

Anonymity in social media

Saptarshi Ghosh
and Mainack Mondal

CS 60017
Autumn 2019



So far we covered

- Advertising systems
 - Why should we bother about them?
 - How these ad interfaces can be abused?
- Preserving privacy in social media from background actors
 - K-anonymity
 - Differential Privacy
- Preserving privacy in social media from other users
 - Privacy definitions
 - Exposure control

Now let's go a bit more into the what of privacy

- Recall westin's defintion
 - the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others.
- He defined four states of privacy

Now let's go a bit more into the what of privacy

- Recall westin's defintion
 - the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others.
- He defined four states of privacy
 - Solitude: being free from observation by others

Now let's go a bit more into the what of privacy

- Recall westin's defintion
 - the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others.
- He defined four states of privacy
 - Solitude: being free from observation by others
 - Intimacy: limiting communication within a small group

Now let's go a bit more into the what of privacy

- Recall westin's defintion
 - the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others.
- He defined four states of privacy
 - Solitude: being free from observation by others
 - Intimacy: limiting communication within a small group
 - Reserve: desire to limit information disclosure to others and others respecting that desire

Now let's go a bit more into the what of privacy

- Recall westin's defintion
 - the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others.
- He defined four states of privacy
 - Solitude: being free from observation by others
 - Intimacy: limiting communication within a small group
 - Reserve: desire to limit information disclosure to others and others respecting that desire
 - Anonymity: freedom from identification and from surveillance

Now let's go a bit more into the what of privacy

- Recall westin's defintion
 - the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others.
- He defined four states of privacy
 - Solitude: being free from observation by others
 - Intimacy: limiting communication within a small group
 - Reserve: desire to limit information disclosure to others and others respecting that desire
 - Anonymity: freedom from identification and from surveillance

**Anonymous social media site:
whisper** (based on Correa et al.'s ICWSM'15 paper)

Anonymous Online Social Media sites

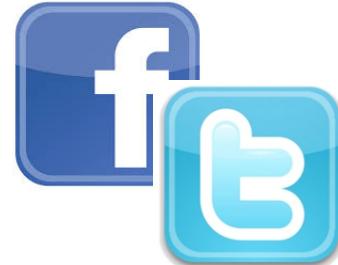
Traditional Online social networks are very popular

284 million monthly active Twitter users

500 million tweets per day

These social networks are non-anonymous

Requires unique usernames and profile information



However **anonymous social networks are emerging**

Whisper, secret.ly, Yik Yak



No unique usernames or profile information

Whisper: example of Anonymous social media site

Whisper: popular anonymous OSN

Launched in March 2012

2M+ users, 2.5B+ page views

Age group 17-28 years

Anonymous social media site

No unique username

No profile information

Post messages via mobile

Read-only web interface

Posts are called **whispers**



Whisper: example of Anonymous social media site

Whisper: popular anonymous OSN

Launched in March 2012

2M+ users, 2.5B+ page views

Age group 17-28 years

Anonymous social media site

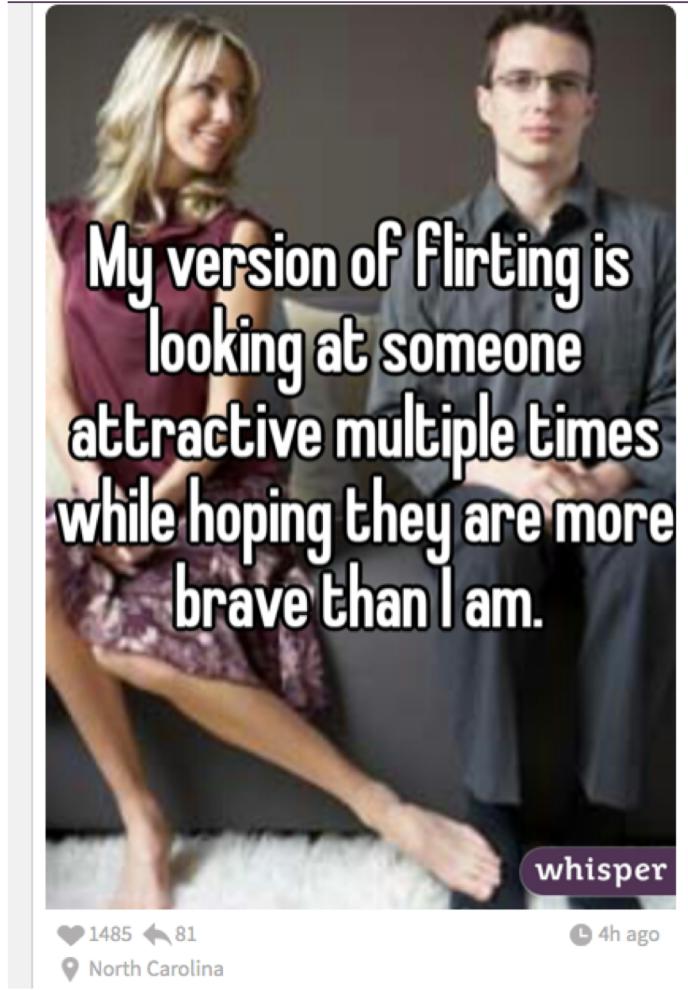
No unique username

No profile information

Post messages via mobile

Read-only web interface

Posts are called **whispers**



Characterizing anonymous social media content

These sites are very new and growing fast

Very little work on how users use this platform

Are these sites different than the existing non-anonymous social media sites?

No investigation before us into the type of content posted

Our experimental platform: Whisper

We collected data from Whisper for 3+ months

July 1 – November 17 2014

20 Millions whispers in our dataset

Texts posted by users

Pictures on which the texts are posted

Locations of 66.3% posts

We categorize whispers along four dimensions

Questions

Sensitivity: How anonymity sensitive is content posted on anonymous media sites?

Content: What types of content are posted on anonymous social media?

Audience: Do demographic variables play a role in content anonymity sensitivity?

Language: What are the linguistic differences between content posted on anonymous versus non-anonymous social media?

Questions

Sensitivity: How anonymity sensitive is content posted on anonymous media sites?

Content: What types of content are posted on anonymous social media?

Audience: Do demographic variables play a role in content anonymity sensitivity?

Language: What are the linguistic differences between content posted on anonymous versus non-anonymous social media?

How to measure anonymity sensitivity of a content?

Challenge

Anonymity sensitivity for same content differs for different users

Solution

We define **Anonymity Sensitivity Score (AS Score)**

The probability that users would call a piece of text anonymous

If 3 out of 10 users call a piece of text anonymous then AS Score is 0.3

AS Score varies from 0.0 to 1.0

Higher AS Score means more sensitive

Computing AS Scores of social media content

We picked 500 random tweets, 500 random whispers

Asked 100 AMT workers to label each text anonymous or not

What is AMT?

Amazon Mechanical Turk: A crowdsourcing platform

Crowdsourcing

What is AMT?

collaborative approach for acquiring annotated corpora

The requester floats jobs (identify objects in a picture)

The workers do those jobs

The requesters pay the workers via AMT platform

Crowdsourcing: Why

Paid for crowdsourcing can be 33% cheaper than in-house employees when applied to tasks such as tagging and classification (Hoffmann, 2009)

Participants (workers) are paid money to complete easy tasks
(HIT = Human Intelligence Task)

Contributors are extrinsically motivated through economic incentives

Most NLP/AI/ML projects use crowdsourcing : AMT and CrowdFlower

Great for labelling data (does this post contain sensitive data?)

Can even get survey responses (Do you want to change privacy of your facebook posts?)

Computing AS Scores of social media content: Continued

We picked 500 random tweets, 500 random whispers

Asked 100 AMT workers to label each text

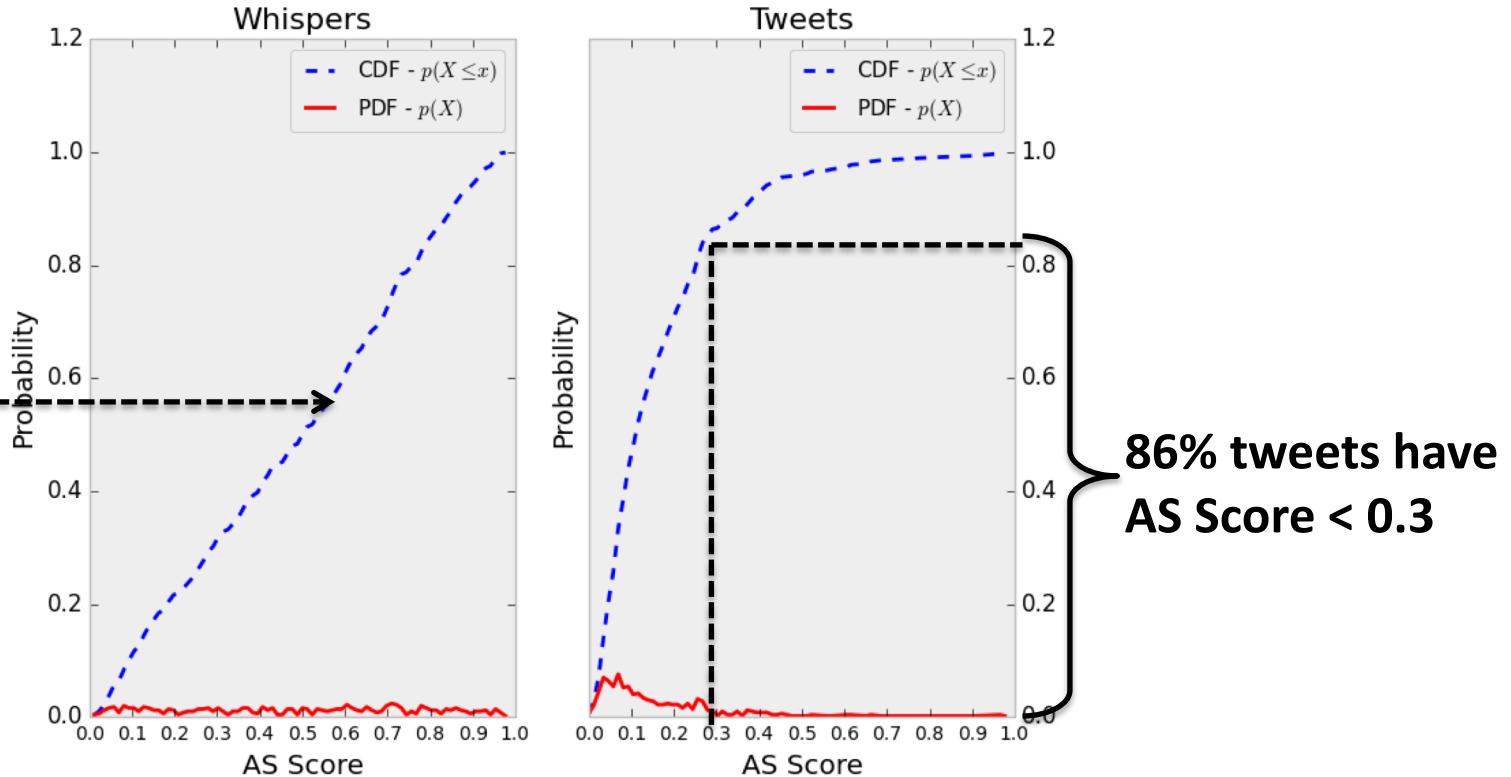
Question: if you would have made this post would you post it anonymously or non-anonymously?

Computed AS Score for each of the text

AS Score	Sample text
0.0	absolutely roastin today!!!! im dying lol x
0.5	Is it to late to join a beer softball league?
1.0	I suck at sexting! 19y female.

Are the AS Scores of anonymous and non-anonymous social media posts different?

similar fraction of whispers for all AS Scores



Anonymous media content have higher AS Scores (overall)

Not all anonymous media content is equally sensitive

Categorization dimensions

Sensitivity: How anonymity sensitive is content posted on anonymous media sites?

Anonymous media content are more sensitive

Content: What types of content are posted on anonymous social media?

Audience: Do demographic variables play a role in content anonymity sensitivity?

Language: What are the linguistic differences between content posted on anonymous versus non-anonymous social media?

Categories for classifying content

We create 9 categories inspired by Whisper categories

We asked 5 AMT workers to classify 500 random whispers

93.8% Whispers are captured by the 9 provided category

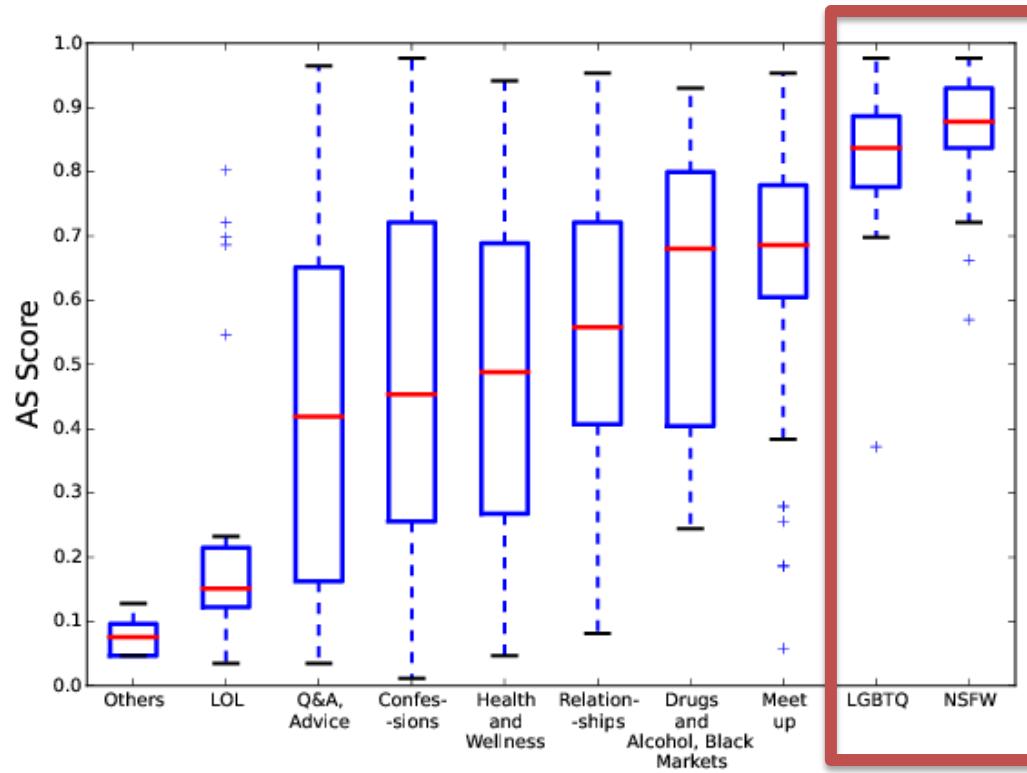
Category	Example
Confessions	I told everyone I loved my job as a lifeguard when I actually hated it, I just liked being tan
Relationships	My mother wants me converse more with her. Yet when I try, the discussion turns to me being fat, useless and pathetic. She wonders why I don't speak to her.
Meetup	any ladies 18+ want a new snap friend? 30/m here...
QnA/Advice	is it okay that I'm a virgin at 17?
NSFW	boobs <3
LGBTQ	If you're trans and reading this, have a hug from me.
Health and Wellness	Having depression and anxiety really sucks... :(
LOL	Each time I see a whisper on the popular page I think damn why didn't I think of that
Drugs and alcohol, black markets	I smoke weed everyday

What type of content is posted in whisper?

Category	percentage of whispers
Confessions	56.8
Relationships	24.7
Meetup	17.8
QnA/Advice	14.4
NSFW	7.1
LGBTQ	5.0
Health and Wellness	5.0
LOL	4.8
Drugs and alcohol, black markets	2.5

Most of anonymous social media content is classified as confessions, relationships, meet up or Q&A/Advice.

Are these different types of content equally sensitive?



Posts in NSFW and LGBTQ are significantly more anonymity sensitive than others

Categorization dimensions

Sensitivity: How anonymity sensitive is content posted on anonymous media sites?

Anonymous media content are more sensitive

Content: What types of content are posted on anonymous social media?

Mostly confessions, relationships, meetup, advices

Audience: Do demographic variables play a role in content anonymity sensitivity?

Language: What are the linguistic differences between content posted on anonymous versus non-anonymous social media?

How demographically diverse is our AMT worker base?

Our AMT workers voluntarily submitted data about demographics

Some basic demographic statistics

We have workers mainly from US

Median age of our users: 32

male-female ratio same as in whole US (0.95)

39.5% of our AMT workers have graduated from college

Our AMT workers have a wide income range

Median income 27k USD per annum

Correlation between user demographics and content sensitivity

We check the difference between the AS Scores assigned by different demographic groups of our AMT workers

Key findings

Gender affects anonymity sensitivity only for **NSFW content**

Older users are more anonymity sensitive than younger users

College Education increases anonymity sensitivity

Categorization dimensions

Sensitivity: How anonymity sensitive is content posted on anonymous media sites?

Anonymous media content are more sensitive

Content: What types of content are posted on anonymous social media?

Mostly confessions, relationships, meetup, advices

Audience: Do demographic variables play a role in content anonymity sensitivity?

Yes, gender, age, education effects sensitivity

Language: What are the linguistic differences between content posted on anonymous versus non-anonymous social media?

Linguistic differences between anon. and non-anon. content

We sample 100,000 random whispers and tweets

We used Linguistic Inquiry and Word Count (LIWC) - hierarchical linguistic lexicon with 70 categories

biological, positive/negative emotion, Personal Concerns

Key findings

Whispers have **more personal, social and informal** words than tweets

Words in Whisper communicate **more negative emotions** than tweets

Whispers express **more wants, needs and wishes**

Can we automatically suggest user what is a possible anon social media content based on linguistic characteristics?

Can we detect anonymous social media content using language?

We tried to build a classifier using LIWC labels as features

Our train and test set consists of 100K random whispers and 100K random tweets

Used 10-fold cross validation to validate classifier

Our classifier achieved a F1 score of 72.88%

A good first step, but we can possibly improve it in future

**Anonymous social media site:
4chan and /b/** (slides taken from Bernstein et al.)



4chan and /b/: An Analysis of Anonymity and Ephemerality in a Large Online Community

Michael Bernstein, Andrés Monroy-Hernández, Drew Harry, Paul André, Katrina Panovich and Greg Vargas

>>



Anonymous 07/20/11(Wed)15:30:05 No.339760588

mit computer science and ai lab
mit media lab
university of southampton

Identity Representation

pseudonyms



Mark Zuckerberg

Pseudonyms and multiple accounts demonstrate “a lack of integrity”



Google

“The web means
the end of forgetting.”
— New York Times

Identity Representation

Archiving Strategy



Mark Zuckerberg

Pseudonyms and multiple
accounts demonstrate
“a lack of integrity”

Identity Representation



Millen 2000

Archives are highly valued and expected to be read by newcomers.

Archiving Strategy



Millen + Patterson 2003

Real names and pseudonyms “promote trust, cooperation and accountability.”

Identity Representation



What happens in a
large scale
anonymous, ephemeral
community?

Archiving Strategy



Ma + Agarwal 2007

Strong identity and permanent archives correlate with satisfaction and participation

Identity Representation



How does such a community produce so much of the Internet's shared culture?

Archiving Strategy



Ma + Agarwal 2007

Strong identity and permanent archives correlate with satisfaction and participation

Identity Representation



4chan /b/

Archiving Strategy



Ma + Agarwal 2007

Strong identity and permanent archives correlate with satisfaction and participation

Our Goal

Use 4chan /b/ as a lens
to understand the concepts
of anonymity and ephemerality.

Our Argument

Anonymity and ephemerality support community behaviors that generate internet culture.

File : [1310254071.jpg](#) (255 KB, 772x679)



Anonymous 07/09/11(Sat)19:27:51 No.339793221 [Reply]

Your fortune: Bad Luck

http://www.youtube.com/watch?v=4IC4hYaoJJA&feature=feedrec_grec_index
<http://www.youtube.com/watch?v=pnqSwWcdZRw&NR=1>

This is why, when someone talks about how fucked up /b/ used to be and how it was sicker and far more evil.... you can tell they are newfags.

Truth

107 posts and 85 image replies omitted. Click Reply to view.

>> **Anonymous** 07/09/11(Sat)21:08:51 No.339807752

Your fortune: Outlook good

[>>339807618](#)

yeah but the links are dead...

>> **Anonymous** 07/09/11(Sat)21:09:25 No.339807847

Your fortune: Better not tell you now

[>>339807618](#)

dude everyone has pics of that chick.. not the one kissing the guy, the other girl

>> **Anonymous** 07/09/11(Sat)21:10:03 No.339807955

File [1310260203.jpg](#) (191 KB, 600x440)



Your fortune:
Bad Luck

[>>339807666](#)

File : [1310259037.png](#) (69 KB, 378x378)



Anonymous 07/09/11(Sat)20:50:37 No.339804918 [Reply]

Aspie with Narcissistic Personality Disorder here

what's your "excuse" for sitting at home browsing /b/ on a saturday night?

19 posts and 1 image reply omitted. Click Reply to view.

>> **Anonymous** 07/09/11(Sat)21:08:04 No.339807630

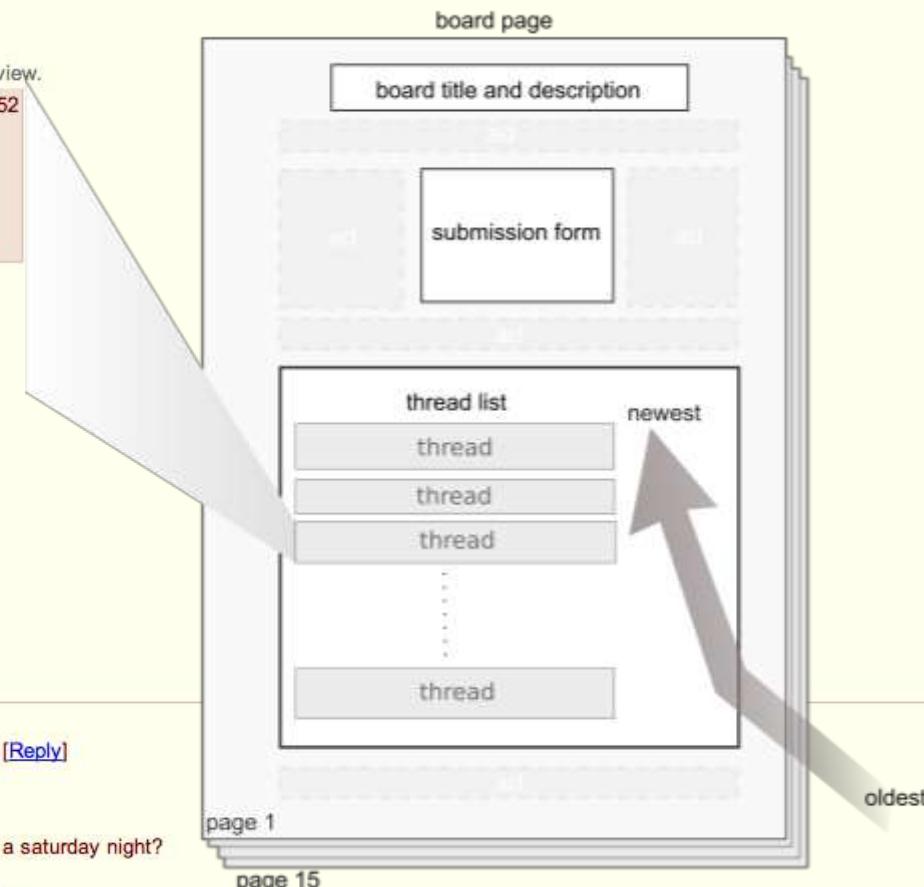
[>>339807257](#)

Punk my good sir.

I'm sure you're an excellent frontman

>> **Anonymous** 07/09/11(Sat)21:08:16 No.339807665

I could care less what day of the week it is. I have to work at 8am tomorrow.



4chan's Impact

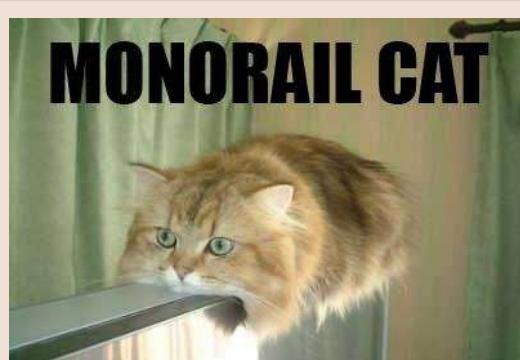
>>



moot!Ep8pui8Vw2 07/20/11(Wed)15:30:05 No.339760588

over seven million users

>>



Anonymous 07/20/11(Wed)15:30:23 No.339760

prolific meme factory

- > lolcats
- > rickrolling
- > chocolate rain
- > pedobear

>>



Anonymous 07/20/11(Wed)15:30:43 No.339823

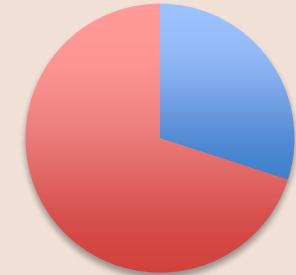
highly visible off-site activities

- > electing moot Time Person of the Year
- > hacktivist group Anonymous

/b/, the “Random” board

The most active board

>>



Anonymous 07/20/11(Wed)15:30:05 No.339760588

30% of all traffic

>>



moot!Ep8pui8Vw2 07/20/11(Wed)15:30:05 No.339760588

“the lifeforce of the website”

where the “rowdiness and lawlessness” happen

4chan /b/ and Ephemerality

>>



Anonymous 07/20/11(Wed)15:30:05 No.339760588

/b/ deletes content as soon as newer content arrives

>>



Anonymous 07/20/11(Wed)15:30:05 No.339760597

fast-paced and chaotic experience

Twitter's design suggests ephemerality

4chan /b/ and Anonymity

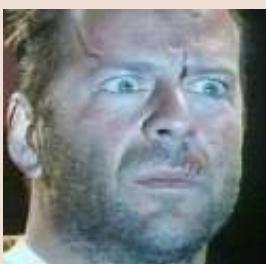
>>



Anonymous 07/20/11(Wed)15:30:05 No.339760588

posts are fully anonymous by default
rarely contain pseudonyms or other identity signals

>>



Anonymous 07/20/11(Wed)15:30:23 No.339760597

traditional reputation systems become unworkable

Formspring.me is based on anonymous feedback
Blizzard's shift to real names sparked an uprising

Related work

>>



Anonymity 07/20/11(Wed)15:30:05 No.339760588

>>



Sites with (discouraged) anonymous commenting:
Usenet [Donath 1999], Slashdot [Lampe+Resnick 2004]

Removing social cues makes communication
impersonal [Short, Williams, Christie 1976]
and cold [Hiltz, Johnson, Turoff 1986], and anonymity
directly undermines credibility [Rains 2007].

>>



Ephemerality 07/20/11(Wed)15:30:05 No.339760588

>>



Technology does not let us forget
[Grudin 2002; Mayer-Schonberger 2009]

ITT: In This Talk

Content

Analysis of threads on 4chan /b/

Ephemerality

Tempo and deletion dynamics

Anonymity

Practices around identity

Content Analysis

Content Analysis Method

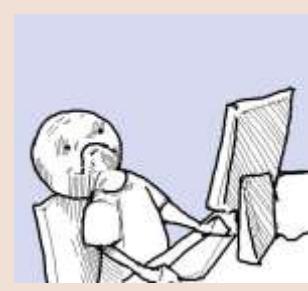
>>



Anonymous 07/20/11(Wed)15:31:04 No.3397256562

collected 598 thread-starting posts over ten days

>>



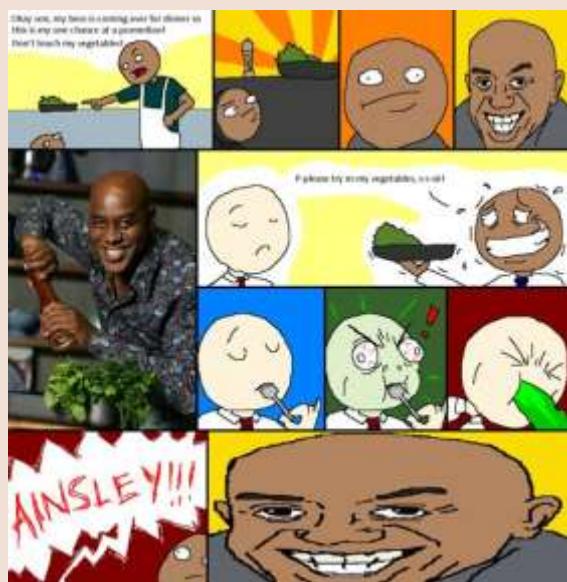
Anonymous 07/20/11(Wed)15:30:05 No.339760588

rounds of iterative coding

developed nine categories to describe thread themes

Themed 28%

>>



Anonymous 07/20/11(Wed)15:30:05

ITT [“In this thread”] we only post stuff we have laughed at so hard we had tears

Sharing Content 19%

>>



Anonymous 07/20/11(Wed)15:30:05

This guy is a hero. :)

[http://www.youtube.com/\[xxxxxx\]](http://www.youtube.com/[xxxxxx])

Question or Advice 10%

>>



Anonymous 07/20/11(Wed)15:30:05

Soup /b/

Recently I've been hanging out with a girl a lot, we're both in college. I spend the night at her place all the time and we kiss and whatnot. Problem is, she just broke up with her ex [...] and I know she's not over him. I really like her but I dunno what to do, so what do /b/?

>>



Sharing Personal Information 9%

U JELLY? ["Jealous?"] This is me suiting up at my formal looking fucking brilliant, then there is you fags sitting back and watching.

>>



Discussion 8%

Hi Anonymous! So ive started this game called League of Legends a few days ago. [...] Is there anyone here who also plays it? Lets talk about it!

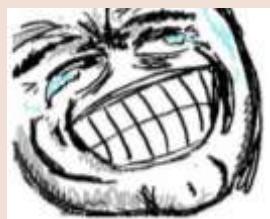
>>



Request for Item 8%

anyone has a pic of that star wars battle tank with the german insignia shooped ["photoshopped"] on it? in return tits

>>



Request for Action 7%

Make a group saying [name] is awesome on face book.
DO IT FAGGOTS

>>



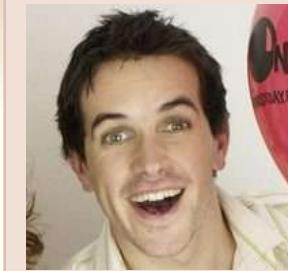
Discussing /b/ 5%

Heidi Ho there /b/ I'm a Newfag and now that i've been here all Summer I was wondering if i need a letter of recommendation From a Registered OldFag?

Study 1

Ephemerality

Ephemerality on 4chan



Anonymous 07/20/11(Wed)15:30:05
cool paper bro

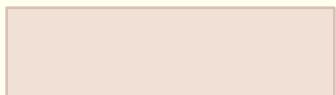
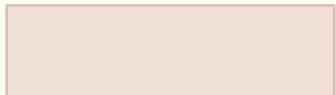
Ephemerality on 4chan



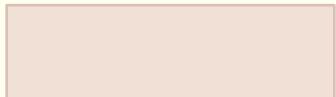
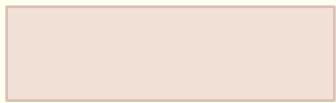
Ephemerality on 4chan



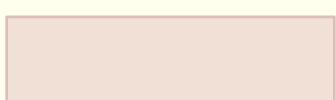
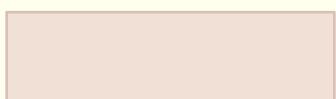
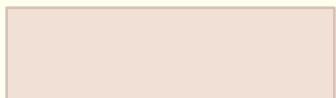
Ephemerality on 4chan



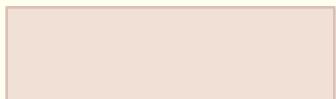
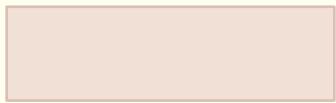
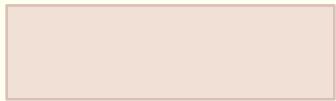
Ephemerality on 4chan



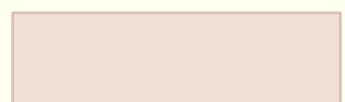
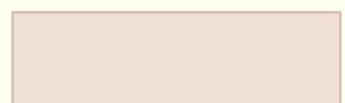
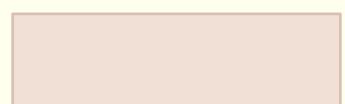
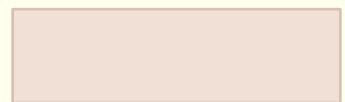
Ephemerality on 4chan



Ephemerality on 4chan



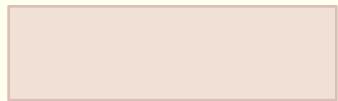
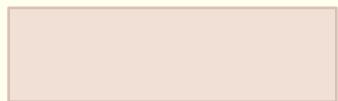
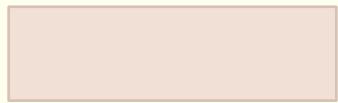
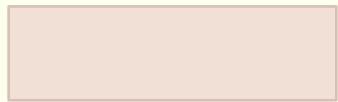
Ephemerality on 4chan



15 threads per page
15 pages

225 threads on /b/

Ephemerality on 4chan



404

Ephemerality on 4chan

Bumping



Anonymous 07/20/11(Wed)15:30:05

cool paper bro

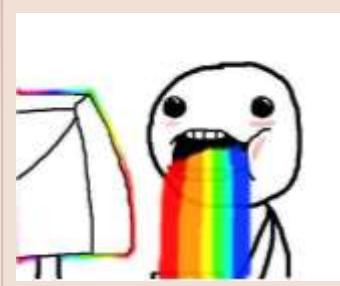
>>

Anonymous

p values or GTFO

Data Collection

>>



Anonymous 07/20/11(Wed)15:30:05 No.339760588

complete dataset of /b/ activity over two weeks
5,576,096 posts in 482,559 threads

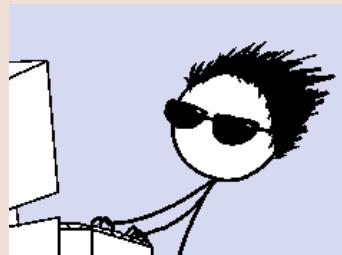
>>

Anonymous 07/20/11(Wed)15:30:34 No.339734523

recreate /b/'s thread position history using
post creation timestamps

Entire Lifetime

>>



Thread Lifetime 07/20/11(Wed)15:30:34 No.3397345

median thread lifetime: **3.9 minutes**
min: 28 seconds
max: 6.2 hours

>>



43% of threads get no replies
median thread has original post and one reply

>>

ASK A
PAGAN
ANYTHING

Longest-Lived Thread 07/20/11(Wed)15:30:34

go on do your worst (or alternatively actually get my respect and actually ask something useful)

>>

Short-Lived Thread 07/20/11(Wed)15:30:34 No.33973453984

I hope you're glad the captcha is gone, woo yeah, let's all be random lolz!

First Page

>>



Why first page? 07/20/11(Wed)15:30:34

threads get most of their visibility on the first page

reminder: bumped threads are returned to the first page

>>

Total time on first page 07/20/11(Wed)15:30:34

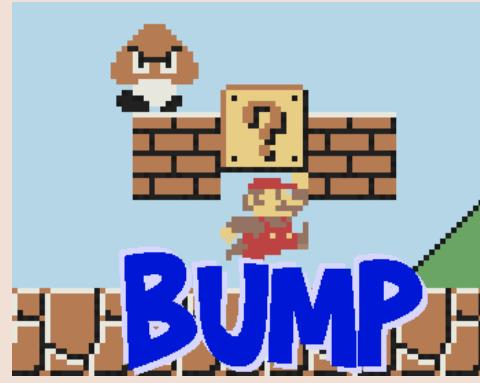
median time on first page: **5 seconds**

min: less than 1 second

max: 37 minutes

Bumping and Sage

>>



bump 07/20/11(Wed)15:30:34 No.33973454556

can reply with deliberate intent to bring the thread to the first page and keep it from 404ing

2.16% of posts contained “bump”, “bamp”, etc. or similar inflections

>>



sage 07/20/11(Wed)15:30:34 No.33973454556

“sage” in email field leaves a reply without bumping the thread: “burying” a thread

0.77% of posts were sages

>>

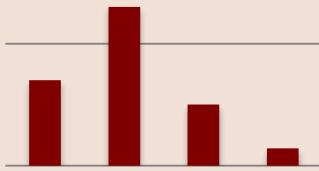
Interpretation 07/20/11(Wed)15:30:34

4chan members play within the bounds of ephemerality and manipulate it

Ephemerality Discussion

>>

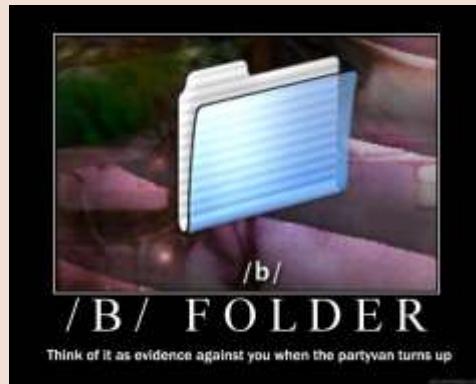
Volume comparisons 07/20/11(Wed)15:30:34



35,000 threads and 400,000 /b/ posts per day
vs. 7,100 digg threads per day
25,000 usenet posts per day
65,000 youtube videos per day

[Szabo+Huberman 2010, newsadmin.com]

>>



Alternate Archives 07/20/11(Wed)15:30:34

users save memorable content
to /b/ folders on their hard drives

4chanarchive.org for “epic” threads

Ephemerality Discussion

>>



Role in meme creation 07/20/11(Wed)15:30:34

fast pace leads to continuously fresh content
no history may mitigate “rich get richer” phenomena

>>



Raising participation? 07/20/11(Wed)15:30:34

counterintuitively, ephemerality may actually
be increasing participation in /b/

participating in a thread is the best way
to keep it from 404ing

Study 2

Anonymity

Anonymity on 4chan

The screenshot shows a 4chan post submission interface. On the left, there are several input fields: Name, E-mail, Subject, Comment, Verification, File, and Password. The Verification field contains the text "Merad immune". Below these fields is a CAPTCHA image with the text "Merad immune" and a text input field next to it. To the right of the CAPTCHA is the reCAPTCHA logo with the tagline "stop spam, read books.". Two large red arrows point from the right side of the slide towards the CAPTCHA area.

Name	<input type="text"/>
E-mail	<input type="text"/>
Subject	<input type="text"/>
Comment	<input type="text"/>
Verification	<input type="text"/> Merad immune
File	<input type="button" value="Choose File"/> No file chosen
Password	<input type="password"/> (Password used for file deletion)

Supported file types are: GIF, JPG, PNG
Maximum file size allowed is 2048 KB.
Images greater than 250x250 pixels will be thumbnailled.
Read the rules and FAQ before posting.
 このサイトについて - 翻訳
Currently 2866 unique user posts.

>> **Anonymous** 07/20/11(Wed)15:30:34 No.33973454556
[comment text]

>> **OP** 07/20/11(Wed)15:30:34 No.33973454556
[post using an unverifiable name]

>> **username!Oo43raDvH61** 07/20/11(Wed)15:30:34 No.33973454556
[post using a tripcode, /b/'s only identity mechanism]

Identity Markers

>>



Anonymous 07/20/11(Wed)15:30:34 No. 234028393

90.1% of posts were credited to “Anonymous”

>>

Gómez et al. 2008 07/20/11(Wed)15:30:34 No. 23402839

closest comparison in the literature:

18.6% of Slashdot comments are anonymous

>>

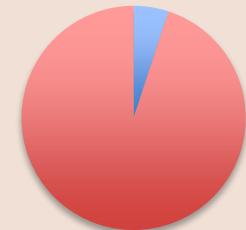
Runners-up 07/20/11(Wed)15:30:44 No. 23402839323

David 1.06% (part of a 4chan meme)

– 1.06%

OP 0.24% (claiming to be the Original Poster of the thread)

>>



Tripcodes 07/20/11(Wed)15:30:34 No. 234028393

5% of posts used tripcodes

An Argument for Anonymity

>>

Anonymous 07/20/11(Wed)15:30:34 No. 2832974298

typical online community narrative is pro-identity
but, anonymity and disinhibition can have positive effects

>>



Anonymous 07/20/11(Wed)15:33:23

masking failure makes experimentation safer
creativity
memes
group ownership of creative content

>>

Anonymous 07/20/11(Wed)15:30:34 No. 29038240923

anonymity can provide a cover for intimate conversations

Soup /b/ Recently I've been hanging out with a girl a lot, we're both in college. I spend the night at her place all the time and we kiss and whatnot. Problem is, she just broke up with her ex [...] and I know she's not over him. I really like her but I dunno what to do, so what do /b/?

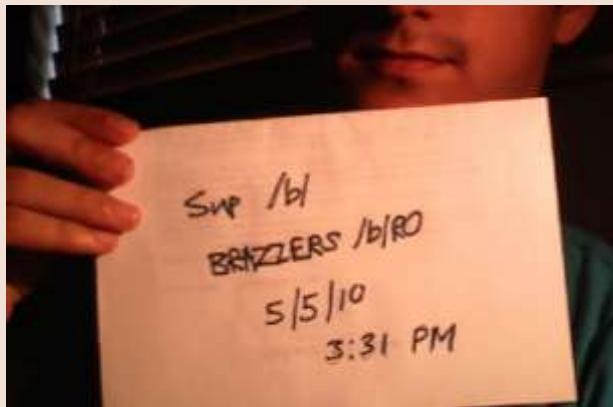
Identity without Accounts

>>

Anonymous 07/20/11(Wed)15:30:34 No. 2832974298

alternative identity signaling

>>



timestamping in photos
suggests authenticity

>>



slang helps delineate
group membership

4chan 'dialect' is nontrivial

required familiarity
with current trends and memes

Ingroup and Outgroup

>>

Anonymous 07/20/11(Wed)15:30:34 No. 2832974298

Triforcing as a classic identity barrier

>>

Anonymous 08/30/09(Sun)17:59:38 No.156666179



>>

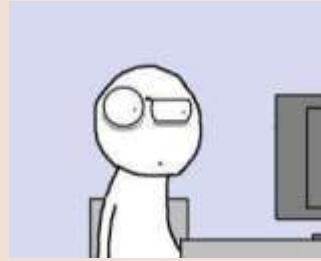
Anonymous 07/03/10(Sat)10:32:29 No.247423XXX

sweet im not a newfag anymore



TL;DR: Conclusion

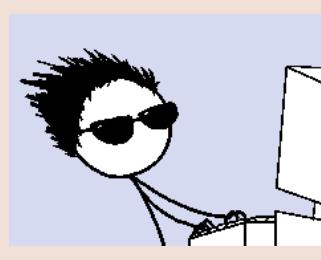
>>



Content 07/20/11(Wed)15:30:05 No.339760588

dominated by playful exchanges of images and links

>>



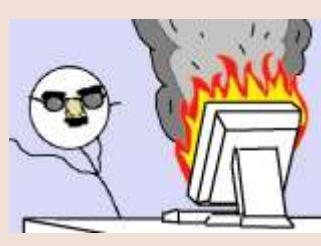
Ephemerality 07/20/11(Wed)15:30:05 No.339760588

median thread spends 5 seconds on the first page

expires completely in 5 minutes

content deletion drives many of /b/'s dynamics

>>



Anonymity 07/20/11(Wed)15:30:05 No.339760588

over 90% of posts are made completely anonymously

contrast to pseudonymous and fully identified sites

5% of posts use tripcode identity mechanisms

Come join the lulz

>>

Anonymous 07/20/11(Wed)15:30:34 No. 2832974298

research follows the data
but this turns a blind eye to important communities and movements

our chanthropologist toolkit can help you collect a dataset

>>

Anonymous 07/20/11(Wed)16:23:44 No. 2234534879

meme creation and spread
relationship to online activism and off-site activities



Updated about a month ago by Dori. Added about a year ago by Drew.

MIT Research Paper on 4chan & /b/

In May 2011, a group of scholars at the Massachusetts Institute of Technology and University of Southampton published a [research paper](#) titled "4chan and /b/: An Analysis of Anonymity and Ephemerality in a Large Online Community"^[5] analyzing various environmental factors and user behaviors of the Anonymous community, particularly in the notorious /b/ board^[6].

4chan and /b/: An Analysis of Anonymity and Ephemerality in a Large Online Community

Michael S. Bernstein¹, Andrés Monroy-Hernández¹, Drew Harry¹,

Paul André², Katrina Panovich¹ and Greg Vargas¹

¹Massachusetts Institute of Technology

Cambridge, MA 02139

{msbernst, amonroy, dharry, kp, gvargas}@mit.edu

²University of Southampton

SO17 1BJ, United Kingdom

pa2@ecs.soton.ac.uk

chan and /b/. emerality in a Large

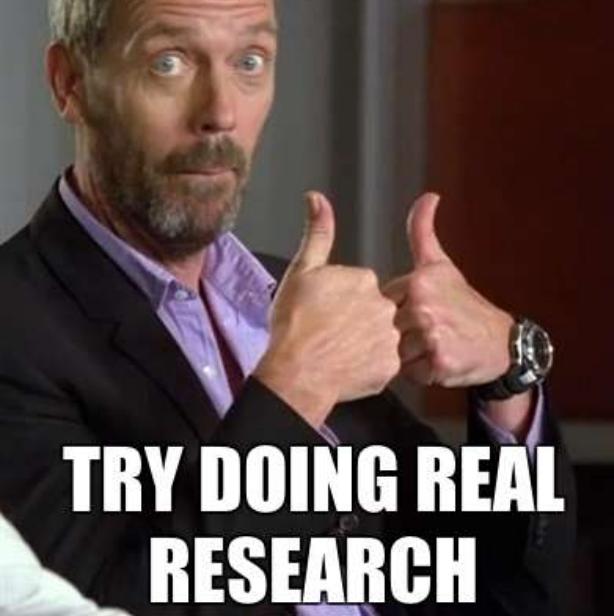
essage board teaches us

AUTHOR



I AM DISAPPOINT

COOL PAPER BRO



TRY DOING REAL
RESEARCH

DIDN'T CONTROL FOR SOLAR FLARES IN
YOUR ANALYSIS

REJECT

SUBMIT PAPER

REFRESH EMAIL EXPECTING REVIEWS



I USED THE FOUR CHAN
ONCE

EXPERTISE = 4

4chan and /b/
An Analysis of Anonymity and Ephemerality in a Large Online Community

Abstract

We present two studies of ephemerality and anonymity online using the popular discussion board *4chan* at 4chan.org, where users often pseudonymously post as anonymous or under various handles. Researchers and practitioners have often argued that user identity and data permanency are central tools in the design of online communities; we explore how *4chan* succeeds despite being based on entirely anonymity and extremely short-lived posts.

We hypothesize that the lack of user identity and data permanency in *4chan* may be a key factor in its success. In our first study, we find that users are more likely to post in *4chan* than in other online communities because they feel that their posts will not be remembered. In our second study, we find that users are more likely to post in *4chan* than in other online communities because they feel that their posts will not be remembered.

Our results support these hypotheses. We find that users are more likely to post in *4chan* than in other online communities because they feel that their posts will not be remembered. We also find that users are more likely to post in *4chan* than in other online communities because they feel that their posts will not be remembered.

Keywords: anonymity, ephemerality, online communities, 4chan, /b/

Introduction

Researchers and practitioners have often argued that user identity and data permanency are central features to the design of online communities. Real names and pseudonyms can help “protect” users from being “accosted” (Miller 2003), whereas anonymous users will make communication impersonal and undermine credibility (Giltz, Johnson and Tureff 1998; Rains 2007). Influential industry players like Facebook argue that pseudonyms and multiple identities show a “lack of integrity” (Kirkpatrick 2010). Similarly, data permanency is the norm. Quora states “you can’t control your past with the hope of ‘erasing’ national memory” (Ackerman 1994); communities highly value their searchable archives (Miller 2000), and search engines make content findable years after it is created (Rosen 2010).

While some argue that anonymous contributions and ephemeral participation online can be desirable (Larue Copyright © 2011, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

In this paper, we perform a content analysis and two derivative studies of *4chan*, focusing on anonymity and ephemerality. In our first study, we use a dataset of 5.5 million posts

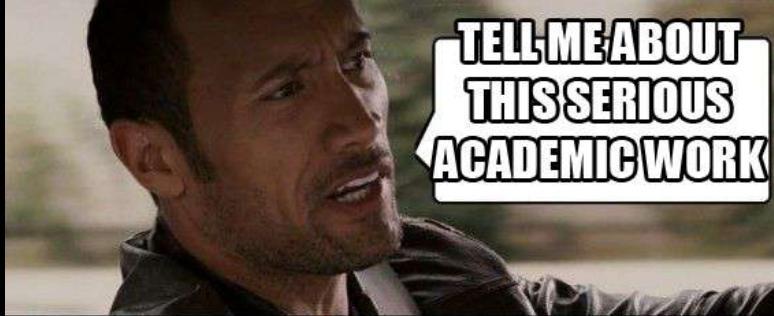
and Resnick 2004; Ren, Kraut, and Kiesler (forthcoming) 2004). We have a limited understanding of how anonymity and ephemerality contribute to the success of *4chan*, especially at a large scale. In this paper we analyze one such community that claims over seven million users: the imageboard *4chan* (Poole 2010). We focus on *4chan*'s first and most popular board, the “random” board */b/* (Figure 1).

Our goal is to use *4chan* as a lens to understand the concepts of anonymity and ephemerality online. First, unlike many other online communities, posts on *4chan* are temporary. This makes *4chan* an ideal research subject for studying the effects of ephemeral data on user behavior. Second, *4chan* is a well-known and well-studied imageboard, making it an ideal testbed for fast-paced, high-volume interactions. While there are many theoretical models of user behavior, few have been applied to this type of study.

Readers may know of *4chan* and */b/* from their influence on internet culture and media coverage of their off-site activities. *4chan* is a prolific meme factory, originating popular memes like LOLcats (Rolleff 2007) and nekkidling (Lekach 2009). Additionally, some *4chan* and */b/* members have participated in highly visible offline activism. These activists include coordinating *Time Magazine* poll to elect its founder the “World’s Most Inflammatory Person” (Time Staff 2009) and participating in hacktivist group “Anonymous.” Anonymous has executed highly visible protests of the Church of Scientology (Coleman 2011) and DDoS attacks against Mastercard and Paypal in support of WikiLeaks (Wright 2010). Some *4chan* and */b/* users, while minors have brought the same positive media attention (Blepy-Warrin 2006), harassment and activism have often been negatively and sensationalized covered in the media (e.g., “*4chan* is the ‘internet hate machine,’ says Phil Shponcer on Fox News [2010], and many media outlets have seized on the image of the *4chan* user as a ‘nerd’” Poole 2010). Much of this coverage is in response to *4chan*’s distinctive culture – a culture that merits critical analysis beyond the scope of this paper. Here, we focus on the board’s design choices of anonymity and ephemerality, and how they may support or enhance this culture.

In this paper, we perform a content analysis and two derivative studies of *4chan*, focusing on anonymity and ephemerality. In our first study, we use a dataset of 5.5 million posts





Acknowledgments: danah boyd, Amy Bruckman, Gabriella Coleman, Danyel Fisher, Benjamin Mako Hill, Alex Leavitt, Rob Miller, Lisa Nakamura, Chris Schmandt, Christina Xu and Sarita Yardi



Reactions on /b/

>>

Find it funny or weird 07/20/11(Wed)15:30:34 No. 2832974298

ITT: /b/ gets trolled by the MIT

I bet the FBI paid MIT to study us

That's the britfags' tax money at work lol, paying scientists to fap over Emily Watson fakes

Lol at how formally they present what goes on here. Like newt gingrich is some isolated tribe in the rainforest with unique and primitive customs.

>>

Engaged with the content 07/20/11(Wed)15:30:34 No. 2832974298

skimmed the whole paper. interesting. i especially like seeing the breakdown of what types of threads are posted most. I would like to see the themed section broken down ie. 20% CP 8% furfag 17% raff u roose 2% fingerbox 5% sup my name is john and you're all a bunch of candy-ass s

This is... This is actually pretty damn interesting.

I think something is lost to the analysts here. The fact that one of the most popular /b/ activities is "pissing into the sea of piss".

Time of Day

>>

primarily an after-work, after-school phenomenon

