

# Technology as Enabler of Fake News and a Potential Tool to Combat it


prof.dr. Žiga Turk  
[ziga.turk@gmail.com](mailto:ziga.turk@gmail.com)

Research commissioned by Policy Department A on  
the request of IMCO Committee

# Content

- news circulation process has changed
  - in the past it was edited by professionals
  - now it is edited by everyone
- “fake news” circulates largely due to social networks
  - 30-40%
- technically it is possible to interfere with circulation of “fake news”
  - but cannot be stopped entirely
- conclusion
  - yes we can but should we

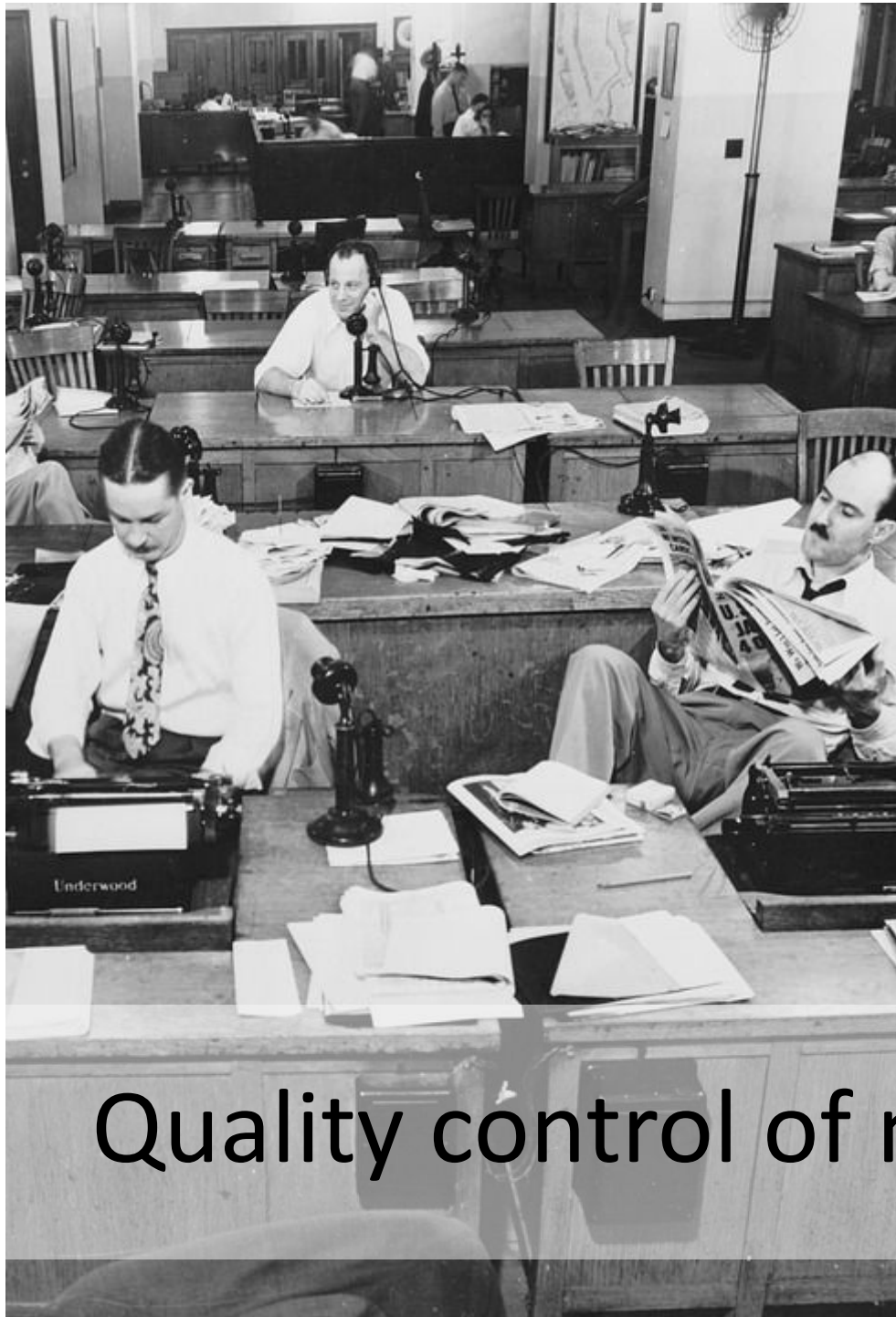
# The news process



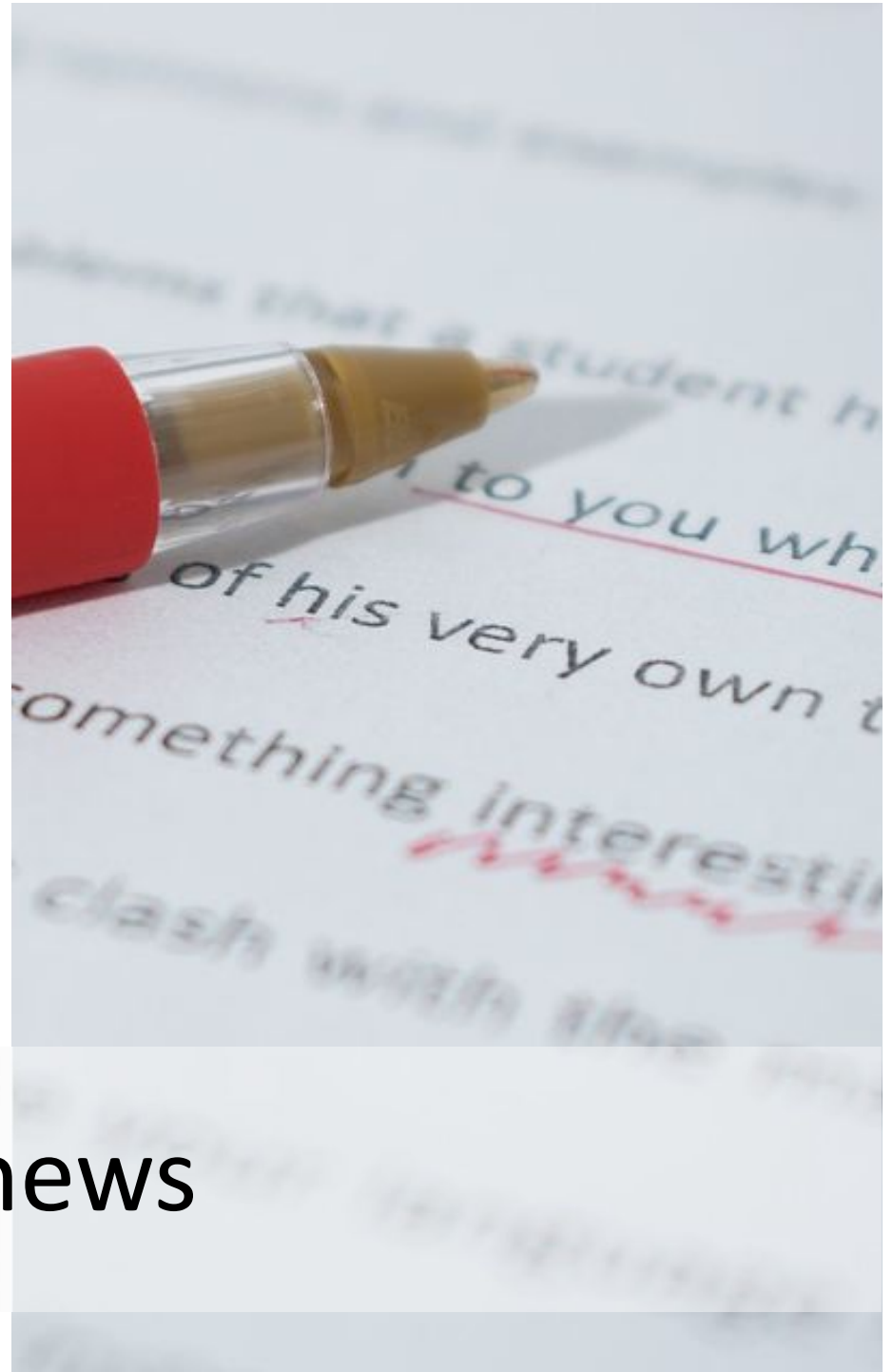
Travel is to make a journey or to have an adventure to somewhere by bicycle, train, airplane, car, motorcycle, or boat. It could be an exploration to somewhere new planned or unplanned to meet new people, new things and new places. There are different types of adventures waiting for you to explore.

There are lots of places to explore. Places could be urban or suburban. Some people love to be with nature to free their minds and refresh their souls, but some like to be in the city. You will get lots of benefits such as exploring new culture.

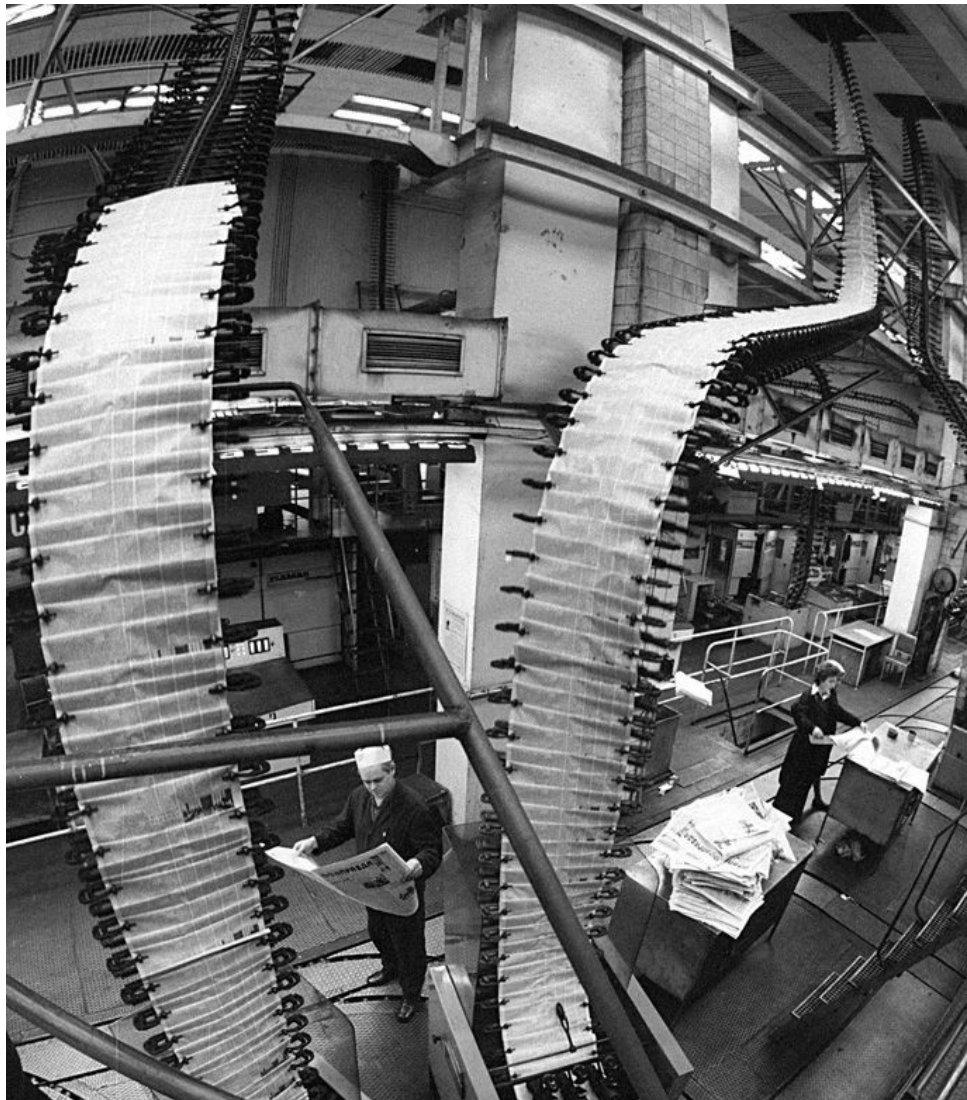
Create news



Quality control of news







Publication of news



**The New York Times.**  
**ARMISTICE SIGNED. END OF THE WAR!**  
**BERLIN SEIZED BY REVOLUTIONISTS;**  
**NEW CHANCELLOR BEGS FOR ORDER;**  
**OUSTED KAISER FLEES TO HOLLAND**

**EXTRA**  
**Breaking**

**NEWS**

Amplification of news

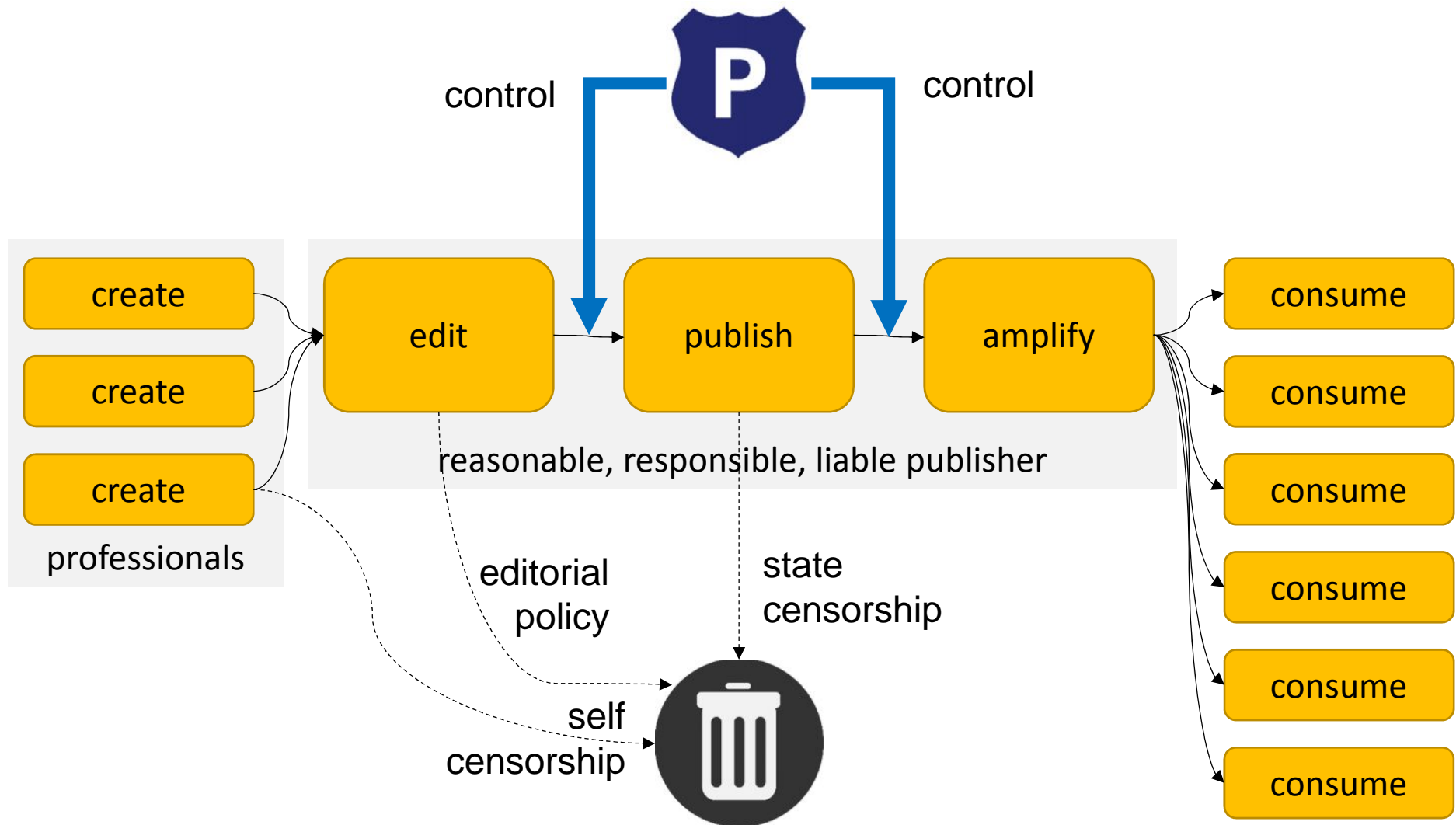


Consumption of news



# The news process schematically

# Traditional news process



# Enters the Internet

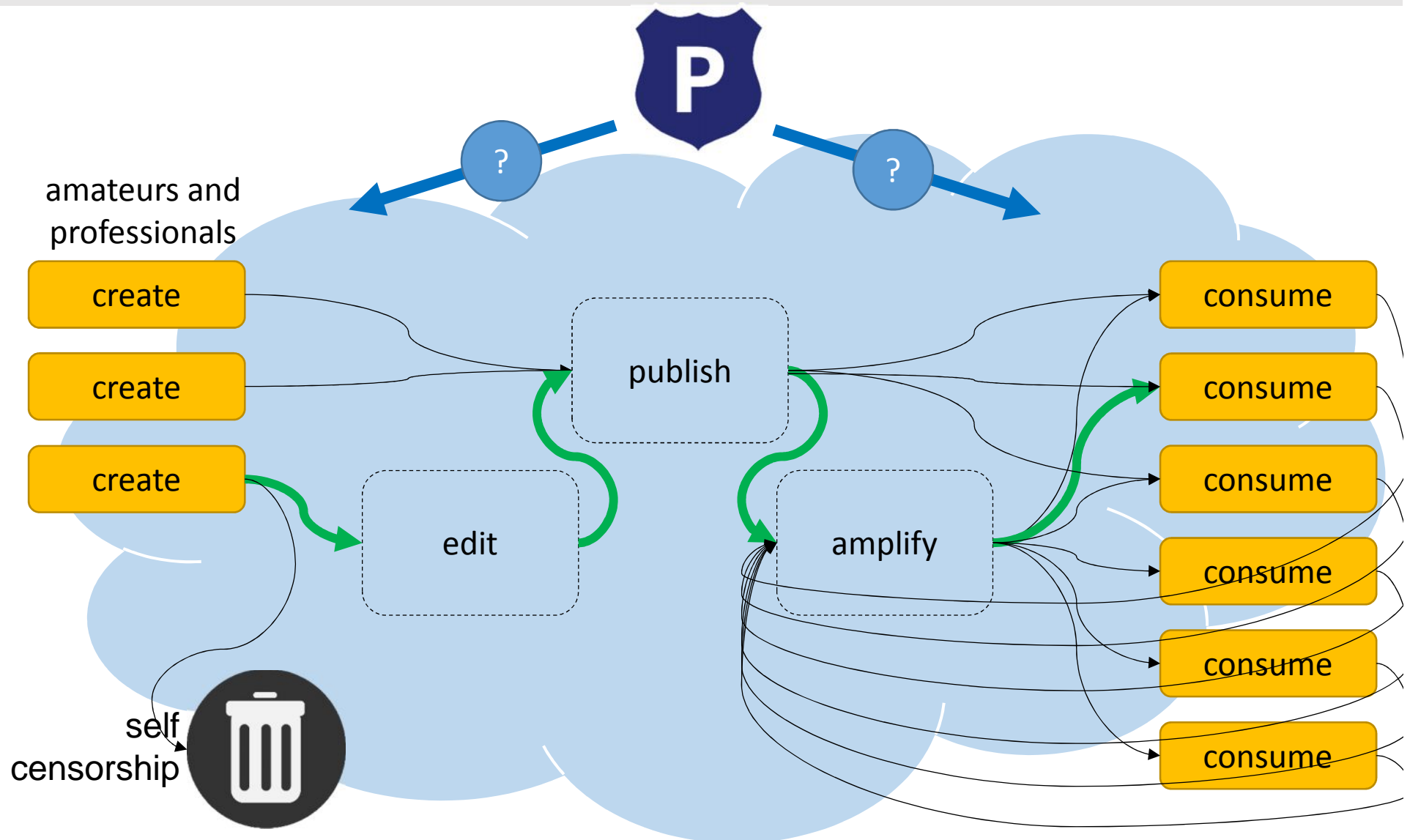
- anyone can publish
- anyone can amplify
- anyone can edit
- editors are gone
- no trashcan



source: wikipedia



# Internet-age news process

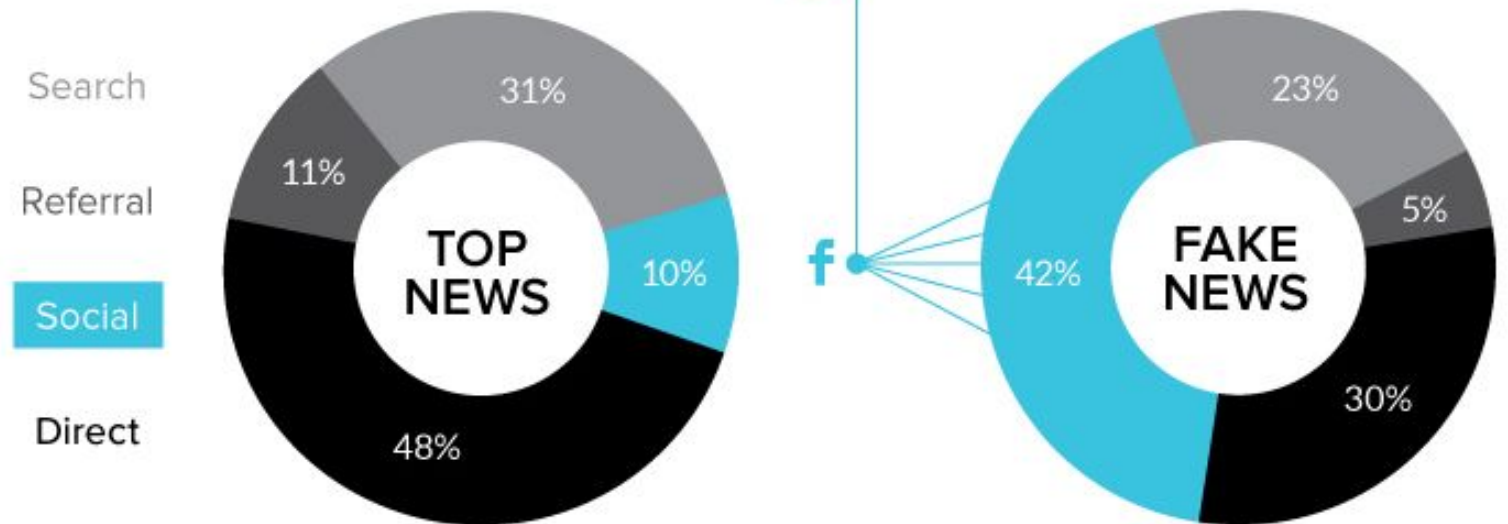


Technology for “fake news” dissemination

# Social media is main driver for “fake news”

- 10% of readers of top news come via social media
- 40% of readers of “fake news” come via social media

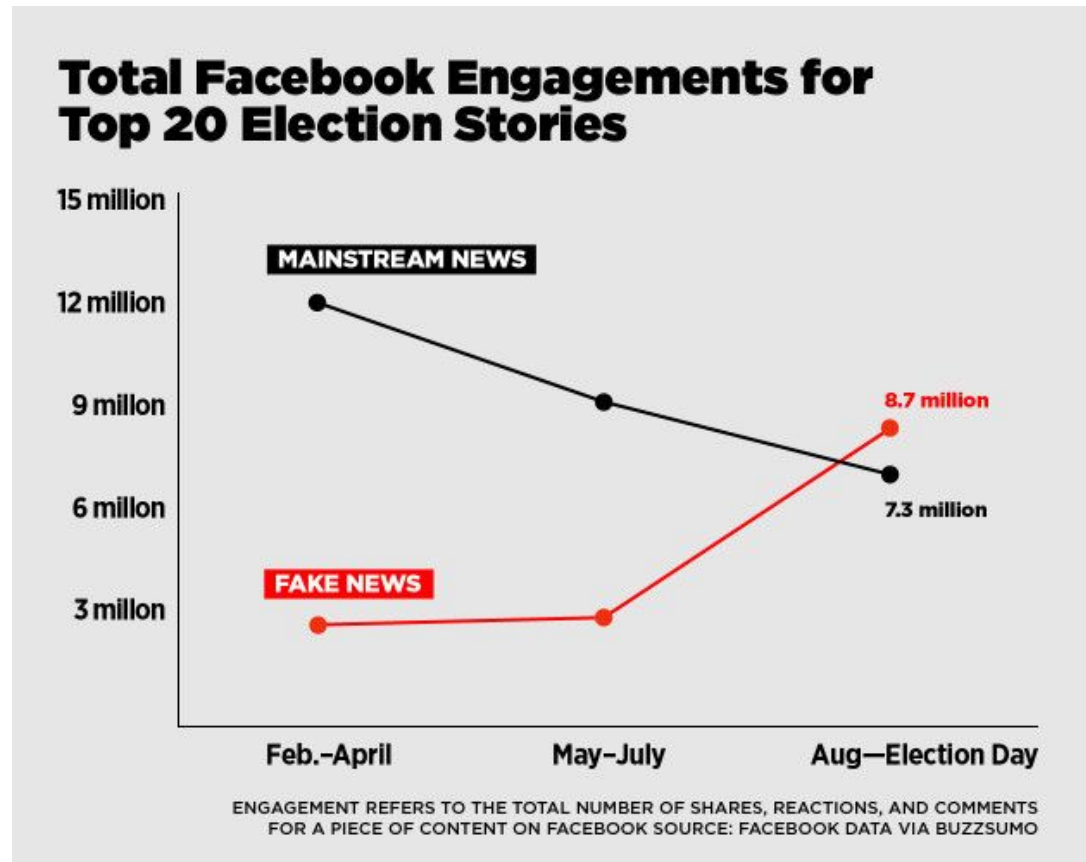
## TRAFFIC SOURCES





# Fake and real news have about the same engagement on social media

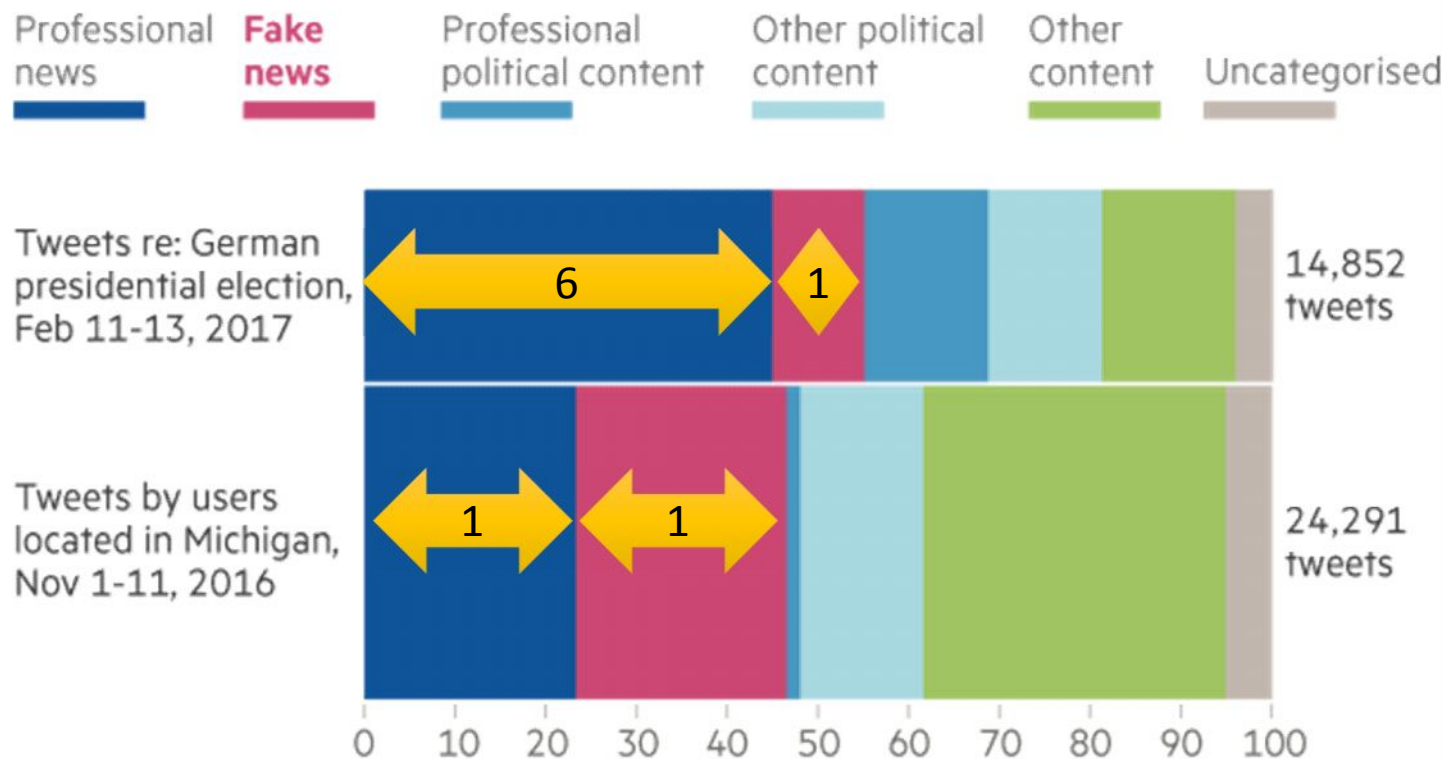
- this is 10% for real news
- but 40% for “fake news”



# In Europe “real” beats “fake” 6:1, in US it is 1:1

University of Oxford studies show differing rates at which  
“fake news” is shared on Twitter

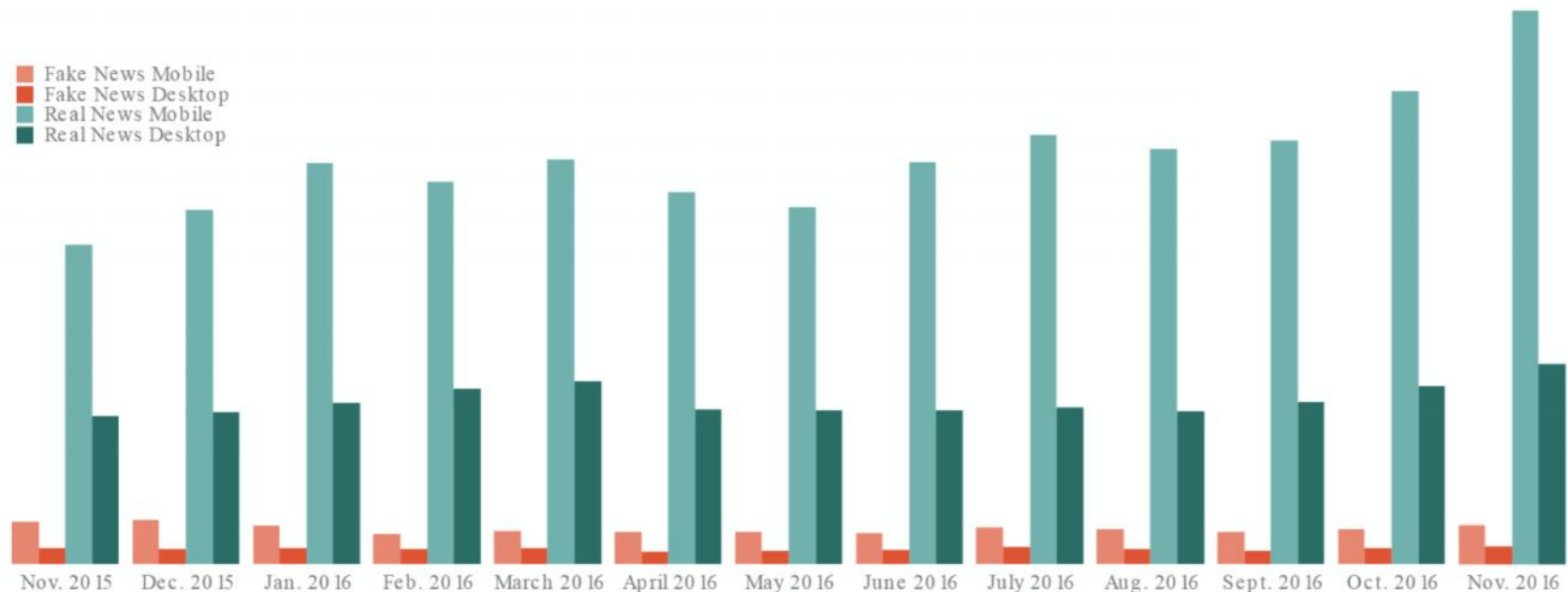
Percentage of links shared



# “Fake news” (orange) audience is 10% of real news (green)

## Columbia Journalism Review.

The real news audience even spiked in October and November 2016, while the fake news audience stayed the same.



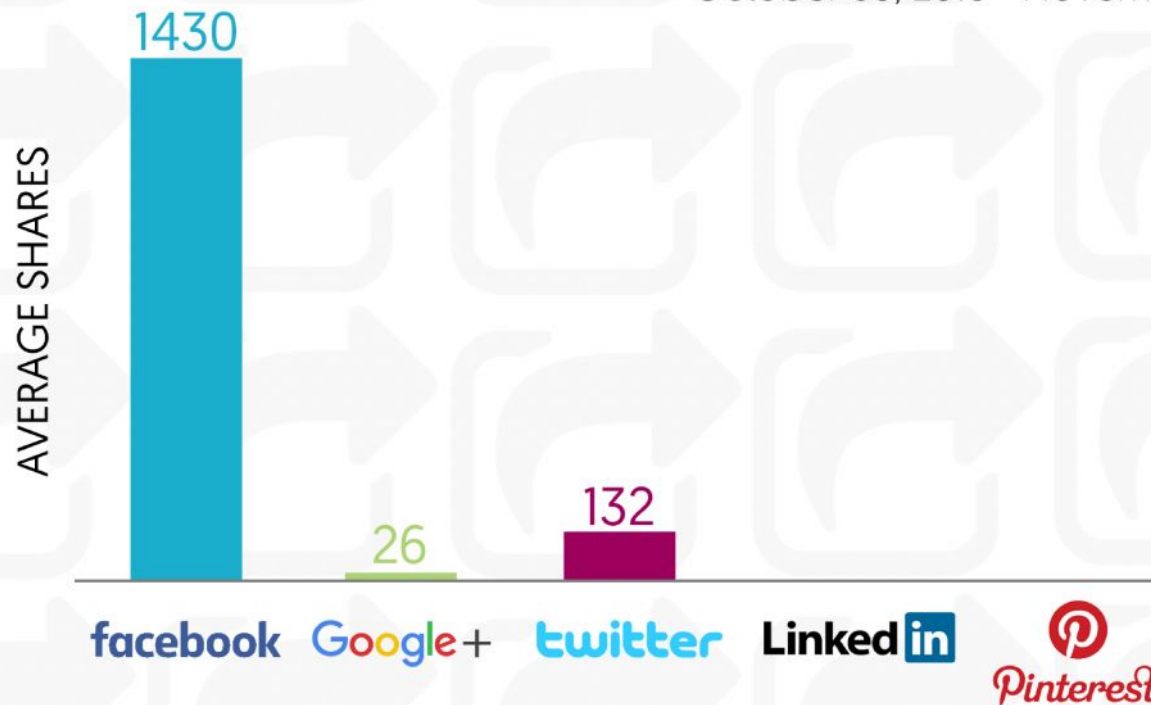


# For „fake news“ main source is Facebook

## Average “Pizzagate” Shares By Network

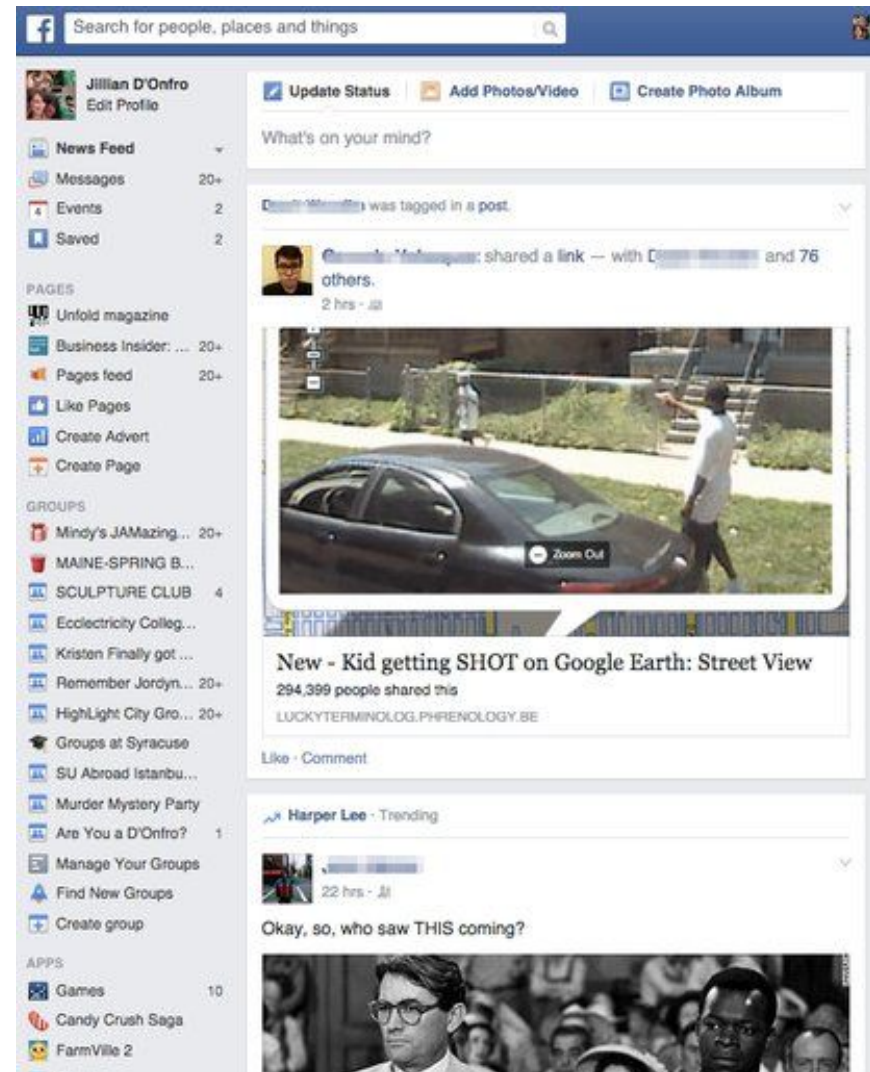
Including Both Fake And Real News Articles

October 30, 2016 - November 21, 2016



# On Facebook users see 10% of what they „subscribe” to

- Facebook decides which 10%
- Facebook chose to remove news from feed
  - hurting 10% of real news traffic
  - hurting 40% of fake news traffic
- Both unhappy!

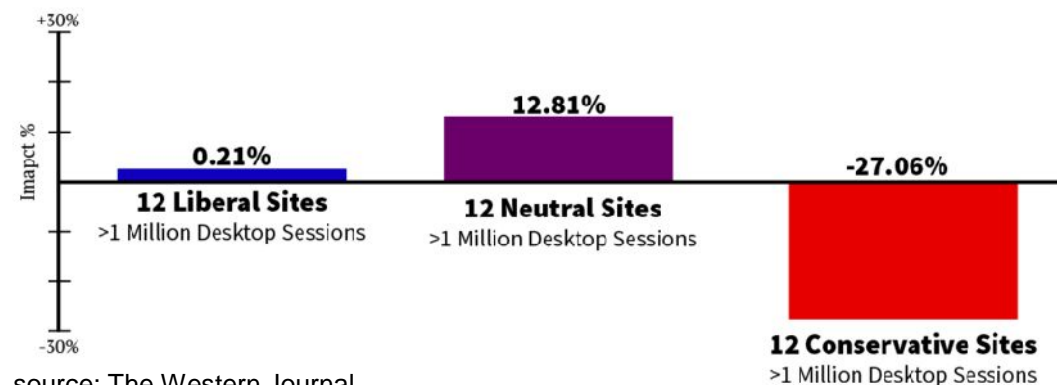


# How they choose the 10% is key issue for stopping “fake news”!

- Facebook’s goal: „keep audience at site and engaged“
- Algorithm is opaque
- Algorithm is changing
- Algorithm is unfair?



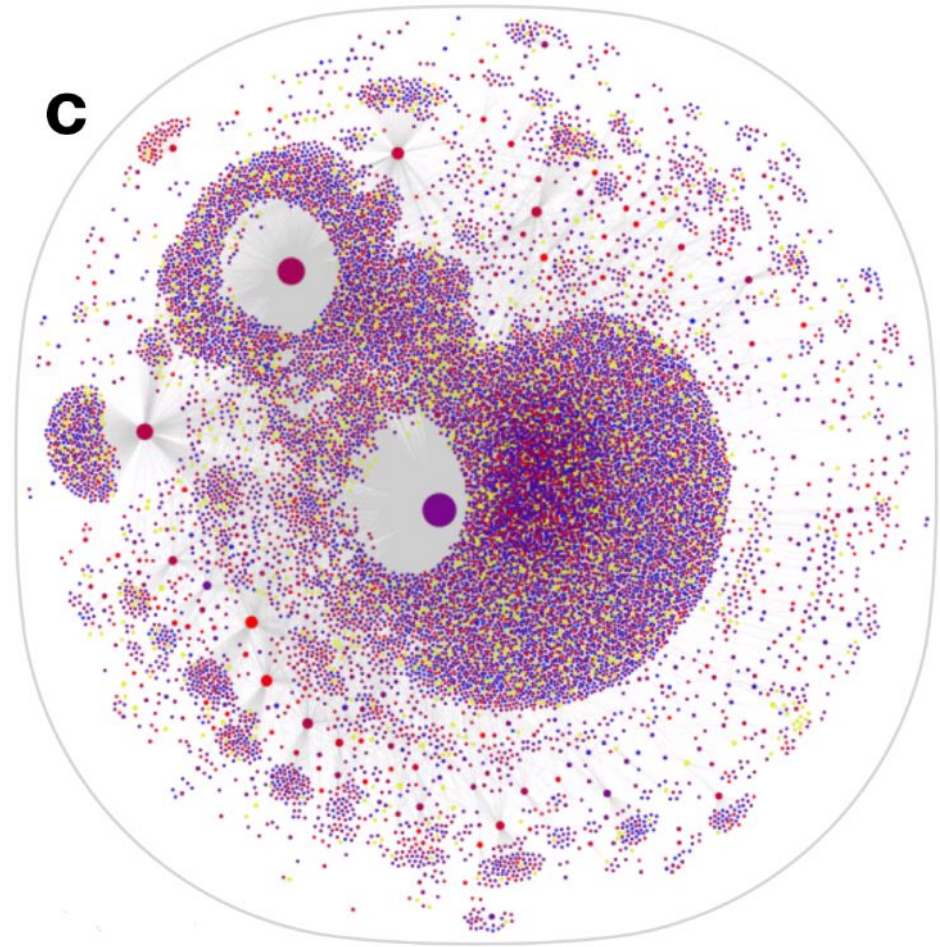
## Average Facebook Algorithm Impact





# Bots spread fake news on Twitter. And real news.

- “We find evidence that social bots play a disproportionate role in spreading and repeating misinformation.” ([source](#)).
- “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” ([source](#))



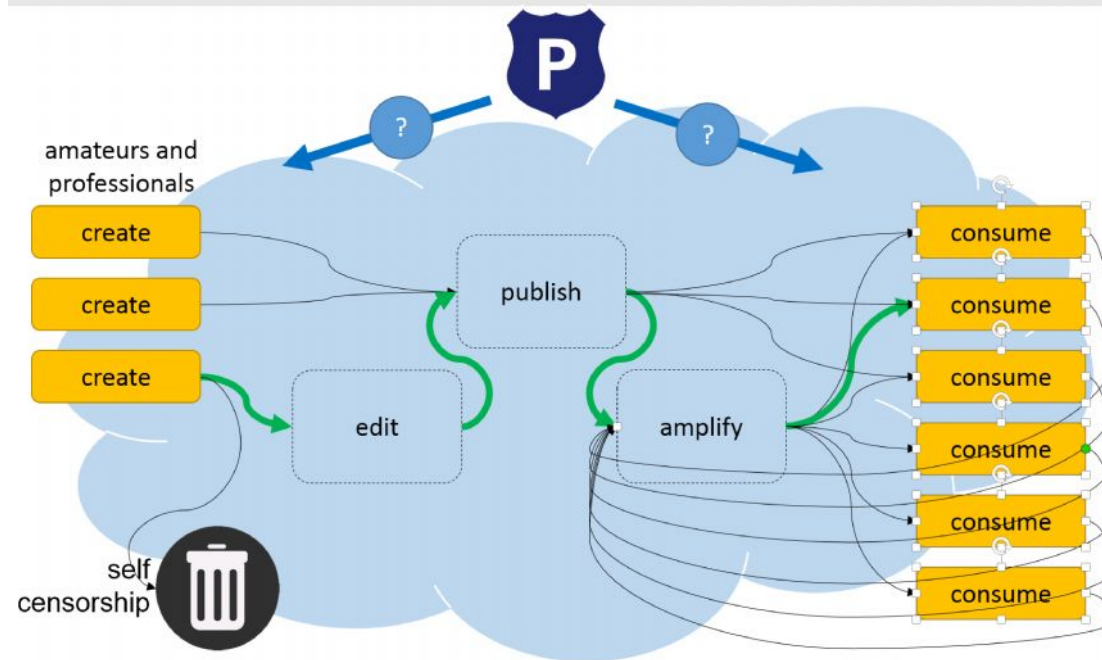
Propagation of fake news: Nodes and links represent Twitter accounts and retweets of the claim, respectively. Node size indicates account influence, measured by the number of times an account is retweeted. Node color represents bot score, from blue (likely human) to red (likely bot).

Can we stop  
“fake news”?

# Yes we can!

- with author
- with publisher
- with platform
- with amplifier
- with reader

## Internet-age news process



# Stopping authors to publish “fake news”


- demonetize
  - deny advertising at fake news sites
- threaten prison
  - e.g. Polish anti-holocaust law
- lock them up
  - e.g. Turkey

The Rubin Report (talk show) Dave Rubin +3

## What was YouTube's reason for demonetizing the Rubin Report conversation between Dave Rubin, Jordan Peterson and Ben Shapiro?

Answer Request Follow 22 Comment Downvote

1 Answer

Jeff Franz-Lien, Sr.  Answered Feb 2

### What was YouTube's reason for demonetizing the conversation between Dave Rubin, Jordan Peterson and Ben Shapiro?

According to chatter on Twitter, YouTube demonetized “controversial” content, not just Rubin or other right-wing speakers. Others who were hit too, including Dr. Jordan Peterson.


Rubin confirms this in the video, saying he is losing big bucks (Dave said he was losing \$100,000) for offensive content. Who is to blame? YouTube and its advertisers.

FINANCIAL TIMES myFT

Vodafone Group PLC + Add to myFT

### Vodafone to block its ads appearing on offensive websites

Telecoms group working with Google, Facebook, WPP to avoid hate speech and 'fake news'



© Bloomberg



# Stopping “fake news” when published

- deny hosting
- deny listing in DNS directory
- block internet traffic to “fake news” site



Within 48 hours of the  
website the Daily Star



Protesters wave Catalan separatist flags out:  
Catalan authorities have accused t  
websites aimed to spread informal  
independence referendum.  
The move is part of a crackdown c  
movement to prevent the vote, wd  
ballot papers and the deployment  
Catalonia.



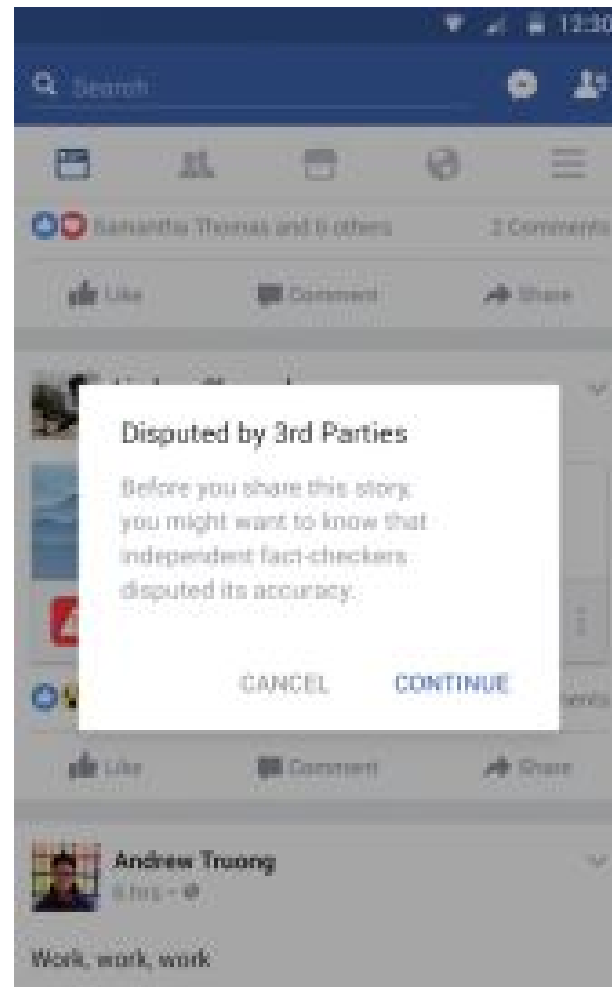
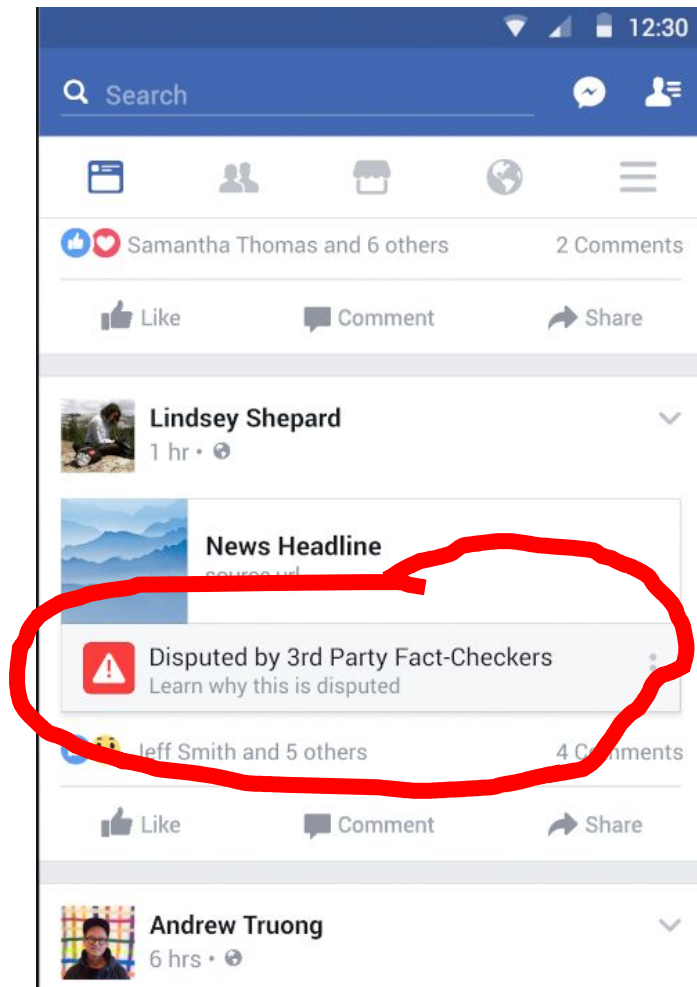
Bloomberg News  
Updated on December 1, 2017, 4:19 AM GMT+1  
From Bloomberg QuickTake

China's online population of 731 million gets a highly restricted internet, one that doesn't include access to Google, Facebook, YouTube or the New York Times. There's little coverage of the 1989 student protests in Tiananmen Square. Even Winnie the Pooh got temporarily banned. China is able to control such a vast ocean of content through the largest system of censorship in the world, aptly known as the Great Firewall of China. It's a joint effort between government monitors and the technology and telecommunications companies that are compelled to enforce the state's rules. The stakes go beyond China, which is setting an example that other authoritarian countries can imitate.

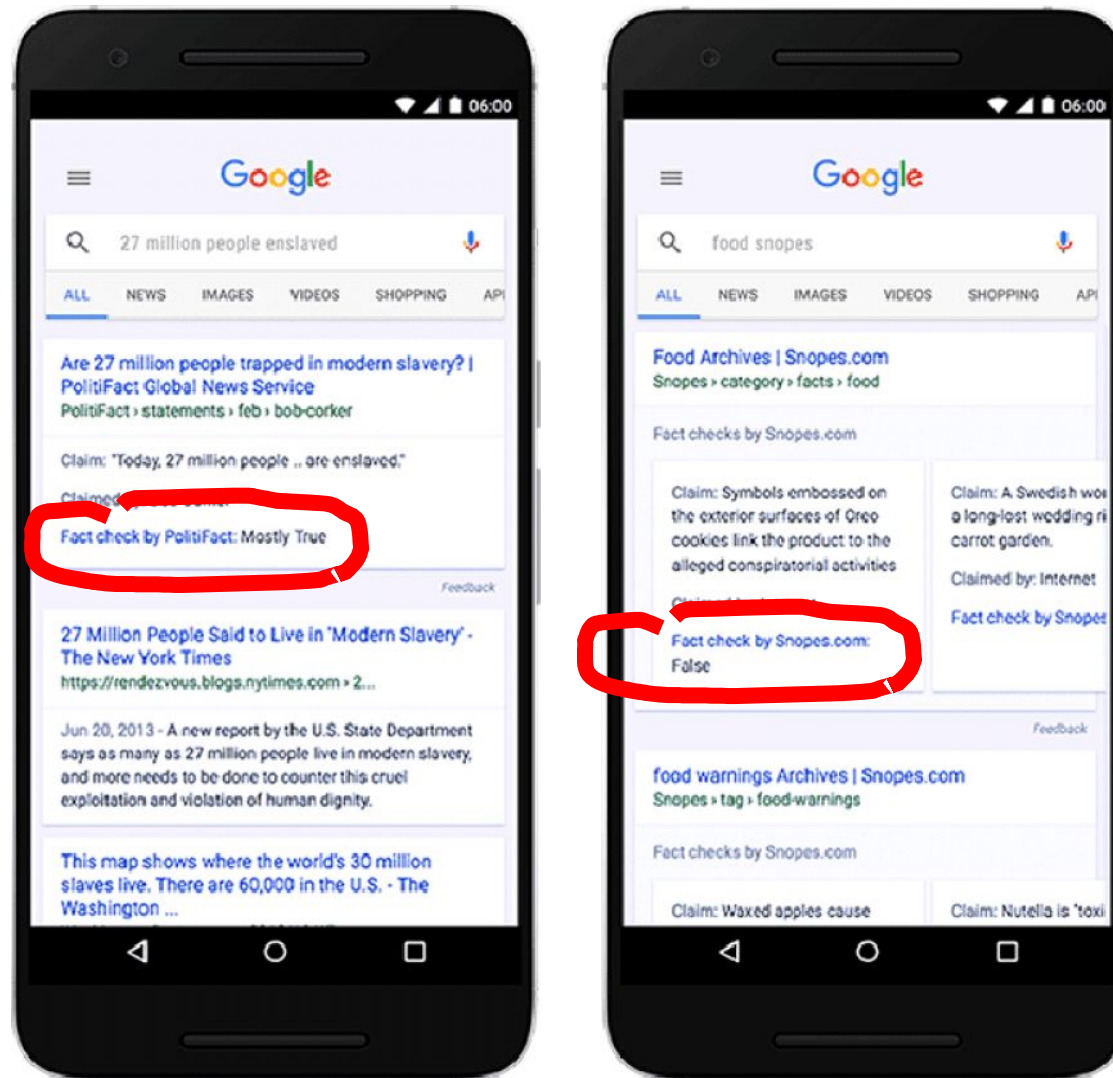
# Stopping fake news at platforms

- public platforms
  - Facebook, Google, Twitter, Baidu ...
  - label, augment, hide or remove “fake news”
- private messaging platforms
  - Viber, WeChat, Snapchat
  - listen to private conversations?!
- needs collaboration of platforms

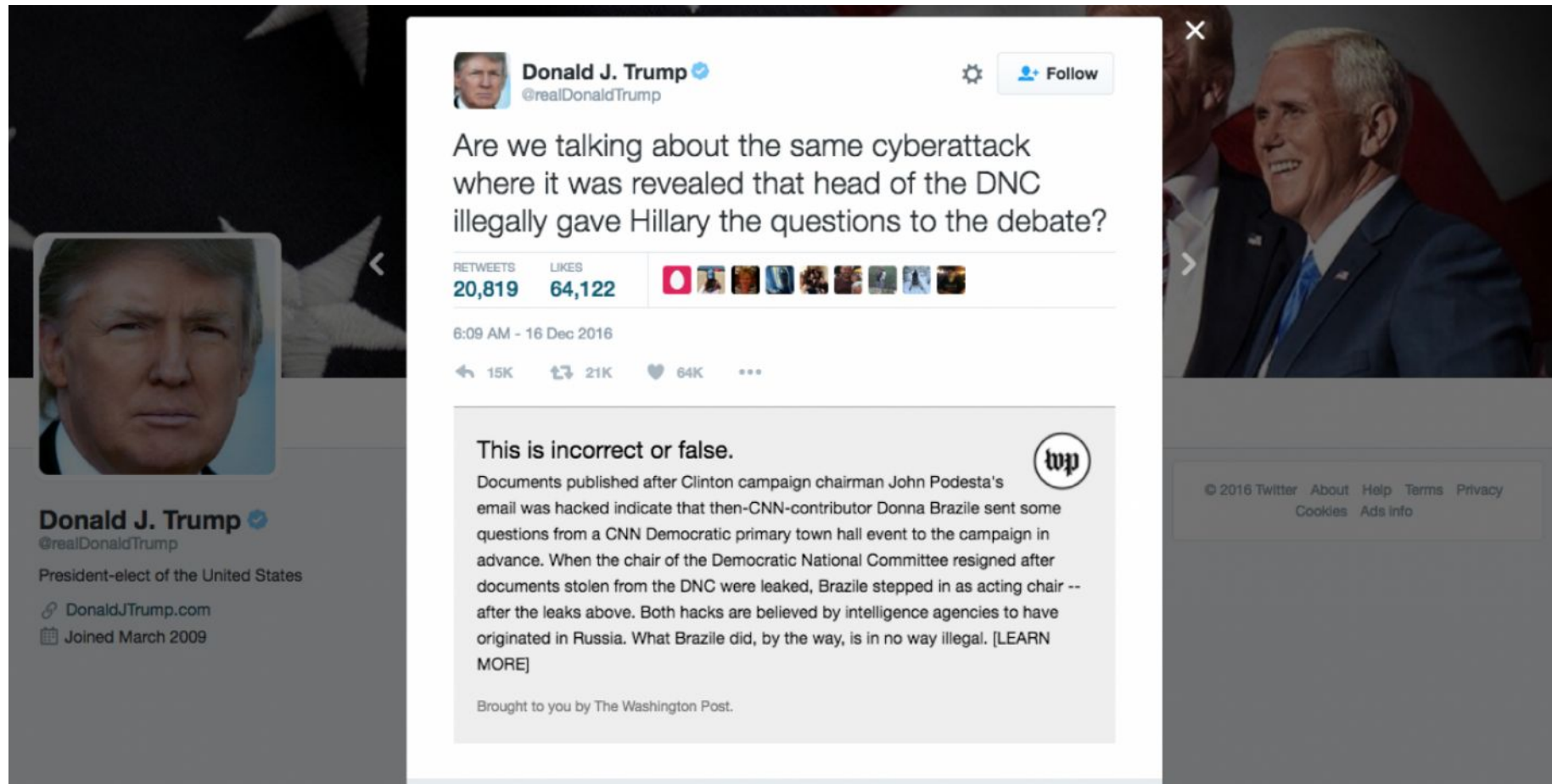
# Label „fake news“ on Facebook



# Label “fake news” on Google



# Label “fake news” on ~~Twitter~~ browser



The image is a screenshot of a Twitter interface. On the left is the profile of Donald J. Trump, showing his name, handle @realDonaldTrump, and a bio identifying him as the President-elect of the United States. The main content is a tweet from Donald J. Trump asking a question about a cyberattack and the DNC. The tweet has 20,819 retweets and 64,122 likes. Below the tweet is a fact-check overlay from The Washington Post (wp) stating that the information is incorrect or false. The overlay provides a detailed explanation involving John Podesta's hacked email and Donna Brazile's role as acting chair of the DNC. The background of the interface shows a blurred image of Mike Pence.

**Donald J. Trump** @realDonaldTrump

Are we talking about the same cyberattack where it was revealed that head of the DNC illegally gave Hillary the questions to the debate?

RETWEETS 20,819 LIKES 64,122

6:09 AM - 16 Dec 2016

15K 21K 64K

**This is incorrect or false.**

Documents published after Clinton campaign chairman John Podesta's email was hacked indicate that then-CNN-contributor Donna Brazile sent some questions from a CNN Democratic primary town hall event to the campaign in advance. When the chair of the Democratic National Committee resigned after documents stolen from the DNC were leaked, Brazile stepped in as acting chair -- after the leaks above. Both hacks are believed by intelligence agencies to have originated in Russia. What Brazile did, by the way, is in no way illegal. [LEARN MORE]

Brought to you by The Washington Post.

© 2016 Twitter About Help Terms Privacy Cookies Ads info



# Hide “fake news” by downranking – pushing down on search results



## Facebook now flags and down-ranks fake news with help from outside fact checkers

Josh Constine @JoshConstine / Dec 15, 2016

Comment



Snopes, FactCheck.org, Politifact, ABC News, and AP will help Facebook make good on four of the six promises Mark Zuckerberg made about fighting fake news without it becoming “the arbiter of truth.” It will make fake news posts less visible, append warnings from fact checkers to fake news in the feed, make reporting hoaxes easier and disrupt the financial incentives of fake news spammers.



REUTERS World Business Markets Politics TV

Myanmar

Energy & Environment

Brexit

North Korea

Earnings

Futuro o

TECHNOLOGY NEWS NOVEMBER 23, 2017 / 3:15 PM / 4 MONTHS AGO

## Russia says Google down-ranking Sputnik, RT would be censorship

Reuters Staff

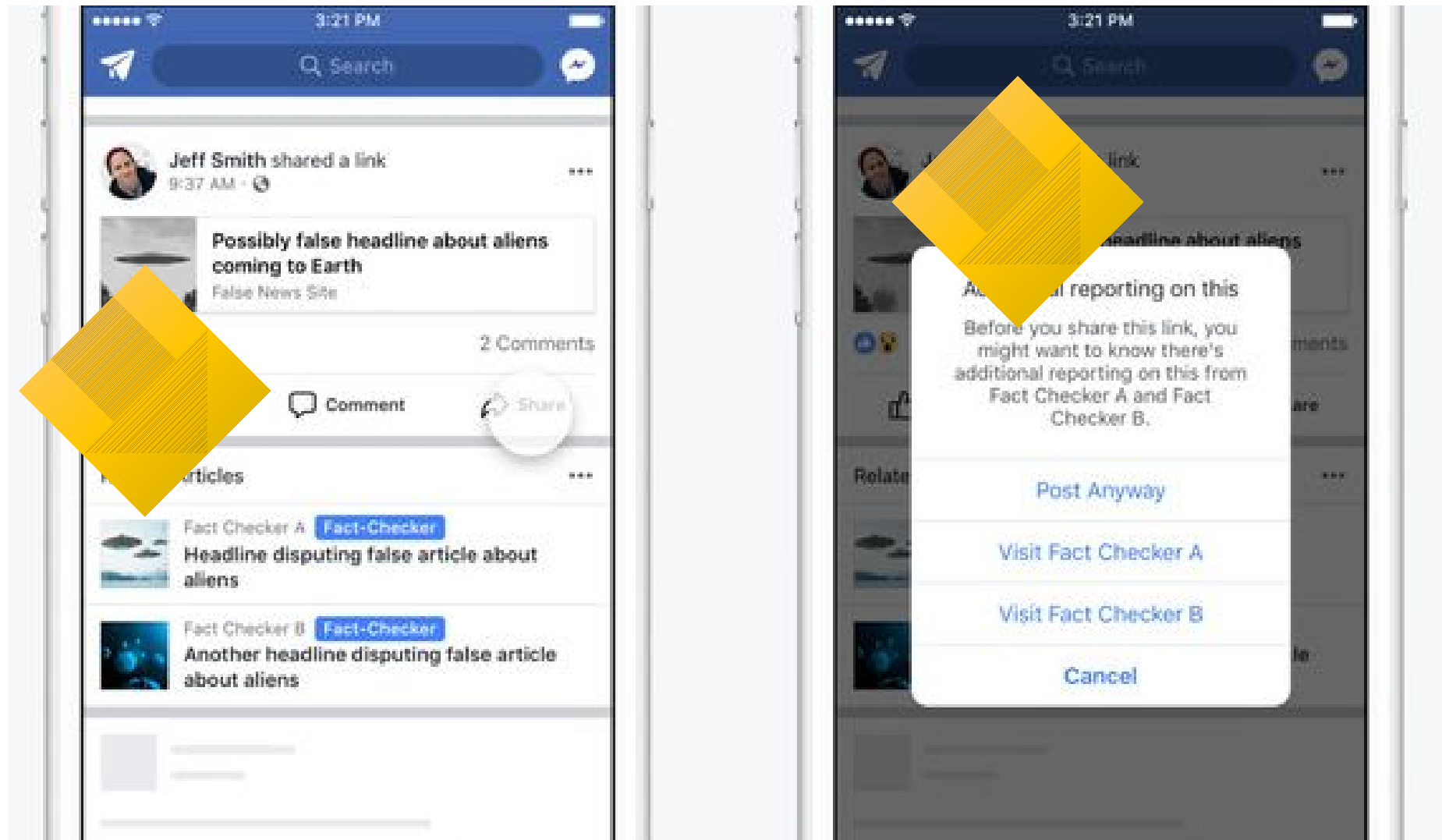
1 MIN READ



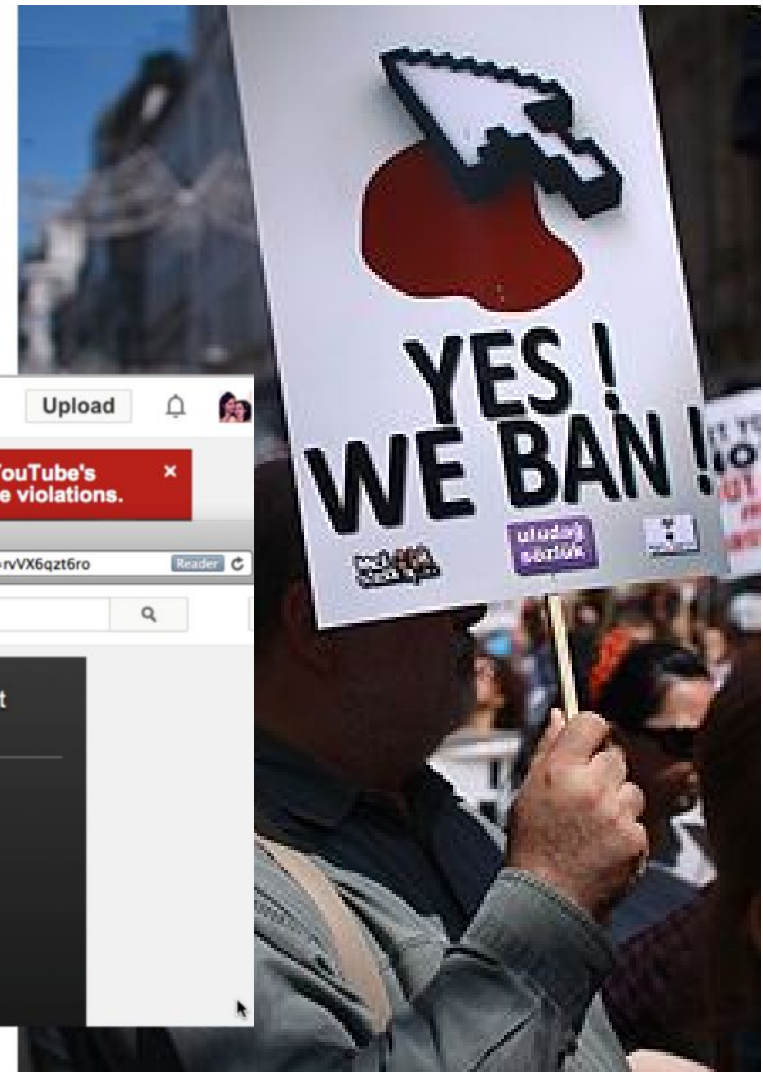
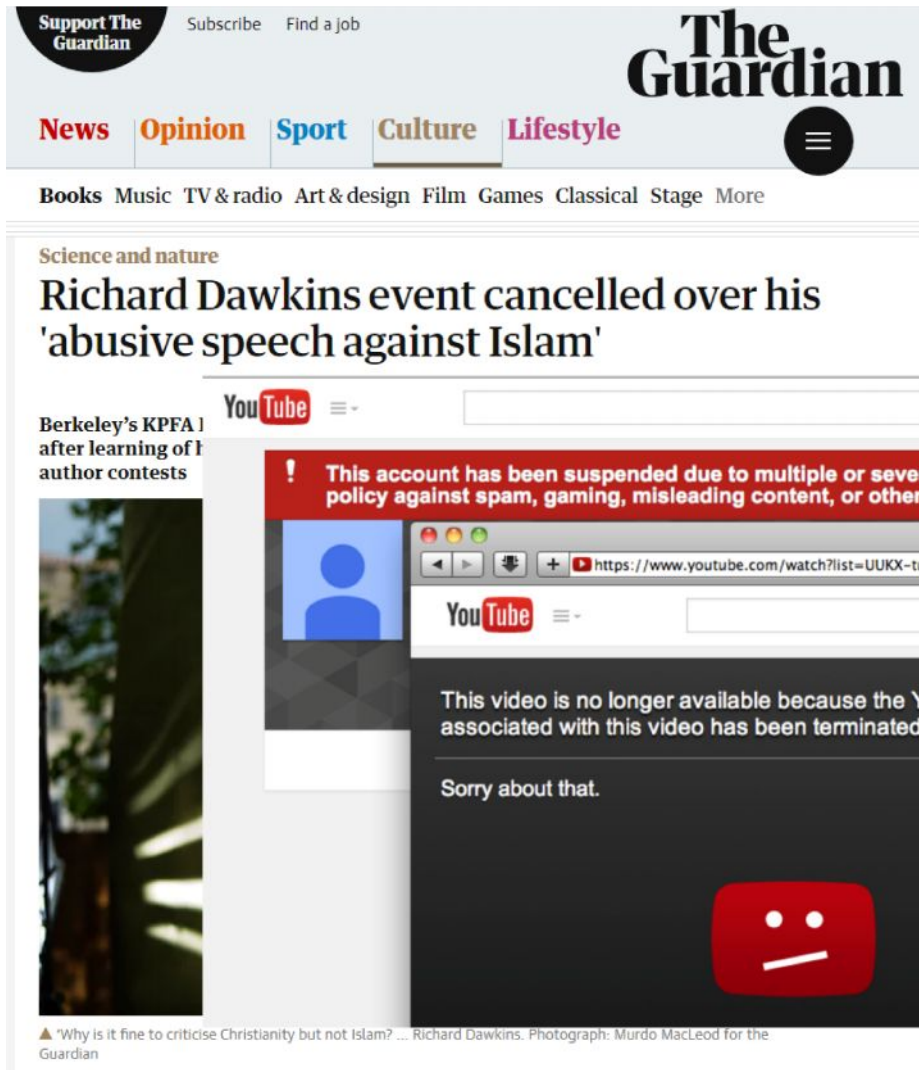
MOSCOW (Reuters) - Russia's foreign ministry said on Thursday that moves by Alphabet Inc's (GOOGL) Google to place articles from Russian news outlets Sputnik and Russia Today lower in search results would amount to censorship.



# Augment “fake news” with links to real news, warn before sharing



# Remove “fake news” - deplatform



# Stop bots!



**twitter**

 **Verified Account** BETA

**What is a verified account?**

To prevent identity confusion, Twitter is experimenting (beta testing) with a 'Verified Account' feature. We're working to establish authenticity with people who deal with impersonation or identity confusion on a regular basis. Accounts with a  are the real thing!



**JOHN DOE**

e: john.doe@miiCard.com  
t: 07900 999 000  
my.miiCard.com/card/johndoe

**miiCard**

 **Assured** Identity Assured at  
14-Jun 2013 @ 02:12:50 (UTC)



[Save To Contacts](#)

This miiCard account was created on Thursday, February 23, 2012 and has been active for 523 days.

John Doe  
Identity  
Tuesday  
Email

# Stopping “fake news” in private chat, email – the next frontier

- What's Up
- Facebook Chat
- Snapchat
- Gmail
- Signal
- Telegram
- Viber
- can be done, but!?



## How China's highly censored WeChat and Weibo fight fake news ... and other controversial content

As debate rages over Facebook's role in spreading false information on US elections, we look at how Chinese social media sites control such content shared online

PUBLISHED : Friday, 16 December, 2016, 4:01pm  
UPDATED : Monday, 12 June, 2017, 11:19pm

COMMENTS: 7





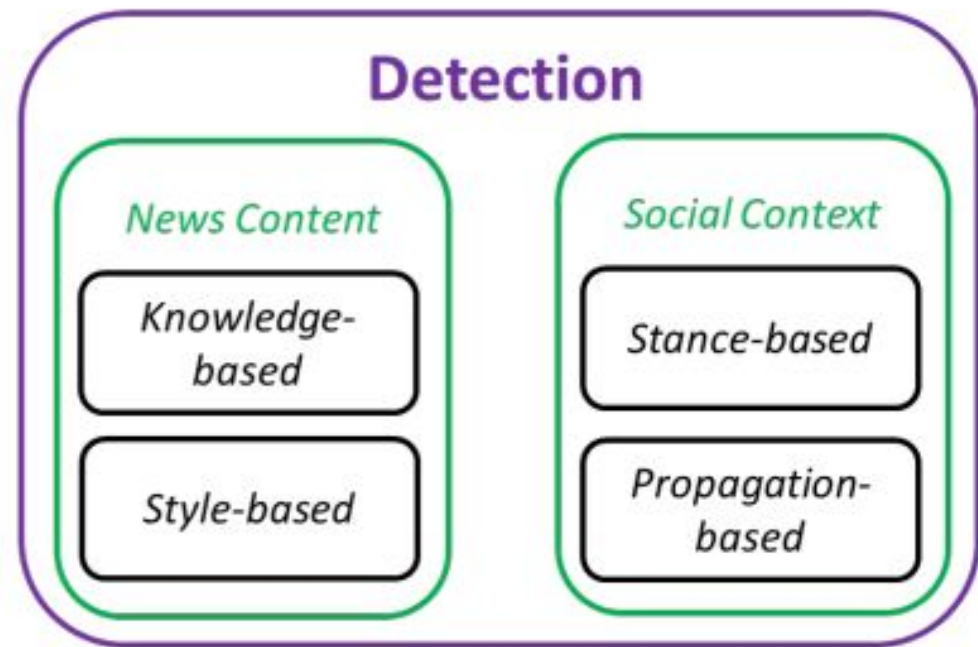
Key element: how to  
identify “fake news”?

# Fake news can be identified by humans, machines or combination

- Human
  - human work by fact checkers
  - fact-checking industry appearing, looking for business model
- Human-machine teamwork
  - machine identifies potentially fake stories
  - double-checked by human
- Machine only
  - artificial intelligence

# Two strategies for automatic identification of “fake news”

- From content
  - facts
  - style
- From social
  - stance
  - propagation



Shu, Kai, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. "Fake news detection on social media: A data mining perspective." *ACM SIGKDD Explorations Newsletter* 19, no. 1 (2017): 22-36.

# Content analysis is hard because ...

- ... needs to understand content
  - few news pieces are blatantly false
- style?
  - authors will adapt



**POLITICAL  
LANGUAGE**  
is designed to make lies  
sound truthful and  
murder respectable,  
and to give an  
appearance of solidity  
to pure wind.

George Orwell

# Artificial Intelligence can help. But “Computer did it” is a fallacy

[SUBSCRIBE](#)

SCIENTIFIC AMERICAN

English ▾ Carl  Sign In | Register 

THE SCIENCES MIND HEALTH TECH SUSTAINABILITY EDUCATION VIDEO PODCASTS BLOGS STORE


COMPUTING


## Could AI Be the Future of Fake News and Product Reviews?


An experiment using a computer algorithm to create deceptive Yelp reviews was disturbingly successful, and could point to bigger problems as artificial intelligence matures


By Larry Greenemeier on October 16, 2017

**SHARE**

 SHARE 431


 TWEET

 COMMENT

 EMAIL

TEM.SIMONITE BUSINESS 09.19.17 11:04 AM

### HUMANS CAN'T EXPECT AI TO JUST FIGHT FAKE NEWS FOR THEM




YENPITSU NEMOTO/GETTY IMAGES


**GET WIRED MAGAZINE**


Don't Let the Future Leave You Behind.


[SUBSCRIBE NOW](#)


#### MOST POPULAR

 **BUSINESS**  
Theranos and Silicon Valley's 'Fake It Till You Make It' Culture  
KRISTIN GRUFFITH


 **SCIENCE**  
Maybe Nobody Wants Your Space Internet  
KARAH SCOLLES

 **SECURITY**  
Today's Debate Over Online Porn Laws Started Decades Ago  
LOUISE MATSARIS

 MORE STORIES



## FAKE NEWS DETECTOR AI

Ex: nytimes.com 

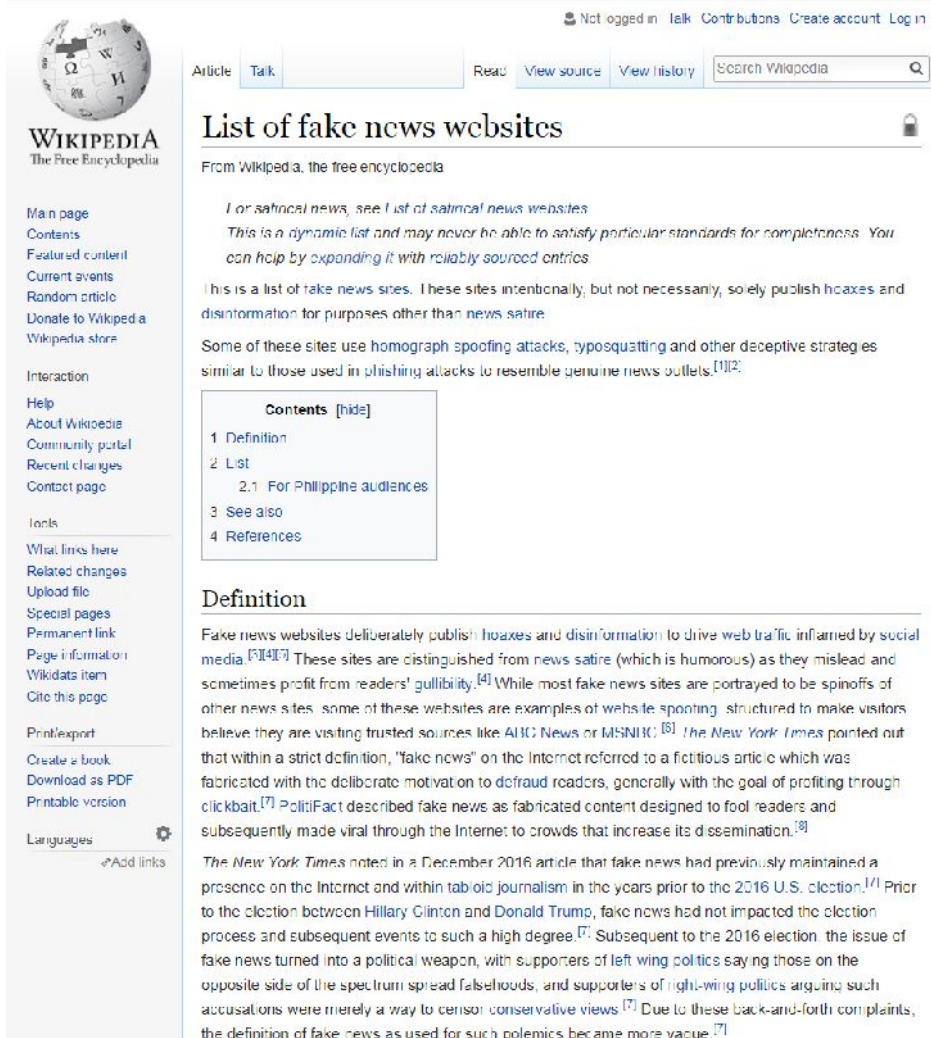
Detect fake news sites using the power of artificial intelligence!

We analyze websites to see if they are similar to known fake news sites using a neural network. The same technology is used to power other artificial intelligence applications, like Siri and self-driving cars!



# New source is an easy identifier of “fake news”

- what site is the origin of news
- shortlist of “fake news” sites



The screenshot shows the Wikipedia article titled "List of fake news websites". The page includes the standard Wikipedia header with the logo, navigation tabs (Article, Talk), and a search bar. The left sidebar contains various navigation links such as "Main page", "Contents", "Featured content", "Current events", "Random article", "Donate to Wikipedia", "Wikipedia store", "Interaction", "Help", "About Wikipedia", "Community portal", "Recent changes", "Contact page", "Tools", "What links here", "Related changes", "Upload file", "Special pages", "Permanent link", "Page information", "Wikidata item", "Cite this page", "Print/export", "Create a book", "Download as PDF", "Printable version", "Languages", and "Add links".

The main content area of the article includes a summary from Wikipedia, a note about the list's dynamic nature, and a definition of fake news websites. The definition states that fake news websites deliberately publish hoaxes and disinformation to drive web traffic, often inspired by social media. It also mentions that these sites are distinguished from news satire and that they sometimes profit from readers' gullibility. The article further discusses the impact of fake news on the 2016 U.S. election and the subsequent events.

**Contents [hide]**

- 1 Definition
- 2 List
  - 2.1 For Philippine audiences
- 3 See also
- 4 References

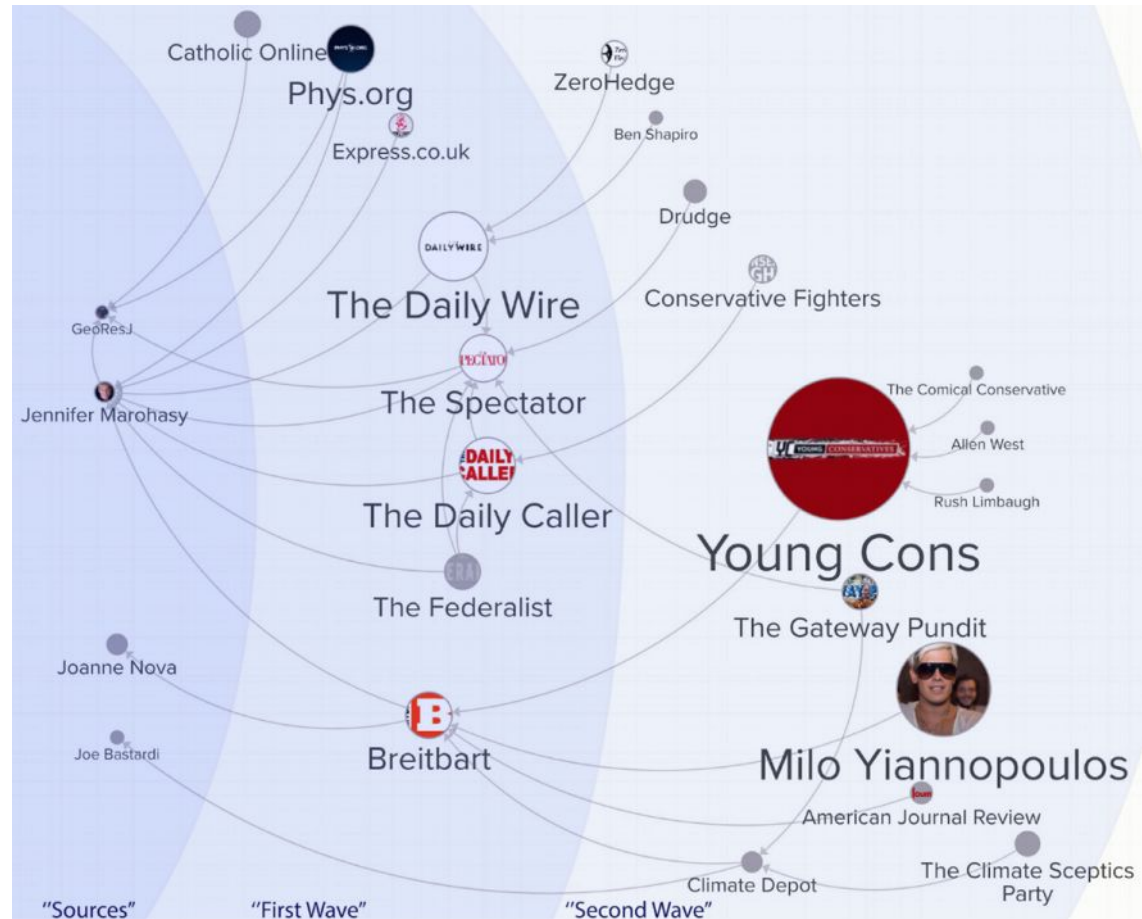
**Definition**

Fake news websites deliberately publish hoaxes and disinformation to drive web traffic inspired by social media.<sup>[5][4][7]</sup> These sites are distinguished from news satire (which is humorous) as they mislead and sometimes profit from readers' gullibility.<sup>[4]</sup> While most fake news sites are portrayed to be spinoffs of other news sites, some of these websites are examples of website spoofing, structured to make visitors believe they are visiting trusted sources like ABC News or MSNBC.<sup>[6]</sup> *The New York Times* pointed out that within a strict definition, "fake news" on the Internet referred to a fictitious article which was fabricated with the deliberate motivation to defraud readers, generally with the goal of profiting through clickbait.<sup>[7]</sup> PolitiFact described fake news as fabricated content designed to fool readers and subsequently made viral through the Internet to crowds that increase its dissemination.<sup>[8]</sup>

*The New York Times* noted in a December 2016 article that fake news had previously maintained a presence on the Internet and within tabloid journalism in the years prior to the 2016 U.S. election.<sup>[1]</sup> Prior to the election between Hillary Clinton and Donald Trump, fake news had not impacted the election process and subsequent events to such a high degree.<sup>[7]</sup> Subsequent to the 2016 election, the issue of fake news turned into a political weapon, with supporters of left-wing politics saying those on the opposite side of the spectrum spread falsehoods, and supporters of right-wing politics arguing such accusations were merely a way to censor conservative views.<sup>[7]</sup> Due to these back-and-forth complaints, the definition of fake news as used for such polemics became more vague.<sup>[7]</sup>

# Social context tells a lot about trustworthiness of news

- “influencers”
- propagation patterns
- deployment of bots
- need for speed

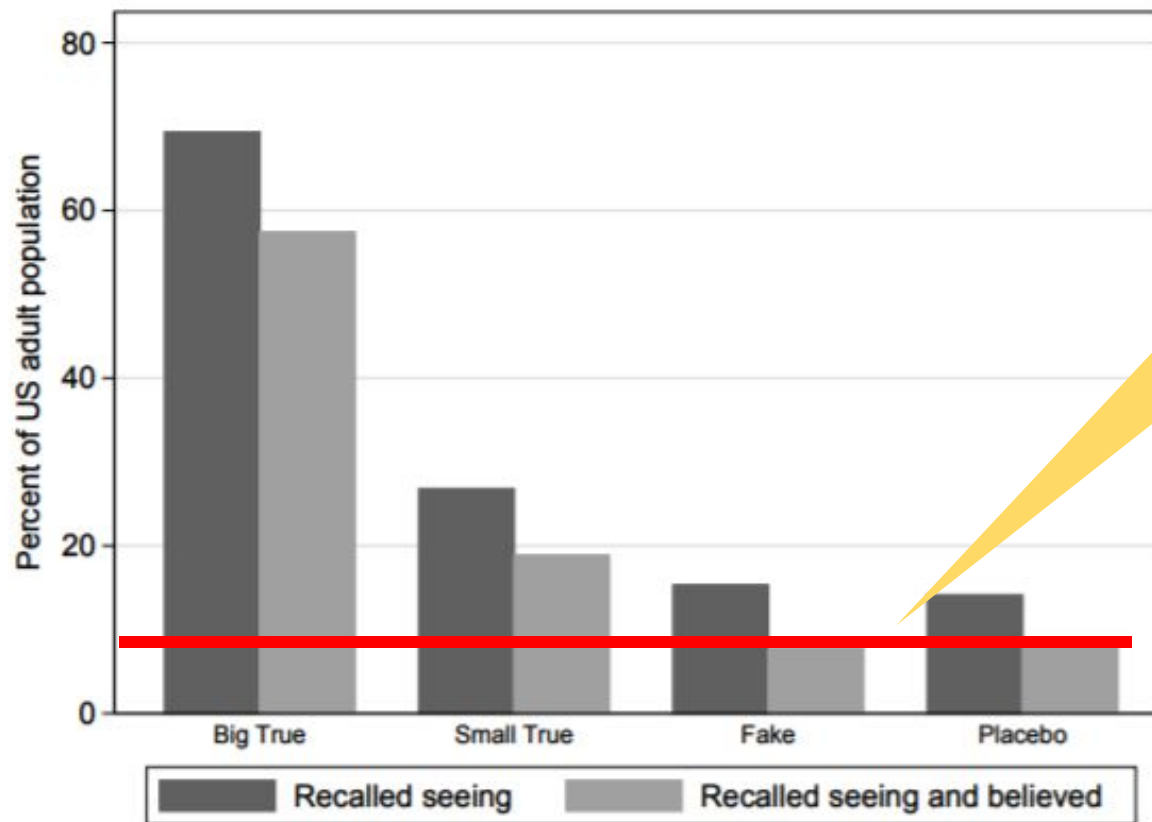


Propagation of Fake climate change story

Stop at consumption

# Media literate user is the last and best line of defense

Figure 5: Percent of US Adult Population that Recall Seeing or Believed Election News



people remember and believe “fake news” about as much as placebo news (non existent news)

**“Available evidence suggests that for now the influence of fake news is limited”.**

Hunt Allcott & Matthew Gentzkow, 2017. "[Social Media and Fake News in the 2016 Election](#)," Journal of Economic Perspectives, vol 31(2), pages 211-236.

# Conclusions



# Yes we can. But may we?

- gatekeepers are gone
  - only guerrilla war on fake news is possible
- technology to disturb circulation of fake news exists
  - most efficient if platforms cooperate
  - some approaches are clearly interfering with free speech
- users seem to be quite capable at identifying fake news
  - too early to blame election results on fake news

# The End

For more information:

[ziga.turk@gmail.com](mailto:ziga.turk@gmail.com)

@zigaTurkEU