

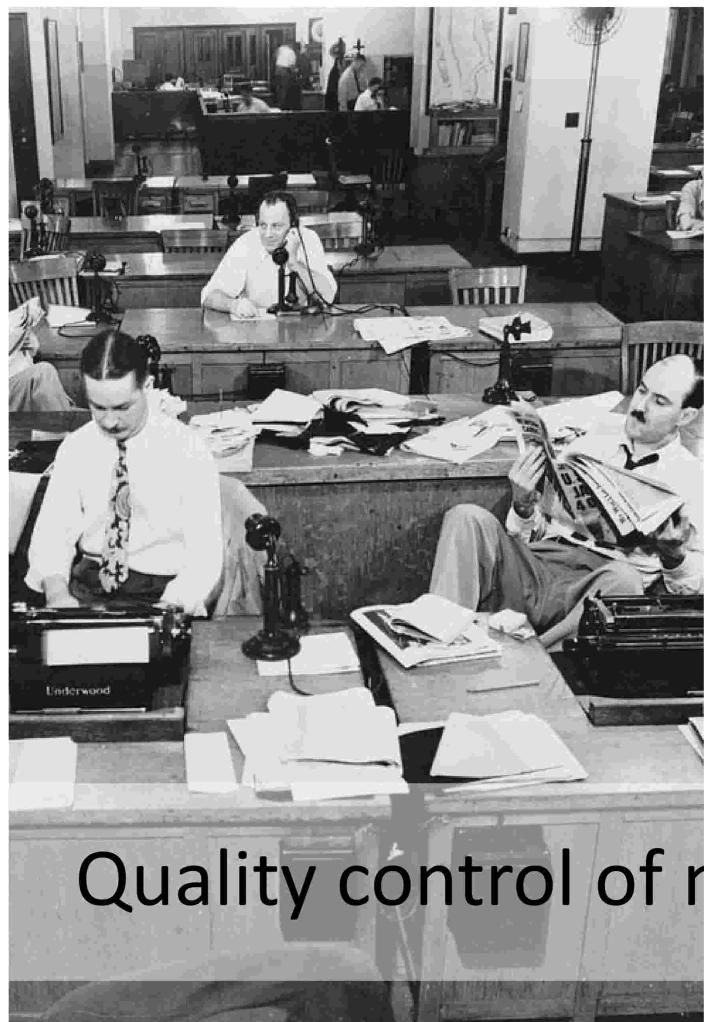
# Content

- 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

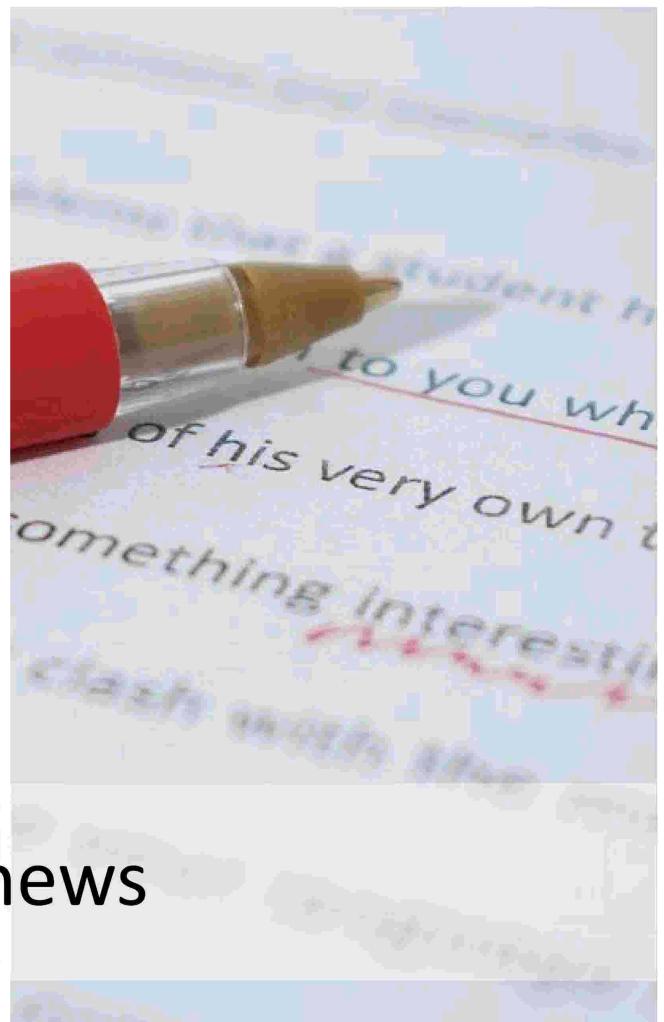
# The news process

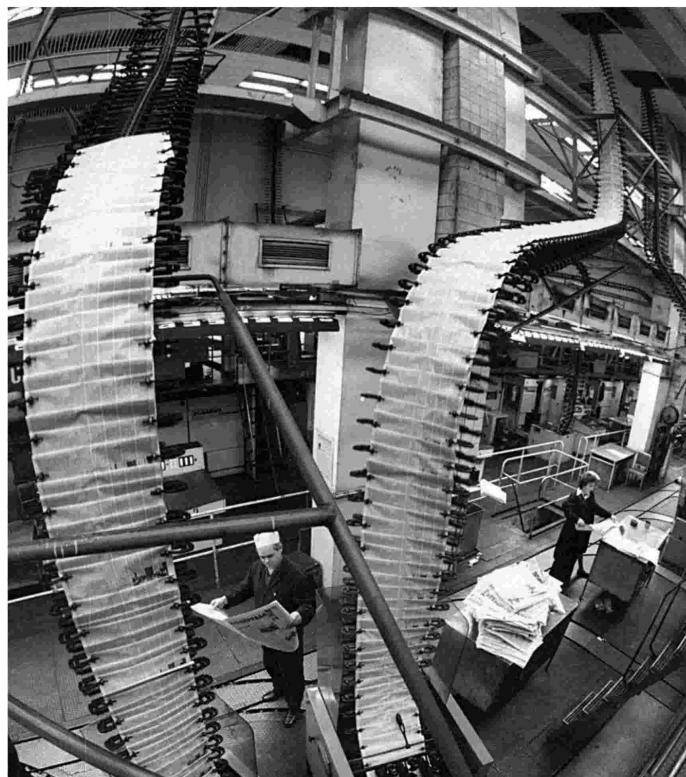


Create news



Quality control of news





Publication of news



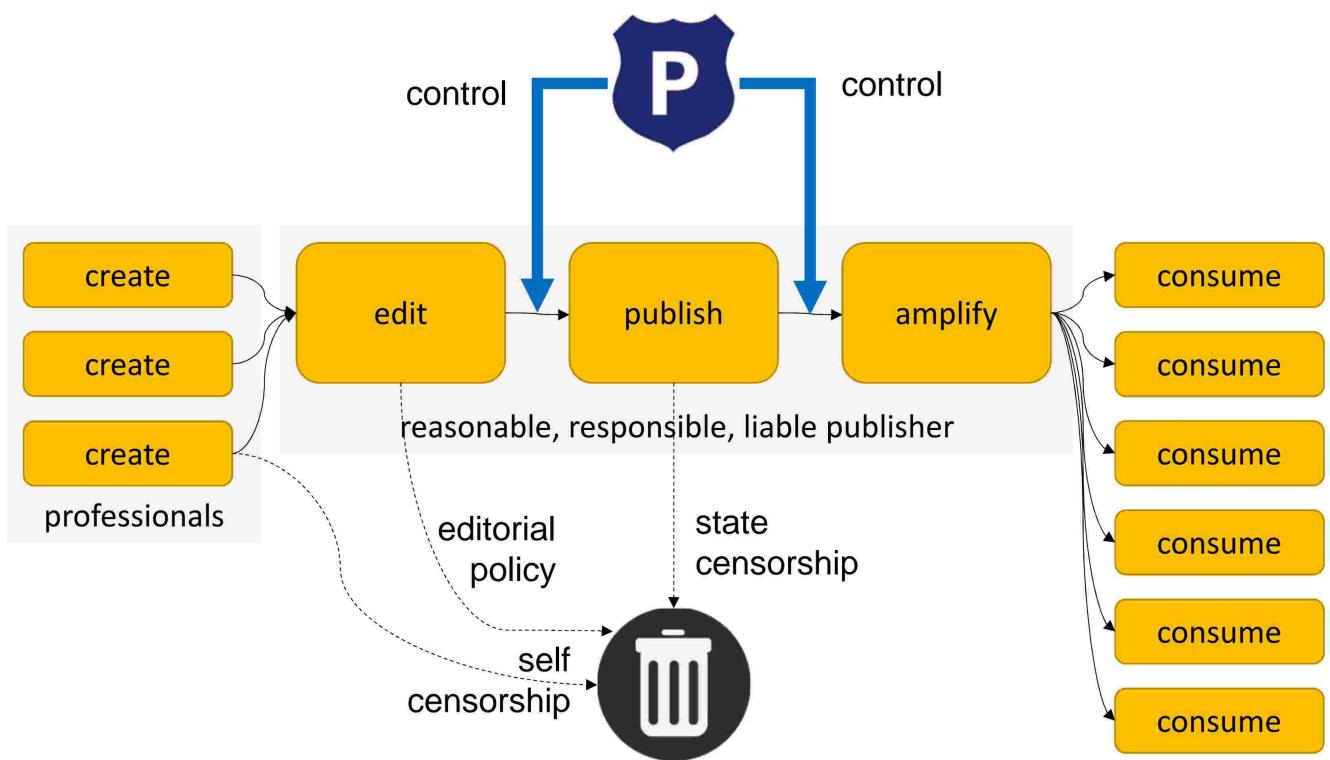
Amplification of news



Consumption of news

# The news process schematically

# Traditional news process



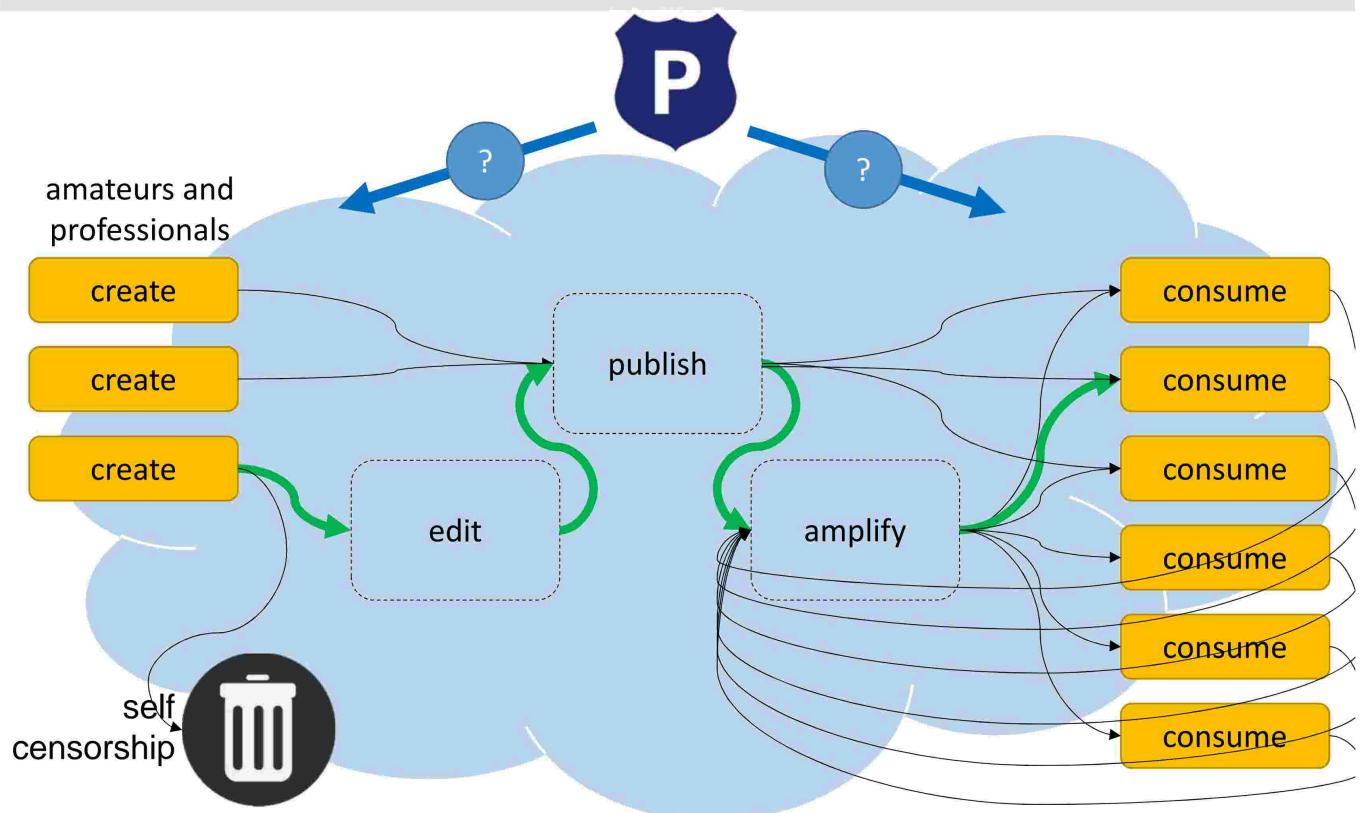
## Enters the Internet

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



source: wikipedia

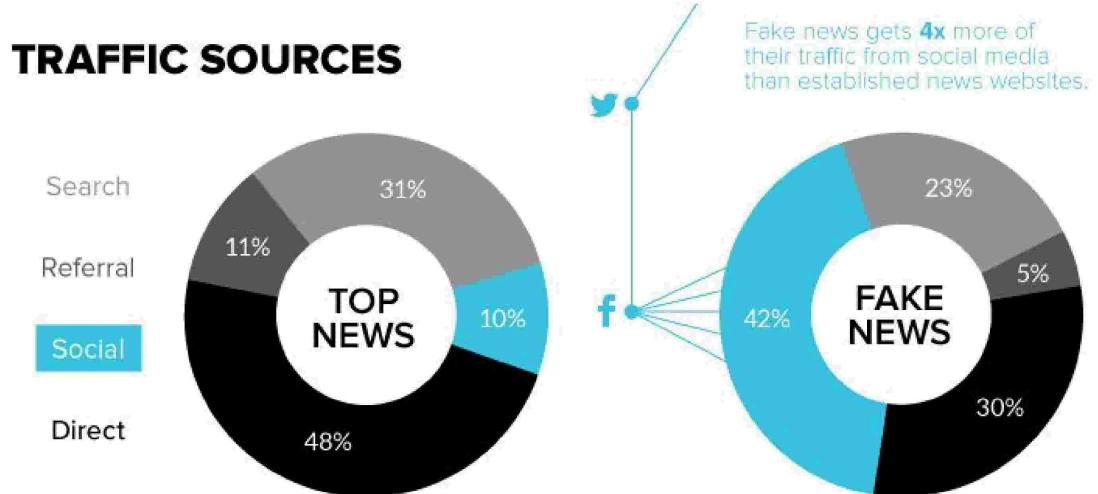
# Internet-age news process



Technology for “fake news” dissemination

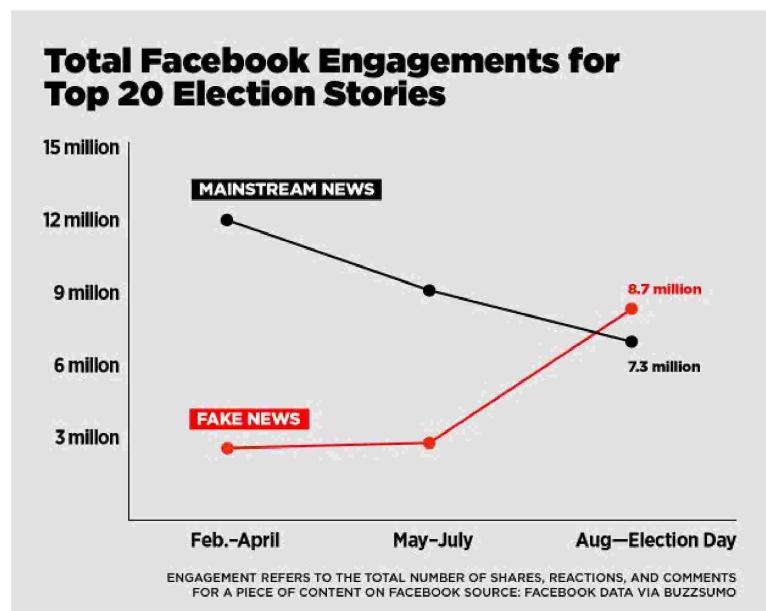
# Social media is main driver for “fake news”

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



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

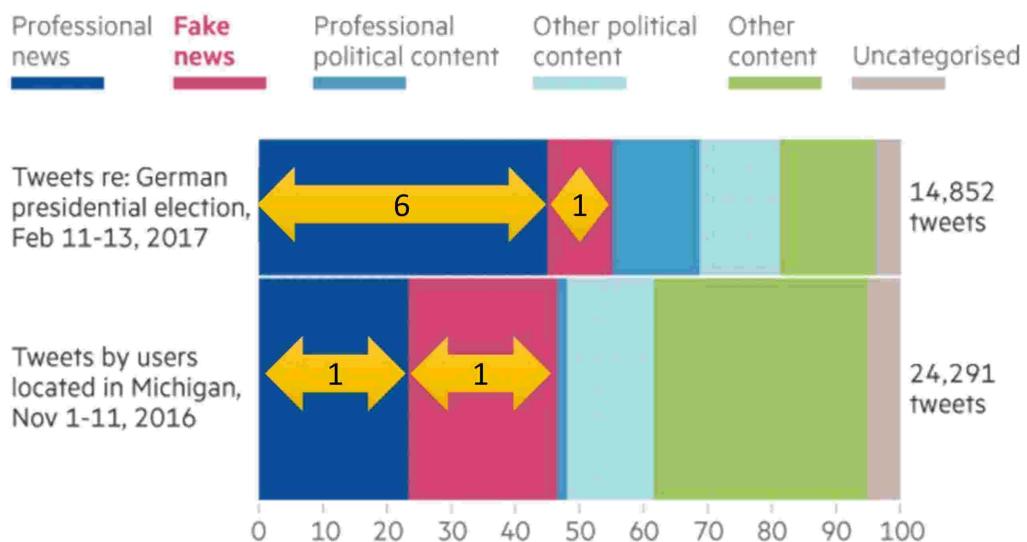
- this is 10% for real news
- but 40% for “fake news”



# 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



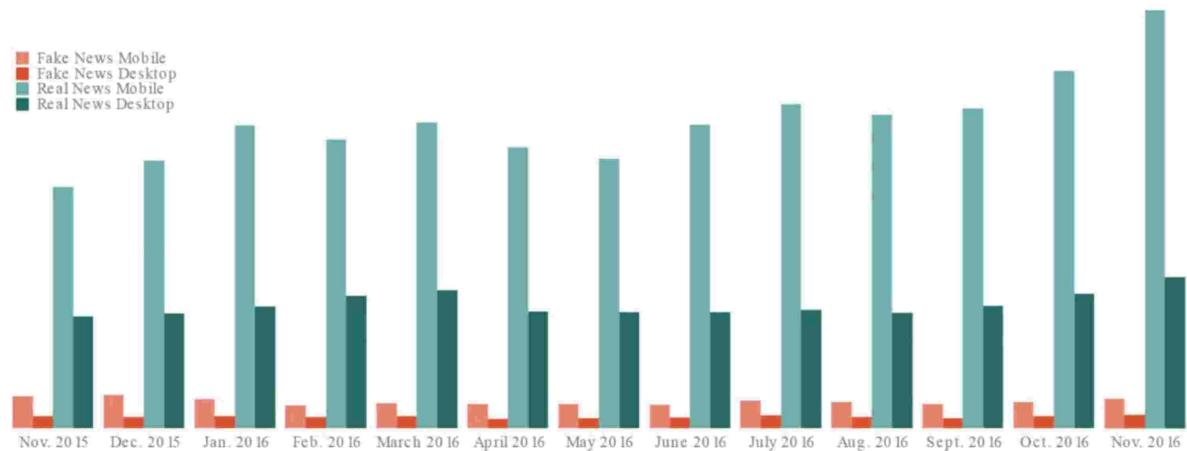
FT graphic: David Blood Source: Oxford Internet Institute



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

**Columbia Journalism Review.**

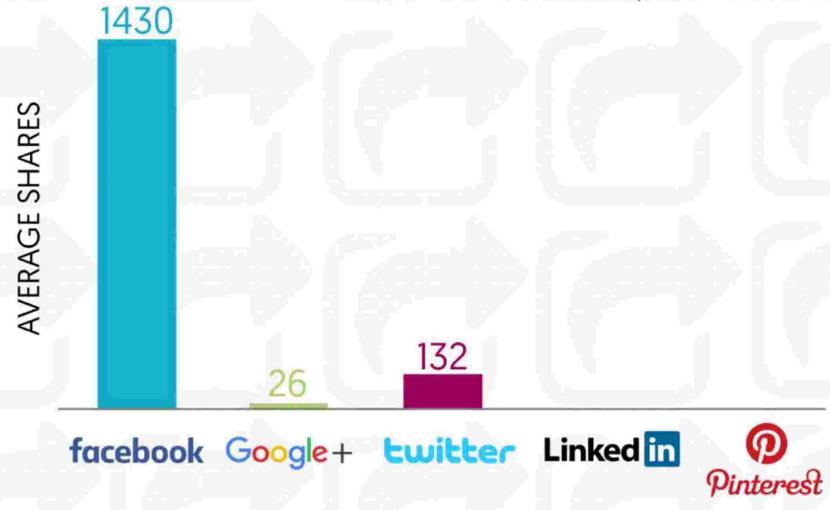
The real news audience even spiked in October and November 2016, while the fake news audience stayed the same.



# For „fake news“ main source is Facebook

## Average “Pizzagate” Shares By Network Including Both Fake And Real News Articles

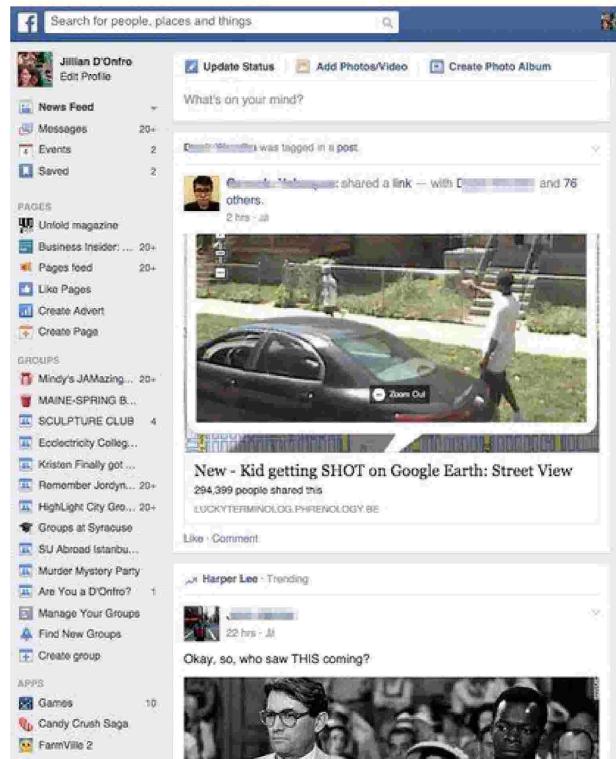
October 30, 2016 - November 21, 2016



MEDIAMATTERS  
FOR AMERICA

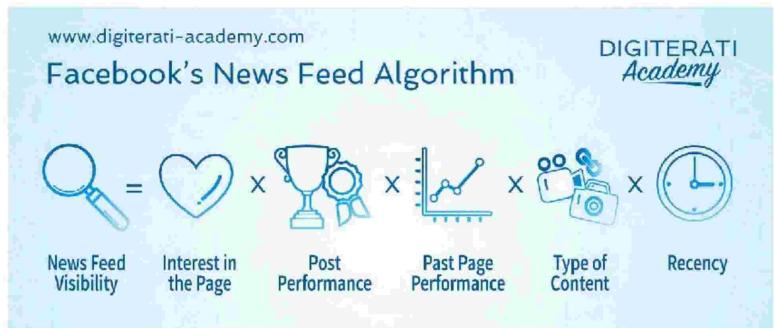
# On Facebook users see 10% of what they „subscribe” to

- 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!

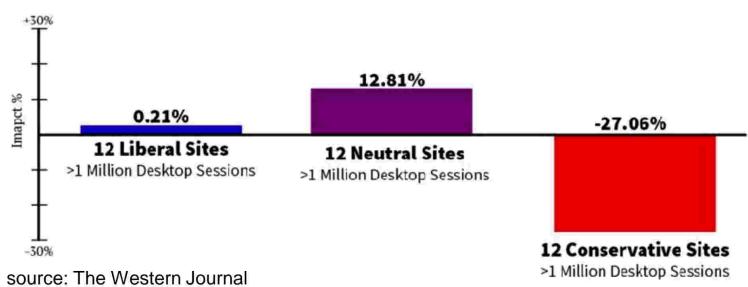


# How they choose the 10% is key issue for stopping “fake news”!

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

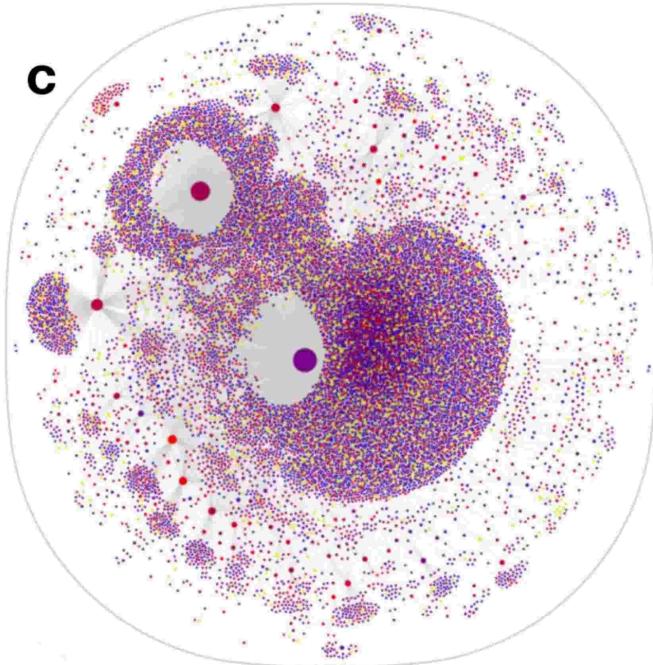


Average Facebook Algorithm Impact



# 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](#))



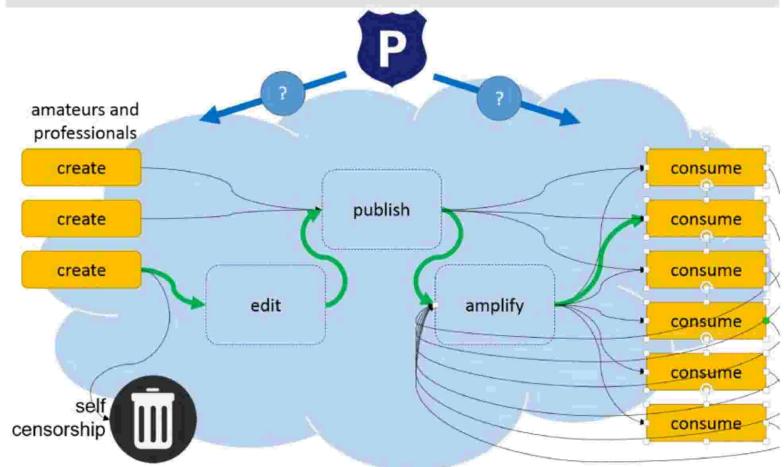
Propagation of fake news: Nodes and links represent Twitter accounts and retweets of the claim, respectively. Node size indicates account influence, measured by the number of times an account is retweeted. Node color represents bot score, from blue (likely human) to red (likely bot).

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 FINANCIAL TIMES myFT

Jeff Franz-Lien, Strat Vodafone Group PLC + Add to myFT  
Answered Feb 2 Vodafone to block its ads appearing on offensive websites

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

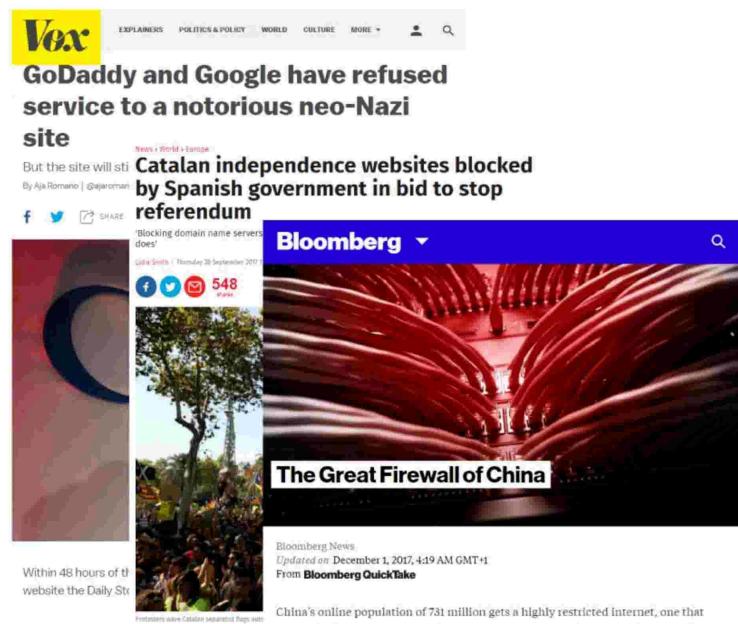
According to chatter on social media, Vodafone is one of several companies demonetizing “controversial” content on YouTube. Not just Rubin or other right-wing hosts, but others who have been hit too, including Donald Trump.

Rubin confirms this in the video above, noting that he’s losing big bucks (Dave says \$100,000 per video) because of his controversial content. Who is responsible for this? It appears to be YouTube and its advertising partners.



# Stopping “fake news” when published

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

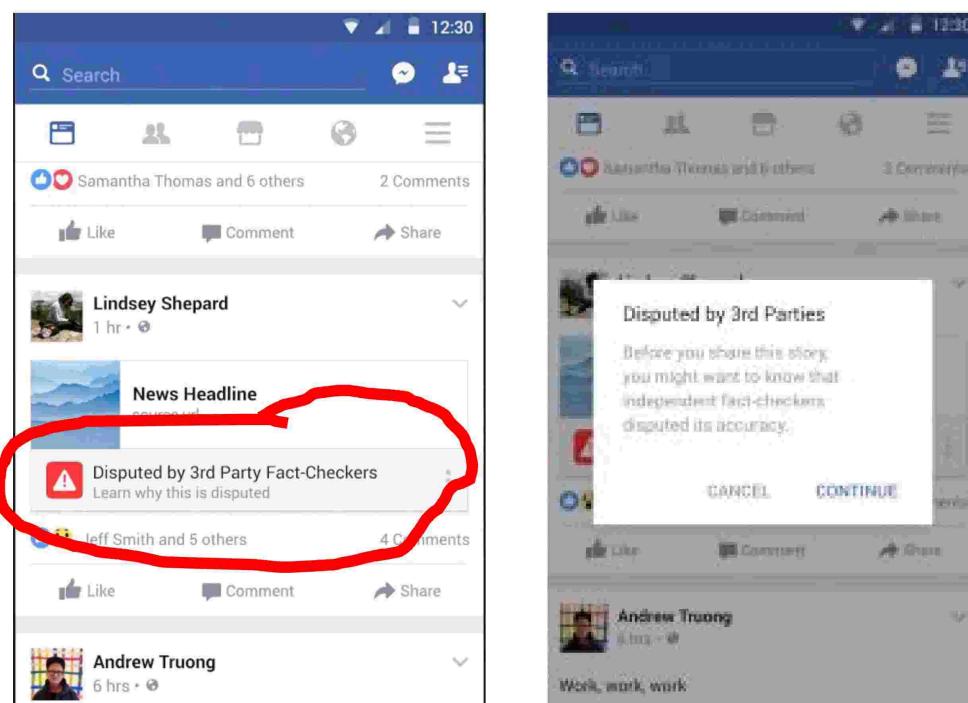


China's online population of 731 million gets a highly restricted internet, one that doesn't include access to Google, Facebook, YouTube or the New York Times. There's little coverage of the 1989 student protests in Tiananmen Square. Even Winnie the Pooh got temporarily banned. China is able to control 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.

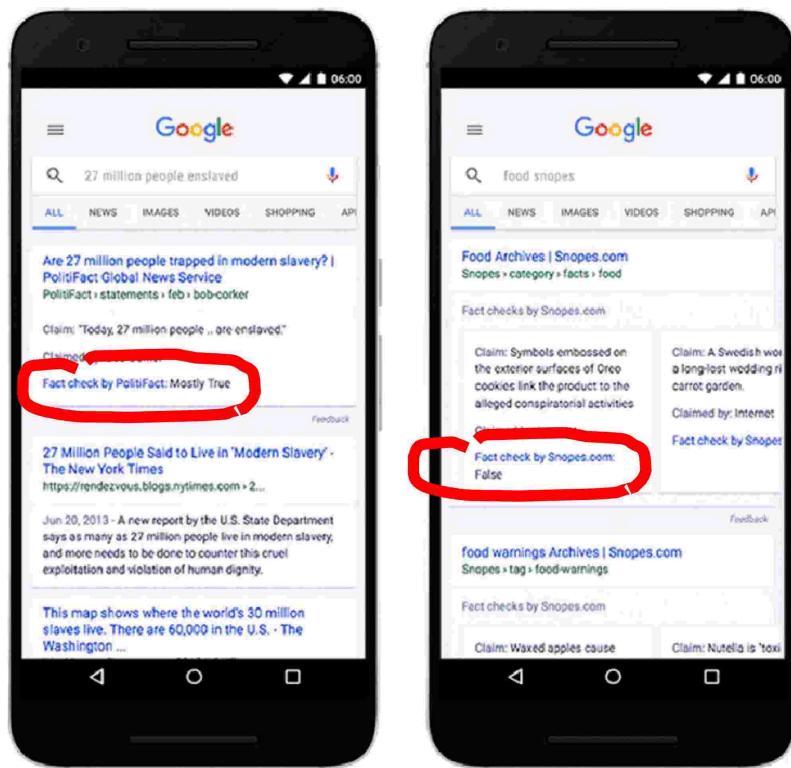
## Stopping fake news at platforms

- 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 shows a Twitter browser window. On the left is the profile of Donald J. Trump (@realDonaldTrump), featuring a portrait of him and his bio: "President-elect of the United States". In the center is a tweet from his account:

**Donald J. Trump**   
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

The tweet was posted at 6:09 AM - 16 Dec 2016. Below the tweet is a fact-check from The Washington Post:

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.

On the right side of the screen, there is a large image of an older man, likely John Podesta, smiling. At the bottom right of the screen, there is a footer for The Washington Post.

# 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 Future

TECHNOLOGY NEWS NOVEMBER 23, 2017 / 9:15 PM / 4 MONTHS AGO

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

Reuters Staff

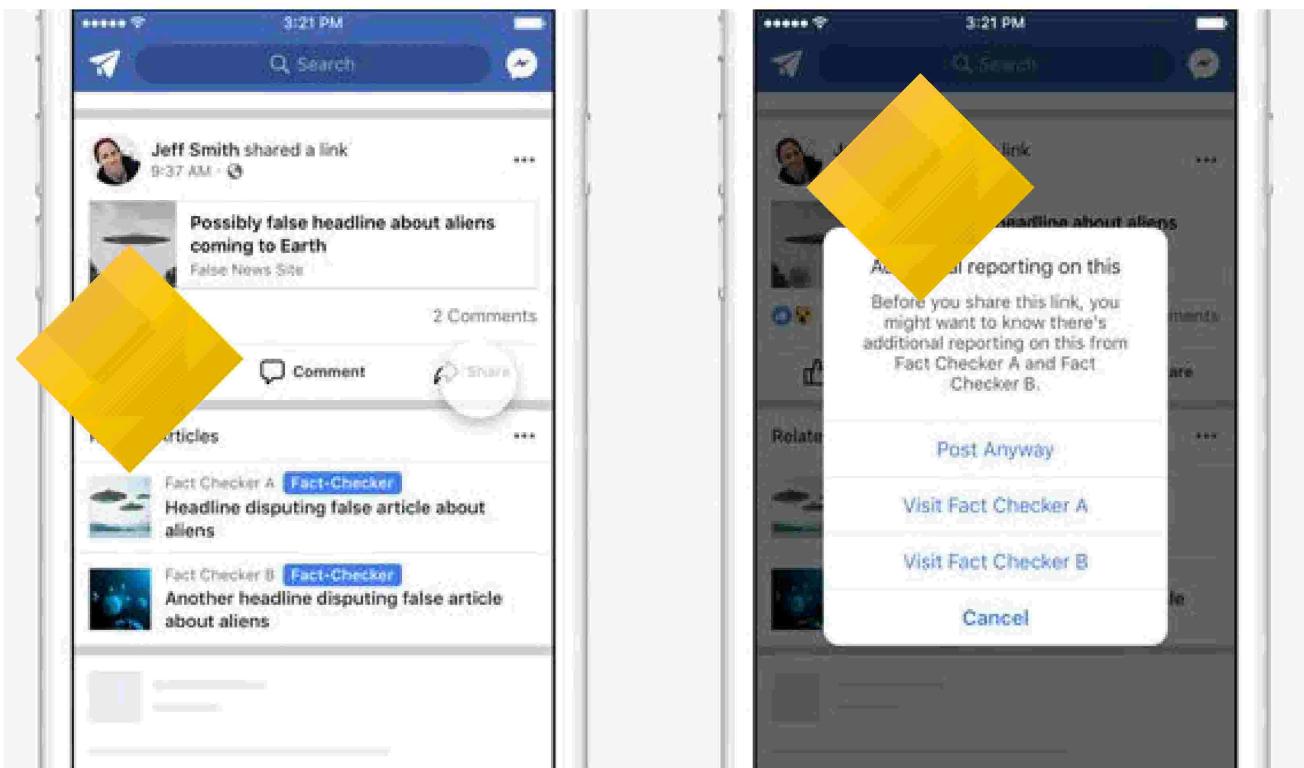
1 MIN READ



MOSCOW (Reuters) - Russia's foreign ministry said on Thursday that moves by Alphabet Inc's (GOOGL.O) 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

The Guardian website header: Support The Guardian, Subscribe, Find a job, News, Opinion, Sport, Culture, Lifestyle, Books, Music, TV & radio, Art & design, Film, Games, Classical, Stage, More.

Richard Dawkins event cancelled over his 'abusive speech against Islam'

Berkeley's KPFA after learning of author contests

This account has been suspended due to multiple or severe violations of YouTube's policy against spam, gaming, misleading content, or other Terms of Service violations.

This video is no longer available because the YouTube account associated with this video has been terminated.

Sorry about that.

▲ "Why is it fine to criticise Christianity but not Islam? ... Richard Dawkins. Photograph: Murdo MacLeod for the Guardian

YES! WE BAN!

# Stop bots!



The image shows a screenshot of the Twitter 'Verified Account' beta testing page. At the top, the Twitter logo is visible. Below it, a blue checkmark icon is followed by the text 'Verified Account' and the word 'BETA' in small letters. A large green button labeled 'Get Started' is prominent. To the left of the button, there's a question 'What is a verified account?' and a detailed explanation about preventing identity confusion through beta testing.



# 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

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

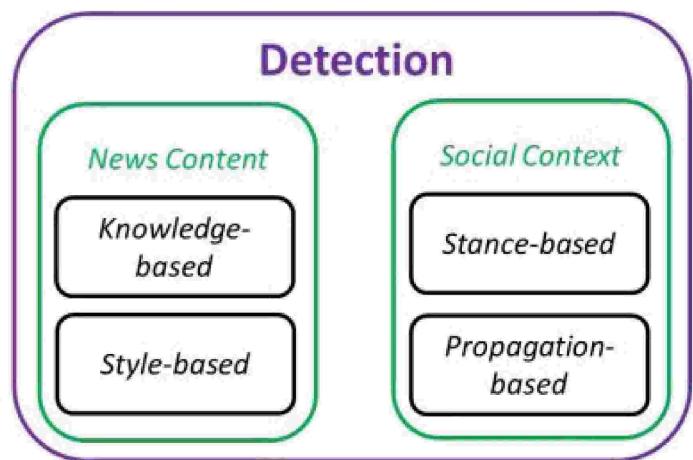
# Two strategies for automatic identification of “fake news”

- 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 ▾ Cart  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 Jan 03, 2017

SHARE

 SHARE 421

 TWEET

 COMMENT

 EMAIL

THE SCIENCE BACKBONE OF FAKE NEWS  
**HUMANS CAN'T EXPECT AI TO JUST FIGHT FAKE NEWS FOR THEM**



© VENDETTA NEWS/GETTY IMAGES



MOST POPULAR



ARTICLES

Theater and Science: What'sFake & Told You Made It' Culture

BY SARAH ERENTS



SCIENCE

Maybe Robotics Wants Your Smart Internet

BY SARAH ERENTS



TECH

Tinder's Debate Over Online Porn Laws Started Decades Ago

BY JESSICA MATTAICIO

 HOME 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!

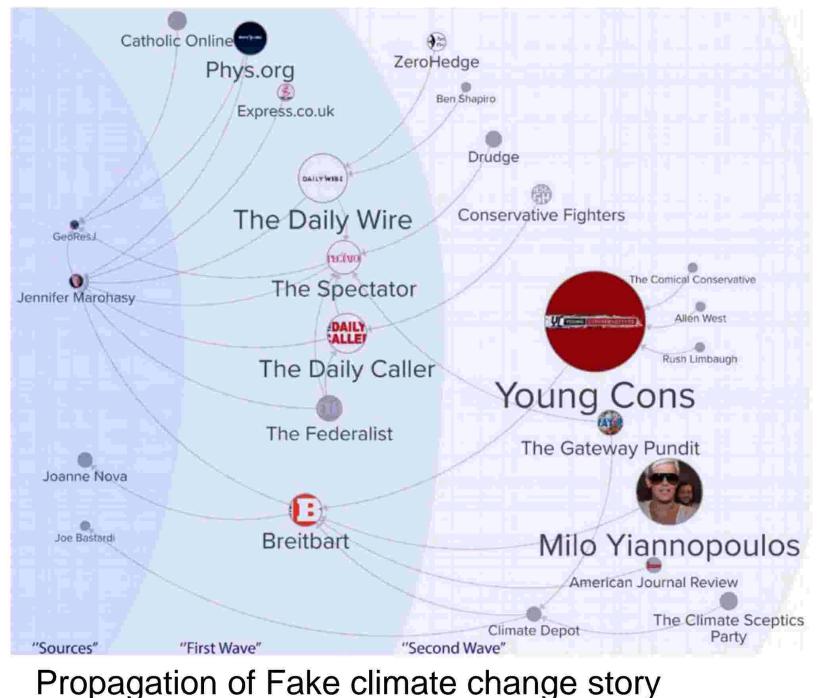
# New source is an easy identifier of “fake news”

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

The screenshot shows a Wikipedia article titled "List of fake news websites". The page header includes the standard Wikipedia navigation bar with links for Article, Talk, Read, View source, View history, and Search Wikipedia. The main content area is titled "List of fake news websites" and is described as "From Wikipedia, the free encyclopedia". A note at the top states: "For satirical news, see [list of satirical news websites](#). This is a dynamic list and may never be able to satisfy particular standards for completeness. You can help by expanding it with reliably sourced entries." Below this, a sub-note says: "This is a list of fake news sites. These sites intentionally, but not necessarily, soley publish hoaxes and disinformation for purposes other than news satire." A sidebar on the left contains links for Main page, Contents, Featured content, Current events, Random article, Donate to Wikipedia, Wikipedia store, Help, About Wikipedia, Community portal, Recent changes, Contact page, Tools, Wikipedia links, 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, and Languages. A "Contents [hide]" box on the right lists sections: 1 Definition, 2 List (with a sub-section 2.1 For Philippine audiences), 3 See also, and 4 References. The main text area discusses the definition of fake news websites, noting their deliberate publication of hoaxes and disinformation to drive web traffic, often inflamed by social media. It highlights how these sites are distinguished from news satire (which is humorous) and how they sometimes profit from readers' gullibility. The text also mentions the 2016 U.S. election and the role of fake news in that election.

# Social context tells a lot about trustworthiness of news

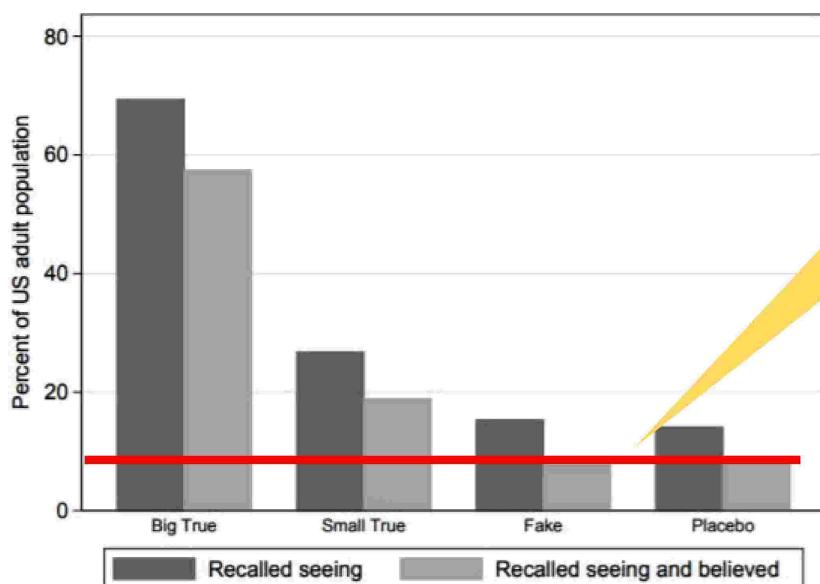
- “influencers”
- propagation patterns
- deployment of bots
- need for speed



**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