Computational Methods to Mitigate and Measure Online Aggression
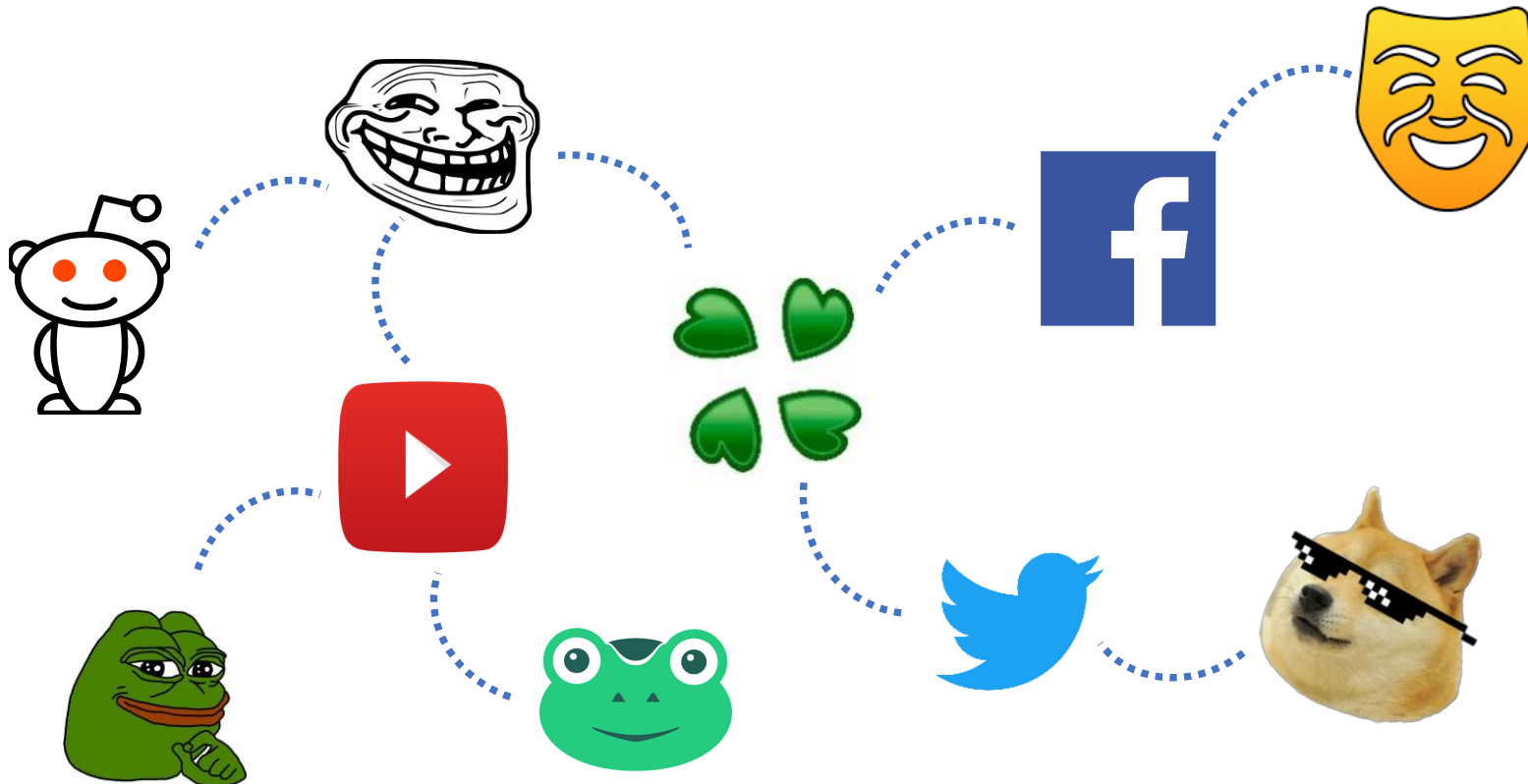
Gianluca Stringhini
Boston University
gian@bu.edu
@gianluca_string

WARNING: Some of this content might be upsetting

Alt-Right

Emilee Reid <emileereid4@gmail.com>
Yesterday, 16:45
Stringhini, Gianluca ⌄

👍 ↩ Reply all | ⌄

Flag for follow up. Completed on 28 April 2017.

Hello,

This recent election has been a shock to my system. I can't believe I live in a country that would elect a person who ran on a campaign of hatred and bigotry. But if there is a silver lining, it has compelled the rest of us to act. I will no longer be silent. I will fight this bigotry with every fiber of my being.

As a Jew, it's been especially startling to see the rise of Anti-Jewish incidents; this is part of an overall rise in prejudicial hatred. Anti-Jew has gone hand-in-hand with anti-black, anti-Muslim, and anti-whomever.

The coincidence of timing between this trend and the political phenomenon of Donald Trump is not merely coincidence. The political themes that Trump has ridden to the White House play directly upon, and stoke further, the sorts of prejudices that can take violent forms and that may be manifested in overturned gravestones in a cemetery.

We must support our Jewish community, LGBT brothers and sisters, and all others who will be victimized by the Trump regime.

With that being said...I am reaching out to find out if any universities are doing work to help stop the rise of hate speech online?? We have to start somewhere. I'd like to explore joining a group of activists who can explore fighting against this hate speech.

Can you please direct me further?

Thank you,
Emilee Reid

Alt-Right

Emilee Reid <emileereid4@gmail.com>
Yesterday, 16:45
Stringhini, Gianluca ⌄

alexgoldmann@airmail.cc | ⌄
Yesterday, 23:26
gianluca@cs.ucsb.edu ⌄

> You forwarded this message on 29/04/2017 08:44

Dear Professor Stringhini,
I have a quick question regarding Community Studies Project 12 where we
have to post against hatespeech. Please email me back as soon as you get
this email. Thank you.

-Alex Goldmann

With that being said...I am reaching out to find out if any universities are doing work to help stop the rise of hate speech online?? We have to start somewhere. I'd like to explore joining a group of activists who can explore fighting against this hate speech.

Can you please direct me further?

Thank you,
Emilee Reid

Alt-Right

ER

alex...

Yester...

gianlu...

You are a retarded f█████t

JD **Joel Dunn** <joelcdunn@gmail.com>
👍 ↩ Reply all | ∨

Fri 09/06, 23:20

Stringhini, Gianluca ∨

Inbox

Flag for follow up. Completed on 10 June 2017.

A) you're in IT but are trying to be a sociology major.

B) you're clearly trying to overblow and be uber-dramatic about what 4chan is.

C) 4chan is people having discussions amongst themselves. Since you don't like or agree with the subject matter or the views of those who spend time communicating on 4chan (because you're an obvious HIV+ fa█████you are attempting to persuade people to believe that free speech is bad, racism is a completely invalid human reaction to danger, etc.

D) the harder people like you try to silence those who you disagree with the more ibvious you make it to people that it is in fact you who are wrong.

E) F███ you. You're an idiot. Enjoy wasting all of your time wrapped and coddled by academia and then getting raped and murdered by a n████.

You're a terrible educator.

Dear

I have

have

this e

-Alex

Yo

Can you please direct me further?

Thank you,
Emilee Reid

Community Studies Project 12:

Over the last four courses we have shifted from community activism in real world scenarios to internet activism in support for minorities, refugees and non-documented citizens. In collaboration with 14 other universities from across the world we will be focusing the remaining lessons on online activism as a whole.

The past 8 lessons has been on methods to address "hate groups" and privilege power structures in the United States, Canada, Scandinavia and Europe. This will be an online activism activity in which you utilize the core methods **(Silence, Shut Down & Self Support)** in fighting one of the largest online hate communities that affects many countries: **4chan.org/pol/** (/pol/ for short).

As we covered in lesson **9c**, /pol/ stands for "Politically Incorrect". Unlike communities such as Stormfront, ATS and smaller outlets for rhetoric from the "Alt-Right"; /pol/ experiences a large amount of internal division from its members. You can review lesson **9c** to get more information on /pol/ (and other online communities) if you need more information.

Please make sure you have downloaded the required software for **Windows** or **Mac** in order to communicate with your group. Your group will be made up of six students. As this is a collaboration with other universities you will be randomly partnered with students. (Note: There may be language barriers within groups. We recommend using free translation tools available online. For a list please contact me.)

Please be forewarned that the content on /pol/ can be **very offensive** and **disturbing**. You may feel free to take Safe Breaks while discussing activism on this website.
Please stick to the following subjects:

LGBTQ+ Issues
Gender Issues
Women's Rights
Non-White Racial Issues
White Privileges and Power
Islam
Marxism (references are available in your **references** *tab)*
Hate Speech
Alt-Right Offenses

*We ask that you stay away from Trump related topics as you will be in a multi-national group and Trump related topics may not apply to them.

Do not expect to "win" any arguments, and it would be recommended to avoid long arguments. Due to the structure of /pol/ a good technique would be to stick to the core methods. I will provide examples below. Please stick to /pol/ as the rest of the site is not politically motivated.
Make sure you visit /pol/ (**www.4chan.org/pol/**) and not /po/ as it is not politically motivated.

A good example of utilizing the core method would be:
Start a "thread" involving one of the subjects above.
Have your group partners reply, while each starts their own "thread" about the same issue and you reply to it.
This will essentially flood their limited amount of subjects and, much like protesting in the real world, take down their ability to reply, discuss or "reason".

Attached is a **tutorial video** on how to start a "thread" on 4chan along with useful tips.

We will be reviewing methods in class the following week and altering methods as a whole, sharing this information between all participating classes.

PLEASE AVOID BEING BANNED.
If you are banned please contact me at **edu**

Community Studies Project 12:

Over the last four courses we have shifted from community activism in real world scenarios to internet activism in support for minorities, refugees and non-documented citizens. In collaboration with 14 other universities from across the world we will be focusing the remaining lessons on online activism as a whole.

The past 8 lessons has been on methods to address "hate groups" and privilege power structures in the United States, Canada, Scandinavia and Europe. This will be an online activism activity in which you utilize the core methods **(Silence, Shut Down & Self Support)** in fighting one of the largest online hate communities that affects many countries: **4chan.org/pol/** (/pol/ for short).

As we covered in lesson **9c**, /pol/ stands for "Politically Incorrect". Unlike communities such as Stormfront, ATS and smaller outlets for rhetoric from the "Alt-Right"; /pol/ experiences a large amount of internal division from its members. You can review lesson **9c** to get more information on /pol/ (and other online communities) if you need more information.

Please make sure you have downloaded the required software for **Windows** or **Mac** in order to communicate with your group. Your group will be made up of six students. As this is a collaboration with other universities you will be randomly partnered with students. (Note: There may be language barriers within groups. We recommend using free translation tools available online. For a list please contact me.)

Please be forewarned that the content on /pol/ ... activism on this website.
Please stick to the following subjects:

LGBTQ+ Issues
Gender Issues
Women's Rights
Non-White Racial Issues
White Privileges and Power
Islam
Marxism (references are available in your refe...
Hate Speech
Alt-Right Offenses

*We ask that you stay away from Trump relate...

Do not expect to "win" any arguments, and it would be recommended to avoid long arguments. Due to the structure of /pol/ a good technique would be to stick to the core methods. I will provide examples below. Please stick to /pol/ as the rest of the site is not politically motivated.
Make sure you visit /pol/ (**www.4chan.org/pol/**) and not /po/ as it is not politically motivated.

A good example of utilizing the core method would be:
Start a "thread" involving one of the subjects above.
Have your group partners reply, while each starts their own "thread" about the same issue and you reply to it.
This will essentially flood their limited amount of subjects and, much like protesting in the real world, take down their ability to reply, discuss or "reason".

Attached is a **tutorial video** on how to start a "thread" on 4chan along with useful tips.

We will be reviewing methods in class the following week and altering methods as a whole, sharing this information between all participating classes.

PLEASE AVOID BEING BANNED.
If you are banned please contact me at g_____@_____edu

s the following week and altering methods

e at **giancula@cs.ucsb.edu**

Community Studies Project 12:

Over the last four courses we have shifted from community activism in real world scenarios to internet activism in support for minorities, refugees and non-documented citizens. In collaboration with 14 other universities from across the world we will be focusing the remaining lessons on online activism as a whole.

The past 8 lessons has been on methods to address "hate groups" and privilege power structures in the United States, Canada, Scandinavia and Europe. This will be an online activism activity in which you utilize the core methods **(Silence, Shut Down & Self Support)** in fighting one of the largest online hate communities that affects many countries: **4chan.org/pol/** (/pol/ for short).

Anonymous (ID: AMaV12jQ) 🇺🇸 04/28/17(Fri)17:16:08 No.123210970 ▶ >>123211305 >>123211335 >>123211559

the email belongs to Gianluca Stringhini.

this is his Twitter: https://twitter.com/gianluca_string?ref_src=twsrc%5Egoogle%7Ctwcamp%5Eserp%7Ctwgr%5Eauthor

one of his tweets: https://twitter.com/gianluca_string/status/786205595745550336

one of his projects: https://arxiv.org/abs/1610.03452?platform=hootsuite

youtube video: https://www.youtube.com/watch?v=tS9Lkpj9DF4 [Embed]

if you dont think this is real, you are retarded.

Start a "thread" involving one of the subjects above.
Have your group partners reply, while each starts their own "thread" about the same issue and you reply to it.
This will essentially flood their limited amount of subjects and, much like protesting in the real world, take down their ability to reply, discuss or "reason".

Attached is a **tutorial video** on how to start a "thread" on 4chan along with useful tips.

We will be reviewing methods in class the following week and altering methods as a whole, sharing this information between all participating classes.

PLEASE AVOID BEING BANNED.
If you are banned please contact me at ████████@██████edu

Community Studies Project 12:

Over the last four courses we have shifted from community activism in real world scenarios to internet activism in support for minorities, refugees and non-documented citizens. In collaboration with 14 other universities from across the world we will be focusing the remaining lessons on online activism as a whole.

The past 8 lessons has been on methods to address "hate groups" and privilege power structures in the United States, Canada, Scandinavia and Europe. This will be an online activism activity in which you utilize the core methods **(Silence, Shut Down & Self Support)** in fighting one of the largest online hate communities that affects many countries: **4chan.org/pol/** (/pol/ for short).

Anonymous (ID: AMaV12jQ) 🇺🇸 04/28/17(Fri)17:16:08 No.123210970 ▶ >>123211305 >>123211335 >>123211559

the email belongs to Gianluca Stringhini.

this is his Twitter: https://twitter.com/gianluca_string?ref_src=twsrc%5Egoogle%7Ctwcamp%5Eserp%7Ctwgr%5Eauthor

one of his tweets: https://twi

one of his projects: https://a

youtube video: https://www.youtube.com/watch?v=tS9LKpJ9DF4 [Embed]

if you dont think this is real, you are retarded.

Anonymous ID:GntBZ4OU Fri 28 Apr 2017 14:11:49 No.123192369 🇺🇸 Report
Quoted By: >>123192640

https://twitter.com/gianluca_string?ref_src=twsrc%5Egoogle%7Ctwcamp%5Eserp%7Ctwgr%5Eauthor

lets cuck this dude. hard

Start a "thread" involving one of the subjects above.
Have your group partners reply, while each starts their own "thread" about the same issue and you reply to it.
This will essentially flood their limited amount of subjects and, much like protesting in the real world, take down their ability to reply, discuss or "reason".

Attached is a **tutorial video** on how to start a "thread" on 4chan along with useful tips.

We will be reviewing methods in class the following week and altering methods as a whole, sharing this information between all participating classes.

PLEASE AVOID BEING BANNED.
If you are banned please contact me at ██████@████████edu

Community Studies Project 12:

Over the last four courses we have shifted from community activism in real world scenarios to internet activism in support for minorities, refugees and non-documented citizens. In collaboration with 14 other universities from across the world we will be focusing the remaining lessons on online activism as a whole.

The past 8 lessons has been on methods to address "hate groups" and privilege power structures in the United States, Canada, Scandinavia and Europe. This will be an online activism activity in which you utilize the core methods **(Silence, Shut Down & Self Support)** in fighting one of the largest online hate communities that affects many countries: **4chan.org/pol/** (/pol/ for short).

Anonymous (ID: AMaV12jQ) 🇺🇸 04/28/17(Fri)17:16:08 No.123210970 ▶ >>123211305 >>123211335 >>123211559

the email belongs to Gianluca Stringhini.

this is his Twitter: https://twitter.com/gianluca_string?ref_src=twsrc%5Egoogle%7Ctwcamp%5Eserp%7Ctwgr%5Eauthor

Anonymous (ID: V8P9sS4w) 🇺🇸 04/28/17(Fri)16:26:15 No.123205663 ▶ >>123211235

Dear Professor,

This recent election has been a shock to my system. I can't believe I live in a country that would elect a person who ran on a campaign of hatred and bigotry. But if there is a silver lining, it has compelled the rest of us to act. I will no longer be silent. I will fight this bigotry with every fiber of my being.

As a Jew, it's been especially startling to see the rise of Anti-Jewish incidents; this is part of an overall rise in prejudicial hatred. Anti-Jew has gone hand-in-hand with anti-black, anti-Muslim, and anti-whomever.

The coincidence of timing between this trend and the political phenomenon of Donald Trump is not merely coincidence. The political themes that Trump has ridden to the White House play directly upon, and stoke further, the sorts of prejudices that can take violent forms and that may be manifested in overturned gravestones in a cemetery.

We must support our Jewish community, LGBT brothers and sisters, and all others who will be victimized by the Trump regime.

With that being said...I am reaching out to find out if any universities are doing work to help stop the rise of hate speech online?? We have to start somewhere. I'd like to explore joining a group of activists who can explore fighting against this hate speech.

Can you please direct me further?

if you dont think this is real, you are retarded.

Start a "thread" involving one of the subjects above.
Have your group partners reply, while each starts their own "thread" about the same issue and you reply to it.
This will essentially flood their limited amount of subjects and, much like protesting in the real world, take down their ability to reply, discuss or "reason".

Attached is a *tutorial video* on how to start a "thread" on 4chan along with useful tips.

We will be reviewing methods in class the following week and altering methods as a whole, sharing this information between all participating classes.

PLEASE AVOID BEING BANNED.
If you are banned please contact me at ▓▓▓▓▓@▓▓▓▓edu

- CLASS PREP -

Community Studies Project 12:

Over the last four courses we have shifted from community activism in real world scenarios to internet activism in support for minorities, refugees and non-documented citizens.

Alt-Right

Emilee Reid <emileereid4@gmail.com>
Yesterday, 16:45
Stringhini, Gianluca ⌄

👍 | 🔄 Reply all | ⌄

Flag for follow up. Completed on 28 April 2017.

Hello,

This recent election has been a shock to my system. I can't believe I live in a country that would elect a person who ran on a campaign of hatred and bigotry. But if there is a silver lining, it has compelled the rest of us to act. I will no longer be silent. I will fight this bigotry with every fiber of my being.

As a Jew, it's been especially startling to see the rise of Anti-Jewish incidents; this is part of an overall rise in prejudicial hatred. Anti-Jew has gone hand-in-hand with anti-black, anti-Muslim, and anti-whomever.

The coincidence of timing between this trend and the political phenomenon of Donald Trump is not merely coincidence. The political themes that Trump has ridden to the White House play directly upon, and stoke further, the sorts of prejudices that can take violent forms and that may be manifested in overturned gravestones in a cemetery.

We must support our Jewish community, LGBT brothers and sisters, and all others who will be victimized by the Trump regime.

With that being said...I am reaching out to find out if any universities are doing work to help stop the rise of hate speech online?? We have to start somewhere. I'd like to explore joining a group of activists who can explore fighting against this hate speech.

Can you please direct me further?

Thank you,
Emilee Reid

PLEASE AVOID BEING BANNED.
If you are banned please contact me at ...@...edu

Community Studies Project 12:

Over the last four courses we have shifted from community activism in real world scenarios to internet activism in support for minorities, refugees and non-documented citizens.

Alt-Right

Emilee Reid <emileereid4@gmail.com>
Yesterday, 16:45
Stringhini, Gianluca ⌄

👍 ↻ Reply all | ⌄

Flag for follow up. Completed on 28 April 2017.

Anonymous (ID: X911YZYH) 🇺🇸 04/28/17(Fri)16:53:18 No.123208401 ▶

>>123205259
That's not him, he was only a grad student at UCSB and now works at UCL.

We must support our Jewish community, LGBT brothers and sisters, and all others who will be victimized by the Trump regime.

With that being said...I am reaching out to find out if any universities are doing work to help stop the rise of hate speech online?? We have to start somewhere. I'd like to explore joining a group of activists who can explore fighting against this hate speech.

Can you please direct me further?

Thank you,
Emilee Reid

PLEASE AVOID BEING BANNED.
If you are banned please contact me at ⬛⬛⬛@⬛⬛⬛edu

(partially obscured background email)

Dear Profes...
This recent...
As a Jew, it's...
The coincide...
overturned g...
We must su...
With that be...
Can you ple...

Anonymous

Community Studies Project 12:

Over the last four courses we have shifted from community activism in real world scenarios to internet activism in support for minorities, refugees and non-documented citizens.

Alt-Right

Emilee Reid <emileereid4@gmail.com>
Yesterday, 16:45
Stringhini, Gianluca ⌄

Reply all | ⌄

Flag for follow up. Completed on 28 April 2017.

Anonymous (ID: X911YZYH) o.123208401 ▶

>>123205259
That's not him, he was o   nd now works at UCL.

This is what we call a **raid**

We must support our Jewish community, LGBT brothers and sisters, and all others who will be victimized by the Trump regime.

With that being said...I am reaching out to find out if any universities are doing work to help stop the rise of hate speech online?? We have to start somewhere. I'd like to explore joining a group of activists who can explore fighting against this hate speech.

Can you please direct me further?

Thank you,
Emilee Reid

PLEASE AVOID BEING BANNED.
If you are banned please contact me at ⬛⬛⬛@⬛⬛⬛edu

"Raids" are an emerging problem

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

By Kristen V. Brown

FilmMagic

| f | SHARE | | TWEET | t | TUMBLR | | EMAIL |

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

Misogyny, death threats and a mob of trolls: Inside the dark world of video games with Zoe Quinn - target of #GamerGate

News > UK > Crime

Police warn of homophobic 4chan cyber attack on LGBT+ Pride month celebrations

HIV fundraising event forced offline by 110 spam bots

"Raids" are an emerging problem

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

By Kristen V. Brown

FilmMagic

| f | SHARE | 🐦 | TWEET | t | TUMBLR | ✉ | EMAIL |

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

Misogyny, death threats and a mob of trolls: Inside the dark world of video games with Zoe Quinn - target of #GamerGate

...phobic 4chan ...T+ Pride month celebrations

HIV fundraising event forced offline by 110 spam bots

We see attacks like these happening 2-5 times a day

How can computer security people help?
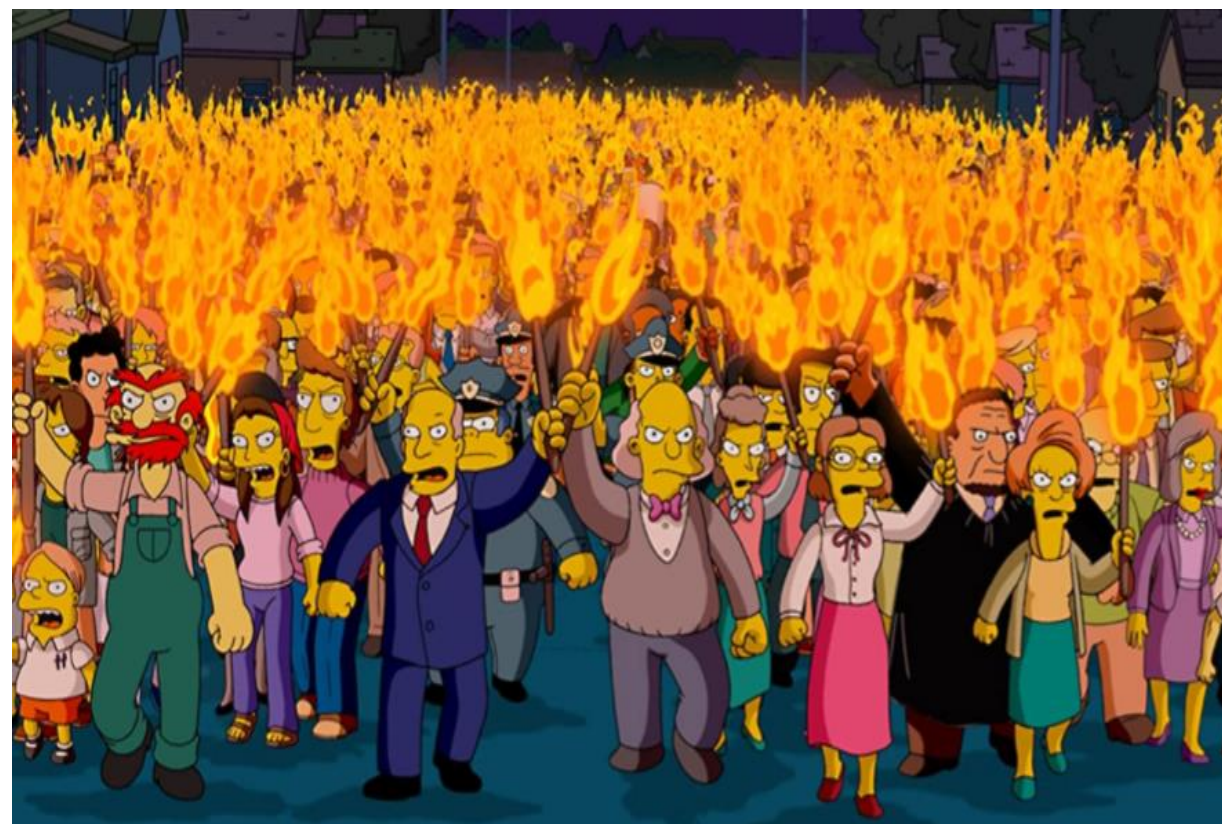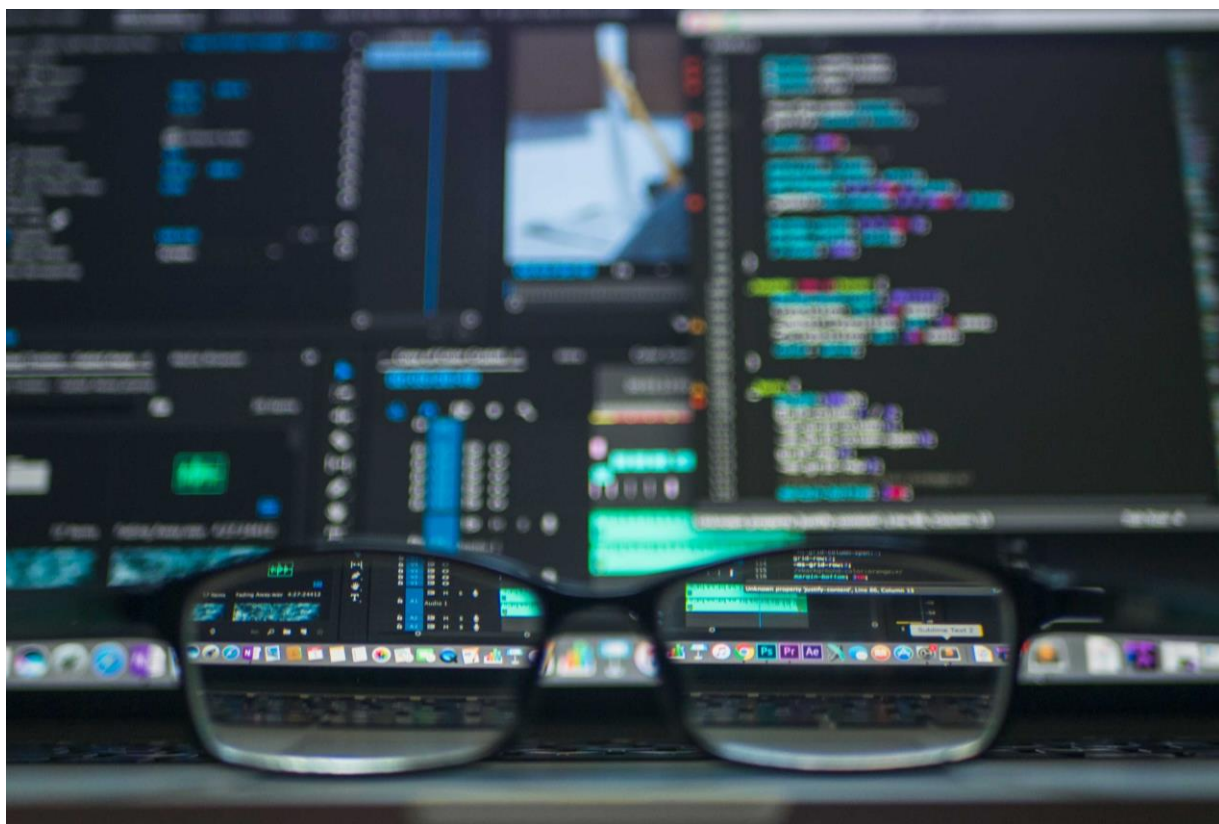
What are we going to see today?

What is 4chan?

- An image board

- Conversations grouped into threads

- Fixed number of threads alive at a given time

- New replies bump a thread to the top of the page

- **Anonymous**

- **Ephemeral** (median thread lifetime: 47 min)

Over a period of 3.5 years, we collected 134M posts

Paper + dataset (ICWSM 2020)

Threat model

Call for attack   Coordination   Delivery   Harm

Organization thread

Raids against YouTube videos

YouTube is the domain most often linked on the
Politically Incorrect board (93k over 3 months)

Anecdotally, we know of many raids directed from
the board towards YouTube

Raids against YouTube videos

YouTube is the domain most often linked on the Politically Incorrect board (93k in our dataset)

Anecdotally, we know of many raids directed from the board towards YouTube

Identifying raided videos on YouTube

Videos showing a high degree of synchronization with 4chan threads tend to attract comments that contain more hate speech
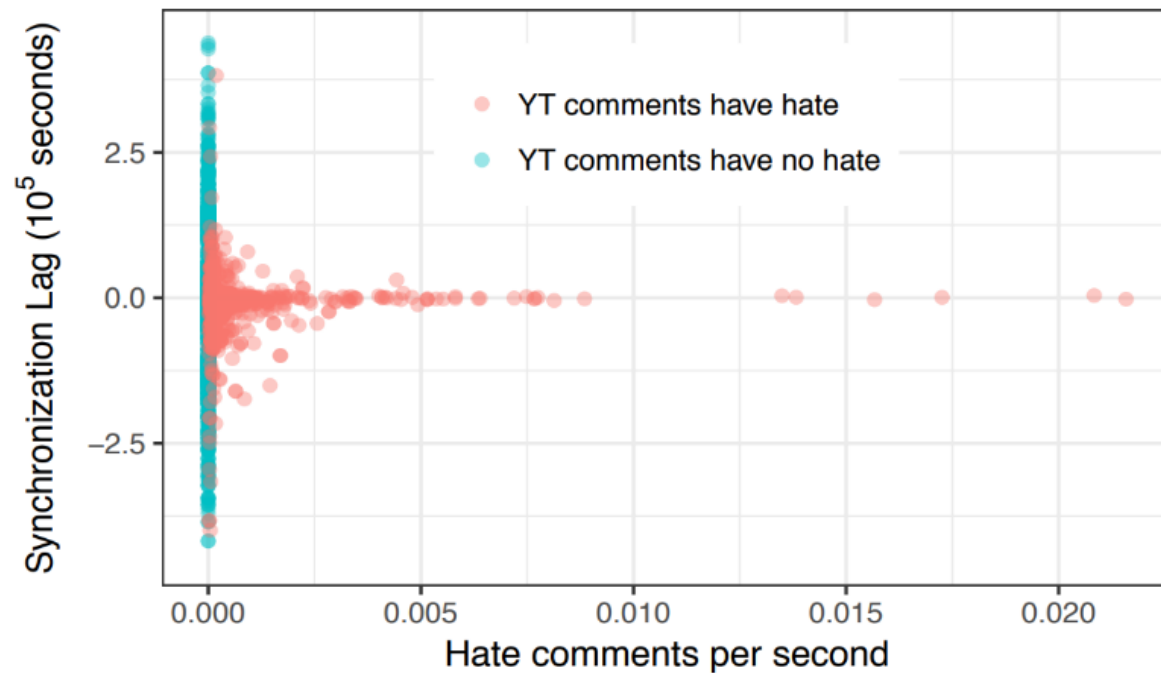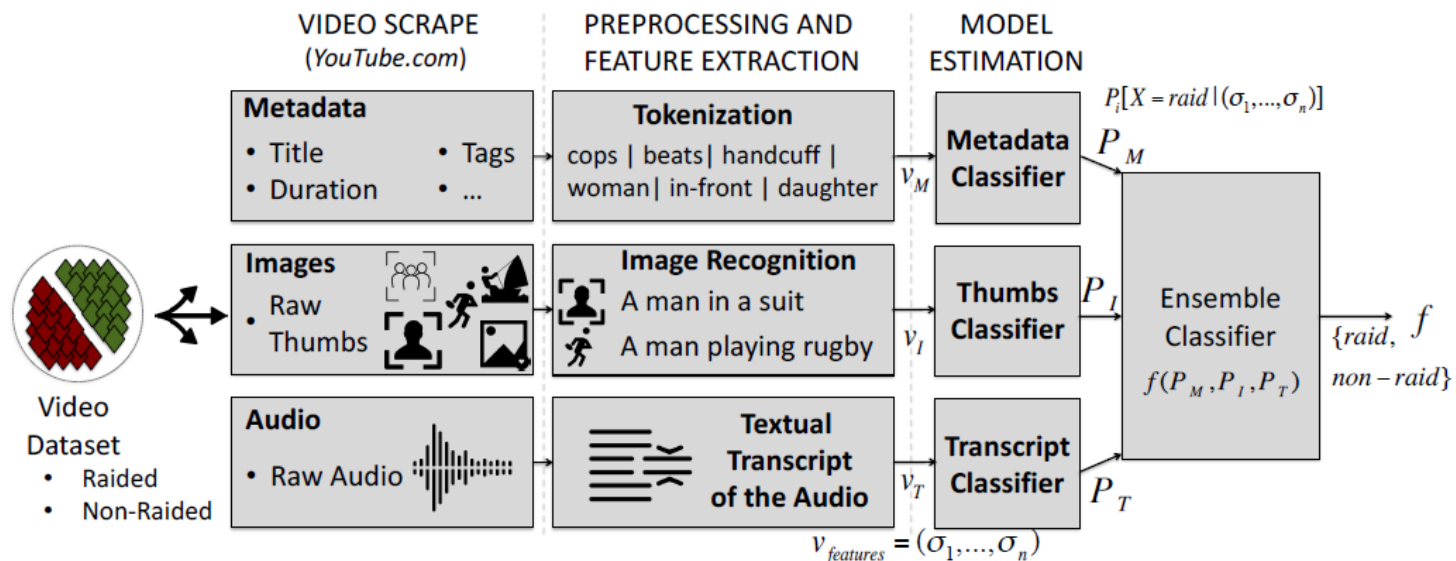
What to do about raids?

Our measurement allowed us to collect a ground truth dataset of raided videos

Can we predict which videos will get raided and help moderators?

Our approach can determine the likelihood that a YouTube video will be raided at upload time with AUC 0.94, allowing for proactive moderation.
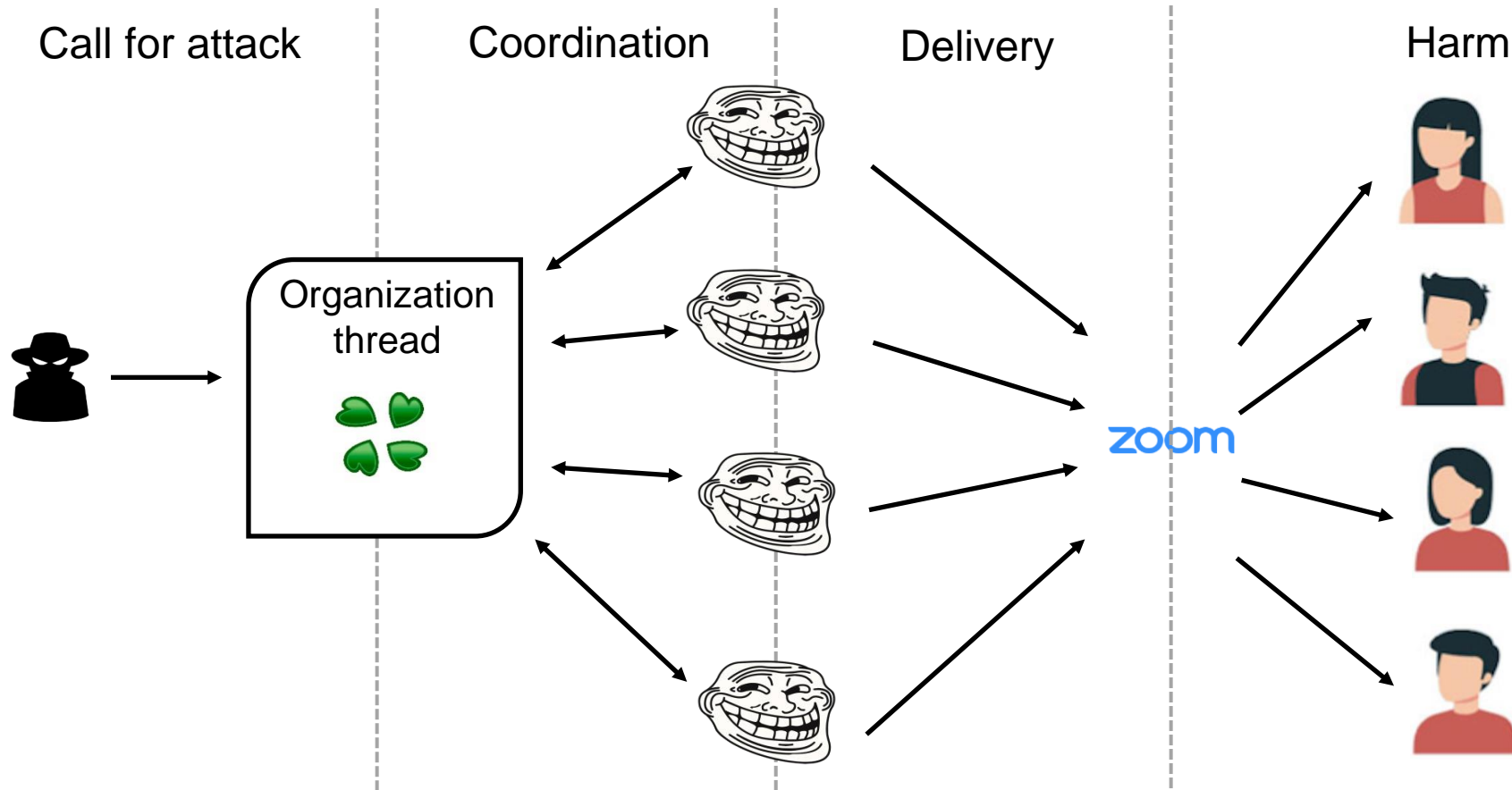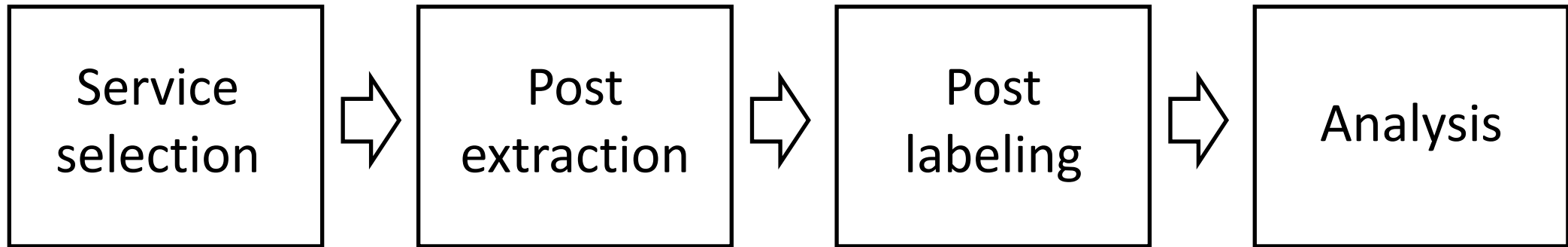
Our entire lives have moved online

The latest threat - zoombombing

Call for attack   Coordination   Delivery   Harm

Organization thread

Measuring zoombombing

```
┌─────────────┐    ┌─────────────┐    ┌─────────────┐    ┌─────────────┐
│   Service   │ ⇨  │    Post     │ ⇨  │    Post     │ ⇨  │  Analysis   │
│  selection  │    │  extraction │    │   labeling  │    │             │
└─────────────┘    └─────────────┘    └─────────────┘    └─────────────┘
```

10 meeting services
- Zoom
- Meet
- Hangouts
- Webex
- Teams
- …

- 434 threads on 4chan
- 3,500 tweets

218 calls for zoombombing

Main findings

74% calls for zoombombing target online classes

Main findings

74% calls for zoombombing target online classes

70% are called by insiders

Main findings

74% calls for zoombombing target online classes

70% are called by insiders

93% target meetings happening in real time
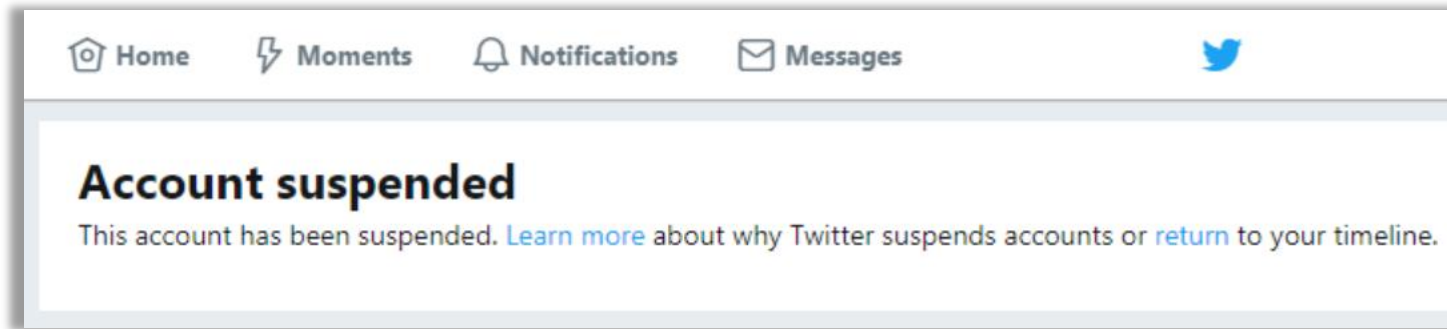
We need to rethink our defense mechanisms

- Passwords don't help
- Waiting rooms don't help

Only two services allow one-time unique links for each participant

Does deplatforming work?

⊙ Home	⚡ Moments	🔔 Notifications	✉ Messages	🐦

Account suspended

This account has been suspended. Learn more about why Twitter suspends accounts or return to your timeline.

People can move to alternative an unmoderated communities
- Will they become more toxic?
- Will they get radicalized?
- Will they come back under false identities?

Conclusion

Looking at cross-platform activity can help identify raids early

Defense mechanisms need to be developed against insiders

More research on deplatforming is needed

Thank you!

gian@bu.edu

@gianluca_string