



# CS 6474/CS 4803

# Social Computing:

# Credibility

*Munmun De Choudhury*

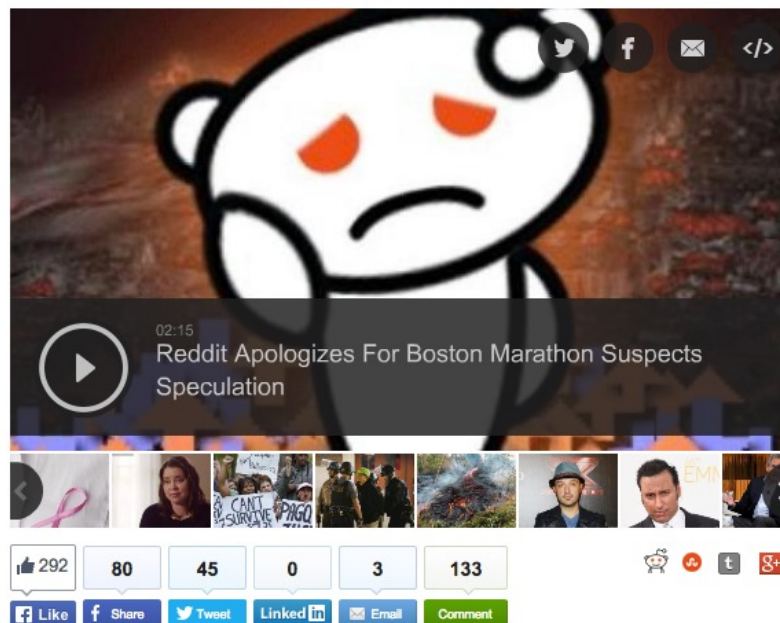
[munmund@gatech.edu](mailto:munmund@gatech.edu)

Week 9 | March 7, 2022

# Reddit Apologizes For Speculating About Boston Marathon Suspects

The Huffington Post | By Katherine Bindley

Posted: 04/22/2013 5:18 pm EDT | Updated: 04/23/2013 7:31 pm EDT



Speculation on social media last week over who was responsible for the [bombing at the Boston Marathon](#) produced its own set of innocent victims: the falsely accused.

Reddit -- which was fiercely criticized for its ["Findthebostonbombers"](#) thread that called out specific people standing near the scene as suspects -- has now issued a [public apology](#) to those whose names were dragged through the mud.

"We all need to look at what happened and make sure that in the future we do everything we can to help and not hinder crisis situations," the statement reads.

#### Sponsored Links



**Odd Trick Fights Diabetes**  
"Unique" Proven Method To Control Blood Sugar In 3 Weeks. Watch Video.  
Smart-Consumer-Lifestyles.com

**Know Where You Stand**

The message, which was published in a blog post on Monday, continues by mentioning the serious ramifications of the rampant speculation that occurred on the site.

Mar 4, 2022, 10:11am EST | 435 views

## Navigating Social Media In Wartime – Is The Truth Still The First Casualty Of War?



**Peter Suci** Contributor ⓘ  
Social Media

Follow



Listen to article 8 minutes



## THE WALL STREET JOURNAL.

English Edition ▼ | Print Edition | Video | Podcasts | Latest Headlines

TECH | FAMILY & TECH: JULIE JARGON

## How to Tell News Fact from Fiction, Even During a War

People have been sharing information about the war in Ukraine on social media without verifying it. News-literacy tactics taught in school can benefit many of us

Can social media  
information credibility  
be inferred  
automatically?



# Information Credibility on Twitter

Carlos Castillo<sup>1</sup>

Marcelo Mendoza<sup>2,3</sup>

Barbara Poblete<sup>2,4</sup>

{chato,bpoblete}@yahoo-inc.com, marcelo.mendoza@usm.cl

<sup>1</sup>Yahoo! Research Barcelona, Spain

<sup>2</sup>Yahoo! Research Latin America, Chile

<sup>3</sup>Universidad Técnica Federico Santa María, Chile

<sup>4</sup>Department of Computer Science, University of Chile

## ABSTRACT

We analyze the information credibility of news propagated through Twitter, a popular microblogging service. Previous research has shown that most of the messages posted on Twitter are truthful, but the service is also used to spread misinformation and false rumors, often unintentionally.

On this paper we focus on automatic methods for assessing the credibility of a given set of tweets. Specifically, we analyze microblog postings related to “trending” topics, and classify them as credible or not credible, based on features extracted from them. We use features from users’ posting and re-posting (“*re-tweeting*”) behavior, from the text of the posts, and from citations to external sources.

We evaluate our methods using a significant number of human assessments about the credibility of items on a recent sample of Twitter postings. Our results shows that there are measurable differences in the way messages propagate, that

directly from smartphones using a wide array of Web-based services. Therefore, Twitter facilitates real-time propagation of information to a large group of users. This makes it an ideal environment for the dissemination of breaking-news directly from the news source and/or geographical location of events.

For instance, in an emergency situation [32], some users generate information either by providing first-person observations or by bringing relevant knowledge from external sources into Twitter. In particular, information from official and reputable sources is considered valuable and actively sought and propagated. From this pool of information, other users synthesize and elaborate to produce derived interpretations in a continuous process.

This process can gather, filter, and propagate information very rapidly, but it may not be able to separate true information from false rumors. Indeed, in [19] we observed that immediately after the 2010 earthquake in Chile, when



# Which features are better?

## Faking Sandy: Characterizing and Identifying Fake Images on Twitter during Hurricane Sandy

Aditi Gupta\*, Hemank Lamba\*\*, Ponnurangam Kumaraguru\*, Anupam Joshi†

\*Indraprastha Institute of Information Technology, Delhi, India

\*\*IBM Research Labs, Delhi, India

†University of Maryland Baltimore County, Maryland, USA

{aditig, pk}@iiitd.ac.in, helamba1@in.ibm.com, joshi@cs.umbc.edu

### ABSTRACT

In today's world, online social media plays a vital role during real world events, especially crisis events. There are both positive and negative effects of social media coverage of events, it can be used by authorities for effective disaster management or by malicious entities to spread rumors and fake news. The aim of this paper, is to highlight the role of Twitter, during Hurricane Sandy (2012) to spread fake images about the disaster. We identified 10,350 unique tweets containing fake images that were circulated on Twitter, during Hurricane Sandy. We performed a characterization analysis, to understand the temporal, social reputation and influence patterns for the spread of fake images. Eighty six percent of tweets spreading the fake images were retweets, hence very few were original tweets. Our results showed that top thirty users out of 10,215 users (0.3%) resulted in 90% of the retweets of fake images; also network links such as follower relationships of Twitter, contributed very less (only 11%) to the spread of these fake photos URLs. Next, we used classification models, to distinguish fake images from

content is posted on OSM, not all of the information is of good quality with respect to the event, like it may be fake, incorrect or noisy. Extracting good quality information is one of the biggest challenges in utilizing information from OSM. Over last few years, people have highlighted how OSM can be used to help in extracting useful information about real life events. But, on the other hand, there have been many instances which have highlighted the negative effects on content on online social media on real life events. The information shared and accessed on social media such as Twitter, is in real-time, the impact of any malicious intended activity, like spreading fake images and rumors needs to be detected and curbed from spreading immediately. Such false and incorrect information can lead to chaos and panic among people on the ground. Since detecting whether images posted are fake or not, using traditional image analysis methods, can be highly time and resource consuming, we explore the option of using Twitter specific features, like the content of the tweet and the user details, in identifying fake images from real.

Castillo et al or Gupta et al do not exploit the wealth of information embedded in the network structure of a user.

# Measuring User Credibility in Social Media

Mohammad-Ali Abbasi and Huan Liu

Computer Science and Engineering, Arizona State University

`Ali.abbasi@asu.edu, Huan.liu@asu.edu`

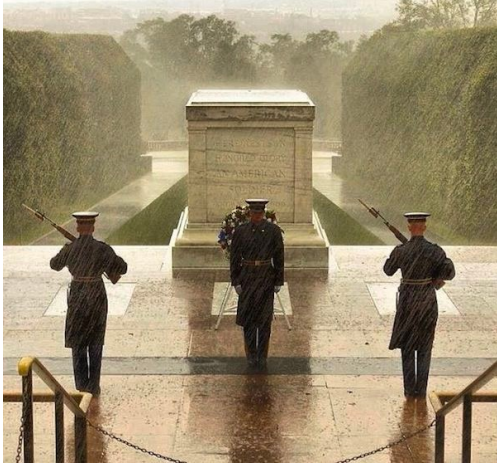
**Abstract.** People increasingly use social media to get first-hand news and information. During disasters such as Hurricane Sandy and the tsunami in Japan people used social media to report injuries as well as send out their requests. During social movements such as Occupy Wall Street (OWS) and the Arab Spring, people extensively used social media to organize their events and spread the news. As more people rely on social media for political, social, and business events, it is more susceptible to become a place for evildoers to use it to spread misinformation and rumors. Therefore, users have the challenge to discern which piece of information is credible or not. They also need to find ways to assess the credibility of information. This problem becomes more important when the source of the information is not known to the consumer.

In this paper we propose a method to measure user credibility in social media. We study the situations in which we cannot assess the credibility of the content or the credibility of the user (source of the information) based on the user's profile. We propose the *CredRank* algorithm to mea-



The supposed  
objectivity of credibility

# Which of these pictures of Hurricane Sandy are Real and which ones are fake?



Amazing picture of hurricane #Sandy decending in New York  
[pic.twitter.com/3mMhCbNq](https://pic.twitter.com/3mMhCbNq)



29 Oct 12

Reply Retweet Favorite



DC Maryland Virginia  
@DMVFollowers



Follow

McDonalds in Virginia Beach flooded.

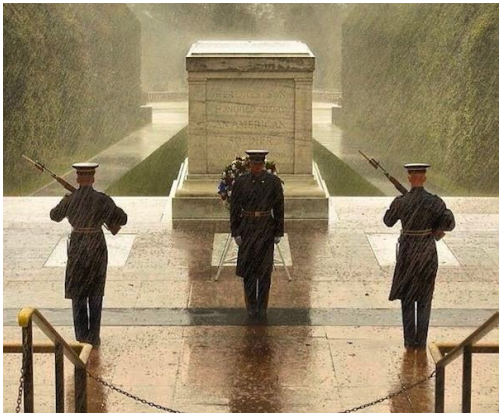
[pic.twitter.com/FZBoCydM](https://pic.twitter.com/FZBoCydM)

Reply Retweet Favorite





# 7 Fake Hurricane Sandy Photos You're Sharing on Social Media



29 Oct 12    Reply    Retweet    Favorite

**DC Maryland Virginia**  
@DMVFollowers



McDonalds in Virginia Beach flooded.  
[pic.twitter.com/FZBoCydM](http://pic.twitter.com/FZBoCydM)  
Reply Retweet Favorite







COMMUNICATIONS OF THE ACM

# Credibility and Computing Technology

By Shawn Tseng, B. J. Fogg

Communications of the ACM, May 1999, Vol. 42 No. 5, Pages 39-44

10.1145/301353.301402

[Comments](#)

VIEW AS:					SHARE:							
----------	--	--	--	--	--------	--	--	--	--	--	--	--

For most of computing's brief history, people have held computers in high regard. A quick review of the popular culture from the past few decades reflects people's general confidence in computing systems. In cinema and literature, computers are often portrayed as infallible sidekicks in the service of humanity. In the consumer realm, computer-based information and services have been marketed as better, more reliable, and more credible sources of information than humans. Consider, for example, computerized weather prediction, computerized automotive analysis, and so-called computer dating. In these and other areas, the public has generally been led to believe that if a computer said it or produced it, it was believable.

But like many aspects of our human society, computers seem to be facing a credibility crisis. Due in part to the popularization of the Internet, the cultural myth of the highly credible computer may soon be history. Although healthy skepticism about computers can be a good thing, if the pendulum swings too far in this direction, computers especially with respect to Web-based content could be viewed as among the least credible information sources, rivaling TV infomercials and supermarket tabloids for such dubious distinction.

What is credibility? What makes computers credible? And what can we, as computer professionals, do to

# Tweeting is Believing?

## Understanding Microblog Credibility Perceptions

One limitation of the work is that their current recruitment method does not include certain demographics that consume tweets, like teenagers or adults without a college degree; education may matter

The paper focused on a rather well-educated and specialized group of participants, and that it failed to contrast results of this population and a more general population



## PDF File Viewer | Download Now

Download Newest Version Now | Excellent Reviews thedocpdfconverter.com

DOWNLOAD



Health ▾

Tech ▾

Enviro ▾

Society ▾

Quirky ▾

## Science News

from research organizations

### Finnish school students outperform US students on 'fake news' digital literacy tasks

Date: May 2, 2019

Source: University of Turku

**Summary:** A recent study revealed students at an international school in Finland significantly outperformed US students on tasks which measure digital literacy in social media and on-line news. The researchers suggest this may be due to the Finnish and International Baccalaureate curricula's different way of facilitating students' critical thinking skills compared to the US system and curriculum.

Share: [f](#) [t](#) [p](#) [in](#) [✉](#)

#### RELATED TOPICS

##### Mind & Brain

- > [K-12 Education](#)
- > [Creativity](#)
- > [Numeracy](#)
- > [Educational Psychology](#)

##### Computers & Math

- > [Educational Technology](#)
- > [Computers and Internet](#)
- > [Photography](#)

#### FULL STORY

A recent study revealed students at an international school in Finland significantly outperformed U.S. students on tasks which measure digital literacy in social media and online news. The researchers suggest this may be due to the Finnish and International Baccalaureate curricula's different way of facilitating students' critical thinking skills compared to the US system and curriculum. The results of this study were published in the *Journal of Research in International Education* in April.



# **Fake images: The effects of source, intermediary, and digital media literacy on contextual assessment of image credibility online**

new media & society  
2019, Vol. 21(2) 438–463

© The Author(s) 2018

Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/1461444818799526

[journals.sagepub.com/home/nms](https://journals.sagepub.com/home/nms)



**Cuihua Shen** 

University of California, Davis, USA

**Mona Kasra**

University of Virginia, USA

**Wenjing Pan, Grace A Bassett  
and Yining Malloch**

University of California, Davis, USA

**James F O'Brien**

University of California, Berkeley, USA

# Finding and Assessing Social Media Information Sources in the Context of Journalism

Nicholas Diakopoulos<sup>1</sup>, Munmun De Choudhury<sup>2</sup>, Mor Naaman<sup>1</sup>

<sup>1</sup>Rutgers University – School of Communication and Information, <sup>2</sup>Microsoft Research  
diakop@rutgers.edu, munmund@microsoft.com, mor@rutgers.edu

**Tottenham Riots**  
402 sources sharing 551 tweets matching "tottenhamriots" or "tottenham"

**Search** (enter search terms here) **Search** **Sort** # times retweeted

**Show Sources** (showing 8 of 10 sources loaded)  
All Ordinary People Journalists / Bloggers Organizations Uncategorized Eyewitnesses

**Show Tweets** All Exclude RTs Images & Videos

**1.1 Filters**  
**1.2 Location**  
**1.3 User Type**  
**1.4 Eyewitness Icon**  
**1.5 Friends' Locations**  
**1.6 Network Sketch**  
**1.7 Entities**

**Daniel Carr, @daniel\_carr** (2 years, 3 months old)  
Myself in 160 characters: Schizophrenic. Also a criminologist  
London, United Kingdom  
Jouno/Blogger 41 RTed 56 Klout

**31 Tweets**  
#tottenham #tottenhamriots Fire near Bruce Grove Station, larger one towards Lordship Lane Aug. 6, 2011, 11:27 p.m.  
#tottenham #tottenhamriots @MrsCheddies by Bruce Grove I mean north of previous fires, on High Rd towards Lordship Lane Aug. 6, 2011, 11:24 p.m.  
@hackneyhive yeah around that area there are 2 fires, one small now, one very large #tottenham #tottenhamriots

**Aidan Rowe, @Aidan\_Rowe** (1 year, 2 months old)  
Post-punk, proto-utopian, anarchist, activist, musician, blogger, student, failed comedian.  
<http://redwriters1.blogspot.com>  
Ordinary Person 23 RTed 49 Klout

**5 Tweets**  
"Why couldn't the people in #Tottenham just have held a nice dignified protest for us to ignore?" - Liberals #tottenhamriots Aug. 7, 2011, 12:49 a.m.  
Any reports of arrests? #tottenham #tottenhamriots Hope everyone is safe. #acab Aug. 7, 2011, 12:06 a.m.  
Anyone using the words "mindless", "hooligans" or "thugs" is a racist and an idiot. #tottenham #tottenhamriots

**Network Sketch**  
213 Followers  
163 Following

**Friends' Locations**  
London, GB 34.48% Glasgow, GB 5.17% Manchester, GB 3.45%

**Top Entities Mentioned Historically**  
Bruce Grove, Tottenham Hale, London, BBC, Haha,

**Network Sketch**  
215 Followers  
395 Following

**Friends' Locations**  
Dublin, IE 43.48% London, GB 4.35% Cork, IE 1.74%

**Top Entities Mentioned Historically**  
Oslo, BBC, Dublin, Dermot Mulqueen, Johann Hari,



# Complexities of multimodal data



**Selected Papers of AoIR 2016:  
The 17<sup>th</sup> Annual Conference of the  
Association of Internet Researchers  
Berlin, Germany / 5-8 October 2016**

## **SEEING IS BELIEVING: DO PEOPLE FAIL TO IDENTIFY FAKE IMAGES ON THE WEB?**

Mona Kasra, PhD  
University of Virginia

Cuihua Shen, PhD  
University of California Davis

James O'Brien, PhD  
University of California Berkeley

US Elections 2020 World Environment Soccer US Politics Business **Tech** Science

## The Observer

Artificial intelligence (AI)

Tom Chivers

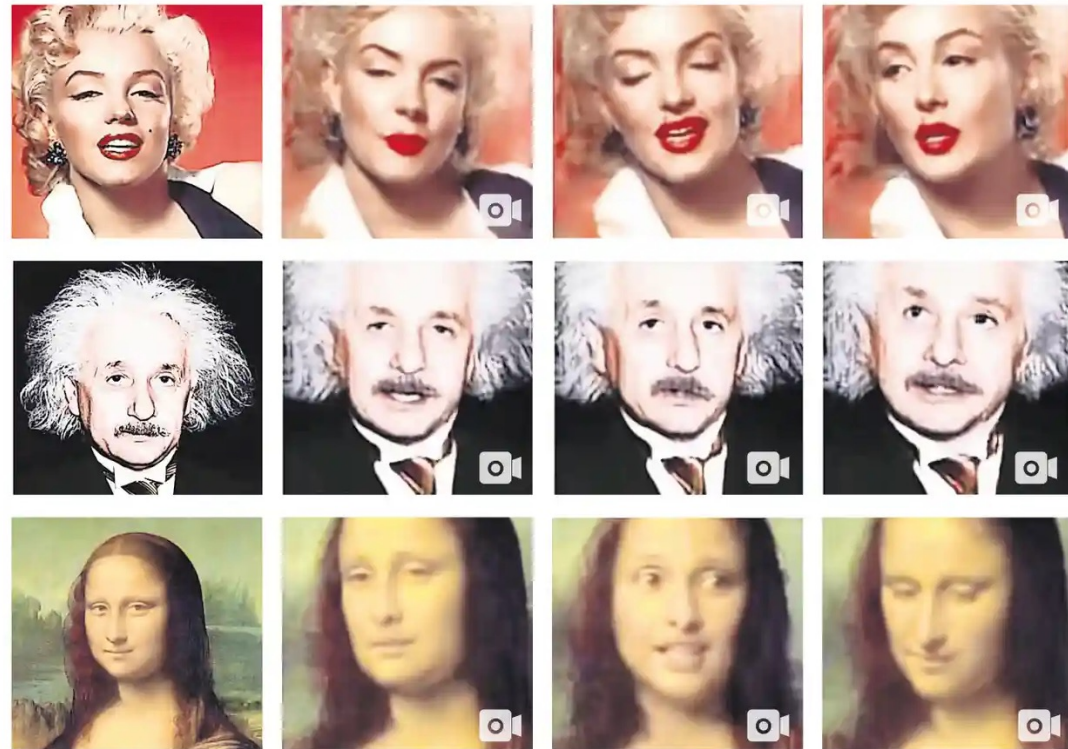
Sun 23 Jun 2019 04:00 EDT



327 60

# What do we do about deepfake video?

Deepfake - the ability of AI to fabricate apparently real footage of people - is a growing problem with implications for us all



▲ Back to life: portraits (first column) and stills from Gan-generated videos of (from top) Marilyn Monroe, Albert Einstein and the Mona Lisa. Photograph: Egor Zakharov/YouTube

There exist, on the internet, any number of videos that show people doing things they never did. Real people, real faces, close to photorealistic footage; entirely unreal events.

US Elections 2020 World Environment Soccer **US Politics** Business Tech Science

Nancy Pelosi

# Real v fake: debunking the 'drunk' Nancy Pelosi footage - video

Source: Various

Fri 24 May 2019 12.38 EDT



218

This footage of Nancy Pelosi was deliberately slowed down to make her look drunk or ill.

Footage of Nancy Pelosi deliberately slowed down ...

Watch later Share Info

Real Fake

facebook.com/PoliticsWatchDog

SUBSCRIBE

0:02 / 0:37

YouTube

**Footage of the Democratic House Speaker was edited to make her appear drunk or unwell, in the latest incident highlighting social media's struggle to deal with disinformation. Compare the original footage with the viral clip**

● **Facebook refuses to delete fake Pelosi video spread by Trump supporters**



**Mark Zuckerberg**

🕒 This article is more than 4 months old

# Doctored video of sinister Mark Zuckerberg puts Facebook to the test

**Last month Facebook declined to remove a manipulated video of Nancy Pelosi even after it was viewed millions of times**

**Luke O'Neil**

🐦 @lukeoneil47

Tue 11 Jun 2019 20.49 EDT



1,174



▲ A misleading clip of Mark Zuckerberg casts the Facebook founder in a sinister light. Photograph: Niall Carson/PA

A doctored video of **Mark Zuckerberg** delivering a foreboding speech has been posted to Instagram, in a stunt that put Facebook's content moderation policies to the test.



# Modulating Video Credibility via Visualization of Quality Evaluations

Nicholas Diakopoulos  
School of Communication and Information  
Rutgers University  
4 Huntington St., New Brunswick, NJ, USA  
diakop@rutgers.edu

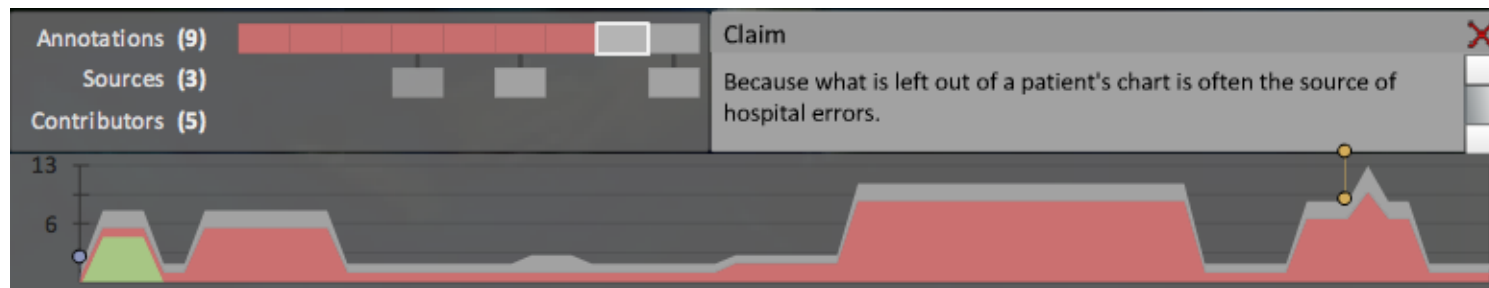
Irfan Essa  
School of Interactive Computing  
Georgia Institute of Technology  
75 5<sup>th</sup> St. NW, Atlanta, GA, USA  
irfan@cc.gatech.edu

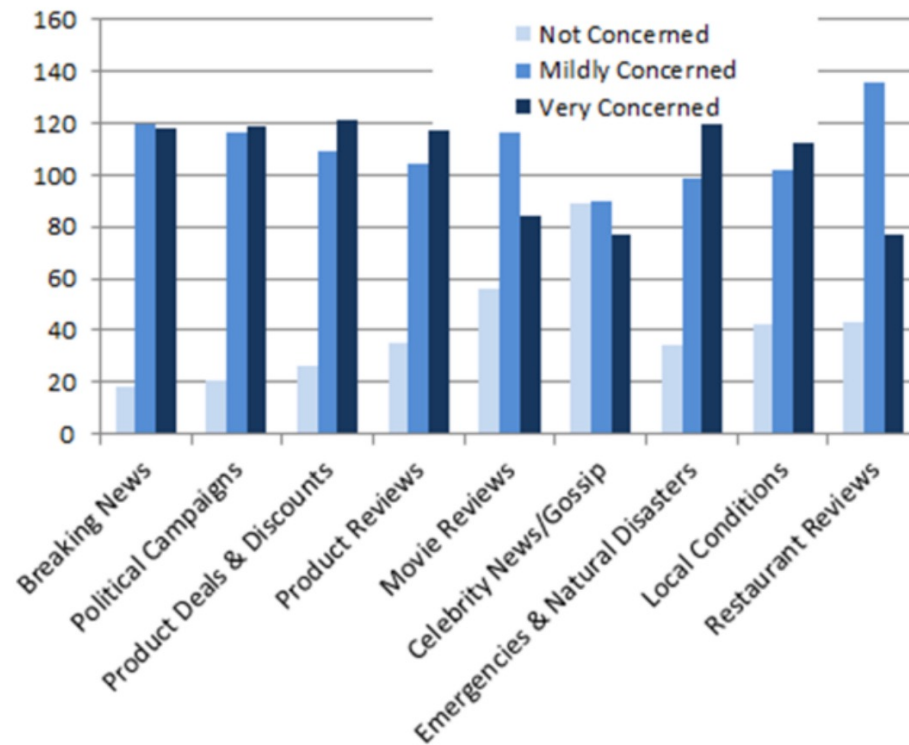
## ABSTRACT

In this work we develop and evaluate a method for the syndication and visualization of aggregate quality evaluations of informational video. We enable the sharing of knowledge between motivated media watchdogs and a wider population of casual users. We do this by developing simple visual cues which indicate aggregated activity levels and polarity of quality evaluations (i.e. positive / negative) which are presented in-line with videos as they play. In an experiment we show the potential of these visuals to engender constructive changes to the credibility of informational video under some circumstances. We discuss the limitations, and future work associated with this approach toward video credibility modulation.

information quality by combing through the media and engaging in fact-checking and re-contextualization of news and other media reports. For high profile video events such as the State of the Union address given by the president of the U.S., there is a considerable demand for this type of watchdogging activity. For instance, recent coverage by news outlets like PBS included annotated transcripts and video snippets showing analysis from experts and journalists<sup>1</sup>. One of the major issues with such analytic presentations as are found on Politifact, Factcheck, and PBS is that, especially for video, the analysis is divorced from the video itself, making the multimedia context difficult to understand in relationship with the textual analysis.

While most methods of watchdogging are labor intensive, another





Morris et al. focus on assessing credibility of news. Would same observations apply to judging credibility of non-real time information?

E.g., health myths and misinformation

# Role of social media in supporting conspiracy theories

## Understanding Anti-Vaccination Attitudes in Social Media

**Tanushree Mitra**<sup>1,2</sup>

<sup>1</sup>Georgia Institute of Technology

tmitra3@gatech.edu

**Scott Counts**<sup>2</sup>

<sup>2</sup>Microsoft Research

counts@microsoft.com

**James W. Pennebaker**<sup>2,3</sup>

<sup>3</sup>University of Texas at Austin

pennebaker@mail.utexas.edu

### Abstract

The anti-vaccination movement threatens public health by reducing the likelihood of disease eradication. With social media's purported role in disseminating anti-vaccine information, it is imperative to understand the drivers of attitudes among participants involved in the vaccination debate on a communication channel critical to the movement: Twitter. Using four years of longitudinal data capturing vaccine discussions on Twitter, we identify users who persistently hold pro and anti attitudes, and those who newly adopt anti attitudes towards vaccination. After gathering each user's entire Twitter timeline, totaling to over 3 million tweets, we explore differences in the individual narratives across the user cohorts. We find that those with long-term anti-vaccination attitudes manifest conspiratorial thinking, mistrust in government, and are resolute and in-group focused in language. New adoptees appear to be predisposed to form anti-vaccination attitudes via similar government distrust and general paranoia, but are more social and less certain than their long-term counterparts. We discuss how this apparent predisposition can interact with social media-fueled events to bring newcomers into the anti-vaccination movement. Given the strong base of conspiratorial thinking underlying anti-vaccination attitudes, we conclude by highlighting the need for alternatives to traditional methods of using authoritative sources such as the government when correcting misleading vaccination claims.

persistent vaccine criticism movement has spread rapidly through social media, a channel often used to disseminate medical information without verification by the expert medical community (Keelan et al. 2010).

Given the increasing reliance on online media for accurate health information and the general growth of social media sites, the attitudes of anti-vaccination advocates risk becoming a global phenomenon that could impact immunization behavior at significant scale (Kata 2010). In fact a controlled study showed that parents opting to exempt children from vaccination are more likely to have received the information online compared to those vaccinating their kids (Salmon et al. 2005). These parents benefit from “herd immunity” in which eradication is achieved by immunizing a critical proportion of the population. However, as internet-fueled misbeliefs drive people to opt out of vaccination, herd immunity is weakened, increasing the chances of a disease outbreak. Thus it is important to understand the underlying characteristics of individuals with anti-vaccination attitudes. What drives people to develop and perpetuate the anti-vaccination movement?

In this paper we explore this question by examining individuals' overt expressions towards vaccination in a social

# Class Exercise

Credibility is, after all, a domain-dependent attribute. Take the example of the COVID-19 anti-vax attitudes. What additional new features would you consider, in addition to the ones raised in Morris et al., that could be useful to allow people to assess credibility of the information? How would you factor in end users' bias in perception of credibility?



# Curating Quality? How Twitter's Timeline Algorithm Treats Different Types of News

Social Media + Society  
July-September 2021: 1–17  
© The Author(s) 2021  
Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/20563051211041648  
[journals.sagepub.com/home/sms](https://journals.sagepub.com/home/sms)  


Jack Bandy  and Nicholas Diakopoulos 

## Abstract

This article explores how Twitter's algorithmic timeline influences exposure to different types of external media. We use an agent-based testing method to compare chronological timelines and algorithmic timelines for a group of Twitter agents that emulated real-world archetypal users. We first find that algorithmic timelines exposed agents to external links at roughly half the rate of chronological timelines. Despite the reduced exposure, the proportional makeup of external links remained fairly stable in terms of source categories (major news brands, local news, new media, etc.). Notably, however, algorithmic timelines slightly increased the proportion of “junk news” websites in the external link exposures. While our descriptive evidence does not fully exonerate Twitter's algorithm, it does characterize the algorithm as playing a fairly minor, supporting role in shifting media exposure for end users, especially considering upstream factors that create the algorithm's input—factors such as human behavior, platform incentives, and content moderation. We conclude by contextualizing the algorithm within a complex system consisting of many factors that deserve future research attention.

How can platforms counterbalance the tension between algorithmic curation of timelines and information quality?

A need for “fact checking systems” that operate outside of the social media ecosystem. But these systems are difficult to build and use. Why?

## Article Menu

[Close](#) ^[Download PDF](#)[Free EPUB View](#)

Accessing resources off campus can be a challenge. Lean Library can solve it

[Full Article](#)

### Content List

[Abstract](#)[The Societal Cost of Misinformation](#)[Origins of Misinformation](#)[From Individual Cognition to](#)

## Misinformation and Its Correction: Continued Influence and Successful Debiasing

[Stephan Lewandowsky](#), [Ullrich K. H. Ecker](#), [Colleen M. Seifert](#), more...

[Show all authors](#) ▾

First Published September 17, 2012

Research Article

[Find in PubMed](#)

<https://doi.org/10.1177/1529100612451018>

[Article information](#) ▾

1018



## Abstract

The widespread prevalence and persistence of misinformation in contemporary societies, such as the false belief that there is a link between childhood vaccinations and autism, is a matter of public concern. For example, the myths surrounding vaccinations, which prompted some parents to withhold immunization from their children, have led to a marked increase in vaccine-preventable disease, as well as unnecessary public expenditure on research and public-information campaigns aimed at rectifying the situation.

We first examine the mechanisms by which such misinformation is disseminated in society, both inadvertently and purposely. Misinformation can originate from rumors but also from works of fiction, governments and politicians, and vested interests. Moreover, changes in the media landscape, including the arrival of the Internet, have fundamentally influenced the ways in which information is communicated and misinformation is spread.