

# Studying jerks on the Internet: A Data-Driven Approach

#### **Emiliano De Cristofaro**



(Thanks to Jeremy Blackburn and Savvas Zannettou for most of the slides)



#### **Outline**

- 1. Hate on and "raids" from fringe communities like 4chan [ICWSM 2017]
- 2. Influence of fringe communities on misinformation [IMC 2017]
- 3. Misuse of Web archiving services [ICWSM 2018]

More work on online hate, cyberbullying, etc: <a href="https://encase.socialcomputing.eu/publications">https://encase.socialcomputing.eu/publications</a>



### What is 4chan?



#### 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.

## What is 4chan?



Boards				filter
Boards  Japanese Culture  Anime & Manga  Anime/Cute  Anime/Wallpapers  Mecha Cosplay & EGL Cute/Male Flash Transportation Otaku Culture Video Games Video Game Generals Pokémon	Interests Comics & Cartoons Technology Television & Film Weapons Auto Animals & Nature Traditional Games Sports Alternative Sports Science & Math History & Humanities International Outdoors	Creative Oekaki Papercraft & Origami Photography Food & Cooking Artwork/Critique Wallpapers/General Literature Music Fashion 3DCG Graphic Design Do-It-Yourself Worksafe GIF	Other Business & Finance Travel Fitness Paranormal Advice LGBT Pony Current News Worksafe Requests Very Important Posts Misc. (NSFW) Random ROBOT9001	Adult (NSFW) Sexy Beautiful Women Hardcore Handsome Men Hentai Ecchi Yuri Hentai/Alternative Yaoi Torrents High Resolution Adult GIF Adult Cartoons Adult Requests
Retro Games	Toys	Quests	Politically Incorrect International/Random Cams & Meetups Shit 4chan Says	Addit Neguests

## What is 4chan?



-6



#### WHERE DID EVERYTHING

# GO SO REICH



## Why Do We Care About 4Chan?

"4chan is niche and irrelevant" – Anonymous WWW'17 Reviewer

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



Leslie Jones hack: Inside the alt-right's racist celebration of Jones' humiliation



4chan may have remotely wiped a Clinton official's iPhone and iPad

Related: 4chan is the worst

VE

#### **BUSINESS**

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

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

#MediasNextTrumpLie:

Inside 4chan's pro-Trump spam campaign to derail the media



GET MIC

**TAYWARD YOUTH** 3/24/16 3:10 PM

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

## /pol/ - Politically Incorrect Board





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

Anonymous (ID: №96TFJC) = 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



## In This Talk: Challenges of Measuring 4chan

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

## **Anonymity & Ephemerality**

#### 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 (More later)

## **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						

#### Methodology:

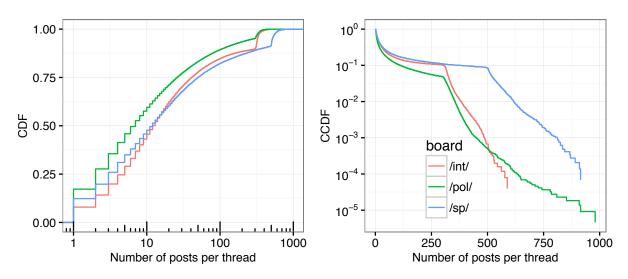
Visit the "catalog"

Take a snapshot every 5 minutes

Once a thread is pruned, retrieve full/final contents from archive

## We're still crawling...

## **Ephemerality: The Bump System**



Limit boards to N live threads

Threads ordered by MRU

A new post in a thread "bumps" it up to the top

Create new thread → Old thread dies

Bump limit

Max times thread can be bumped No discussion will dominate forever

## **Geographic Distribution of Users**



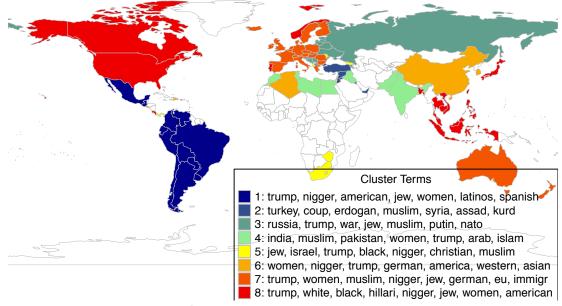
1 09/11/17(Mon)07:17:26

/pol/ users seems well distributed

Native English speaking countries most highly represented

Plenty of other countries really well represented too though!

## **Are Flags Trustworthy?**



Use spectral clustering of the topics that each country posts about

The clusters follow real world socio-political blocks

While flags are not perfect, they seem reasonable

## **Hate Speech?**



Crowdsourced dictionary

Manually filtered a bit

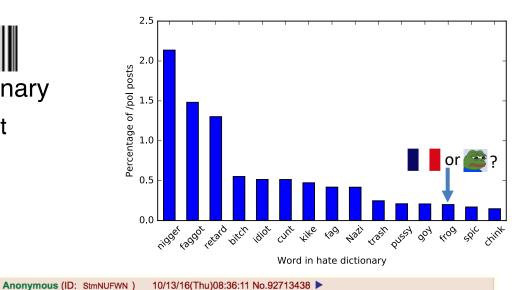
/pol/ by far most hate speech use

/pol/ 12%

/sp/ 7.3%

/int/ 6.3%

Twitter 2.2%



Can't believe I sat and read all this shit, most interesting part was the countries that get the most replies.

Trolls now have the list of the top 10 countries to use to bait.

Also, the fact "trash" is the 10th most popular "hate speech" word, fucking burgers and leafs.

NIGGER NI

#### Raids

Attempts to disrupt another site

Not a DDoS

Disrupts community that calls service home, not the service itself

Raids are a favorite past time on 4chan

"Pool's closed!"

Have become less "funny" and more "scary" lately

It's a socio-technical problem

#### YouTube Raids

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 nigger!"

If raid is taking place:

Peak in YouTube comments while thread alive?

/pol/ thread and YT comments synchronized?

#### Raids



Hans2k4ever 1 month ago

When will the normies learn?

Reply • 181 if 🥍

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

Reply • 13



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!

Reply • if 🏴



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

Reply • 227

View all 16 replies ∨



Szgerle 1 month ago

Not even just your regular nazi but ULTRA NAZI!

Reply • 1



Drakewood 1 month ago

+Potato Jenova A FUCKING LEAF!

Reply · if 🏴

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

Certain boards on 4chan additionally include the



Emiliano De Cristofaro

6:31 / 34:40

11,013 views





**#** 🗆 🔡

Thought Criminal 1 month ago

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

Reply • 169 🍎 🟴



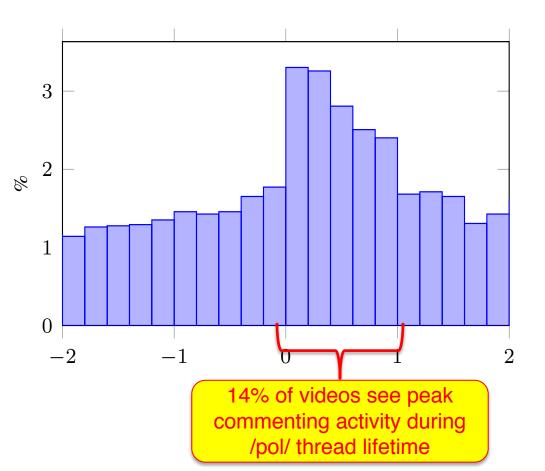








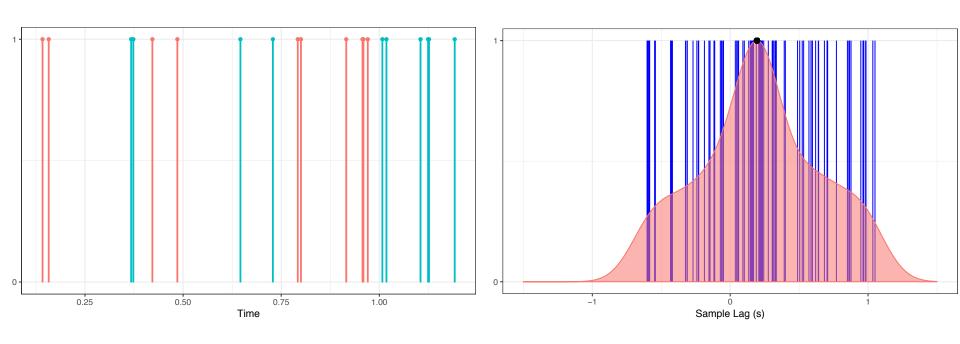
## **Activity Peaks**



YT videos with peaks during 4chan thread

Determined via PDF of commenting timeseries

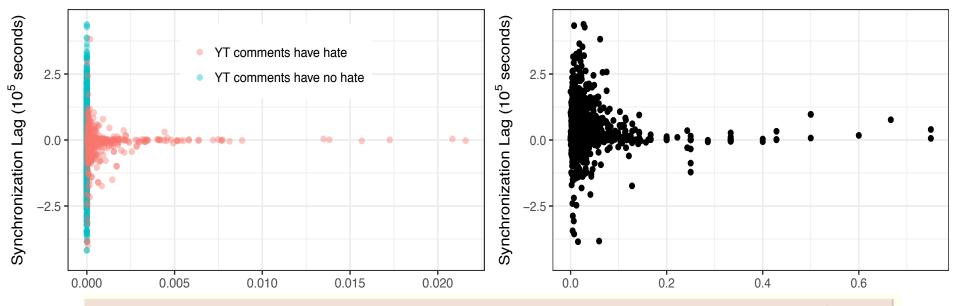
## **Synchronization**



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**



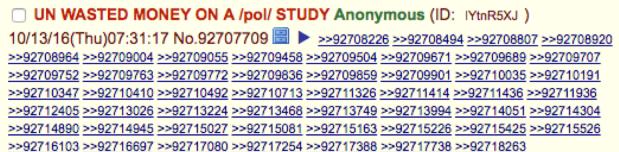
**Anonymous** (ID: DaWZlbeh ) 10/13/16(Thu)20:08:55 No.92772435 ▶

>>92765569

Physicist finalist here, can confirm the maths is legitimate. Holy fuck



and I be married in the first property of the first products of the paper.



https://arxiv.org/pdf/1610.03452v1.pdf

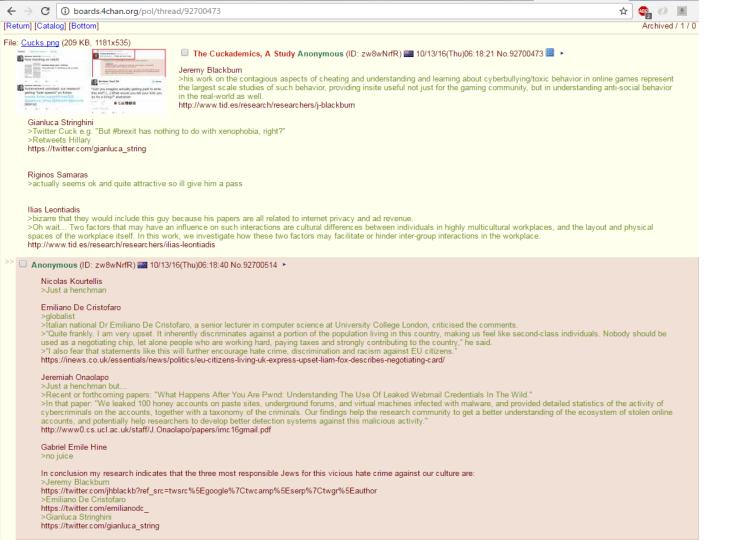
series," was viewed by proper first \$8.000 report, series get in here fellow niggers, post your least rare Pepes



#### Sam Hyde

2016-11-17 at 22:05 UTC

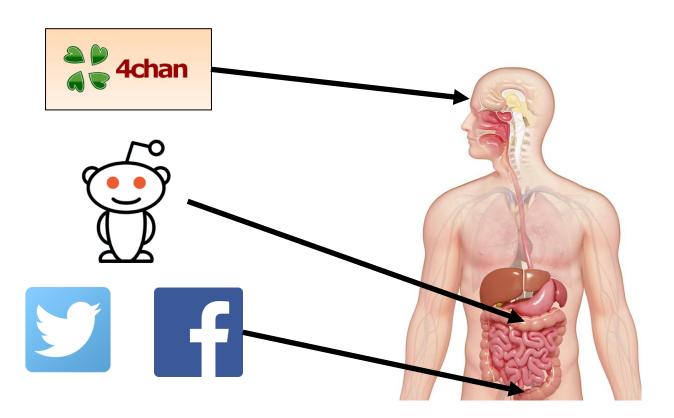
When U.N. sends its Pepes, they're not sending their best. They're not sending Kek, they're not sending rares. They're sending Pepes that have lots of copies, and they're bringing those copies with us. They're bringing tumblr. They're bringing facebook. They're common. And some, I assume, are rare Pepes. 26



#### **Outline**

- 1. Hate on and "raids" from fringe communities like 4chan [ICWSM 2017]
- 2. Influence of fringe communities on misinformation [IMC 2017]
- 3. Misuse of Web archiving services [ICWSM 2018]

## Can We Quantify How Rest of Web is Affected?



# The Information Ecosystem













































SauceNAO 🕹 C58mabBXEAEm6iY.jpg, 46KiB, 497x503

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

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.



54,218 Retweets 148,062 Likes 👸 🎥 🏶 🐞 🚳 🗳

9:54 AM - 3 Mar 2017







Donald J. Trump 📀 @realDonaldTrump

#### #FraudNewsCNN #FNN

4:21 PM - Jul 2, 2017

141,936 🕽 368,969

C 605.362



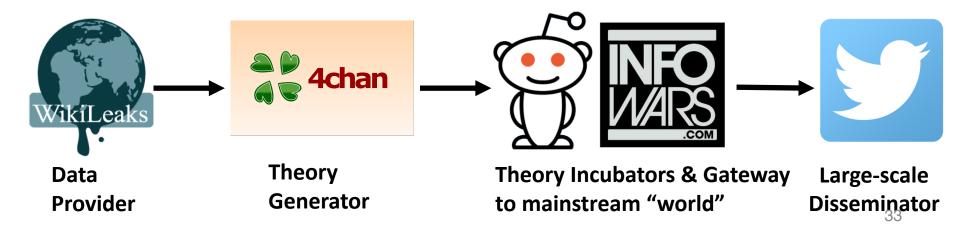
#### Trump's anti-CNN tweet originated from Reddit's largest right-wing extremist forum

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

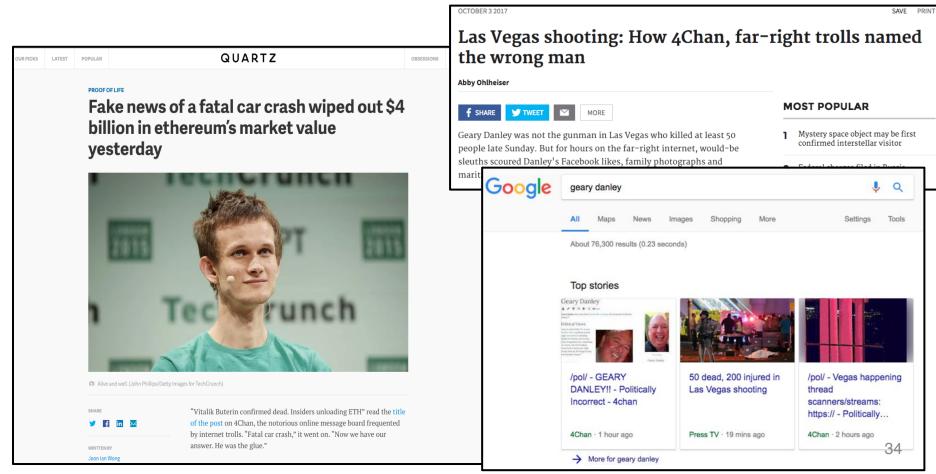
Updated by Aja Romano | @ajaromano | Jul 3, 2017, 5:00pm EDT

## The Pizzagate Conspiracy Theory





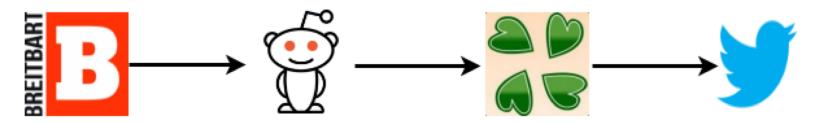
## Oh Hey... It's 4chan Again...



#### Idea...

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

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



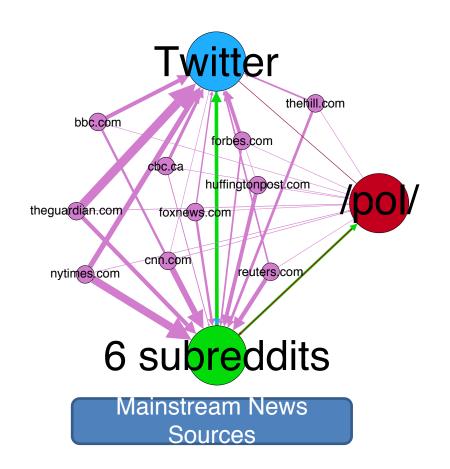
Build a graph with the sequences

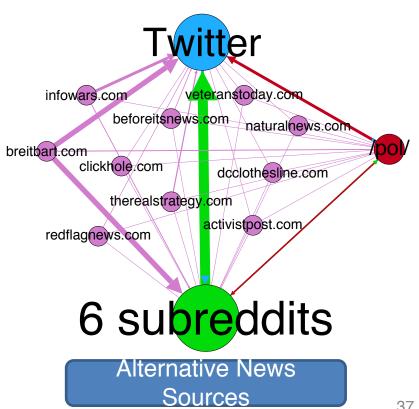
## The Data

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





## **Quantify Influence? Hawkes Processes!**

#### Assume K processes

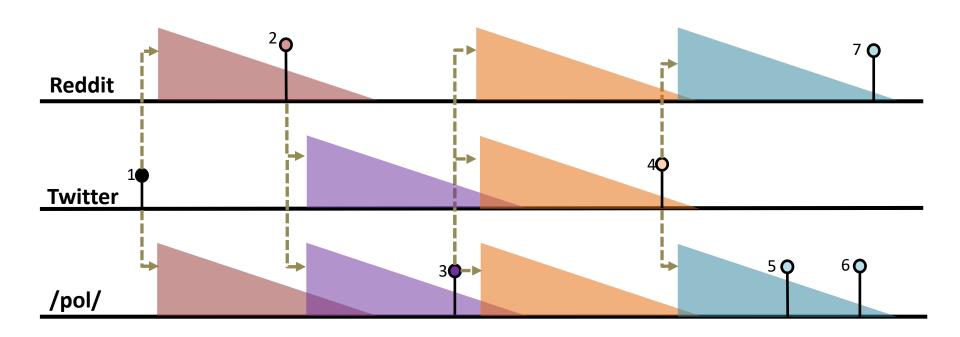
Each with a rate of events (i.e., posting of a URL), called the background rate

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** 

## **Hawkes Processes Example**



## For Our Purposes...

#### Hawkes model with 8 processes

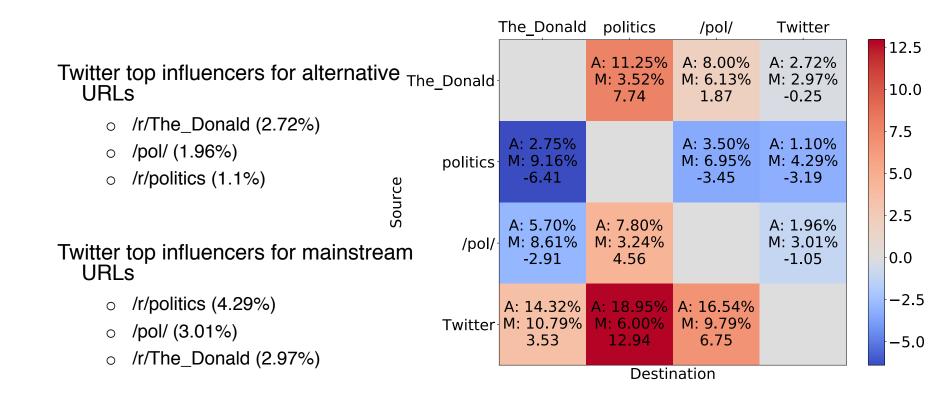
One for each platform/community

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

### What Communities Influence Each Other?





#### **Outline**

- 1. Hate on and "raids" from fringe communities like 4chan [ICWSM 2017]
- 2. Influence of fringe communities on misinformation [IMC 2017]
- 3. Misuse of Web archiving services [ICWSM 2018]

## Web archiving

#### Archive.is

On-demand archiving service

Generates 5-character archive URLs

E.g., <u>www.google.de</u> becomes <u>archive.is/HVbU</u>



#### Wayback Machine

Works mainly through a proactive crawler

Supports also on-demand archival

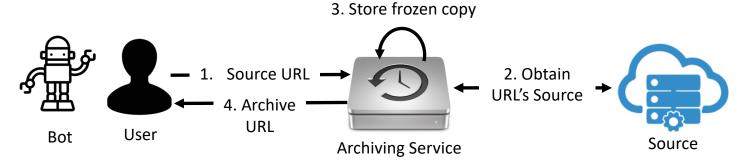
E.g., <u>www.google.com</u> becomes

https://web.archive.org/web/20100205062719/http://www.google.com/

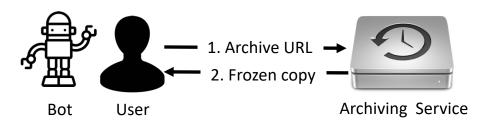


## How do Web archiving services work?

#### A. Archival

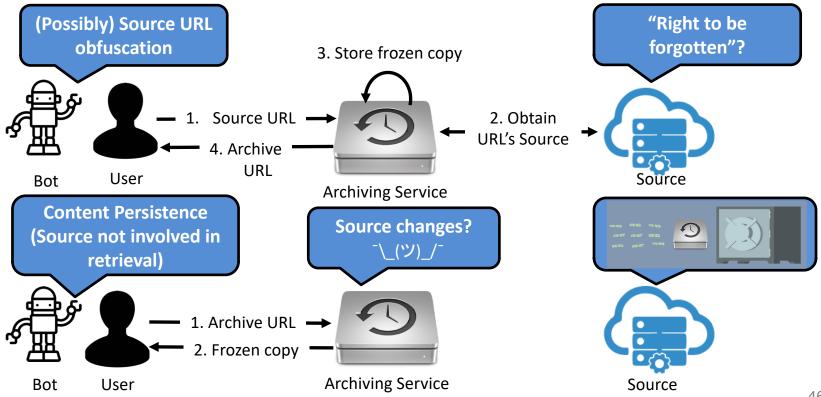


#### **B.** Retrieval





## How do Web archiving services work?



#### **Research Questions**

1. What kind of content get archived, by whom, and why?

2. How are archive URLs disseminated on the Web?

3. Are archiving services misused?









#### **Datasets**

1. Live Feed archive is dataset

2. Archive.is and Wayback Machine URLs posted on Reddit, /pol/, Gab, Twitter

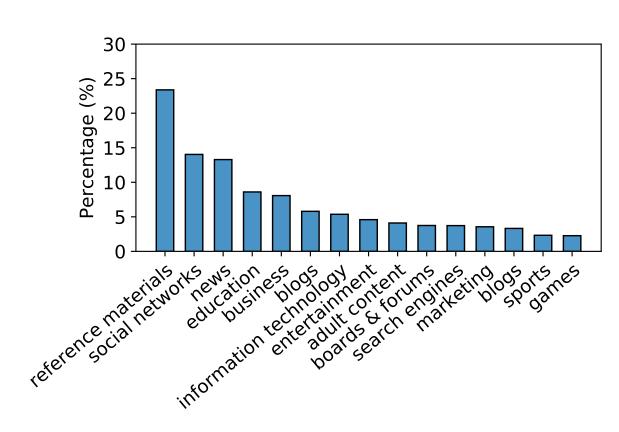
	platform	Archive URLs	Source URLs	Source Domains
Live Feed	Archive.is	21M	20.6M	5.3M
Reddit	Archive.is	310K	291K	15.9K
	Wayback	387K	343K	21K
/pol/	Archive.is	36K	33K	3.9K
	Wayback	2K	2K	0.9K
Gab	Archive.is	5.9K	5.7K	1.3K
	Wayback	0.3K	0.3K	0.2K
Twitter	Archive.is	3.7K	3.6K	0.8K
	Wayback	1.2K	1.2K	0.8K

## What kind of content gets archived?

Get categories for domains using Virus Total API

Live feed dataset has 5.3M domains

Select top 100K domains for categorization



### What kind of content gets archived? (OSNs)



OSN posts among the top two categories in all Web Communities





News among the top five categories in all Web Communities





## Platform-specific censorship

Some domains exclusively shared via archiving services. Why?

8chan and Facebook considered spam (rejected) on /pol/

Users utilize archive.is to circumvent this platform-specific censorship

Evidence of accidental news source censorship because of substitution filters (used for fun)

smh becomes baka (smh.com.au becomes baka.com.au)

# **Original content availability**

	platform	% of unavailable URLs
Reddit	Archive.is	7%
	Wayback	11%
/pol/	Archive.is	18%
	Wayback	34%
Gab	Archive.is	13%
	Wayback	52%
Twitter	Archive.is	24%
	Wayback	51%

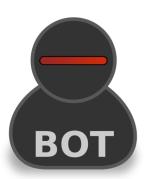
## Who is archiving URLs? Reddit

31% and 82% of archive.is and Wayback Machine URLs, resp., are posted by "SnapshillBot"



44% of archive.is URLs are shared by bots, 85% for Wayback Machine

→ Content persistence within subreddits?



# Speaking of subreddits.....

	Subreddit (archive.is)	(%)	Subreddit (Wayback)	(%)
Trump	The_Donald	24.48%	EnoughTrumpSpam	31.82%
	KotakuInAction	15.83%	MGTOW	7.38%
Trump	EnoughTrumpSpam	12.06%	SnapshillBotEx	7.19%
	MGTOW	3.48%	undelete	5.90%
	undelete Gaming	2.74%	SubredditDrama	5.50%
Drama	SubredditD	2.61%	Drama	5.03%
Drama	Drama	2.33%	Gamingcirclejerk	3.47%
	Gamingcirclejerk	1.57%	ShitAmericansSay	1.63%
Conspiracies	conspiracy	1.44%	TopMindsOfReddit	1.51%
	MensRights	1.12%	TheBluePill	1.25%
	savedyouaclick	1.00%	Buttcoin_1000	1.15%
Politics	politics	0.98%	AgainstHateSubreddits	1.06%
	DerekSmart	0.76%	subredditcancer	0.99%
	ShitAmericansSay	0.75%	The_Donald	0.95%
Politics	PoliticsAll	0.72%	badeconomics	0.75%

### **Ad-Revenue Deprivation**

#### post reply

from AutoModerator [M] via /r/The\_Donald sent 5 days ago

Your submission was removed because it is from washingtonpost.com, which has been identified as a severely anti-Trump domain. Please submit a cached link or screenshot when submitting content from this domain. We recommend using http://www.archive.is for this purpose.

I am a bot, and this action was performed automatically. Please contact the moderators of this subreddit if you have any questions or concerns.

context full comments (1) report block user mark unread reply

## News "Censorship"

13K submissions removed from The\_Donald subreddit

In total 23 "anti-Trump" news sources censored



News Source	# of submissions removed	Percentage
washightonpost.com	3.8K	44%
cnn.com	3.3K	40%
nydailynews.com	1.0K	46%
huffingtonpost.com	0.9K	43%
nationalreview.com	0.7K	45%

# Thank you!







#### **CUCKADEMICS**

#### Primary members



Jeremy Blackburn University of Alabama at Birmingham



Emiliano De Cristofaro University College London



Guillermo Suarez-Tangil King's College London



Nicolas Kourtellis Telefonica Research



Ilias Leontiadis Telefonica Research



Michael Sirivianos Cyprus University of Technology



Gianluca Stringhini University College London

#### **DRAMANAUTS**

#### Students and other affiliates



Tristan Caulfield
University College London



Despoina Chatzakou Aristotle University of



Antigoni Founta Aristotle University of



Gabriel Hine Roma3 University



Jeremiah Onaolapo University College London



Savvas Zannettou Cyprus University of