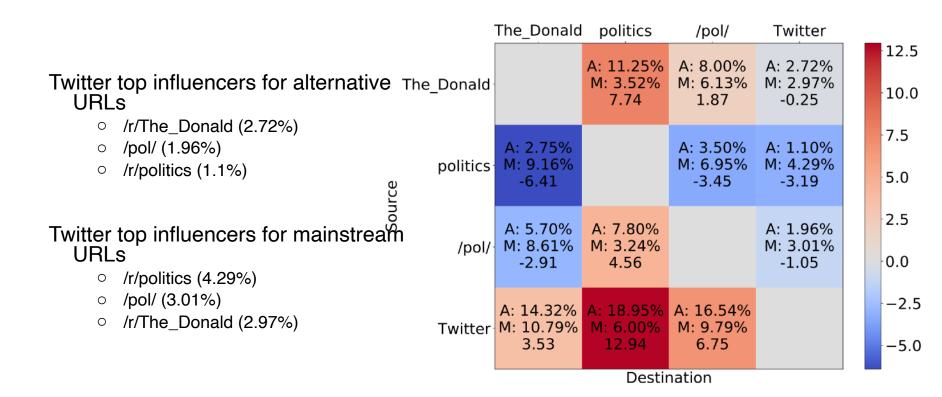
#### Twitter top influencers for alternative URLs

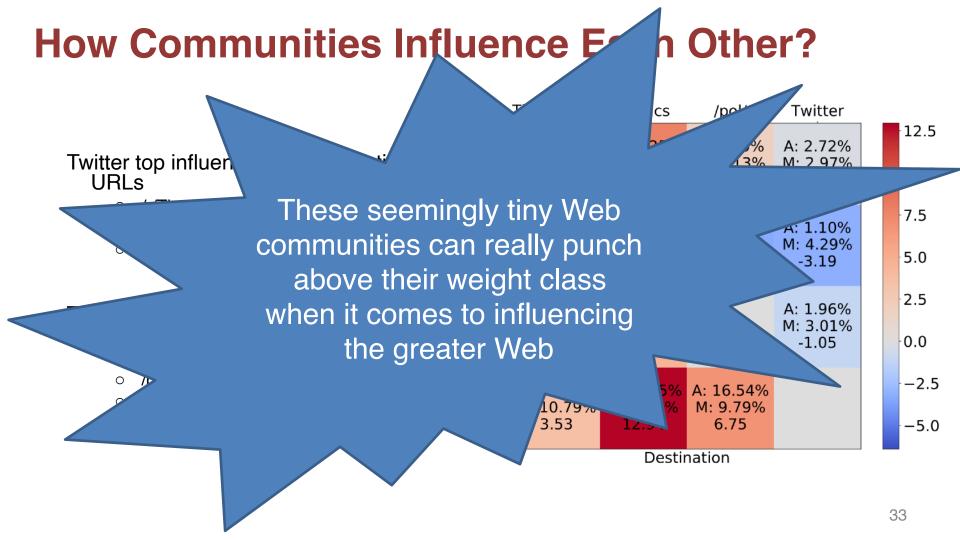
- /r/The\_Donald (2.72%)
- o /pol/ (1.96%)
- /r/politics (1.1%)

#### Twitter top influencers for mainstream URLs

- /r/politics (4.29%)
- o /pol/ (3.01%)
- /r/The\_Donald (2.97%)

#### **How Communities Influence Each Other?**





### Memes are fun!







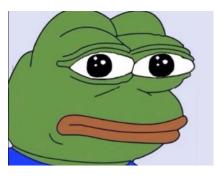
#### WHEN YOUR BOSS THANKS YOU FOR STAYING LATE BUT YOU WERE JUST LOOKING AT MEMES AND LOST TRACK OF TIME

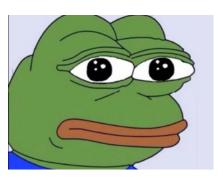




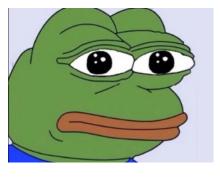
#### Memes are fun!















## Not always though...







### Memes in politics

**Sport** 

Culture

Lifestyle

More ~

The Guardian

US midterms 2018 Business Tech Science

# Annotated: the Trump memes stuck to Cesar Sayoc's van

The vehicle seized by law enforcement in Florida was covered in strange propaganda



▲ Cesar Sayoc's van is seen in Boca Raton, Florida, on 18 October 2018 in this picture obtained from social media. Photograph: Social Media/Ed Kennedy via Reuters

#### most viewed in US



Pittsburgh synagogue shooting: suspect held after at least four killed



'I bet \$500 they are lying': Trump fans sceptical about pipe bomb arrest



'Why is that racist?' Trump greets young black leaders with freewheeling rant



Barack Obama takes aim at Donald Trump for 'making stuff up'



Muncy hits walk-off homer as Dodgers win longest ever World Series game



**Culture** 

Lifestyle

More ~



US midterms 2018 Business Tech Science

# Annotated to Cesar Sa

#### Trump m

Memes have become a popular, and seemingly effective, method to transmit ideology.

Memes have been weaponized

st vic

gh synagogue

sceptical about

oung black leaders heeling rant



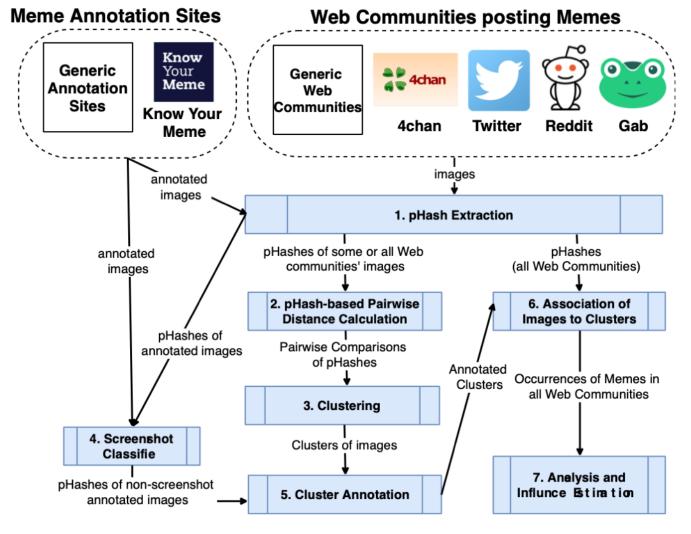


Muncy hits walk-off homer as Dodgers win longest ever World Series game



### Memes processing pipeline

### Memes



### **Datasets**











# of posts	1.4B	1.0B	48M	12M	15K
# of posts with images	242M	62M	13M	955K	15K
# of Images	114M	40M	4M	235K	706K
# of unique pHashes	74M	30M	3M	193K	597K

/pol/		Reddit		Gab		Twitter		
Entry	Posts (%)	Entry	Posts (%)	Entry	Posts (%)	Entry	Posts(%)	
Feels Bad Man/Sad Frog	64,367 (4.9%)	Manning Face	12,540 (2.2%)	Jesusland (P)	454 (1.6%)	Roll Safe	55,010 (5.9%)	
Smug Frog	63,290 (4.8%)	That's the Joke	7,626 (1.3%)	Demotivational Posters	414 (1.5%)	Evil Kermit	50,642 (5.4%)	
Happy Merchant (R)	49,608 (3.8%)	Feels Bad Man/ Sad Frog	7,240 (1.3%)	Smug Frog	392 (1.4%)	Arthur's Fist	37,591 (4.0%)	
Apu Apustaja	29,756 (2.2%)	Confession Bear	7,147 (1.3%)	Based Stickman (P)	391 (1.4%)	Nut Button	13,598 (1,5%)	
Pepe the Frog	25,197 (1.9%)	This is Fine	5,032 (0.9%)	Pepe the Frog	378 (1.3%)	Spongebob Mock	11,136 (1,2%)	
Make America Great Again (P)	21,229 (1.6%)	Smug Frog	4,642 (0.8%)	Happy Merchant (R)	297 (1.1%)	Reaction Images	9,387 (1.0%)	
Angry Pepe	20,485 (1.5%)	Roll Safe	4,523 (0.8%)	Murica	274 (1.0%)	Conceited Reaction	9,106 (1.0%)	
Bait this is Bait	16,686 (1.2%)	Rage Guy	4,491 (0.8%)	And Its Gone	235 (0.9%)	Expanding Brain	8,701 (0.9%)	
I Know that Feel Bro	14,490 (1.1%)	Make America Great Again (P)	4,440 (0.8%)	Make America Great Again (P)		Demotivational Posters	7,781 (0.8%)	
Cult of Kek	14,428 (1.1%)	Fake CCG Cards	4,438 (0.8%)	Feels Bad Man/ Sad Frog	206 (0.8%)	Cash Me Ousside/Howbow Dah	5,972 (0.6%)	
Laughing Tom Cruise	14,312 (1.1%)	Confused Nick Young	4,024 (0.7%)	Trump's First Order of Business (P)	192 (0.7%)	Salt Bae	5,375 (0.6%)	
Awoo	13,767 (1.0%)	Daily Struggle	4,015 (0.7%)	Kekistan	186 (0.6%)	Feels Bad Man/ Sad Frog	4,991 (0.5%)	
Tony Kornheiser's Why	13,577 (1.0%)	Expanding Brain	3,757 (0.7%)	Picardia (P)	183 (0.6%)	Math Lady/Confused Lady	4,722 (0.5%)	
Picardia (P)		Demotivational Posters	3,419 (0.6%)	Things with Faces (Pareidolia)	156 (0.5%)	Computer Reaction Faces	4,720 (0.5%)	
Big Grin / Never Ever	12,893 (1.0%)	Actual Advice Mallard	3,293 (0.6%)	Serbia Strong/Remove Kebab	149 (0.5%)	Clinton Trump Duet (P)	3,901 (0.4%)	
Reaction Images	12,608 (0.9%)	Reaction Images	2,959 (0.5%)	Riot Hipster	148 (0.5%)	Kendrick Lamar Damn Album Cover	3,656 (0.4%)	
Computer Reaction Faces	12,247 (0.9%)	Handsome Face	2,675 (0.5%)	Colorized History	144 (0.5%)	What in tarnation	3,363 (0.3%)	
Wojak / Feels Guy	11,682 (0.9%)	Absolutely Disgusting	2,674 (0.5%)	Most Interesting Man in World	140 (0.5%)	Harambe the Gorilla	3,164 (0.3%)	
Absolutely Disgusting	11,436 (0.8%)	Pepe the Frog	2,672 (0.5%)	Chuck Norris Facts	131 (0.4%)	I Know that Feel Bro	3,137 (0.3%)	
Spurdo Sparde		Pretending to be Retarded	2,462 (0.4%)	Roll Safe	131 (0.4%)	This is Fine	3,094 (0.3%)	
Total 44	15,179 (33.4%)	9.	4,069 (16.7%)	4,	808 (17.0%)	2	49,047 (26.4%)	



/pol/		Reddit		Gab		Twitter	
Entry	Posts (%)	Entry	Posts (%)	Entry	Posts (%)	Entry	Posts(%)
Feels Bad Man/Sad Frog	64,367 (4.9%)	Manning Face	12,540 (2.2%)	Jesusland (P)	454 (1.6%)	Roll Safe	55,010 (5.9%)
Smug Frog	63,290 (4.8%)	That's the Joke	7,626 (1.3%)	Demotivational Posters	414 (1.5%)	Evil Kermit	50,642 (5.4%)
Happy Merchant (R)	49,608 (3.8%)	Feels Bad Man/ Sad Frog	7,240 (1.3%)	Smug Frog	392 (1.4%)	Arthur's Fist	37,591 (4.0%)
Apu Apustaja	29,756 (2.2%)	Confession Bear	7,147 (1.3%)	Based Stickman (P)	391 (1.4%)	Nut Button	13,598 (1,5%)
Pepe the Frog	25,197 (1.9%)	This is Fine	5,032 (0.9%)	Pepe the Frog	378 (1.3%)	Spongebob Mock	11,136 (1,2%)
Make America Great Again (P)	21,229 (1.6%)	Smug Frog	4,642 (0.8%)	Happy Merchant (R)	297 (1.1%)	Reaction Images	9,387 (1.0%)
Angry Pepe	20,485 (1.5%)	Roll Safe	4,523 (0.8%)	Murica	274 (1.0%)	Conceited Reaction	9,106 (1.0%)
Bait this is Bait	16,686 (1.2%)	Rage Guy	4,491 (0.8%)	And Its Gone	235 (0.9%)	Expanding Brain	8,701 (0.9%)
I Know that Feel Bro	14,490 (1.1%)	Make America Great Again (P)	4,440 (0.8%)	Make America Great Again (P)	207 (0.8%)	Demotivational Posters	7,781 (0.8%)
Cult of Kek	14,428 (1.1%)	Fake CCG Cards	4,438 (0.8%)	Feels Bad Man/ Sad Frog	206 (0.8%)	Cash Me Ousside/Howbow Dah	5,972 (0.6%)
Laughing Tom Cruise	14,312 (1.1%)	Confused Nick Young	4,024 (0.7%)	Trump's First Order of Business (P	<ul><li>192 (0.7%)</li></ul>	Salt Bae	5,375 (0.6%)
Awoo	13,767 (1.0%)	Daily Struggle	4,015 (0.7%)	Kekistan	186 (0.6%)	Feels Bad Man/ Sad Frog	4,991 (0.5%)
Tony Kornheiser's Why	13,577 (1.0%)	Expanding Brain	3,757 (0.7%)	Picardia (P)	183 (0.6%)	Math Lady/Confused Lady	4,722 (0.5%)
Picardia (P)	13,540 (1.0%)	Demotivational Posters	3,419 (0.6%)	Things with Faces (Pareidolia)	156 (0.5%)	Computer Reaction Faces	4,720 (0.5%)
Big Grin / Never Ever	12,893 (1.0%)	Actual Advice Mallard	3,293 (0.6%)	Serbia Strong/Remove Kebab	149 (0.5%)	Clinton Trump Duet (P)	3,901 (0.4%)
Reaction Images	12,608 (0.9%)	Reaction Images	2,959 (0.5%)	Riot Hipster	148 (0.5%)	Kendrick Lamar Damn Album Cove	r 3,656 (0.4%)
Computer Reaction Faces	12,247 (0.9%)	Handsome Face	2,675 (0.5%)	Colorized History	144 (0.5%)	What in tarnation	3,363 (0.3%)
Wojak / Feels Guy	11,682 (0.9%)	Absolutely Disgusting	2,674 (0.5%)	Most Interesting Man in World	140 (0.5%)	Harambe the Gorilla	3,164 (0.3%)
Absolutely Disgusting		Pepe the Frog	2,672 (0.5%)	Chuck Norris Facts	131 (0.4%)	I Know that Feel Bro	3,137 (0.3%)
Spurdo Sparde	9,581 (0.7%)	Pretending to be Retarded	2,462 (0.4%)	Roll Safe	131 (0.4%)	This is Fine	3,094 (0.3%)
Total 44	45,179 (33.4%)	9.	4,069 (16.7%)	4	4,808 (17.0%)		249,047 (26.4%)









/pol/		Reddit		Gab		Twitter	
Entry	Posts (%)	Entry	Posts (%)	Entry	Posts (%)	Entry	Posts(%)
Feels Bad Man/Sad Frog	64,367 (4.9%)	Manning Face	12,540 (2.2%)	Jesusland (P)	454 (1.6%)	Roll Safe	55,010 (5.9%)
Smug Frog	63,290 (4.8%)	That's the Joke	7,626 (1.3%)	Demotivational Posters	414 (1.5%)	Evil Kermit	50,642 (5.4%)
Happy Merchant (R)	49,608 (3.8%)	Feels Bad Man/ Sad Frog	7,240 (1.3%)	Smug Frog	392 (1.4%)	Arthur's Fist	37,591 (4.0%)
Apu Apustaja	29,756 (2.2%)	Confession Bear	7,147 (1.3%)	Based Stickman (P)	391 (1.4%)	Nut Button	13,598 (1,5%)
Pepe the Frog	25,197 (1.9%)	This is Fine	5,032 (0.9%)	Pepe the Frog	378 (1.3%)	Spongebob Mock	11,136 (1,2%)
Make America Great Again (P)	21,229 (1.6%)	Smug Frog	4,642 (0.8%)	Happy Merchant (R)	297 (1.1%)	Reaction Images	9,387 (1.0%)
Angry Pepe	20,485 (1.5%)	Roll Safe	4,523 (0.8%)	Murica	274 (1.0%)	Conceited Reaction	9,106 (1.0%)
Bait this is Bait	16,686 (1.2%)	Rage Guy	4,491 (0.8%)	And Its Gone	235 (0.9%)	Expanding Brain	8,701 (0.9%)
I Know that Feel Bro	14,490 (1.1%)	Make America Great Again (P)	4,440 (0.8%)	Make America Great Again (P)	207 (0.8%)	Demotivational Posters	7,781 (0.8%)
Cult of Kek	14,428 (1.1%)	Fake CCG Cards	4,438 (0.8%)	Feels Bad Man/ Sad Frog	206 (0.8%)	Cash Me Ousside/Howbow Dah	5,972 (0.6%)
Laughing Tom Cruise	14,312 (1.1%)	Confused Nick Young	4,024 (0.7%)	Trump's First Order of Business (P)	192 (0.7%)	Salt Bae	5,375 (0.6%)
Awoo	13,767 (1.0%)	Daily Struggle	4,015 (0.7%)	Kekistan	186 (0.6%)	Feels Bad Man/ Sad Frog	4,991 (0.5%)
Tony Kornheiser's Why	13,577 (1.0%)	Expanding Brain	3,757 (0.7%)	Picardia (P)	183 (0.6%)	Math Lady/Confused Lady	4,722 (0.5%)
Picardia (P)	13,540 (1.0%)	Demotivational Posters	3,419 (0.6%)	Things with Faces (Pareidolia)	156 (0.5%)	Computer Reaction Faces	4,720 (0.5%)
Big Grin / Never Ever	12,893 (1.0%)	Actual Advice Mallard	3,293 (0.6%)	Serbia Strong/Remove Kebab	149 (0.5%)	Clinton Trump Duet (P)	3,901 (0.4%)
Reaction Images	12,608 (0.9%)	Reaction Images	2,959 (0.5%)	Riot Hipster	148 (0.5%)	Kendrick Lamar Damn Album Cover	3,656 (0.4%)
Computer Reaction Faces	12,247 (0.9%)	Handsome Face	2,675 (0.5%)	Colorized History	144 (0.5%)	What in tarnation	3,363 (0.3%)
Wojak / Feels Guy	11,682 (0.9%)	Absolutely Disgusting	2,674 (0.5%)	Most Interesting Man in World	140 (0.5%)	Harambe the Gorilla	3,164 (0.3%)
Absolutely Disgusting	11,436 (0.8%)	Pepe the Frog	2,672 (0.5%)	Chuck Norris Facts	131 (0.4%)	I Know that Feel Bro	3,137 (0.3%)
Spurdo Sparde	9,581 (0.7%)	Pretending to be Retarded	2,462 (0.4%)	Roll Safe	131 (0.4%)	This is Fine	3,094 (0.3%)
Total 44	15,179 (33.4%)	94	4,069 (16.7%)	4,	,808 (17.0%)	24	19,047 (26.4%)





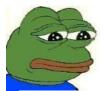








/pol/		Reddit		Gab		Twitter	
Entry	Posts (%)	Entry	Posts (%)	Entry	Posts (%)	Entry	Posts(%)
Feels Bad Man/Sad Frog	64,367 (4.9%)	Manning Face	12,540 (2.2%)	Jesusland (P)	454 (1.6%)	Roll Safe	55,010 (5.9%)
Smug Frog	63,290 (4.8%)	That's the Joke	7,626 (1.3%)	Demotivational Posters	414 (1.5%)	Evil Kermit	50,642 (5.4%)
Happy Merchant (R)	49,608 (3.8%)	Feels Bad Man/ Sad Frog	7,240 (1.3%)	Smug Frog	392 (1.4%)	Arthur's Fist	37,591 (4.0%)
Apu Apustaja	29,756 (2.2%)	Confession Bear	7,147 (1.3%)	Based Stickman (P)	391 (1.4%)	Nut Button	13,598 (1,5%)
Pepe the Frog	25,197 (1.9%)	This is Fine	5,032 (0.9%)	Pepe the Frog	378 (1.3%)	Spongebob Mock	11,136 (1,2%)
Make America Great Again (P)	21,229 (1.6%)	Smug Frog	4,642 (0.8%)	Happy Merchant (R)	297 (1.1%)	Reaction Images	9,387 (1.0%)
Angry Pepe	20,485 (1.5%)	Roll Safe	4,523 (0.8%)	Murica	274 (1.0%)	Conceited Reaction	9,106 (1.0%)
Bait this is Bait	16,686 (1.2%)	Rage Guy	4,491 (0.8%)	And Its Gone	235 (0.9%)	Expanding Brain	8,701 (0.9%)
I Know that Feel Bro	14,490 (1.1%)	Make America Great Again (P)	4,440 (0.8%)	Make America Great Again (P)	207 (0.8%)	Demotivational Posters	7,781 (0.8%)
Cult of Kek	14,428 (1.1%)	Fake CCG Cards	4,438 (0.8%)	Feels Bad Man/ Sad Frog	206 (0.8%)	Cash Me Ousside/Howbow Dah	5,972 (0.6%)
Laughing Tom Cruise	14,312 (1.1%)	Confused Nick Young	4,024 (0.7%)	Trump's First Order of Business (P	) 192 (0.7%)	Salt Bae	5,375 (0.6%)
Awoo	13,767 (1.0%)	Daily Struggle	4,015 (0.7%)	Kekistan	186 (0.6%)	Feels Bad Man/ Sad Frog	4,991 (0.5%)
Tony Kornheiser's Why	13,577 (1.0%)	Expanding Brain	3,757 (0.7%)	Picardia (P)	183 (0.6%)	Math Lady/Confused Lady	4,722 (0.5%)
Picardia (P)	13,540 (1.0%)	Demotivational Posters	3,419 (0.6%)	Things with Faces (Pareidolia)	156 (0.5%)	Computer Reaction Faces	4,720 (0.5%)
Big Grin / Never Ever	12,893 (1.0%)	Actual Advice Mallard	3,293 (0.6%)	Serbia Strong/Remove Kebab	149 (0.5%)	Clinton Trump Duet (P)	3,901 (0.4%)
Reaction Images	12,608 (0.9%)	Reaction Images	2,959 (0.5%)		148 (0.5%)	Kendrick Lamar Damn Album Cover	3,656 (0.4%)
Computer Reaction Faces	12,247 (0.9%)	Handsome Face	2,675 (0.5%)	Colorized History	144 (0.5%)	What in tarnation	3,363 (0.3%)
Wojak / Feels Guy		Absolutely Disgusting	2,674 (0.5%)	Most Interesting Man in World	140 (0.5%)	Harambe the Gorilla	3,164 (0.3%)
Absolutely Disgusting	11,436 (0.8%)	Pepe the Frog	2,672 (0.5%)	Chuck Norris Facts	, ,	I Know that Feel Bro	3,137 (0.3%)
Spurdo Sparde	9,581 (0.7%)	Pretending to be Retarded	2,462 (0.4%)	Roll Safe	131 (0.4%)	This is Fine	3,094 (0.3%)
Total 44	15,179 (33.4%)	9.	4,069 (16.7%)	4	,808 (17.0%)	24	19,047 (26.4%)













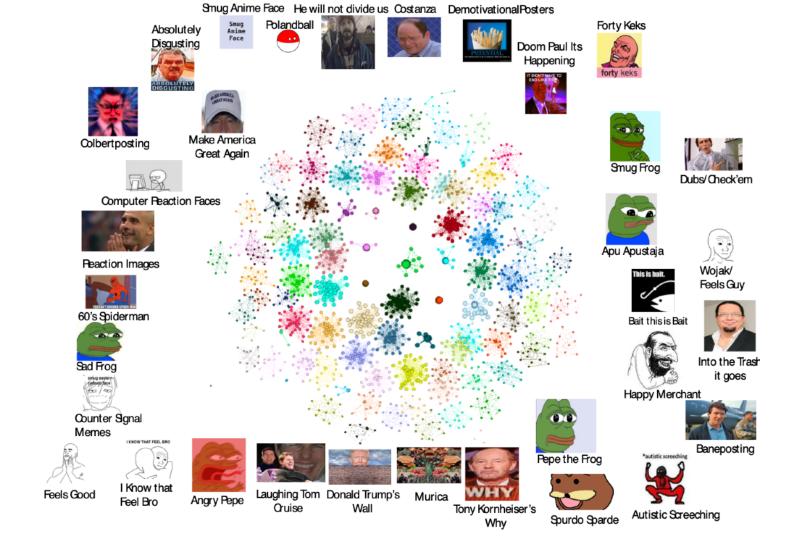


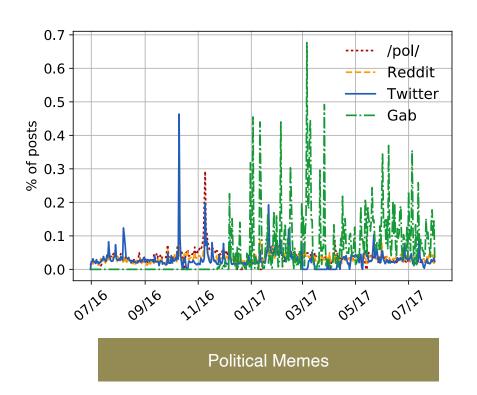


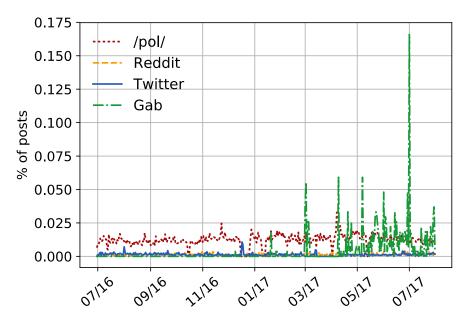
/pol/		Reddit		Gab		Twitter	
Entry	Posts (%)	Entry	Posts (%)	Entry	Posts (%)	Entry	Posts(%)
Feels Bad Man/Sad Frog	64,367 (4.9%)	Manning Face	12,540 (2.2%)	Jesusland (P)	454 (1.6%)	Roll Safe	55,010 (5.9%)
Smug Frog	63,290 (4.8%)	That's the Joke	7,626 (1.3%)	Demotivational Posters	414 (1.5%)	Evil Kermit	50,642 (5.4%)
Happy Merchant (R)	49,608 (3.8%)	Feels Bad Man/ Sad Frog	7,240 (1.3%)	Smug Frog	392 (1.4%)	Arthur's Fist	37,591 (4.0%)
Apu Apustaja	29,756 (2.2%)	Confession Bear	7,147 (1.3%)	Based Stickman (P)	391 (1.4%)	Nut Button	13,598 (1,5%)
Pepe the Frog	25,197 (1.9%)	This is Fine	5,032 (0.9%)	Pepe the Frog	378 (1.3%)	Spongebob Mock	11,136 (1,2%)
Make America Great Again (P)	21,229 (1.6%)	Smug Frog	4,642 (0.8%)	Happy Merchant (R)	297 (1.1%)	Reaction Images	9,387 (1.0%)
Angry Pepe	20,485 (1.5%)	Roll Safe	4,523 (0.8%)	Murica	274 (1.0%)	Conceited Reaction	9,106 (1.0%)
Bait this is Bait	16,686 (1.2%)	Rage Guy	4,491 (0.8%)	And Its Gone	235 (0.9%)	Expanding Brain	8,701 (0.9%)
I Know that Feel Bro	14,490 (1.1%)	Make America Great Again (P)	4,440 (0.8%)	Make America Great Again (P)	207 (0.8%)	Demotivational Posters	7,781 (0.8%)
Cult of Kek	14,428 (1.1%)	Fake CCG Cards	4,438 (0.8%)	Feels Bad Man/ Sad Frog	206 (0.8%)	Cash Me Ousside/Howbow Dah	5,972 (0.6%)
Laughing Tom Cruise	14,312 (1.1%)	Confused Nick Young	4,024 (0.7%)	Trump's First Order of Business (F	<ul><li>192 (0.7%)</li></ul>	Salt Bae	5,375 (0.6%)
Awoo	13,767 (1.0%)	Daily Struggle	4,015 (0.7%)	Kekistan	186 (0.6%)	Feels Bad Man/ Sad Frog	4,991 (0.5%)
Tony Kornheiser's Why	13,577 (1.0%)	Expanding Brain	3,757 (0.7%)	Picardia (P)	183 (0.6%)	Math Lady/Confused Lady	4,722 (0.5%)
Picardia (P)	13,540 (1.0%)	Demotivational Posters	3,419 (0.6%)	Things with Faces (Pareidolia)	156 (0.5%)	Computer Reaction Faces	4,720 (0.5%)
Big Grin / Never Ever	12,893 (1.0%)	Actual Advice Mallard	3,293 (0.6%)	Serbia Strong/Remove Kebab	149 (0.5%)	Clinton Trump Duet (P)	3,901 (0.4%)
Reaction Images	12,608 (0.9%)	Reaction Images	2,959 (0.5%)	Riot Hipster	148 (0.5%)	Kendrick Lamar Damn Album Cove	r 3,656 (0.4%)
Computer Reaction Faces	12,247 (0.9%)	Handsome Face	2,675 (0.5%)	Colorized History	144 (0.5%)	What in tarnation	3,363 (0.3%)
Wojak / Feels Guy	11,682 (0.9%)	Absolutely Disgusting	2,674 (0.5%)	Most Interesting Man in World	140 (0.5%)	Harambe the Gorilla	3,164 (0.3%)
Absolutely Disgusting	11,436 (0.8%)	Pepe the Frog	2,672 (0.5%)	Chuck Norris Facts	131 (0.4%)	I Know that Feel Bro	3,137 (0.3%)
Spurdo Sparde	9,581 (0.7%)	Pretending to be Retarded	2,462 (0.4%)	Roll Safe	131 (0.4%)	This is Fine	3,094 (0.3%)
Total 44	15,179 (33.4%)	9.	4,069 (16.7%)		4,808 (17.0%)	2	49,047 (26.4%)

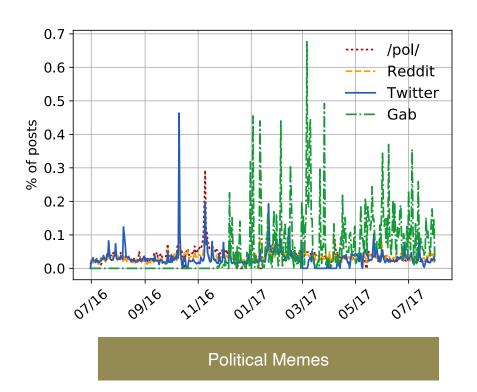
### Who are the most popular people in memes?

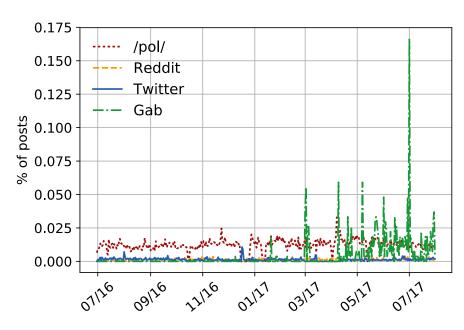
/po	<b>I</b> /	Redo	lit	Gal	b	Twitter	
Entry	Posts (%)	Entry	Posts (%)	Entry	Posts (%)	Entry	Posts(%)
Donald Trump	60,611 (4.6%)	Donald Trump	34,533 (6.1%)	Donald Trump	1,665 (6.1%)	Donald Trump	10,208 (1.3%)
Adolf Hitler	8,759 (0.6%)	Steve Bannon	3,733 (0.6%)	Mitt Romney	455 (1.7%)	Barack Obama	5,187 (0.6%)
Mike Pence	4,738 (0.3%)	Stephen Colbert	3,121 (0.6%)	Bill Nye	370 (1.3%)	Chelsea Manning	4,173 (0.5%)
Jeb Bush	4,217 (0.3%)	Chelsea Manning	2,261 (0.4%)	Adolf Hitler	106 (0.4%)	Kim Jong Un	3,271 (0.4%)
Vladimir Putin	3,218 (0.2%)	Ben Carson	2,148 (0.4%)	Barack Obama	104 (0.4%)	Anita Sarkeesian	2,764 (0.3%)
Alex Jones	3,206 (0.2%)	Bernie Sanders	1,757 (0.3%)	Isis Daesh	92 (0.3%)	Bernie Sanders	2,277 (0.3%)
Ron Paul	3,116 (0.2%)	Ajit Pai	1,658 (0.3%)	Death Grips	91 (0.3%)	Vladimir Putin	1,733 (0.2%)
Bernie Sanders	3,022 (0.2%)	Barack Obama	1,628 (0.3%)	Eminem	89 (0.3%)	Billy Mays	1,454 (0.2%)
Massimo D'alema	2,725 (0.2%)	Gabe Newell	1,518 (0.3%)	Kim Jong Un	87 (0.3%)	Adolf Hitler	1,304 (0.2%)
Mitt Romney	2,468 (0.2%)	Bill Nye	1,478 (0.3%)	Ajit Pai	76 (0.3%)	Kanye West	1,261 (0.2%)
Chelsea Manning	2,403 (0.2%)	Hillary Clinton	1,468 (0.3%)	Pewdiepie	73 (0.3%)	Bill Nye	968 (0.2%)
Hillary Clinton	2,378 (0.2%)	Death Grips	1,463 (0.3%)	Bernie Sanders	71 (0.3%)	Mitt Romney	923 (0.1%)
A. Wyatt Mann	2,110 (0.2%)	Adolf Hitler	1,449 (0.3%)	Alex Jones	70 (0.3%)	Filthy Frank	777 (0.1%)
Ben Carson	1,780 (0.1%)	Mitt Romney	1,294 (0.2%)	Hillary Clinton	59 (0.2%)	Hillary Clinton	758 (0.1%)
Filthy Frank	1,598 (0.1%)	Eminem	1,274 (0.2%)	Anita Sarkeesian	54 (0.2%)	Ajit Pai	715 (0.1%)



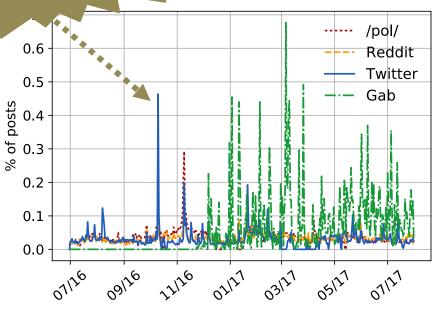


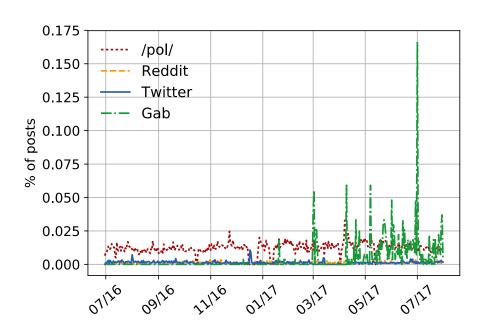




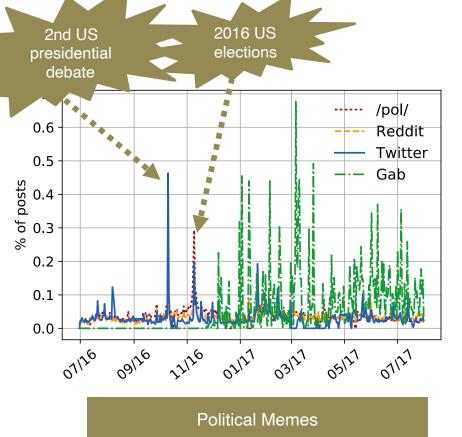


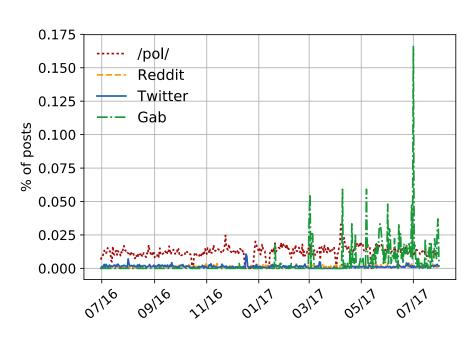
2nd US presidential debate

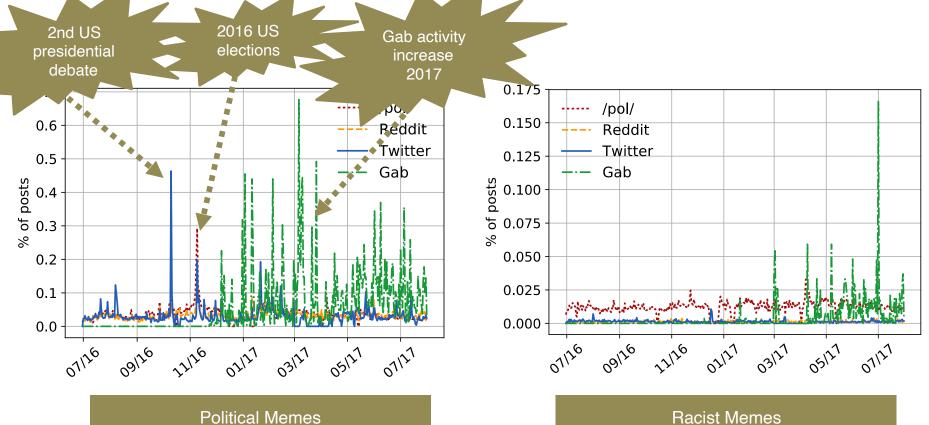


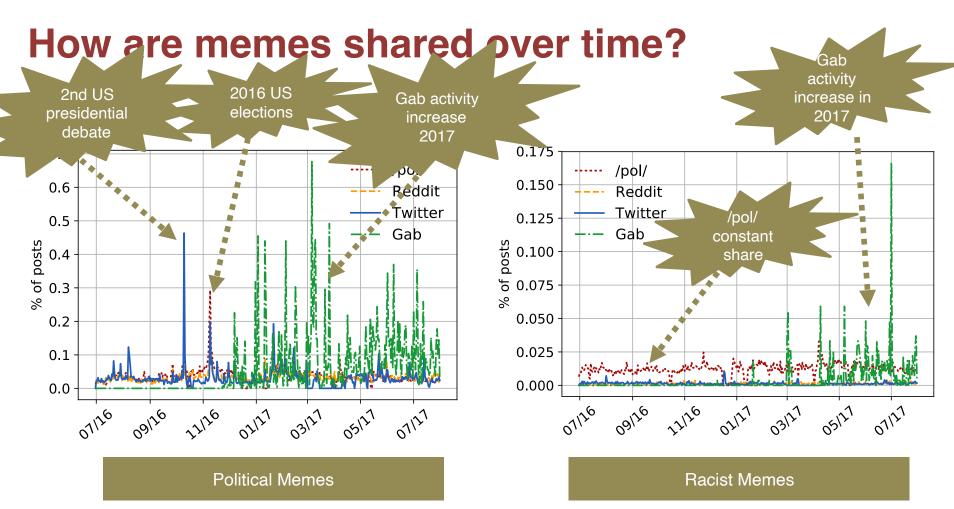


**Political Memes** 









### Communities' influence (racist memes)

/pol/ is most influential in terms of spreading racist memes

	/pol/	Reddit	Twitter	Gab	T_D
/lod/	R: 99.44%	R: 6.89%	R: 3.94%	R: 18.93%	R: 14.17%
	NR: 97.12%*	NR: 3.95%*	NR: 2.93%*	NR: 12.91%	NR: 16.41%*
Reddit	R: 0.29%	R: 88.95%	R: 2.79%	R: 1.25%	R: 9.84%
	NR: 1.28%*	NR: 90.87%*	NR: 4.64%*	NR: 9.42%	NR: 9.02%*
Source	R: 0.15%	R: 1.75%	R: 92.95%	R: 0.66%	R: 1.37%
Twitter	NR: 0.79%	NR: 2.85%*	NR: 90.96%*	NR: 8.16%	NR: 3.66%*
Gab	R: 0.04%	R: 0.65%	R: 0.10%	R: 76.67%	R: 0.08%
	NR: 0.09%	NR: 0.16%	NR: 0.20%	NR: 59.72%	NR: 0.58%
٥٢	R: 0.07%	R: 1.75%	R: 0.23%	R: 2.49%	R: 74.53%
	NR: 0.73%*	NR: 2.18%*	NR: 1.27%*	NR: 9.79%	NR: 70.32%*

Destination

### Communities' efficiency (racist memes)

If we look at the influence normalized to the number of memes posted, the The Donald is most efficient in terms of disseminating memes

	/pol/	Reddit	Twitter	Gab	T_D	Total Ext
/lod/	R: 99.4	R: 0.5	R: 0.3	R: 0.2	R: 0.1	R: 1.1
	NR: 97.1*	NR: 1.5*	NR: 1.4*	NR: 0.4	NR: 0.9*	NR: 4.2
Reddit	R: 4.2	R: 89.0	R: 2.7	R: 0.2	R: 1.4	R: 8.5
	NR: 3.3*	NR: 90.9*	NR: 5.7*	NR: 0.7	NR: 1.3*	NR: 11.1
Source	R: 2.2	R: 1.8	R: 92.9	R: 0.1	R: 0.2	R: 4.4
Twitter	NR: 1.7	NR: 2.3*	NR: 91.0*	NR: 0.5	NR: 0.4*	NR: 4.9
Gab	R: 4.5	R: 4.9	R: 0.7	R: 76.7	R: 0.1	R: 10.3
	NR: 3.0	NR: 2.0	NR: 3.2	NR: 59.7	NR: 1.1	NR: 9.3
O_	R: 7.5	R: 12.1	R: 1.5	R: 2.3	R: 74.5	R: 23.4
	NR: 13.6*	NR: 15.5*	NR: 11.1*	NR: 5.3	NR: 70.3*	NR: 45.6
•			•	Destination		·

#### What about state-level adversaries?

#### What about state-level adversaries?

# Raise your hand if you're under investigation!



















b)



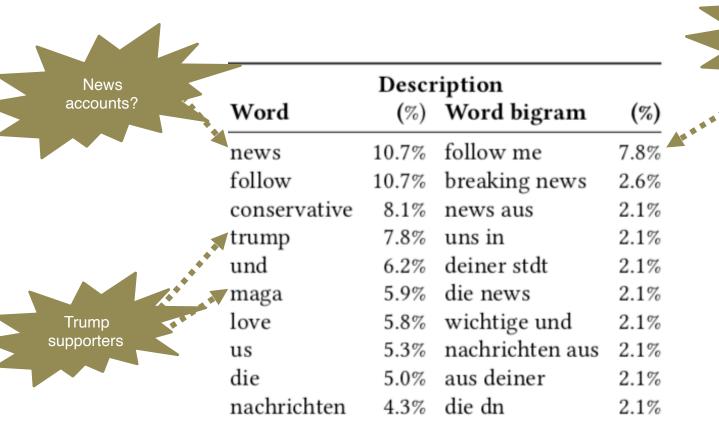
a)

#### **Research Questions**

- How do state-sponsored actors operate and evolve?
- How does the behavior of state-sponsored trolls compare to random users?

 More importantly, what was their influence on the Web with respect to the dissemination of news?

### What are they posing as?



Nudging users to follow them

### Where are they allegedly posting from?



### What hashtags they shared?

Trolls							
Hashtag	(%)	Hashtag	(%)	Hashtag	(%)	Hashtag	(%)
news	7.2%	US	0.7%	iHeartAwards	1.8%	UrbanAttires	0.6%
politics	2.6%	tcot	0.6%	BestFanArmy	1.6%	Vacature	0.6%
sports	2.1%	PJNET	0.6%	Harmonizers	1.0%	mPlusPlaces	0.6%
business	1.4%	entertainment	0.5%	iOSApp	0.9%	job	0.5%
money	1.3%	top	0.5%	JouwBaan	0.9%	Directioners	0.5%
world	1.2%	topNews	0.5%	vacature	0.9%	JIMIN	0.5%
MAGA	0.8%	ISIS	0.4%	KCA	0.9%	PRODUCE101	0.5%
health	0.8%	Merkelmussbleiben	0.4%	Psychic	0.8%	VoteMainFPP	0.5%
local	0.7%	IslamKills	0.4%	RT	0.8%	Werk	0.4%
Black Lives Matter	0.7%	breaking	0.4%	Libertad2016	0.6%	dts	0.4%

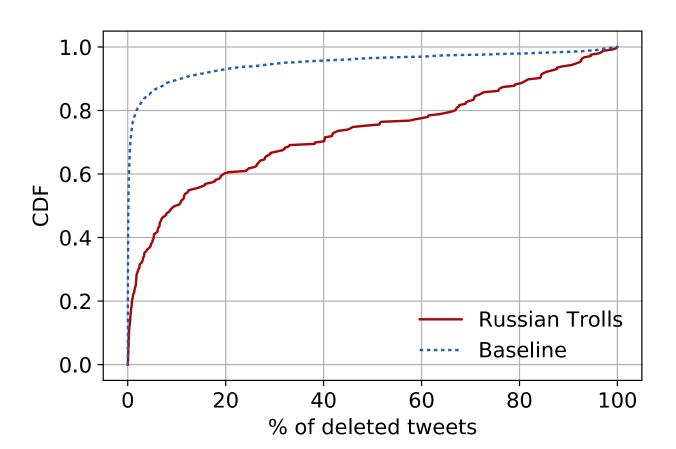
### Do Russian trolls change their screen names?

9% of the Russian troll accounts changed their screen name Up to 4 times per account

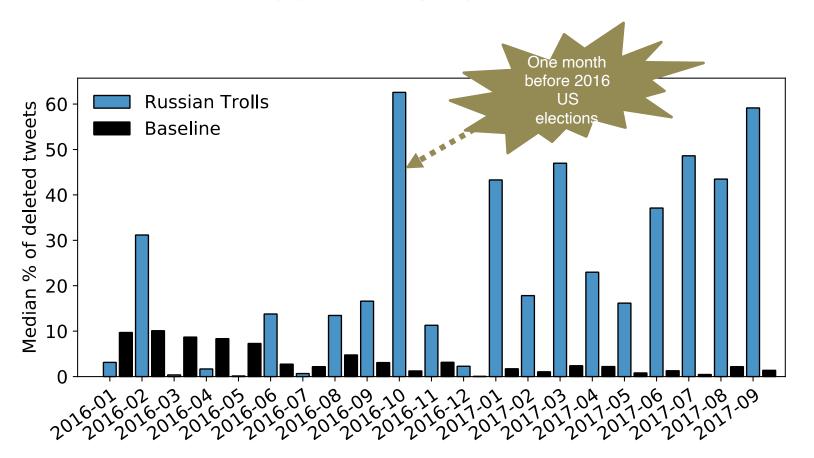
E.g., from "OnlineHouston" to "HoustonTopNews" Clear attempt to pose as local news outlet

In our baseline dataset 19% of the accounts changed their screen name Up to 11 times per account

#### Do Russian trolls delete their tweets?



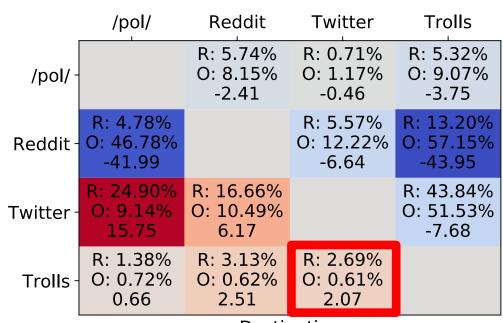
#### When did Russian trolls deleted their tweets?



#### Influence

1K trolls caused 2.6% of all Russian state-sponsored news outlets URLs (i.e., RT) on Twitter's 1%.

1K trolls caused 0.6% of all other news outlets URLs on Twitter's 1%.



Destination

## **Main Findings**

We find differences in the use of the Twitter platform between trolls and random users

Trolls seem to reset their "personas" by changing names and deleting tweets

Particularly influential in spreading Russian state-sponsored URLs on Twitter and other platforms

#### **Data Collection**

Crawling 4chan, Reddit, Gab, Twitter, YouTube Processing, cleaning, storing, indexing, etc.

#### **Data Collection**

Crawling 4chan, Reddit, Gab, Twitter, YouTube Processing, cleaning, storing, indexing, etc.

#### **Data Analysis**

Statistical tools to study distributions, behaviors, etc.

Longitudinal analysis, understanding trends and change points

Influence estimation

Image analysis

Clustering, contextual analysis, unsupervised learning Language modeling, topic extraction Classification and prediction

[....]

**Exploratory & hypotheses-driven** 

#### **Data Collection**

Crawling 4chan, Reddit, Gab, Twitter, YouTube Processing, cleaning, storing, indexing, etc.

#### **Data Sharing**

Everything open source
Easy to use, facilitate qualitative/
social science research

#### **Data Analysis**

Statistical tools to study distributions, behaviors, etc.

Longitudinal analysis, understanding trends and change points

Influence estimation

Image analysis

Clustering, contextual analysis, unsupervised learning
Language modeling, topic extraction
Classification and prediction

[...]

**Exploratory & hypotheses-driven** 

There are still many problems remaining in the broad space

There are still many problems remaining in the broad space We have made a lot of progress into gaining a quantitative understanding

There are still many problems remaining in the broad space

We have made a lot of progress into gaining a quantitative understanding

We now need to deploy proactive systems in production

Real time detection of harassment and online attacks

Actively fight weaponized information

**Policy** 

Accountability

# Thank you!











