

# computational social media

## lecture 3: tweeting

### part 5

daniel gatica-perez

# announcements

## reading #4

S. Vosoughi, D. Roy, S. Aral, The spread of true and false news online, Science, 359, Mar. 2018

## reminder:

we won't meet the next 2 weeks (15.04 and 22.04)

next lecture: 29.04

# project schedule & evaluation

## 1. team building: **DONE**

email the list of your team members on **Week 2: Fri 04.03.2022**

each team will have a designated project mentor

## 2. project pitch : **DONE**

5-minute presentation of your project on **Week 5: Fri 25.03.2022**

**structure:** title, problem, goals, approach, evaluation

## 3. project progress presentation on **Week 10: Fri 29.04.2022**

5-minute presentation per team about project progress

## 4. final project presentation on **Fri 10.06.2022**

talk: 25-minute presentation + 20-minute questions

reserve day from 09:00-16:00

## 5. final project report by **Fri 17.06.2022**

ACM conference paper format (6 pages + references + appendix)

# this lecture



a human-centric view of twitter

1. introduction
2. twitter users & uses
3. understanding large-scale human behavior
4. inferring real-world events & trends
5. spreading information in the real world

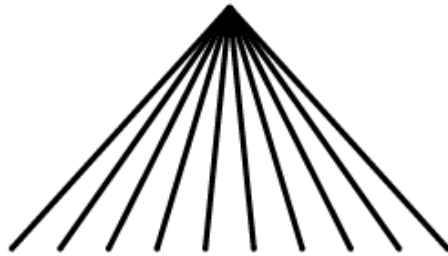
# spreading information in the real world

1. who says what to whom on twitter
2. cascading behavior in networks
- 3. structural virality of online diffusion**
4. twitter and the news

## **3. structural virality of online diffusion**

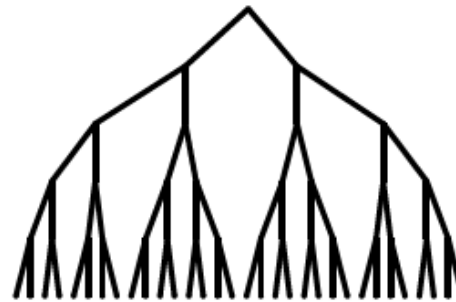
# two mechanisms to reach a large population

A Schematic Depiction of Broadcast vs. Viral Diffusion,  
Where Nodes Represent Individual Adoptions and Edges  
Indicate Who Adopted from Whom



## **Broadcast diffusion:**

“Large burst of adoptions  
from a single parent node”



## **Viral diffusion:**

“Multigenerational branching process  
in which any one node directly “infects”  
only a few others”

“Popular content might have benefited from some possibly  
complicated combination of both mechanisms”

# structural virality

“It quantifies the distinction between broadcast and viral diffusion and allows for interpolation between them, characterizing the structure of observable adoption patterns”

## **cascade size:**

total number of nodes involved in a cascade

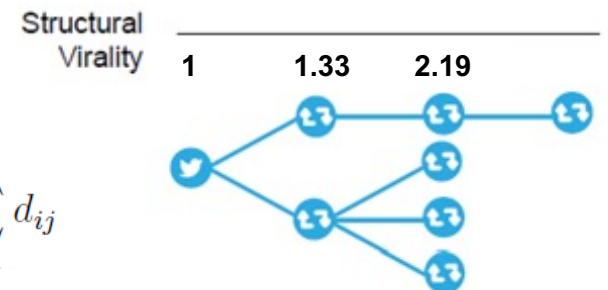
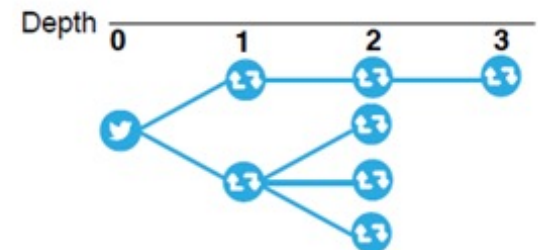
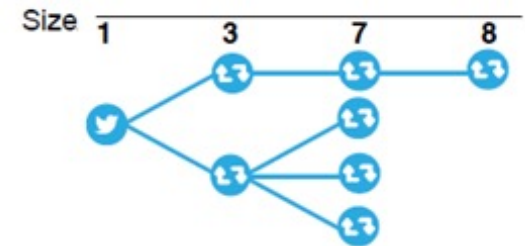
## **cascade depth:**

maximum depth of the nodes in the cascade (depth of a node is the number of edges from the node to the root node)

## **cascade structural virality :**

average distance between all the nodes in the cascade (for  $n > 1$  nodes)

$$v = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n d_{ij}$$





# Twitter dataset and pre-processing

1 billion **diffusion events (cascades)**

## **definition of event:**

- + independent introduction of a content item (URLs to websites)
- + 4 categories: video, image, news stories, and petition
- + all re-postings of the item

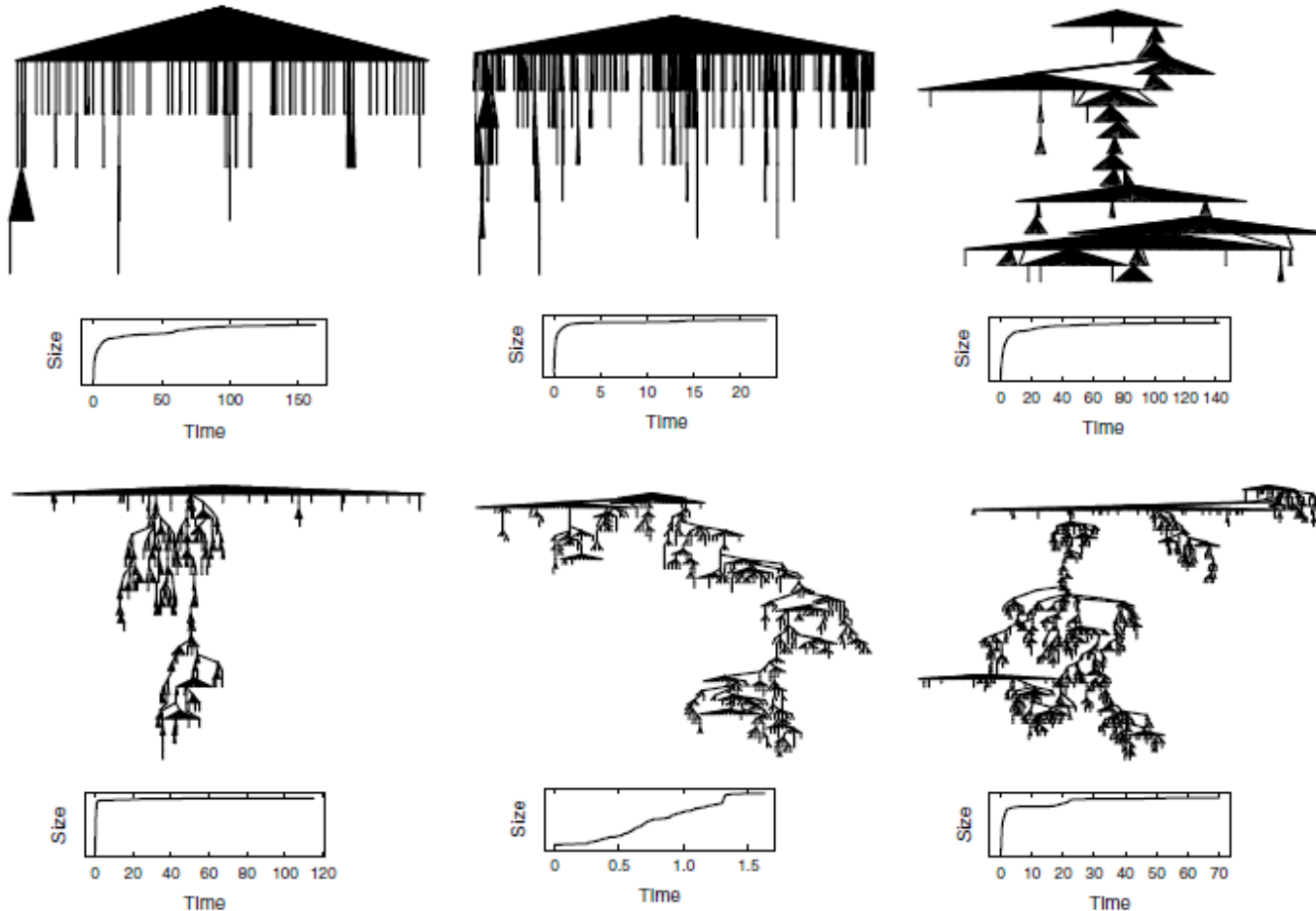
**99% of adoptions** correspond to either to root nodes (cascade size = 1) or immediate followers of root nodes (cascade depth = 1)

# structural diversity

diffusion trees containing at least 100 nodes

number of studied cascades: 219,000 (0.025% of total)

A Random Sample of Cascades Stratified and Ordered by Increasing Structural Virality, Ranging from 2 to 50



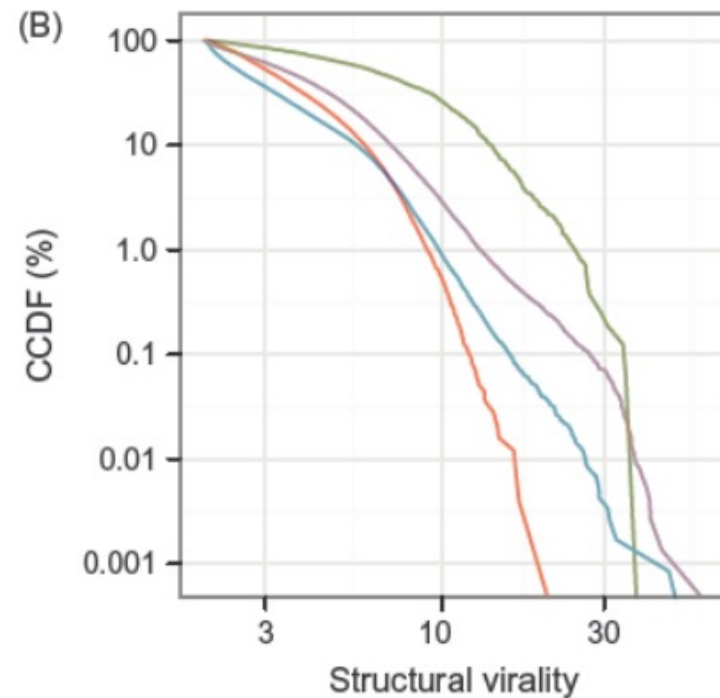
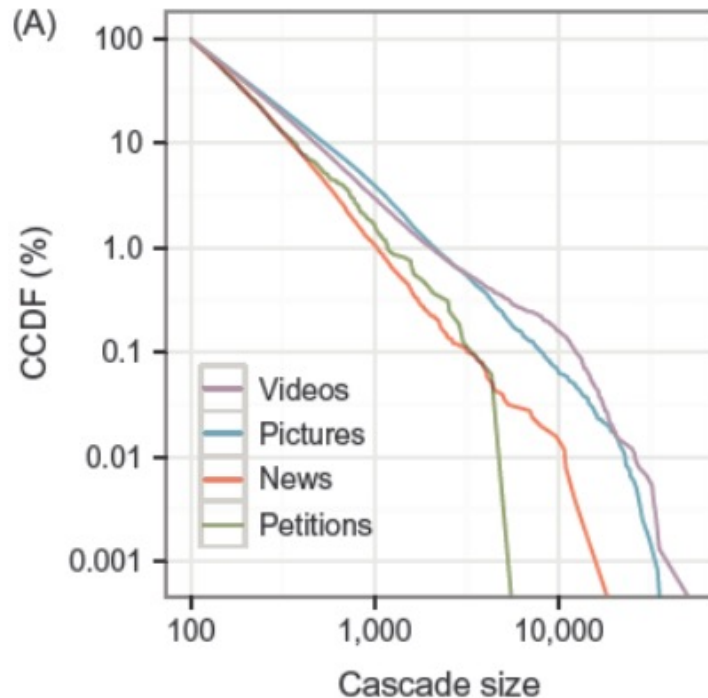
**key result:**  
diversity in structure across the entire range of possible options

*Notes.* For ease of visualization, cascades were restricted to having between 100 and 1,000 adopters. Cumulative adoption curves (i.e., total cascade size over time) are shown below each cascade, with time indicated in hours. For visual clarity, the adoption curves terminate at 99% of the final cascade size.

# cascade size & structural virality for different content

Size and Structural Virality Distributions on a Log-Log Scale for Cascades Containing at Least 100 Adopters, Separated by Domain

CCDF, complementary cumulative distribution function.



**Images & videos:**  
largest cascades  
attract tens of  
thousands of  
retweets

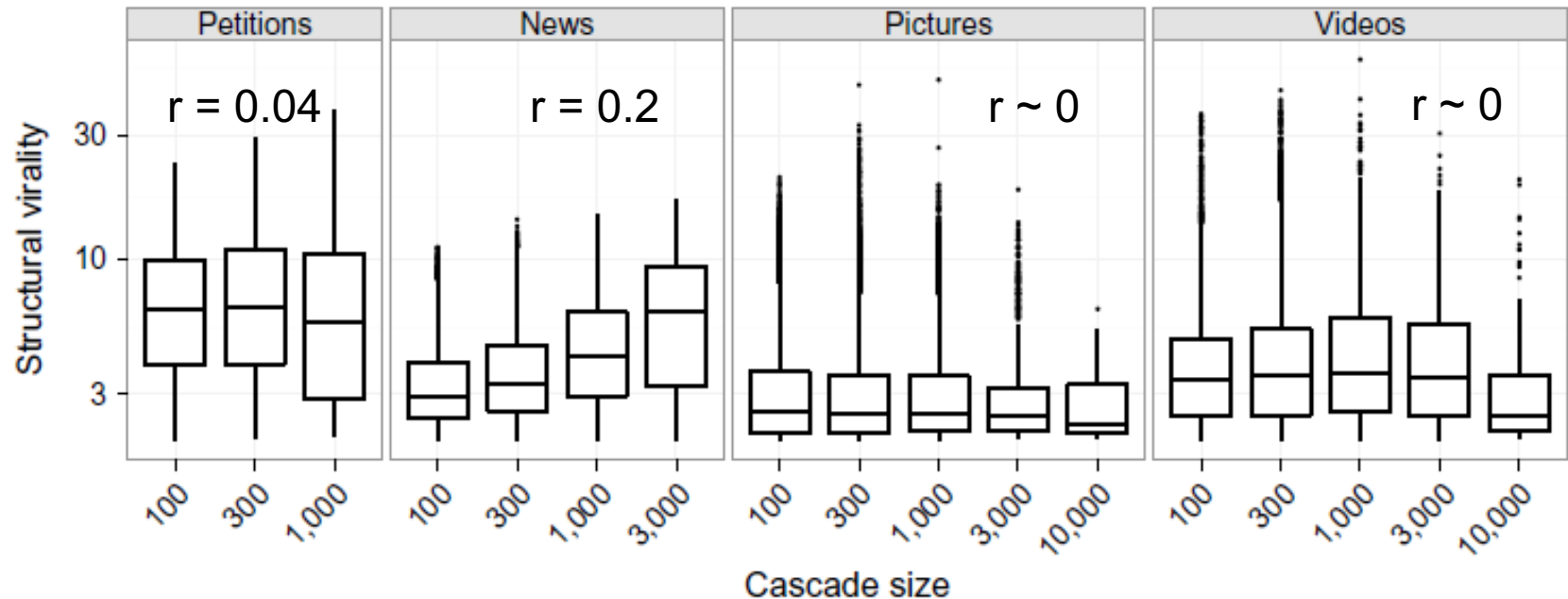
**News** cascades  
are smaller;  
**petitions** are the  
smallest

**Petitions** are  
more structurally  
viral than other  
content

No bimodal  
distribution of  
structural  
virality, rather  
continuous

# links between cascade size and structural virality

Box Plot of Structural Virality by Size on a Log-Log Scale, Separated by Domain



*Note.* Lines inside the boxes indicate median structural virality, whereas the boxes themselves show interquartile ranges.

For **photos**, **videos**, and **petitions**, median structural virality remains relatively invariant with respect to cascade size (no correlation)

For **news**, the correlation between cascade size and structural virality is weak

# spreading information in the real world

1. who says what to whom on twitter
2. cascading behavior in networks
3. structural virality of online diffusion
4. **twitter and the news**

## **4. twitter and the news**

**citizen journalism**

+ Janis Krums was on commuter ferry to rescue passengers of US Airways Flight 1549

+ photo posted it on Twitter

+ shared with 170 followers; spread quickly

+ news networks started interviewing him

+ photo was earliest to appear; made front page of newspapers

+ LA Times: "...among the most striking instances yet of instant **citizen reporting**"

+ Krums's Twitter followers grew to 10,000

+ Hollywood movie in 2016



Twitter's photo sharing site

<http://twitpic.com/135xa> - There's a plane in the Hudson. I'm on the ferry going to pick up the people. Crazy.



15.01.2009



# twitter & citizen journalism

## PROS

- + global coverage
- + real-time coverage
- + experienced journalists don't have to be there
- + lower costs
- + many sources



## CONS

- + journalistic standards & rigor
- + credibility, authenticity
- + is it just crowdsourcing of free stories?

**collective action**



credit (cc): Ramy Raof @ flickr, 11.02.2011, <https://www.flickr.com/photos/ramyraoof/5436339799/>

# Egypt: Timeline of Communication Shutdown during the Revolution

population (2011) ~ 80 million

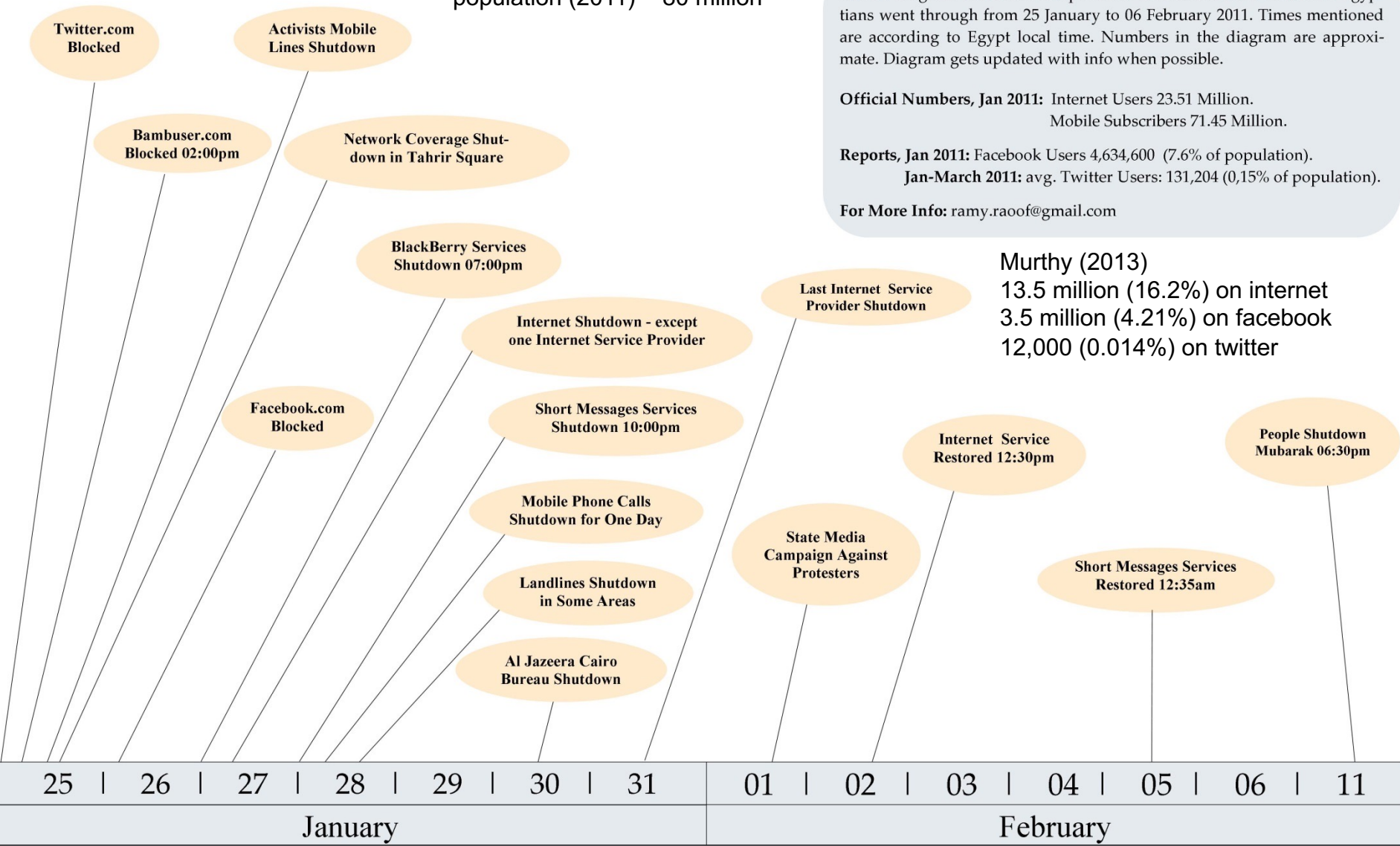
**About:** diagram to illustrate sequence of communications shutdown Egyptians went through from 25 January to 06 February 2011. Times mentioned are according to Egypt local time. Numbers in the diagram are approximate. Diagram gets updated with info when possible.

**Official Numbers, Jan 2011:** Internet Users 23.51 Million.  
Mobile Subscribers 71.45 Million.

**Reports, Jan 2011:** Facebook Users 4,634,600 (7.6% of population).  
**Jan-March 2011:** avg. Twitter Users: 131,204 (0.15% of population).

**For More Info:** ramy.raoof@gmail.com

**Murthy (2013)**  
13.5 million (16.2%) on internet  
3.5 million (4.21%) on facebook  
12,000 (0.014%) on twitter



- + within Egypt, Twitter direct reach was minimal
  - + ... but government's perception was one of a threat
  - + shutting it down legitimized the service and its role
  - + this increased internal awareness (by 03.2011, 131,000 users, 10x increase)
- ...and contributed to media & government reactions externally



- + shutting down the internet had a much larger effect, affecting many more groups of society (e.g. business)

# collective action and the threshold model of diffusion of innovations

situation: participating in a protest against a regime

- + **collective action** produces payoffs only if enough people participate
- + a regime like the one we just discussed tends to **limit communication** among citizens

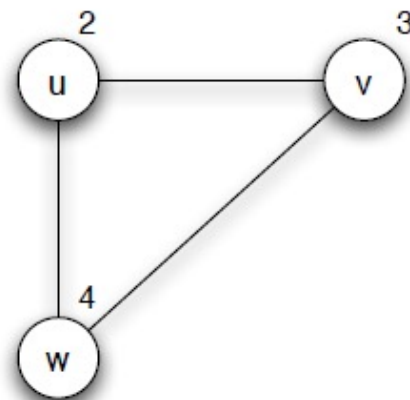
**pluralistic ignorance**: “erroneous estimates about the prevalence of certain opinions of the population at large”

# effect of knowledge on collective action

assume each person has a personal threshold to participate in a protest (at least  $K$  people including the person herself)

assume each node **only** knows the threshold of self and its neighbors

## example 1



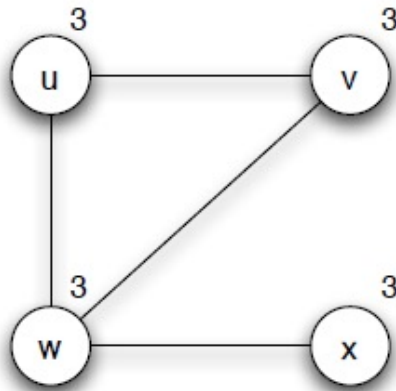
+  $w$  would join if 4 people do; since it only has 2 neighbors (with lower thresholds),  $w$  does not join

+  $v$  knows  $w$ 's threshold, so  $v$  might infer that  $w$  will not join. Since  $v$  requires 3 people to join,  $v$  does not join

+  $u$  requires 2 people, but knows the thresholds of  $w$  and  $v$ , hence it might infer that neither will join, and does not join

+ outcome: none of the nodes join the protest

## effect of knowledge on collective action (2)



### example 2

+  $u$ ,  $v$ , and  $w$  each know that there are three nodes with threshold 3, and this fact is **common knowledge**: each node knows this fact, and each node knows that each node knows it. As a result, all three nodes join

+  $x$  knows the threshold of  $w$ , so it might infer that  $w$  will join, and if so it might join too

+ outcome: all the nodes join the protest

for collective action, this kind of common knowledge is important

social media platforms can be facilitators of this



**disinformation & misinformation**



# AMERICAN VIEWS: TRUST, MEDIA AND DEMOCRACY

A GALLUP/KNIGHT  
FOUNDATION SURVEY

Nationally representative mail survey  
(19,000 U.S. adults aged 18+)

“Americans believe that it is increasingly harder to be a well-informed citizen.

Perceptions of news media are generally negative; perceptions of bias have grown.

Americans are highly concerned about the effects of “fake news” on democracy, but their definitions of “fake news” vary.

They view many newer sources of news positively; less positive about social media.

Concern about the role that tech companies play in news; division on whether they should be regulated.”

January 2018

<https://knightfoundation.org/reports/american-views-trust-media-and-democracy>

# disinformation & misinformation

## **false news**

“news articles that purport to be factual, but contain intentional misstatements of fact with the intention to arouse passions, attract viewership, or deceive.”

## **disinformation**

“inaccurate or manipulated information spread intentionally. This can include false news, or involve more subtle methods, such as feeding inaccurate quotes or stories to innocent intermediaries, or knowingly amplifying biased or misleading information.”

## **misinformation**

“inadvertent or unintentional spread of inaccurate information without malicious intent.”

# disinformation & misinformation (2)

## **false amplifiers**

“coordinated activity by inauthentic accounts with the intent of manipulating political discussion (discouraging specific parties from participating in discussion, or amplifying sensationalistic voices over others).”

## **information (or influence) operations**

“actions taken by governments or organized non-state actors to distort domestic or foreign political sentiment, most frequently to achieve a strategic or geopolitical outcome. They can use a combination of methods, such as false news, disinformation, or false amplifiers aimed at manipulating public opinion.”

SOCIAL SCIENCE

# The spread of true and false news online

Soroush Vosoughi,<sup>1</sup> Deb Roy,<sup>1</sup> Sinan Aral<sup>2\*</sup>

We investigated the differential diffusion of all of the verified true and false news stories distributed on Twitter from 2006 to 2017. The data comprise ~126,000 stories tweeted by ~3 million people more than 4.5 million times. We classified news as true or false using information from six independent fact-checking organizations that exhibited 95 to 98% agreement on the classifications. Falsehood diffused significantly farther, faster, deeper, and more broadly than the truth in all categories of information, and the effects were more pronounced for false political news than for false news about terrorism, natural disasters, science, urban legends, or financial information. We found that false news was more novel than true news, which suggests that people were more likely to share novel information. Whereas false stories inspired fear, disgust, and surprise in replies, true stories inspired anticipation, sadness, joy, and trust. Contrary to conventional wisdom, robots accelerated the spread of true and false news at the same rate, implying that false news spreads more than the truth because humans, not robots, are more likely to spread it.

# Twitter labeling individual messages



**Donald J. Trump**  @realDonaldTrump · 9h

There is NO WAY (ZERO!) that Mail-In Ballots will be anything less than substantially fraudulent. Mail boxes will be robbed, ballots will be forged & even illegally printed out & fraudulently signed. The Governor of California is sending Ballots to millions of people, anyone.....



[Get the facts about mail-in ballots](#)

 22K

 24.3K

 86.7K



**Donald J. Trump**  @realDonaldTrump · 9h

...living in the state, no matter who they are or how they got there, will get one. That will be followed up with professionals telling all of these people, many of whom have never even thought of voting before, how, and for whom, to vote. This will be a Rigged Election. No way!



[Get the facts about mail-in ballots](#)

# Updating our approach to misleading information

By [Yoel Roth](#) and [Nick Pickles](#)

Monday, 11 May 2020 [Twitter](#) [Facebook](#) [LinkedIn](#) [Share](#)

In serving the public conversation, our goal is to make it easy to find credible information on Twitter and to limit the spread of potentially harmful and misleading content. Starting today, we're introducing new labels and warning messages that will provide additional context and information on some Tweets containing disputed or misleading information related to COVID-19.

In March, we [broadened our policy](#) guidance to address content that goes directly against guidance on COVID-19 from authoritative sources of global and local public health information. Moving forward, we may use these labels and warning messages to provide additional explanations or clarifications in situations where the risks of harm associated with a Tweet are less severe but where people may still be confused or misled by the content. This will make it easier to find facts and make informed decisions about what people see on Twitter.

## New labels and warnings

During active conversations about disputed issues, it can be helpful to see additional context from trusted sources. [Earlier this year](#), we introduced a new label for Tweets containing synthetic and manipulated media. Similar labels will now appear on Tweets containing potentially harmful, misleading information related to COVID-19. This will also apply to Tweets sent before today.



**The New York Times**  @nytimes · Nov 6, 2020



Since early Tuesday morning, Twitter has labeled 38% of President Trump's tweets and retweets with warnings that said he made misleading claims about the electoral process, according to a tally by The New York Times.



### Twitter Has Labeled 38% of Trump's Tweets Since Tuesday

The company had said it would closely monitor posts that contain misinformation about the election or prematurely declare an outcome.

[nytimes.com](https://www.nytimes.com)

 270

 1.5K

 8.3K



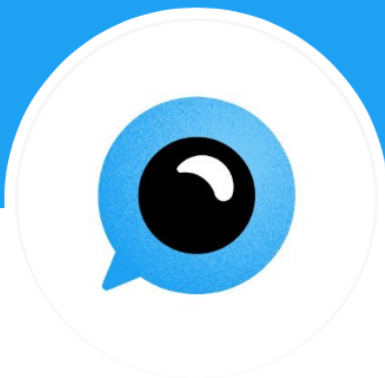
<https://twitter.com/nytimes/status/1324500511937208321>





**Birdwatch** 

99 Tweets



[Follow](#)

**Birdwatch** 

@birdwatch

Empowering the Twitter community to create a better-informed world.

Questions? Tweet/DM us.



[birdwatch.twitter.com](https://birdwatch.twitter.com)



Born January 25



Joined October 2020

0 Following

**26.6K** Followers

<https://twitter.com/birdwatch>

# Introducing Birdwatch, a community-based approach to misinformation

By [Keith Coleman](#)

Monday, 25 January 2021 [Twitter](#) [Facebook](#) [LinkedIn](#) [Link](#)

People come to Twitter to stay informed, and they want credible information to help them do so. We apply labels and add context to Tweets, but we don't want to limit efforts to circumstances where something breaks our rules or receives widespread public attention. We also want to broaden the range of voices that are part of tackling this problem, and we believe a community-driven approach can help. That's why today we're introducing Birdwatch, a pilot in the US of a new community-driven approach to help address misleading information on Twitter.

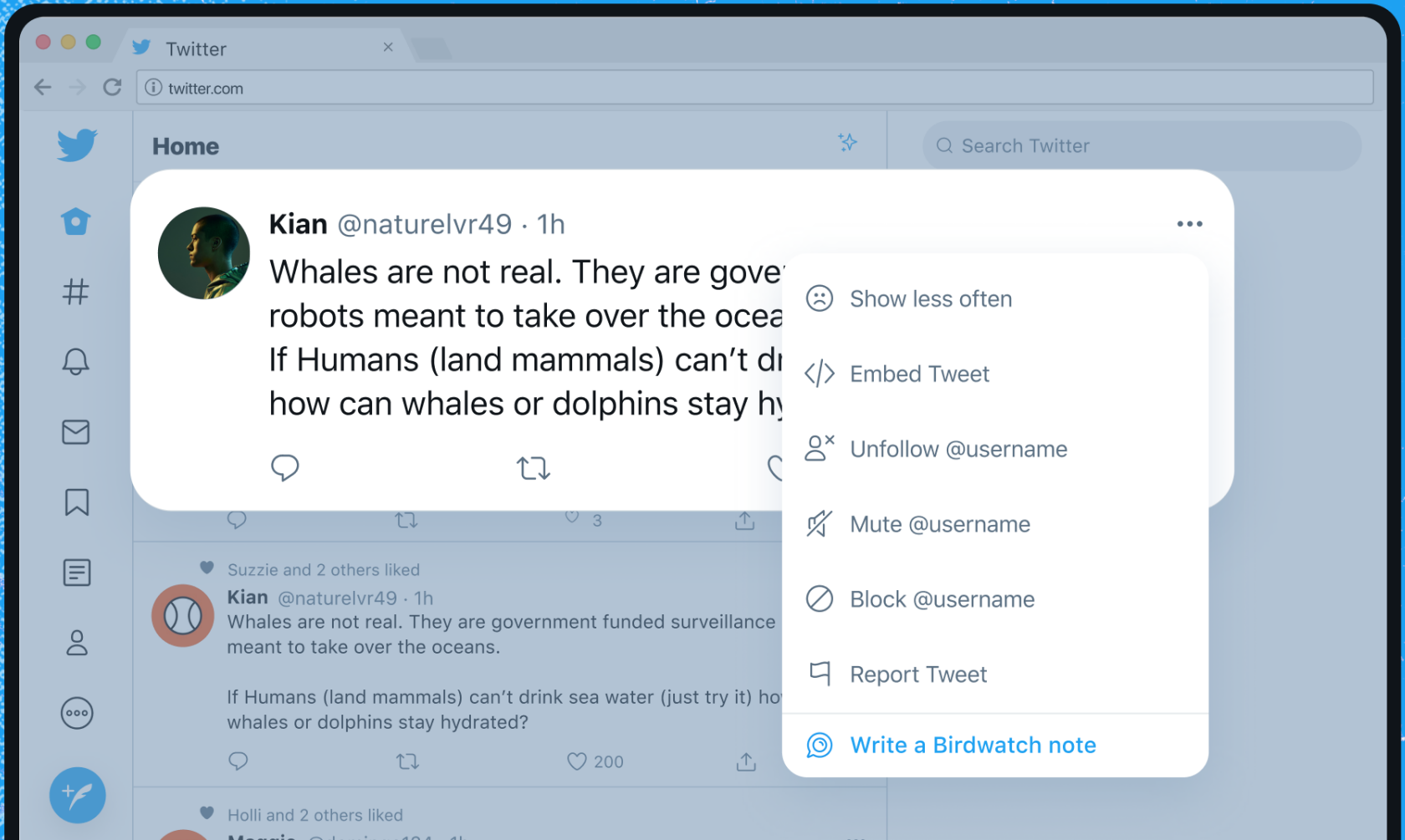
## Here's how it works

Birdwatch allows people to identify information in Tweets they believe is misleading and write notes that provide informative context. We believe this approach has the potential to respond quickly when misleading information spreads, adding context that people trust and find valuable. Eventually we aim to make notes visible directly on Tweets for the global Twitter audience, when there is consensus from a broad and diverse set of contributors.

In this first phase of the pilot, notes will only be visible on a separate [Birdwatch site](#). On this site, pilot participants can also rate the helpfulness of notes added by other contributors. These notes are being intentionally kept separate from Twitter for now, while we build Birdwatch and gain confidence that it produces context people find helpful and appropriate. Additionally, notes will not have an effect on the way people see Tweets or our system recommendations.

**watch video:**

<https://twitter.com/TwitterSupport/status/1353766523664531459>



## 1. “Whales aren’t real”

Here’s a Tweet making misleading claims about whales, followed by helpful and unhelpful Birdwatch notes:



**Birdwatch Example**  
@bwatchexample

Whales are not actually mammals. If Humans (land mammals) can’t drink seawater — just try it! — how can supposed sea mammals like whales stay hydrated?

5:09 PM · Jan 25, 2021

98   Reply   Share this Tweet

[Read 86 replies](#)

### Helpful Birdwatch note

Whales are indeed mammals. Marine mammals are able to “stay hydrated” because their kidneys have evolved to excrete more salt + reclaim more water than humans and many other mammals can. They also obtain water from their food. This is widely documented, for example in <http://reputable...>

- ✓ Easy to understand
- ✓ Directly addresses all of the Tweet’s claims
- ✓ Cites high-quality source, which has robust information on the topic

### Unhelpful Birdwatch note

Just another conspiracy theory from another terrible TV host.

- ✗ Uses inflammatory language
- ✗ Attacks author instead of addressing content of tweet
- ✗ Does not provide additional information about central points made in tweet

# Twitter to show 'Birdwatch' community fact-checks to more users, following criticism

Sarah Perez @sarahintampa / 3:00 PM GMT+1 • March 3, 2022

 Comment

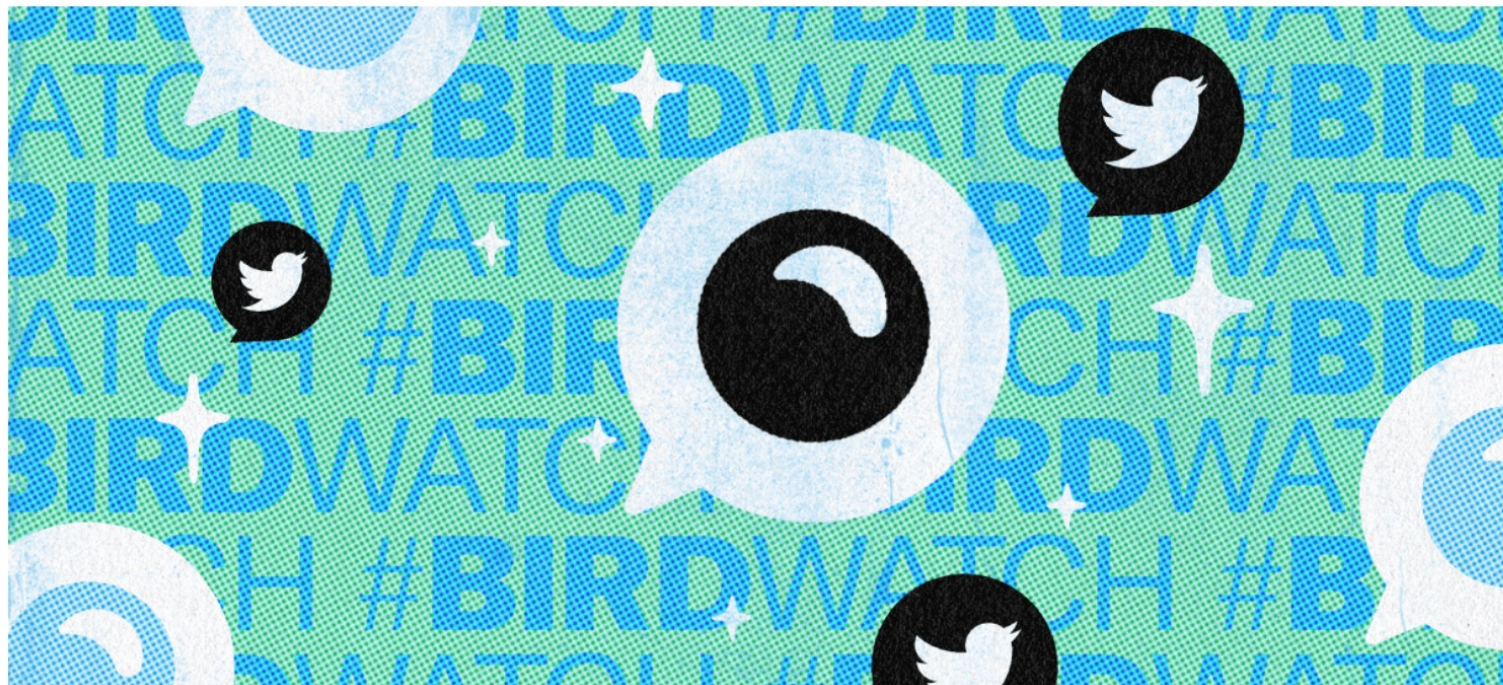


 Image Credits: Twitter

Twitter said it's expanding access to [Birdwatch](#), its community-based fact-checking initiative [first revealed](#) in October 2020. The service so far has been tested by a small group of 10,000 contributors who have invested time in writing and rating notes to add more context to tweets that could be potentially misleading. The company, however, isn't opening up Birdwatch to more contributors, but rather will make those notes viewable and ratable by more U.S. users.

<https://techcrunch.com/2022/03/03/twitter-to-show-birdwatch-community-fact-checks-to-more-users-following-criticism/>

# false news at the scale of the information ecosystem

Analysis on two US-representative sources:

- **Nielsen's** individual-level TV panel (100k people) and desktop web panel (90k people)
- **Comscore's** aggregated app & browser (desktop & mobile) traffic data

Finding #1: Daily US news consumption is small fraction of all consumed media (14% of time)

Finding #2: Most news are consumed via TV (5+ times more than online, measured by time)

Finding #3: Fake news correspond to 0.15% of US daily consumed media (measured by time)

## basic definitions



”**Nielsen TV news consumption:** time spent watching any of 400 programs classified as ‘news’: hard news, magazine news, morning shows, entertainment news, and late-night comedy shows.”

“**Comscore’s mobile & desktop news consumption:** time spent on any articles published on 800 websites that cover hard news: politics, business, and U.S. and international affairs.”

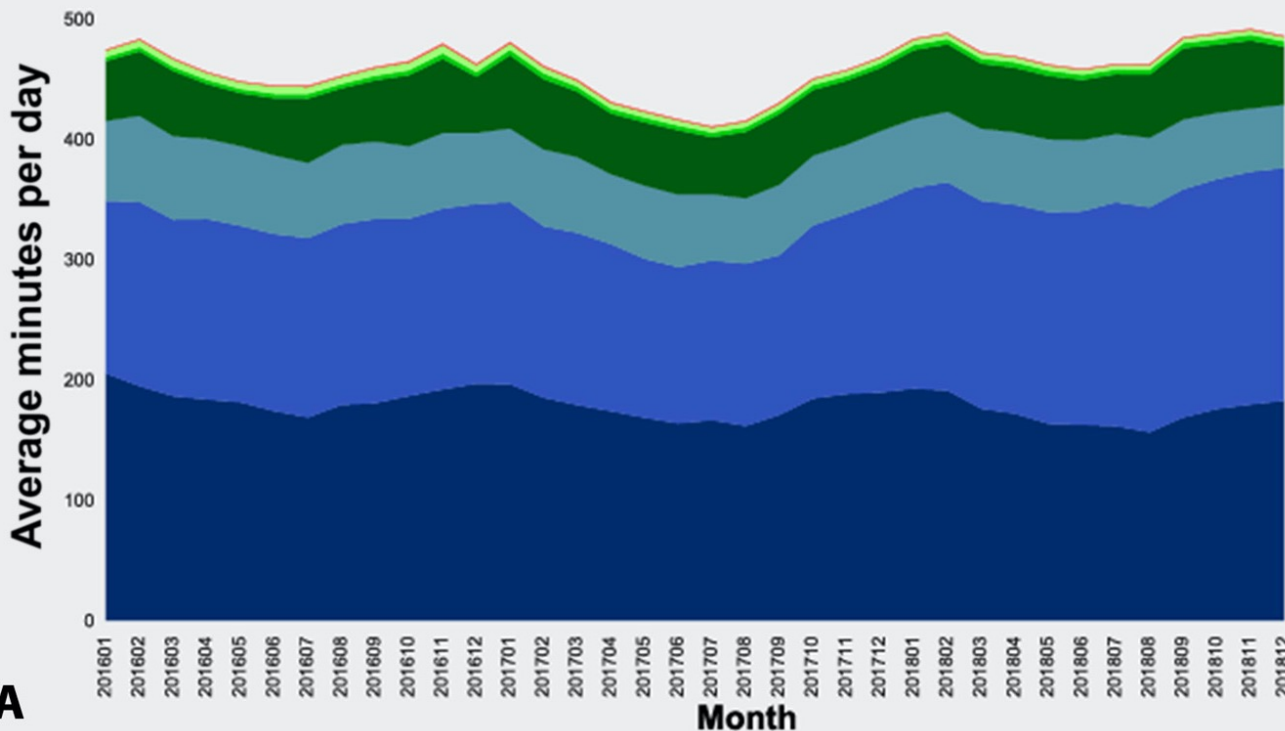
“**Online fake news consumption:** time spent on 98 websites previously identified by researchers, professional fact checkers, and journalists as sources of fake, deceptive, low-quality, or hyperpartisan news. With the exception of YouTube, fake news is defined at the publisher or URL level.”

“**Nielsen’s passive news consumption:** news snippets, images, headlines, and summaries that appear on a newsfeed or search results page but which the user does not click on, for top social media sites (Facebook, YouTube, Twitter, Reddit) and top search engines (Google, Bing, Yahoo).”





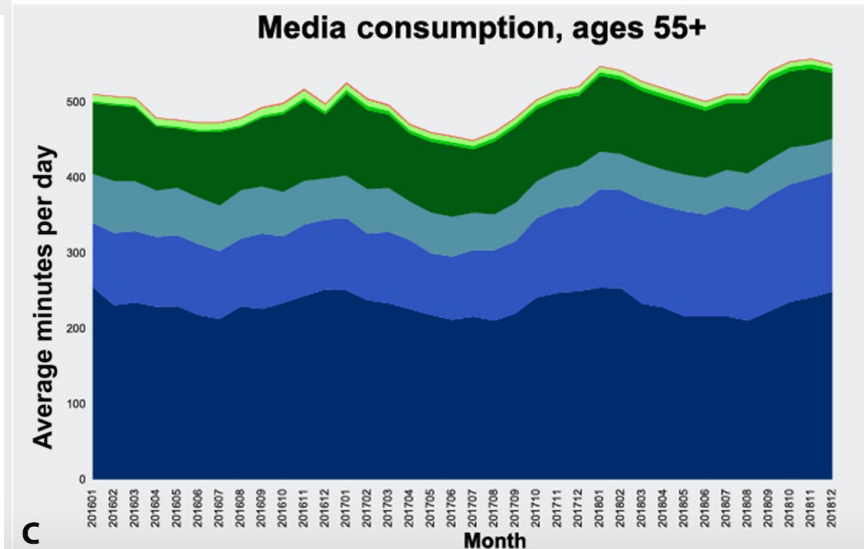
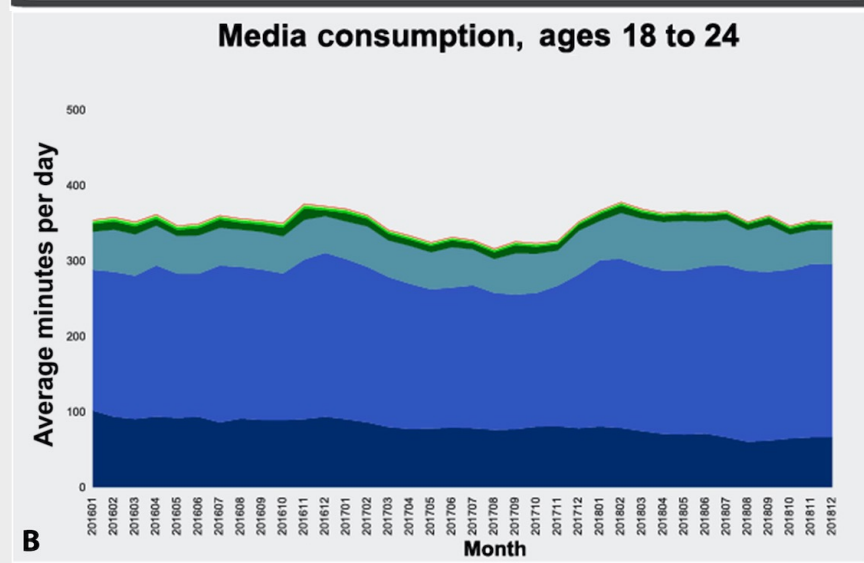
### Media consumption, all (ages 18+)



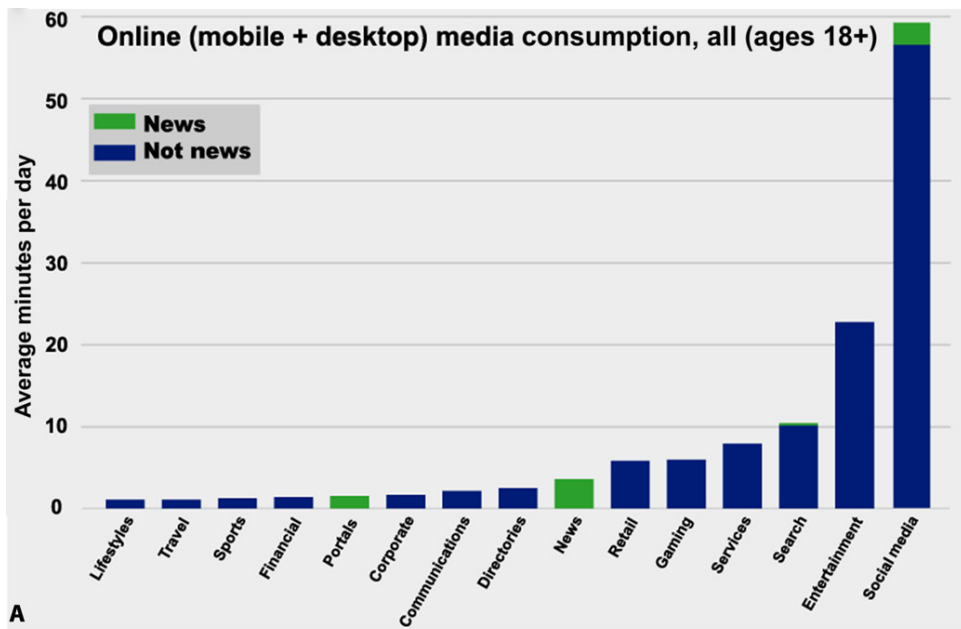
Total daily time:  
460 min/day (7.5h):  
TV, video or music  
streaming, gaming,  
social media, or web  
browsing (desktop or  
mobile)”

Daily US news  
consumption is a  
small fraction of all  
consumed media  
(14% of time)

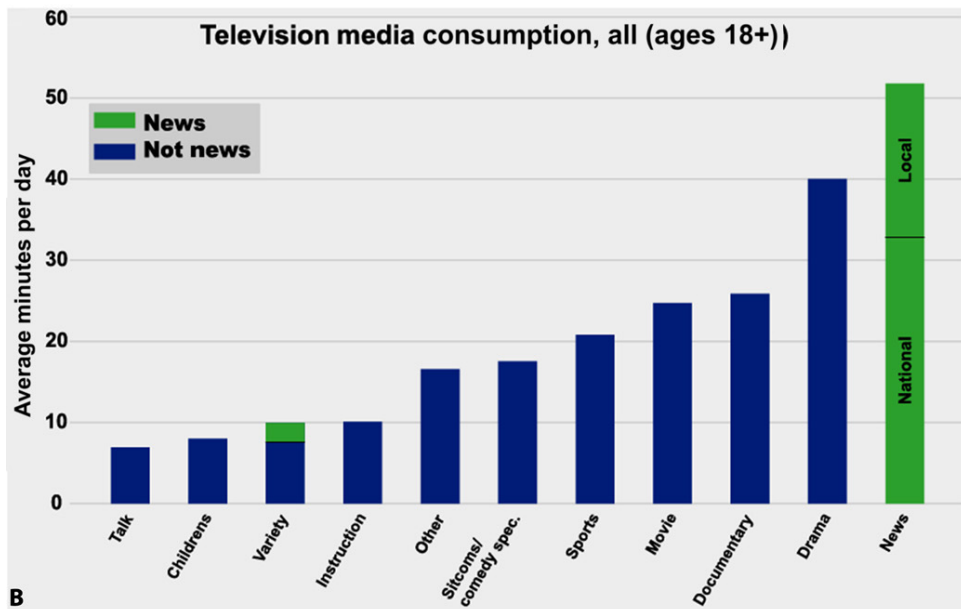
Overall information consumption by platform over time (Jan 2016 - Dec 2018)



Overall information consumption by age & platform over time (Jan 2016 - Dec 2018)



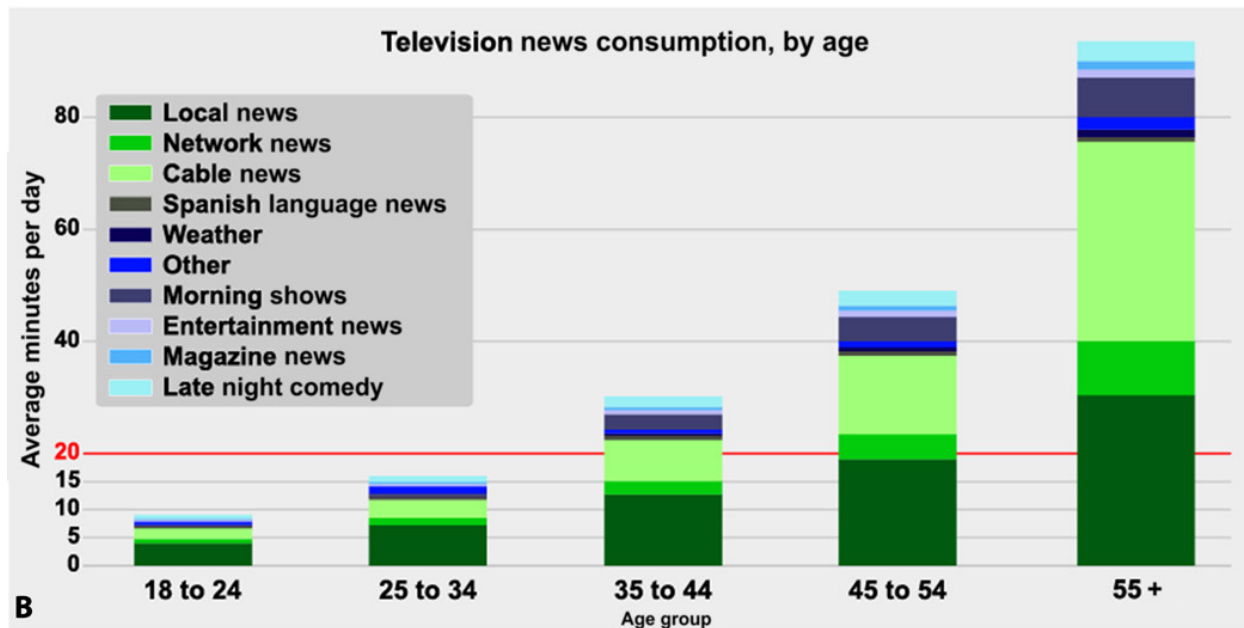
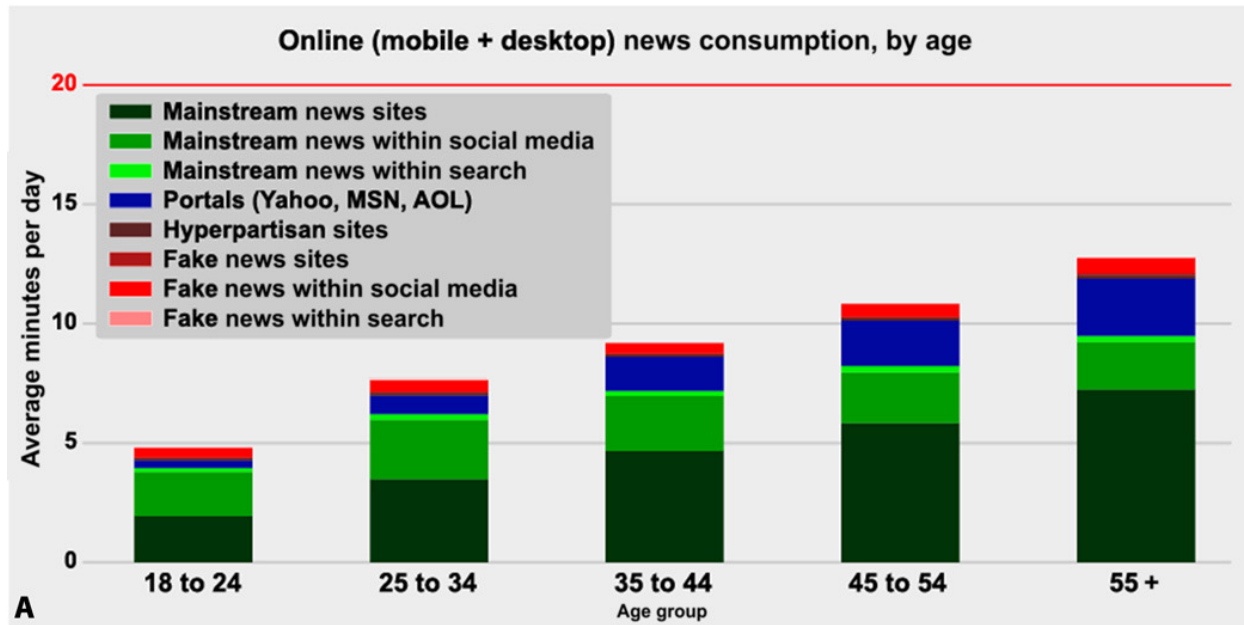
Most news are consumed via **TV** compared to **online**: 54 min vs. 9.7 min (5.5 ratio)



18-24 years-old: 9 min vs. 5 min (1.8 ratio)

55+ years old: 94 min vs. 13 min (7.2 ratio)

**Breakdown of overall media consumption for Online and TV**



## News-only consumption by age

## what is the problem, then?

“Turning to TV, there are no objectively fake news stations of the sort that exist online, i.e., that are exclusively or near exclusively devoted to disseminating deliberate falsehoods while masquerading as legitimate news organizations [...] Nonetheless, misinformation can also manifest itself in regular news programming in the form of selective attention, framing, “spin,” false equivalence, and other forms of bias.”

“Even if its prevalence is low, fake news could be important either because it is **disproportionately impactful** (more impact per minute of exposure) or is **concentrated on certain subpopulations.**”

“The origins of public misinformedness are more likely to lie in the **content of ordinary news** or the **avoidance of news altogether** than in overt fakery [...] Public ignorance or misunderstanding of important political matters could arise out of a combination of (i) ordinary bias and agenda setting in the mainstream media and (ii) the overall low exposure of many Americans to news content in general, especially in written form.”

# what to remember

## structural virality of online diffusion

quantifies distinction between broadcast and viral diffusion

different than cascade size or depth

news, photos, petitions follow different structural virality patterns

## twitter & the news

early days: citizen journalism, collective action

today: misinformation & disinformation

## misinformation & disinformation

built-in risk of socio-technical systems like Twitter

despite progress in NLP, human fact-checking is crucial

exposure time is small at the scale of the info ecosystem, but

what about amplified (or unaccounted for) effects?

**questions?**

daniel.gatica-perez@epfl.ch