Computational Solutions against Fake News: AI vs. DB Approaches

Naeemul Hassan

nhassan@olemiss.edu

University of Mississippi, USA

Dongwon Lee

dongwon@psu.edu

Penn State University, USA

Slide Available for Download

- Tutorial web page

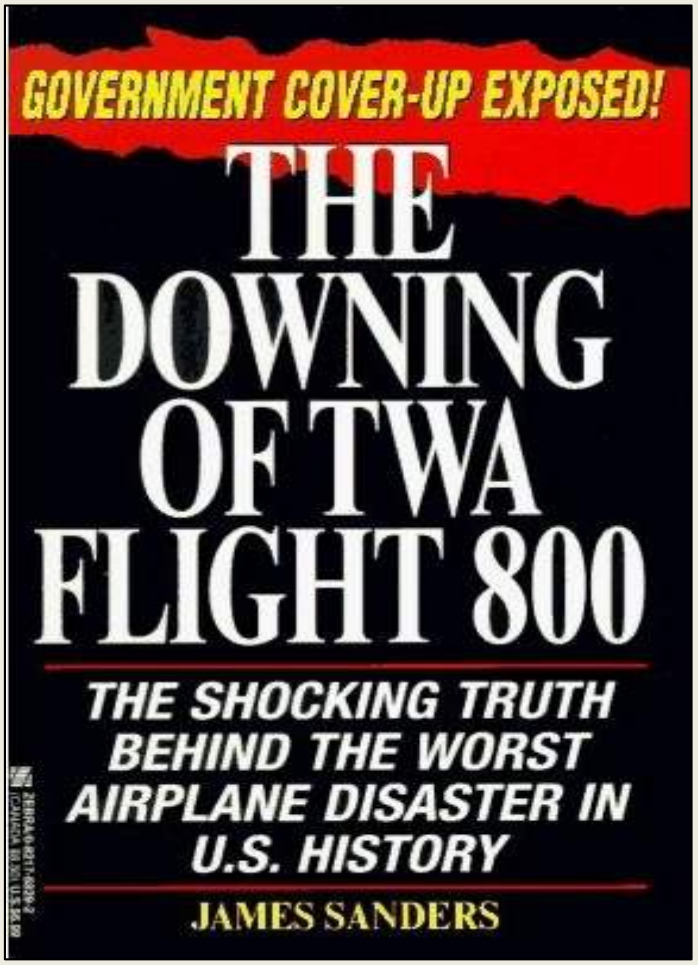
https://goo.gl/jBNW37

Part 1. Introduction

Part 1: Introduction

- What is “Fake News”?
- Issues
- Episode
- Important Questions
- Online Quiz

Fake News is Not New



Hundreds Of Eyewitnesses Know the FBI and CIA Lied!

WE WANT THE NATIONAL TRANSPORTATION SAFETY BOARD TO TELL THE TRUTH

We Saw TWA Flight 800 Shot Down by Missiles

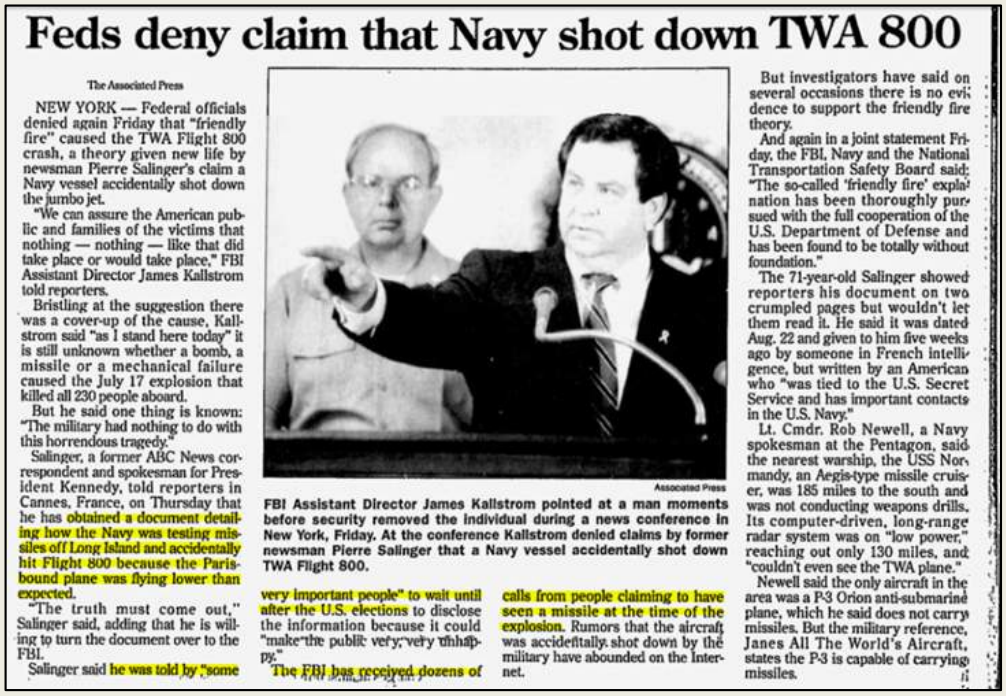
And We Won't Be Silenced Any Longer

We are some of the hundreds of eyewitnesses to the crash of TWA Flight 800 that killed 230 people off the coast of Long Island on July 17, 1996.

We are OUTRAGED that the FBI would not let a single one of us testify at the NTSB's public hearing in December 1997 when it heard testimony on what may have happened to the plane. The FBI feared that our testimony would undermine the video produced by the CIA that was shown on national television to persuade viewers that we all mistook the plane's burning fuel for a missile. They want you to believe that a fuel tank explosion caused the crash. Their herculean efforts failed to find the necessary ignition source.

We are DISGUSTED that for nearly four years the FBI refused to release its hundreds of reports of interviews with eyewitnesses who told them what we saw — the plane being hit by missiles that broke off in mid-air, blew up the fuel tank and sent the plane plummeting into the sea.

And we are SHOCKED at the lengths to which the FBI, the CIA and the NTSB have all gone to discredit and ignore our testimony in order to hide the truth.



Feds deny claim that Navy shot down TWA 800

The Associated Press

NEW YORK — Federal officials denied again Friday that "friendly fire" caused the TWA Flight 800 crash, a theory given new life by newsman Pierre Salinger's claim a Navy vessel accidentally shot down the jumbo jet.

"We can assure the American public and families of the victims that nothing — nothing — like that did take place or would take place," FBI Assistant Director James Kallstrom told reporters.

Bristling at the suggestion there was a cover-up of the cause, Kallstrom said "as I stand here today" it is still unknown whether a bomb, a missile or a mechanical failure caused the July 17 explosion that killed all 230 people aboard.

But he said one thing is known: "The military had nothing to do with this horrendous tragedy."

Salinger, a former ABC News correspondent and spokesman for President Kennedy, told reporters in Cannes, France, on Thursday that he has obtained a document detailing how the Navy was testing missiles off Long Island and accidentally hit Flight 800 because the Paris-bound plane was flying lower than expected.

"The truth must come out," Salinger said, adding that he is willing to turn the document over to the FBI.

Salinger said he was told by "some very important people" to wait until after the U.S. elections to disclose the information because it could "make the public very, very unhappy."

The FBI has received dozens of calls from people claiming to have seen a missile at the time of the explosion. Rumors that the aircraft was accidentally shot down by the military have abounded on the Internet.

Associated Press

FBI Assistant Director James Kallstrom pointed at a man moments before security removed the individual during a news conference in New York, Friday. At the conference Kallstrom denied claims by former newsman Pierre Salinger that a Navy vessel accidentally shot down TWA Flight 800.

But investigators have said on several occasions there is no evidence to support the friendly fire theory.

And again in a joint statement Friday, the FBI, Navy and the National Transportation Safety Board said: "The so-called 'friendly fire' explanation has been thoroughly pursued with the full cooperation of the U.S. Department of Defense and has been found to be totally without foundation."

The 71-year-old Salinger showed reporters his document on two crumpled pages but wouldn't let them read it. He said it was dated Aug. 22 and given to him five weeks ago by someone in French intelligence, but written by an American who "was tied to the U.S. Secret Service and has important contacts in the U.S. Navy."

Lt. Cmdr. Rob Newell, a Navy spokesman at the Pentagon, said the nearest warship, the USS Normandy, an Aegis-type missile cruiser, was 185 miles to the south and was not conducting weapons drills. Its computer-driven, long-range radar system was on "low power," reaching out only 130 miles, and "couldn't even see the TWA plane."

Newell said the only aircraft in the area was a P-3 Orion anti-submarine plane, which he said does not carry missiles. But the military reference, Janes All The World's Aircraft, states the P-3 is capable of carrying missiles.

Definitions of "Fake News"

- Fake news was exclusively used in the satire context
  - Eg, The Daily Show with Jon Stewart
  - Fake news is meant to perceived as unrealistic, while traditional news content is meant to be perceived as realistic [Balmas, 2014]
- News articles that are intentionally and verifiably false, and could mislead readers [Allcott et al, 2017]
- Fake news vs. alt-facts: Alt- facts are different from fake news in that they have no basis in reality
  - Statements are evaluated based on "goodness of fit with a particular ideology, but not facts" [Berghel, 2017]
- Journalism vs. fake news: Gatekeeping– "attempt at exercising reliability, selecting the important over the trivial while avoiding sensationalism" [Bordem et al. 2017]

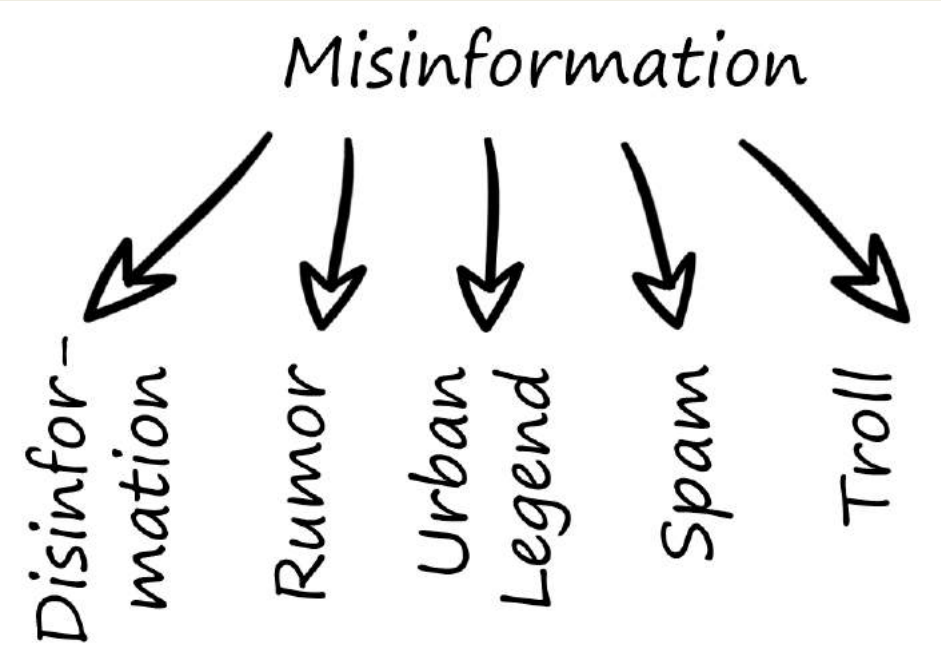
Definitions of "Fake News"

Credit: M. Molina @ PSU

- "The task of categorizing news along a continuum of veracity [intentional deception], with an associated measure of certainty" [Conroy et al., 2015]
- Intentionally false and shared via the Internet [Klein et al., 2017]
- "Everything from malicious stories to political propaganda" [Cohen, 2017]
  - Many articles are written by journalists who write articles using web searches but, with no actual verification [Cohen, 2017]
  - 53.8% of journalists use microblogs (ex. Twitter) to gather information and report from news stories [Willnat, 2014]
- Fake news as a continuum because there is some news that have a biases and do not fall directly into fake news [Novotny, 2017]

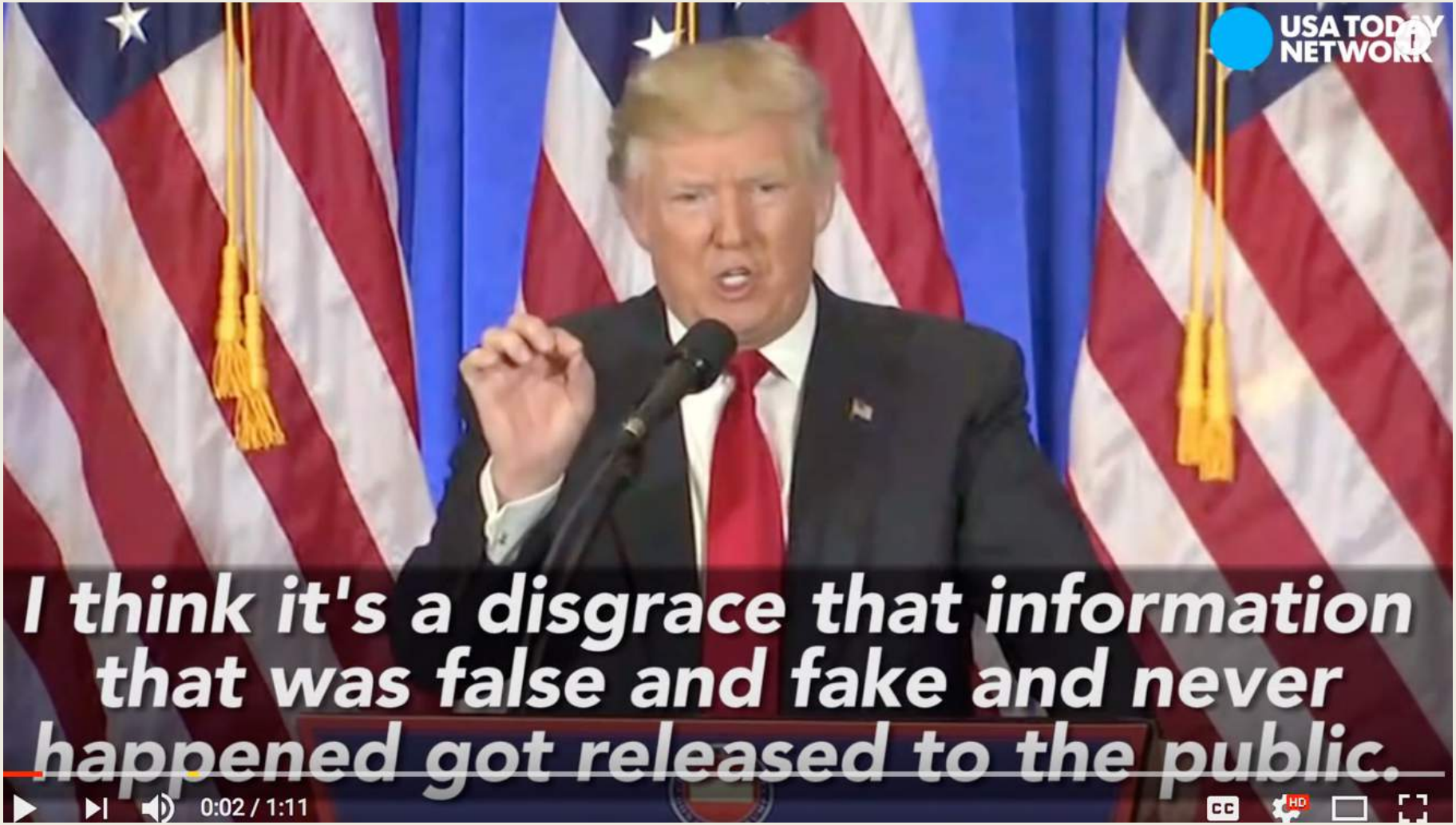
Related Terms

- Misinformation: false information
- Disinformation: deliberately false information
- Rumor: unverified information at the time



https://www.kdnuggets.com/2016/08/misinformation-key-terms-explained.html

So, What is “Fake News”?

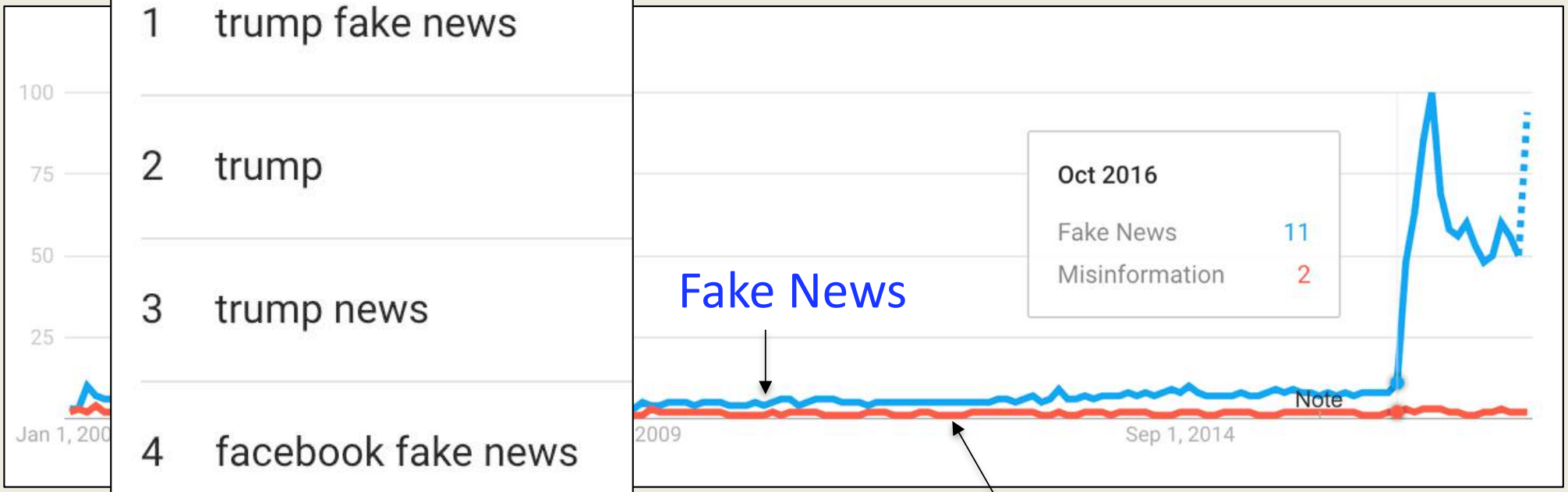


https://www.youtube.com/watch?v=wJxxQM7GxJA

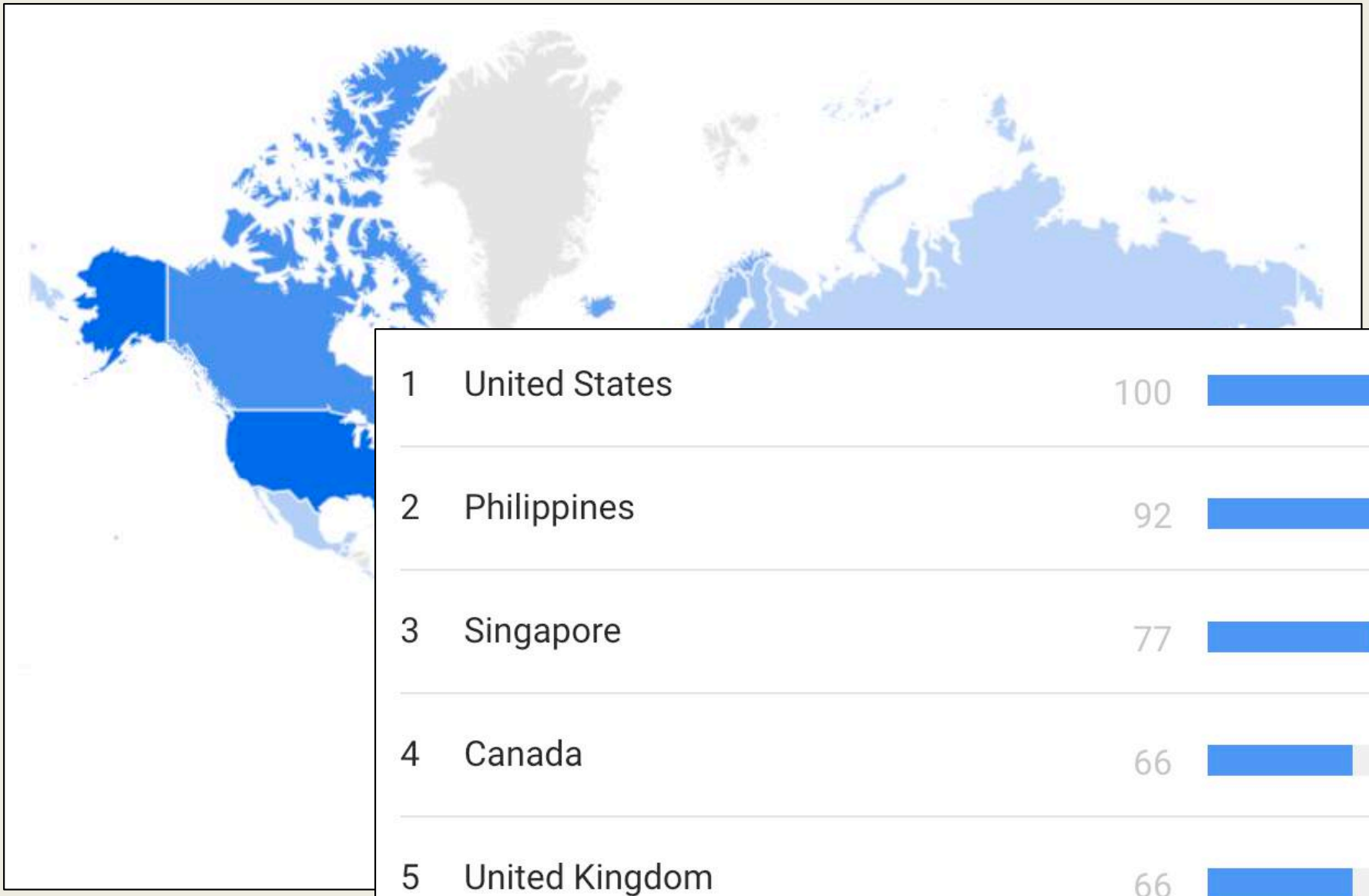
Surge of "Fake News": Google Trend

Related queries

1	trump fake news
2	trump
3	trump news
4	facebook fake news
5	fake news awards

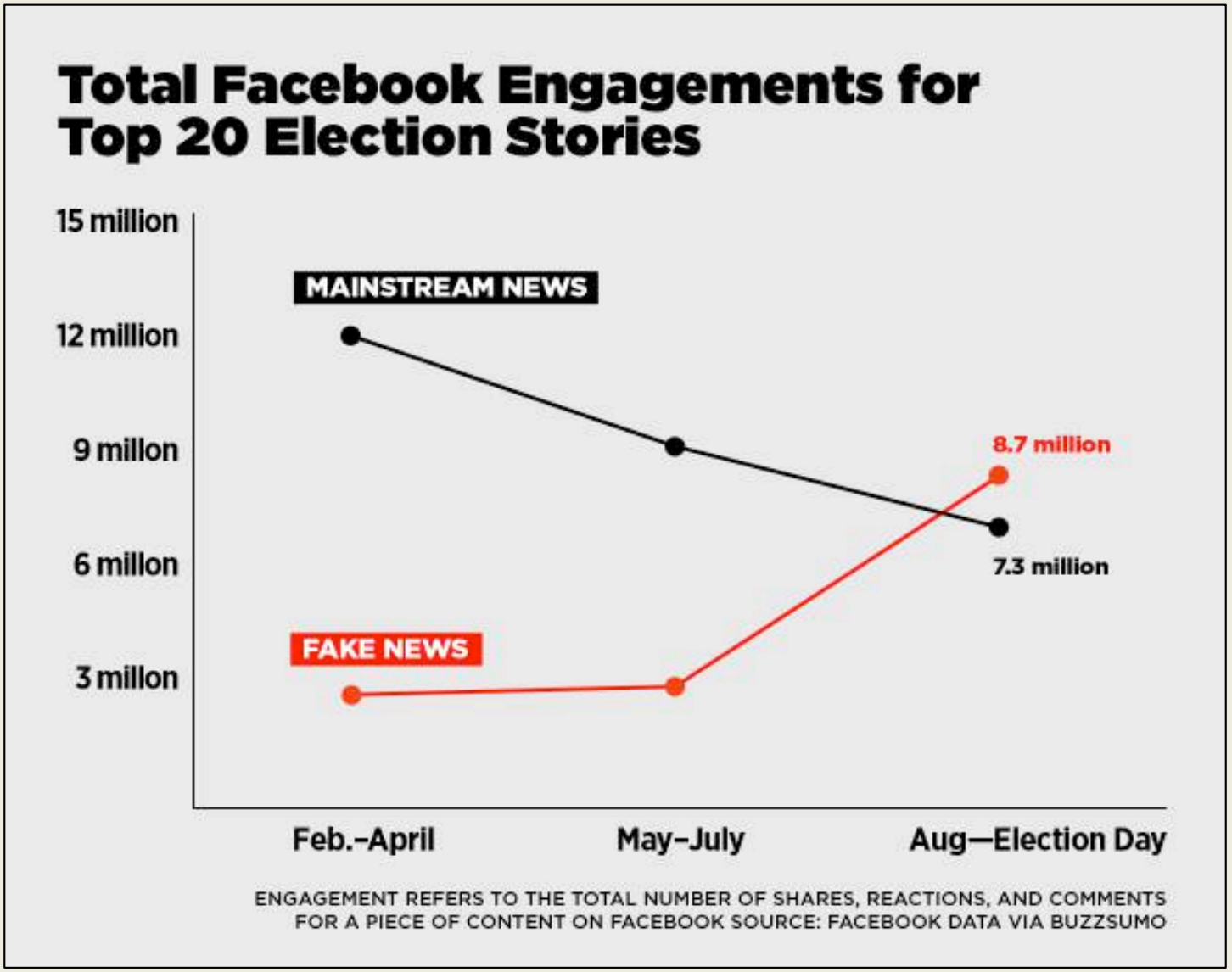


Surge of “Fake News”: Google Trend



1	United States	100
2	Philippines	92
3	Singapore	77
4	Canada	66
5	United Kingdom	66

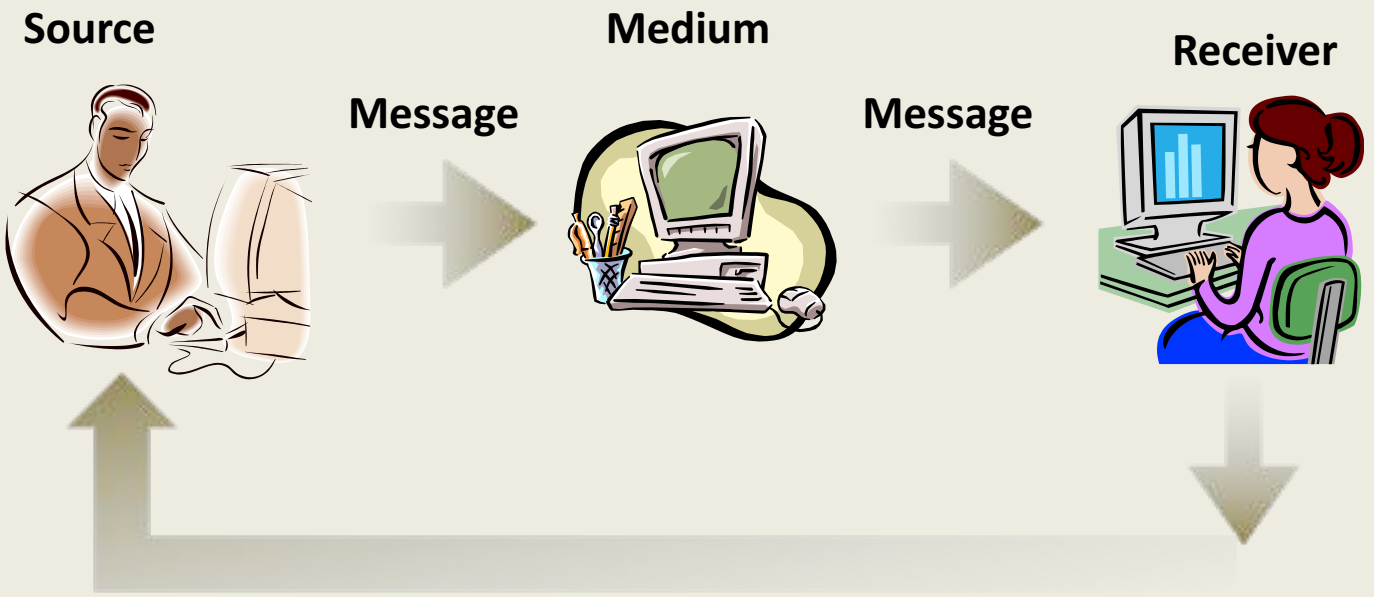
Impact of "Fake News" in 2016



https://www.vox.com/new-money/2016/11/16/13659840/facebook-fake-news-chart

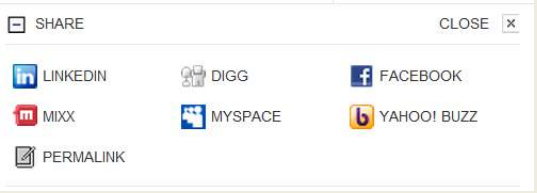
Change: Role of "Receivers"

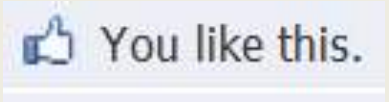
- Interactive Media → fundamental shift in communication
- Receiver becomes the new "sources"

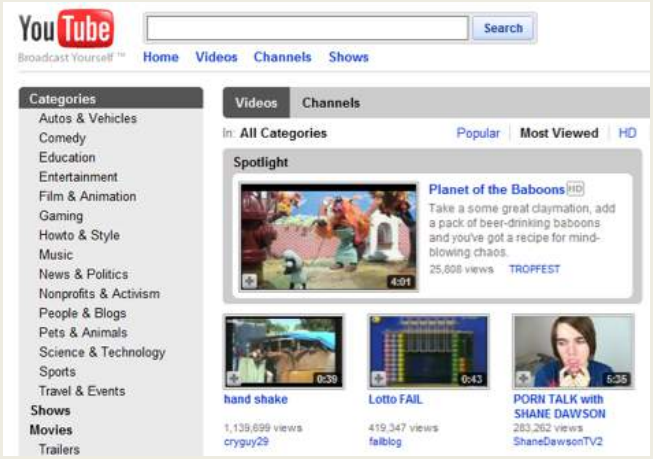


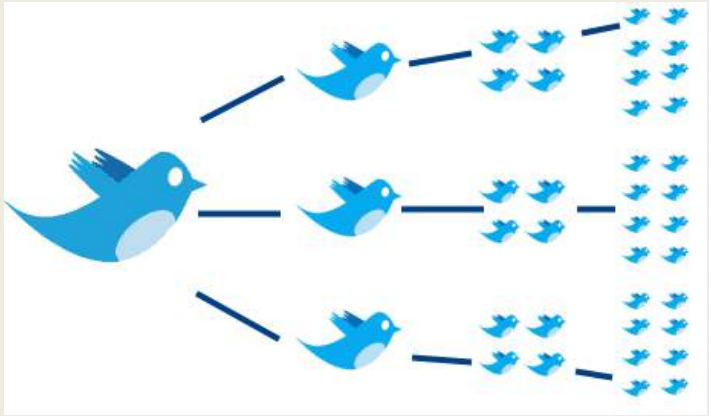
[Shannon and Weaver, 1949; Berlo, 1960; Sundar and Nass, 2001]

Change: Self as Creators and Distributors





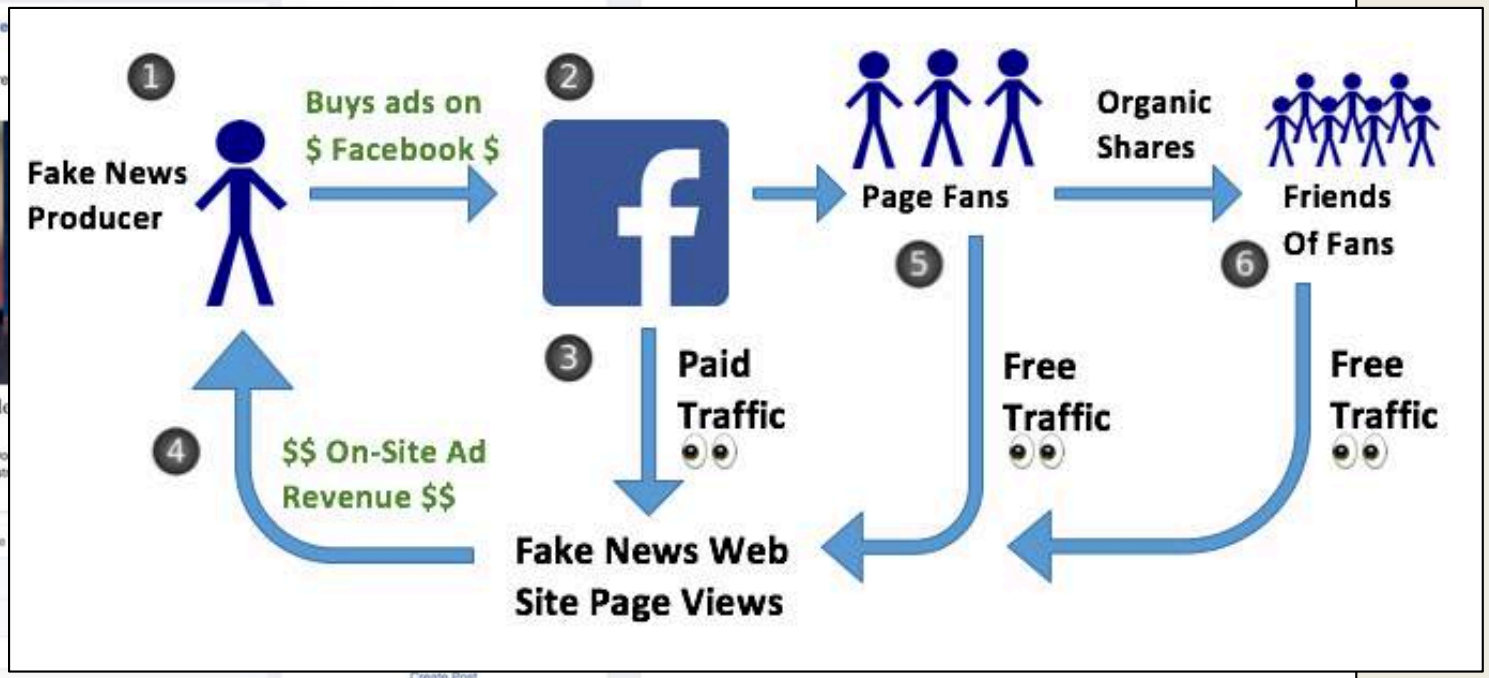




Change: New Incentives

Inside a Fake News Sausage Factory: 'This Is All About Income'

By ANDREW HIGGINS, MIKE McINTIRE and GABRIEL J.X. DANCE NOV. 25, 2016



A fake news article posted by Beqa Latsabidze, a Georgian. "My audience likes Trump," he said. "I don't want to write bad things about Trump. If I write fake stories about Trump, I lose my audience."

https://goo.gl/seKndN @ NY Times

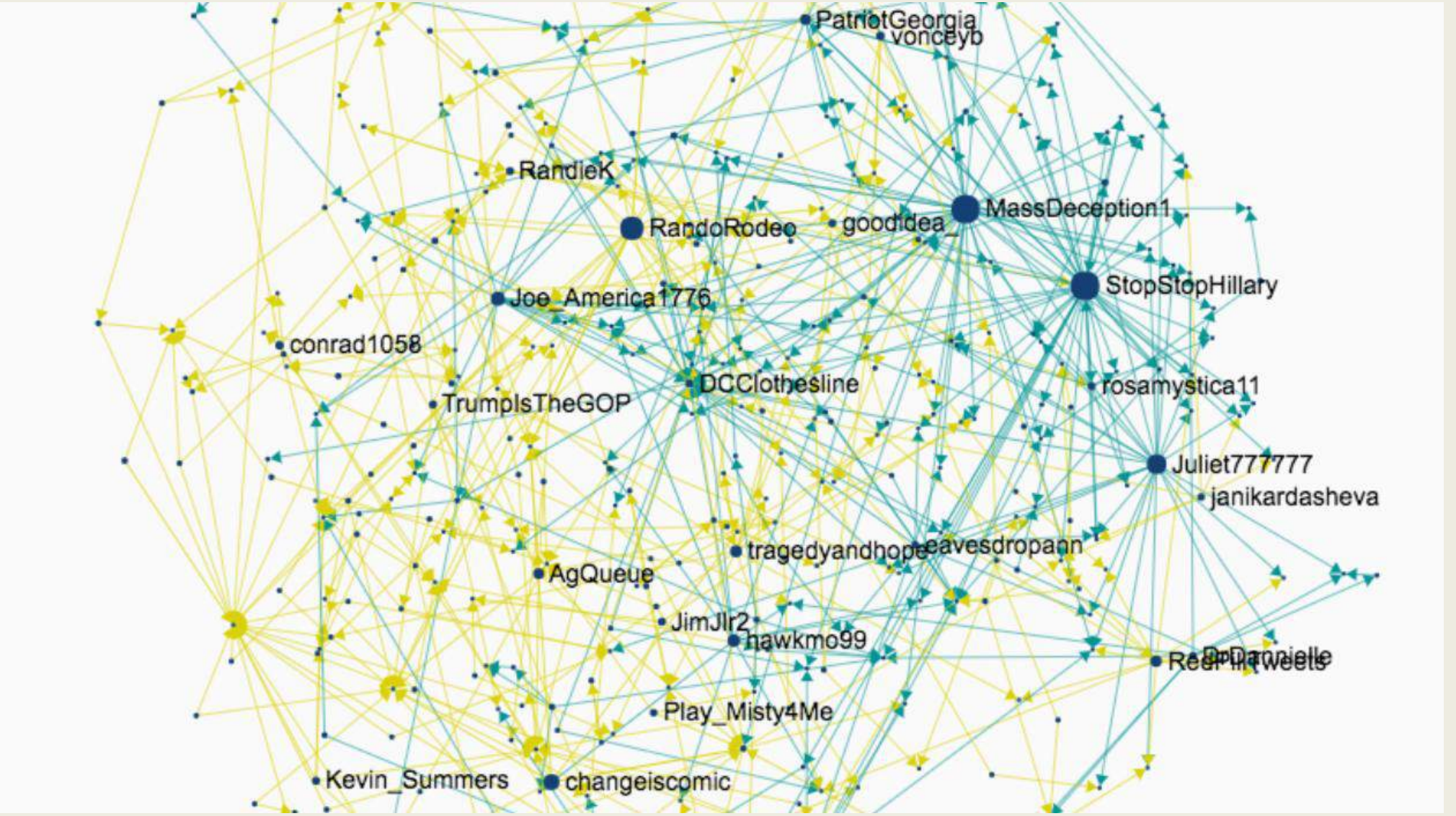
Change: Source Layering

- The reality of source layering:
  - Imagine checking your Facebook news feed and seeing something your friend has shared, a politician's tweet of a newspaper story.
  - Who or What is the Source? Newspaper, Politician, Twitter, Friend or Facebook?

Change: Proximate Sources

- Studies show that only the proximate source matters
- Increasingly, these proximate sources are not news organizations, but
  - Aggregator sites (eg, Comcast, Verizon)
  - Social networking sites (where your friends are sources)

Change: Virality



Credit: Hoaxy @ Indiana U.

Matters in Politics and Beyond

- Conspiracy Theories
- Confirmation Bias
- Selective Exposure
- Selective Perception
- Filter Bubbles
- Internet Credibility
- Media Literacy
- Information Literacy
- Research
- Education

Types of Fake News

Credit: S. Sundar @ PSU

Snopes

- What's New
- Hot 50
- Fact Check
- News
- About
- FAQ
- Contact
- Random

1

Fact Check › Politics

Did an Alabama Church Display a Sign Comparing Roy Moore to Jesus?

5 December 2017 - Living Way Ministries said its pastor did not approve of the sign, but at least two

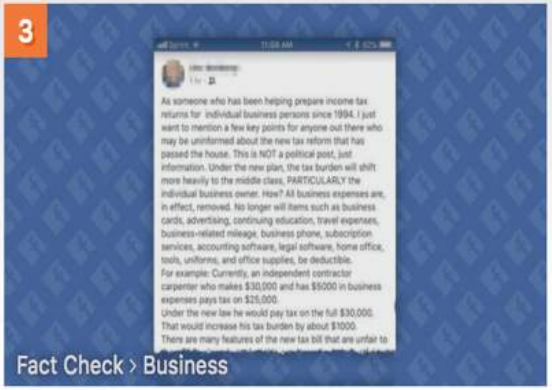
2

Fact Check › Fake News

Did Tokyo Open the First Human Meat Restaurant?

2 December 2017 - Reports that a new restaurant recently opened in Japan is the first in the world to

3



Fact Check › Business

Does the 2017 Tax Reform Bill Eliminate Small Business Deductions?

4 December 2017 - Contrary to rumor, the tax bill does not eliminate deductions for small business expenses,

Real News

Commentary / Feature Writing

Misreporting

Native Advertisement

Professional Political Content

Citizen Journalism

Satire

Polarizing and Sensationalist Content

Fake News / Hoaxes

News › Medical

Does the 2017 Republican Tax Bill Cut Cancer Treatments for Medicare Patients?

4 December 2017 - The Tax Cuts and Jobs Act does not mention the health care program, but critics say cancer patients would be hit hard if the bill passes.

Fact Check › Science

Massive Indian Ocean Earthquake and Tsunami Expected by the End of 2017?

27 September 2017 - An Indian man's prediction is based on his purported psychic abilities and carries no scientific weight.

Fact Check › Fake News

Did a CIA Agent Confess to Killing Bob Marley?

4 December 2017 - A disreputable web site capitalized on an old conspiracy theory about the death of the reggae icon to spin fake news clickbait.

Eg, Real News and Commentary

Real News

Commentary

The New York Times

EUROPE

Suspecting Russian Meddling in 'Brexit' Vote, Lawmaker

By STEPHEN CASTLE OCT. 19, 2017

Arron Banks was a major supporter of Leave.EU, the unofficial campaign for Britain's withdrawal from the European Union. Jack Taylor/Getty Images

LONDON — An opposition lawmaker called on the British government on Thursday to investigate suspicions of Russian interference in Britain's referendum on quitting the European Union, adding that there were "questions" about Arron Banks, one of the chief financial backers of the leave campaign.

Speaking in Parliament, the lawmaker, Ben Bradshaw of the Labour Party, said there was "widespread concern over foreign, and particularly Russian, interference in Western democracies." He asked for assurances that "all the resources spent in the referendum campaign were from permissible

Reports from "official sources"

Adheres to AP style

Lexical and Syntactic structure: Use of past tense reporting

•

The Washington Post

Opinions

He who must be named

President Trump in the Oval Office on Oct. 24. (Jabin Botsford/The Washington Post)

By Ronald A. Klain October 24 at 2:28 PM

For decades, conservatives labored to make their movement more humane. Ronald Reagan put a jovi on conservative policies — more Dale Carnegie than Ayn Rand; George H.W. Bush promised a "kinde gentler" tenure; George W. Bush ran on "compassionate conservatism." As a progressive, I opposed m their specific proposals. But as a country, we benefited from a debate between competing ideas — libe conservative — on how to soften the hard edges of American life and create a more inclusive country.

That was then. Today, we are living the Politics of Mean. In the Trump presidency, with its daily acts cruelty, punching down is a feature, not a bug. And the only thing more disquieting than a president w practices the Politics of Mean are the voters who celebrate it.

President Trump's decision to put some 700,000 "dreamers" in limbo, his snarling statement that athletes protesting policing abuses are "sons of bitches," his opposition to bipartisan efforts to maintain health

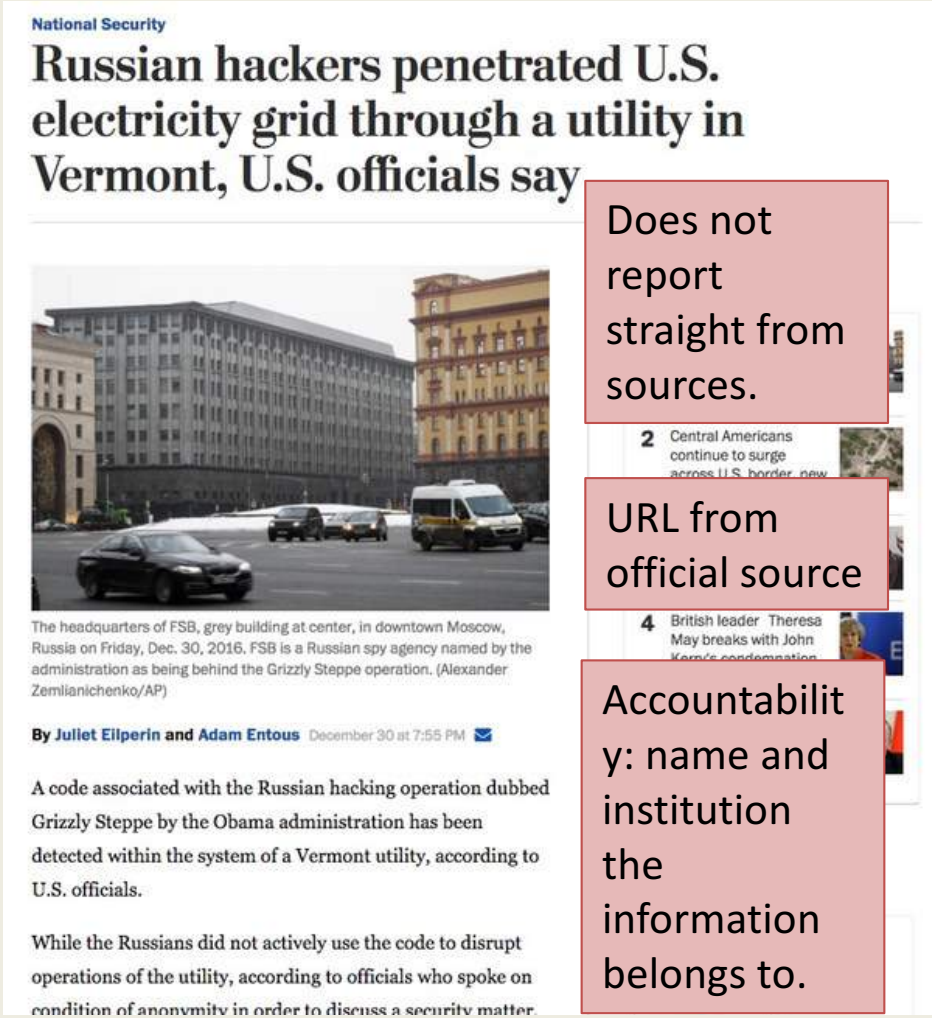
Labeled as "opinion"

Lexical use: "should"

Emotionally charged linguistic markers

Eg, Unintentional Misreporting

- Articles written by journalists and reputable sources, but misreported
- Example: In late 2016, the *Washington Post* ran a fake story about Russian hackers penetrating the electric grid through a utility company
  - The utility company immediately issued a statement saying a malware was found in one laptop, and this was not connected to the rest of the organization



National Security

Russian hackers penetrated U.S. electricity grid through a utility in Vermont, U.S. officials say

The headquarters of FSB, grey building at center, in downtown Moscow, Russia on Friday, Dec. 30, 2016. FSB is a Russian spy agency named by the administration as being behind the Grizzly Steppe operation. (Alexander Zemlianichenko/AP)

By Juliet Eilperin and Adam Entous December 30 at 7:55 PM

A code associated with the Russian hacking operation dubbed Grizzly Steppe by the Obama administration has been detected within the system of a Vermont utility, according to U.S. officials.

While the Russians did not actively use the code to disrupt operations of the utility, according to officials who spoke on condition of anonymity in order to discuss a security matter

Does not report straight from sources.

URL from official source

Accountability: name and institution the information belongs to.

Eg, Satire and Citizen Journalism

Satire

Politics

Trump Comforts Grieving War Widow By Assuring Her He Will Never Die

Yesterday 1:44pm • SEE MORE: POLITICS

WASHINGTON—In an effort to alleviate her sorrow after the tragic event, President Donald Trump reportedly comforted the grieving widow of a fallen soldier Thursday by assuring her that he will never die. "I know things are very difficult for you right now, but I hope you can take some solace in the knowledge that I am going to live on forever," said Trump in a phone call to Deborah Stilling, attempting to sooth the woman who lost her spouse in Afghanistan with promises that he would outlast her, her children, and her children's children. "Yes, your husband is gone, but remember that the commander-in-chief he served will always remain, and no one can ever take that away from you. Dry your eyes; I'm still here." At press time, Trump, who was reportedly unsuccessful in lifting the widow's spirits, offered repeated assurances that he would thrive long after the sun had expanded into a red giant and swallowed the Earth.

- About us section
- Clickbait headline
- Amateur mistakes
- Not fact-checked

Citizen Journalism

GROUND REPORT
world news and opinions from the ground

ABOUT US

TELL YOUR STORY ON GROUNDREPORT; LEARN THE TRADE; GET NOTICED BY MEDIA OUTLETS

TIME FOR ACCOUNTABILITY WITH THE IR… REGIME, ACCORDING TO U.S. REP TED …

jul 19, 2017 | politics, world

For those Iranians who lived through the massacre of 1988, the lack of accountabil… regime is just another indignity in the story of their loved ones. Justice has been de… victims, who professed an ideology different the mullahs. They paid for this ideolog… lives.

"Thirty years ago,…the government, Khomeini, murdered 30,000 people that year because they were political dissidents. They wanted a regime change. Thirty thousand. And here we are almost 30 years later. It's time to have a day of reckoning for those people that were murdered. And it is the responsibility of you young people over here to avenge the murder of those people, your families, 30 years ago," said U.S. Representative Ted Poe during the "Free Iran" rally held in Paris earlier this month.

The initial order for the massacre was based on a fatwa by the Supreme Leader, wh… execution of all prisoners affiliated with the main opposition group People's Mojah… Organization of Iran (PMOI or MEK). While the government has continued to deny t… evidence gathered by International NGOs and the National Council of Resistance of… points to a different reality than the one that the regime would claim.

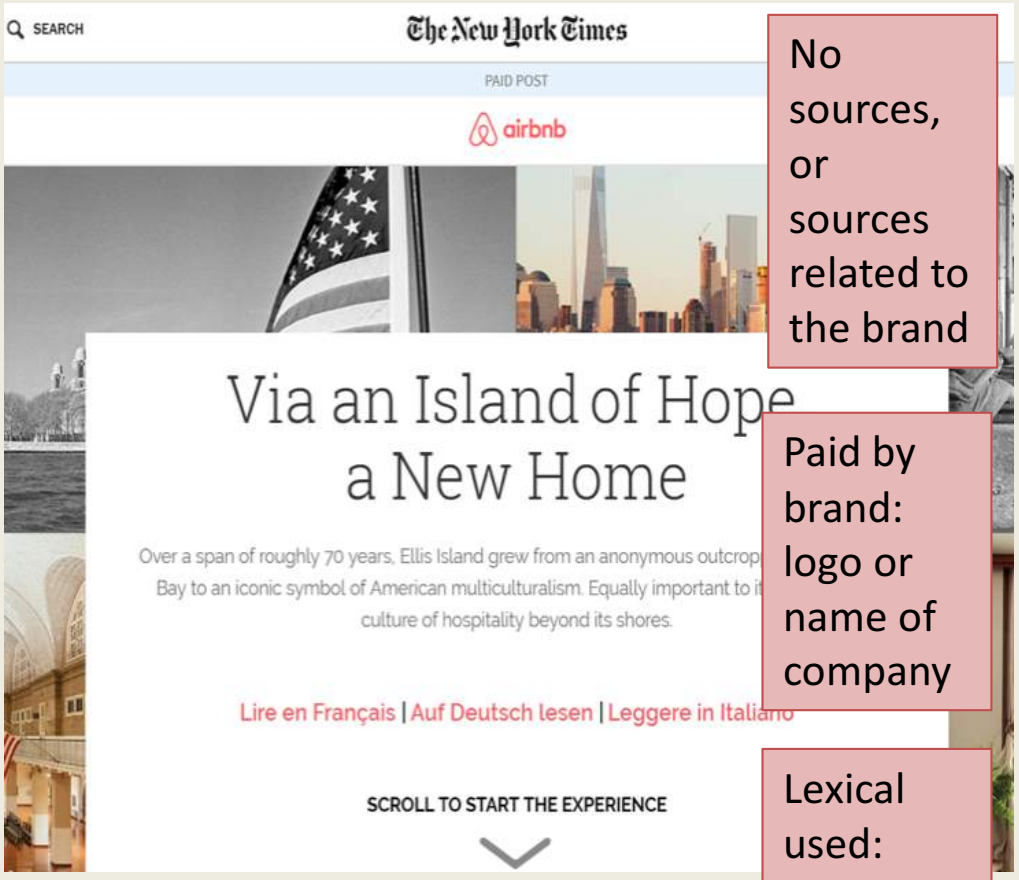
In addition, a tape was released where members of the "Death Commission", charged with carrying out Khomeini's order, was recorded discussing the executions. One of the members of this group, Raisi, ran for the office of President in Iran's last election with Khamenei's approval and backing.

- URL and about us
- Might not follow journalistic practices
- Can be emotionally charged

Eg, Native Ad. & Professional Political Content

Native Ad.

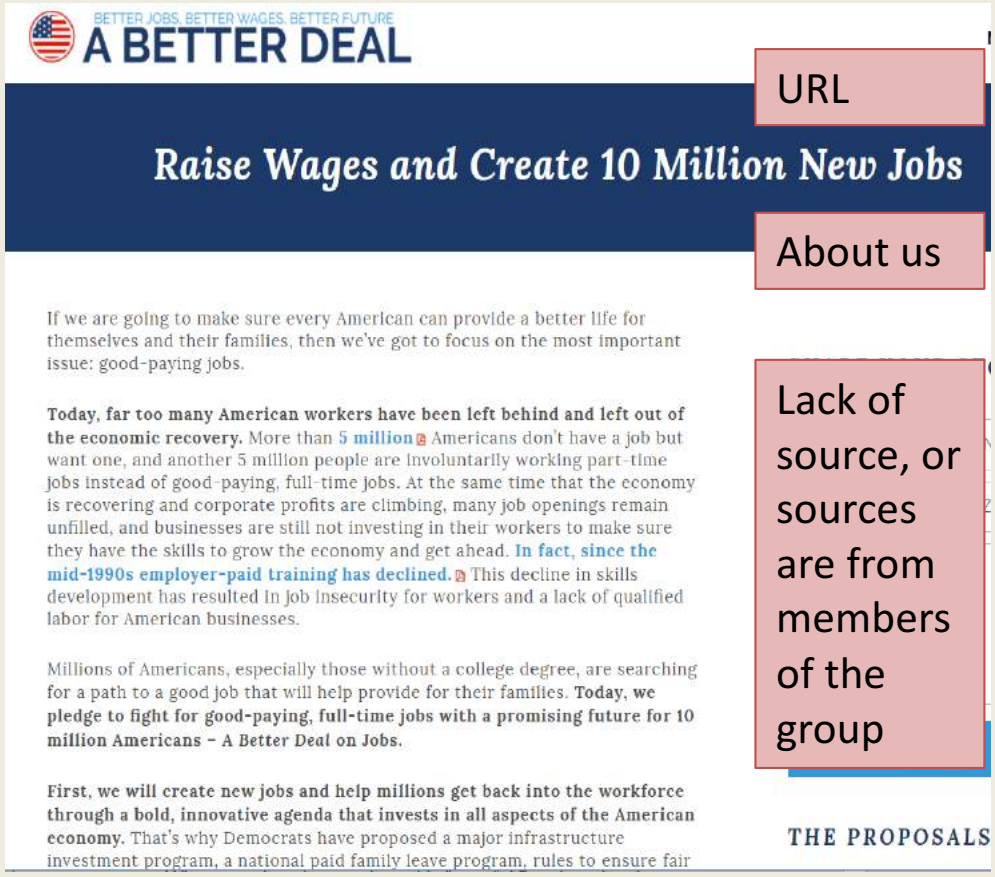
Prof. Political Continent



No sources, or sources related to the brand

Paid by brand: logo or name of company

Lexical used: present tense



URL

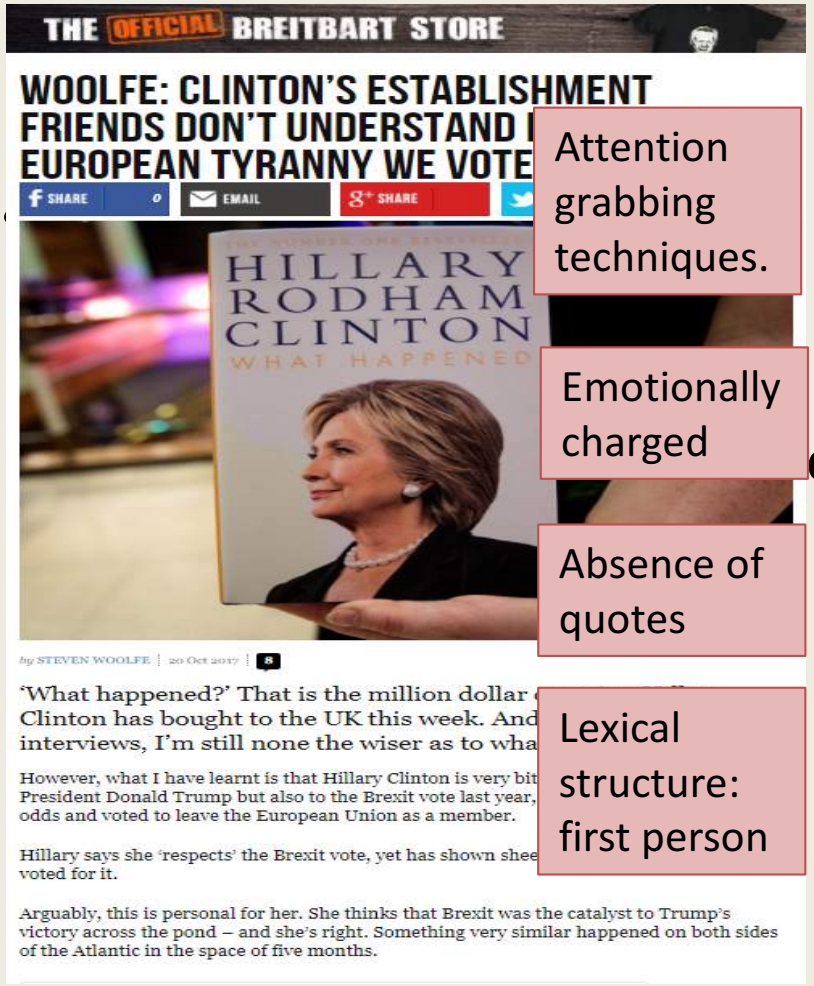
About us

Lack of source, or sources are from members of the group

Eg, Polarizing/Sensational Contents & Fake Sites

Polarizing/Sensational Contents

Fake Sites



Attention grabbing techniques.

Emotionally charged

Absence of quotes

Lexical structure: first person

Emotional linguistic markers, and attention grabbing techniques

Ephemeral

Structural inconsistencies: URL registration

Episode

FOX BUSINESS

MARKETS POLITICS FEATURES FBN TV

MEDIA & A

Professo

for digita

Published August 31, 2017 | Features

Home R

Faculty members ea

TIONS HOME SEARCH The New York Ti

For One Family in Houston, an Overwhelming Start...

See Har Helicopt

The 2017 Volkswagen Golf G

U.S.

Professors Get $300,000 Grant for Digital Fake-News Detector

Salon
@Salon

Seguir

Penn State professors get $300,000 to create fake-news detector ift.tt/2etcod0

18:19 - 31 ago. 2017

upgrUv

CRACKING DOWN

Fake news? Not if 2 PSU profs can help it

90.5 WESA Pittsburgh's NPR News Station

Listen Live · WESA
Only A Game

Politics Science Identity/Justice Education News Series Podcasts Schedule Support Search

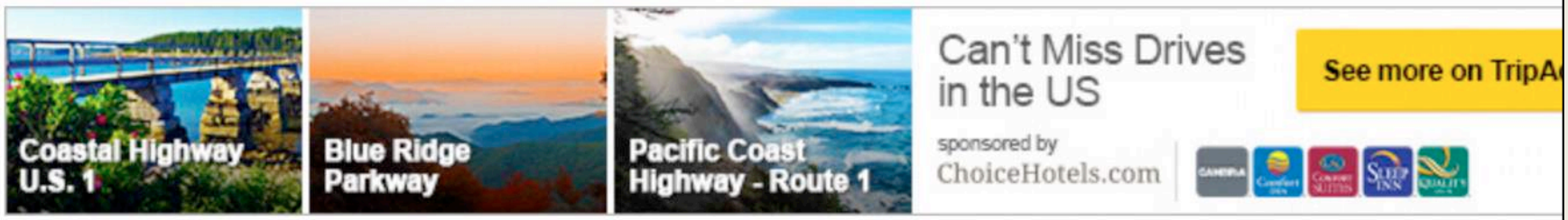
Penn State Professors Want To Build Device That Recognizes Fake News

By ASSOCIATED PRESS • AUG 31, 2017

A pair of professors from Penn State University received a $300,000 grant in an effort to help produce technology — specifically "devices" — that would be able to efficiently detect fake news, according to the Associated Press.

Ep

ADVERTISEMENT



Outgoing Miss America: We're more alike than different

ARKANSAS HIGH SCHOOL FOOTBALL: Full scoreboard + recaps, stats, photos from dozens of games

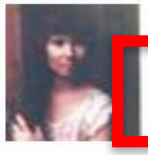
Arkansas Attorney General pledges to end program protecting young people from deportation

The university says information sciences and technology professor Dongwon Lee and communications professor S. Shyam Sundar are working on the project.

Lee says fake news "has been around for decades" but has been "exacerbated" on the Internet and social media platforms.

The professors plan to investigate "characteristic indicators of fake news" and develop complex formulas that will enable digital devices to recognize those indicators and purge stories that contain them.

Episode



Teri Marcum

Lol, a bullshit detector.
thing that keeps our ear
skulls...that's right, it's ca
programming, and regul
very well. Seems to be a
Breitbart and InfoWars.

Like · Reply · 5 · Aug 31,

Joan Hodges

They could start by look
that no fraternity brother

Like · Reply · Aug 31, 2017

3WorldState1 says... August 31, 2017 at 3:17 p.m.

These types of solutions will certainly help online, but I wish we could go back to a time where our FCC wouldn't allow just completely bogus and outright lies to be spewed on our airwaves. This would be pretty easy to enforce. I watch Fox" quite often. And it's their Opinion shows that lie tru their teeth constantly, but the news is just right on the line, many times over the line or constantly insinuating. It can be done. You lie, you pay.

(permalink | suggest removal)

rubythecat96gmailcom says... August 31, 2017 at 5:35 p.m.

3 world when someone watches an opinion show. its just that the person's opinion. I wish they would remove Monday night wrestling cause we know that's not real right ? and how about the "Real" Housewives of Orange County. New York, etc. they need to be removed also. don't you think cause they're not really real same with HGTV I see the same actors pretending to be fake house buyers on other shows lets get them off the air also.

(permalink | suggest removal)

Popsmith says... August 31, 2017 at 5:54 p.m.

Don't forget "Finding Sasquatch" and ghost hunters. They need an "F" rating.

(permalink | suggest removal)

JOHNNYSPINKS says... August 31, 2017 at 8:51 p.m.

Just use the keyword "Trump" "Greatest" "Best"

(permalink | suggest removal)

ARMNAR says... August 31, 2017 at 10:11 p.m.

Just turn on FOX and send me the 300K.

(permalink | suggest removal)

monfnder says... August 31, 2017 at 11:34 p.m.

Another step to regulate free speech and shape the narrative/brainwash the public

Hassan & Lee

Professors Get $300,000 Grant for Digital Fake-News Detector

REPLY REPLY ALL FORWARD

Mark as unread

[redacted]

Sat 9/2/2017 10:40 AM

NSF Assistant Directors

Show all 16 recipients

To: fcordova@nsf.gov; ⇦ NSF Director

Cc: dsellers@nsf.gov; jferrini@nsf.gov; bstone@nsf.gov; fkorsmo@nsf.gov; weward@nsf.gov; jhamos@nsf.gov; vross@nsf.gov; ...

https://www.nytimes.com/aponline/2017/08/31/us/ap-us-odd-fake-news-detector.html

As an American citizen that has supported the NSF, I am surprised that so much money ($300K) is being used to create algorithms for "fake news" "detection".
$300K can be used to support real projects that have a measurable benefit to society such as other types of robotics or medicine.
Chasing unicorns and rabbits called "fake news" is a fantasy that will NEVER change human behavior. Biased media has been purging news stories since the printing press was invented. Psychological biases are how we all tolerate endless streams of input data. Since most forms of media are driven by advertising dollars and donations, truth is less important than the self-sustaining dollars that perpetuate the host media.
When Trump speaks of "fake news", that is nothing more than hostile verbiage designed to discredit his political enemies, of which he has many.
I can author algorithms that sort news stories stored in a complex matrix library containing any number of subjective parameters, tags, flags and origins. Those libraries are created

Episode

NSF FOIA Request [redacted] Proposal #1742702

[redacted]@nsf.gov>

Thu 9/7/2017 6:15 PM

[redacted]

[redacted]

3 FOIA requests since Sep. 2017

[redacted]
September 7, 2017

Dear Tracy Ray (Authorized Organizational Representative),

We have received a Freedom of Information Act (FOIA) request for a copy of National Science Foundation (NSF) funded grant proposal entitled "Training Computers and Humans to Detect Misinformation by Combining Computational and Theoretical Analysis" (Proposal #1742702). The Principal Investigator, Dr. Dongwon Lee, has been copied on this email.

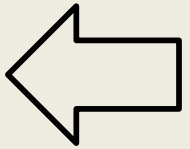
The requester is [redacted]
Records are available to the public on request except for material that is personal, privileged or confidential, or otherwise exempt from disclosure under law. NSF will remove, before disclosure to the requester, personal information (SSN, date/place of birth, individual salaries, bios, pending and non-Federal grants) in the file about yourself or other individuals under Exemption 6 of the FOIA to protect personal privacy.
The FOIA generally prohibits us from withholding records except to the extent that disclosure of proprietary information (trade secrets or commercial or financial information) in the record, if any, that you assert is privileged or confidential. Therefore, I request that you carefully review the proposal, identifying by page number, line or paragraph, the parts you consider confidential. You will need to include a detailed statement explaining how disclosure of this information would harm your organization or benefit your competitors, or why it is otherwise exempt from disclosure under FOIA. Understand that information you submit in response to this request may itself be subject to disclosure. Please do **not** send us a copy of the proposal unless you have significant redactions. **When sending us a proposal with redactions, be sure to encrypt the message or apply the redactions in order to safeguard the information**.
Your response is not an agreement between us that the information you request us to withhold will not be

Important Questions

- What is Fake News?
  - Multi-faceted → concept explication
  - Focus on most impactful fake news facet
    - Biased/opinionated stories are ok
- How to know if solution is good enough?
  - Trust issue
  - Accuracy threshold → Smarter than human
- Virality as key factor
  - Traditional media research looks for intention and credibility
    - No longer simple in internet era
  - How to make people think twice before re-posting fake news?
  - How to label it?

Important Computational Questions

- How does fake news spread?
  - Understand the propagation mechanism of fake news
- Who are the users susceptible to fake news?
  - Who are creators or conveyors of fake news?
- Detect fake news among real news
  - Detect it early



Focus

Related Detection Problems

- Fake News Detection

- Clickbait Detection
- Rumor Detection
- Truth Discovery
- Social Bot Detection
- Social Spammer Detection

Fake News Detection: Human-Oriented

- White/blacklist-Based
  - BS detector: http://bsdetector.tech/
  - Rubtr: http://rbutr.com/
  - ZenMate SafeSearch: https://zenmate.com/products/safesearch/
  - FiB: https://devpost.com/software/fib
  - Media Bias/Fact Check: https://mediabiasfactcheck.com/

- Manual fact-checking

FACTCHECK.ORG
ANNENBERG POLITICAL FACT CHECK

St Petersburg Times
PolitiFact.com
2009 Pulitzer Prize winner

OpenSecrets.org
Center for Responsive Politics

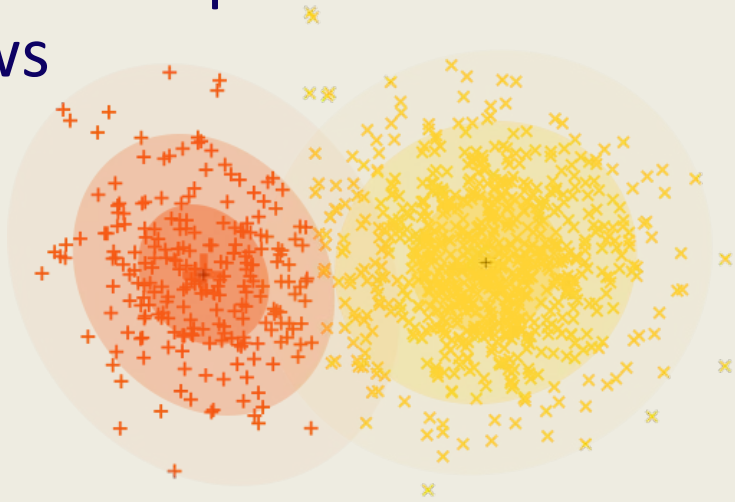
Snopes.com
Rumor Has It

TruthOrFiction.com

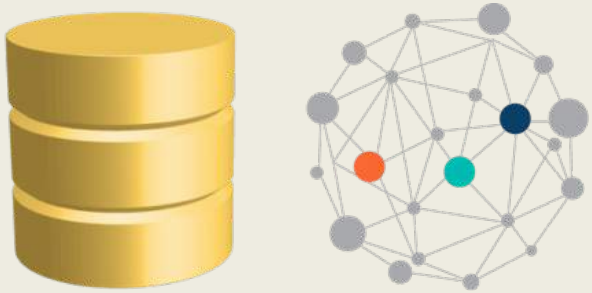
Hoax-Slayer

Computational: AI vs. DB Approaches

- AI Approach
  - Underlying premise: the existence of latent patterns to differentiate fake news from real news
  - Costly to develop labeled samples
  - Can handle fresh fake news better



- DB Approach
  - Underlying premise: if query (= news) to DB/KB of facts yields null, it is fake news
    - Closed-world assumption (CWA)
  - Costly to build/maintain DB/KB
  - May not handle fresh fake news
  - May explain result better



Online Quiz: Fake-O-Meter

- In your browser, go to “Kahoot.it”
- Enter Game PIN, shown on screen
- Enter your nickname (to be shown on screen)

Part 2. AI Approaches

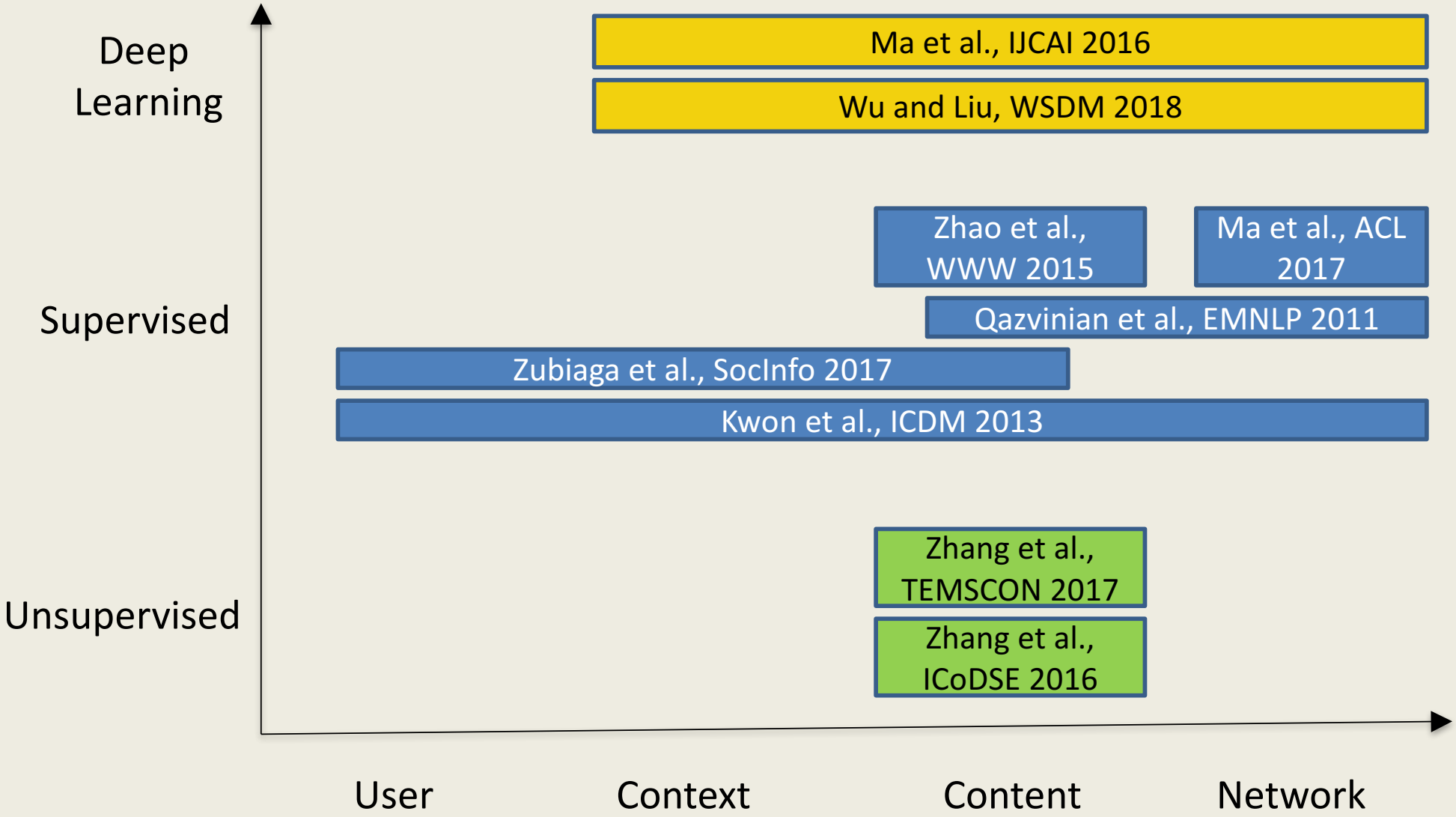
Outline

- Landscape
- Features
- Representative Works
- Post-Detection
- Resources

Landscape of Machine Learning Approaches

- Unsupervised Learning
  - Fake news ≃ outlier
  - Plenty of unlabeled dataset
  - Lower detection accuracy
- Supervised Learning
  - Scarce and unbalanced labeled dataset
  - Hand-crafted feature engineering
- Deep Learning
  - Difficult to explain
  - Post-detection issue

Landscape of AI Approaches



Comprehensive Survey and Other Tutorial

- Shu et al., *Fake News Detection on Social Media: A Data Mining Perspective*, ACM KDD Exploration 19(1), 2017
- Zubiaga et al., *Detection and Resolution of Rumours in Social Media: A Survey*, ACM Computing Survey, in press
- Wu et al., *Mining Misinformation in Social Media: Understanding Its Rampant Spread, Harm, and Intervention*, ICDM tutorial, 2017
  - http://www.public.asu.edu/~liangwu1/ICDM17MisinformationTutorial.html
- Leblay et al., *Computational fact-checking: problems, state of the art, and perspectives*, WWW tutorial, 2018 (coming)

Supervised Approaches

Training

- Supervised Machine Learning
  - **P**: Features from "fake" news
  - **N**: Features from "legit" news
- Feed (**P**, **N**) to ML to build a model ***M***

Testing

- Feed a news story *A* to ***M***
- ***M*** determines if *A* is fake or legit news story

Features: Fake News Example

JANUARY 24, 2018 https://www.facebook.com/thelastlineofdefense/

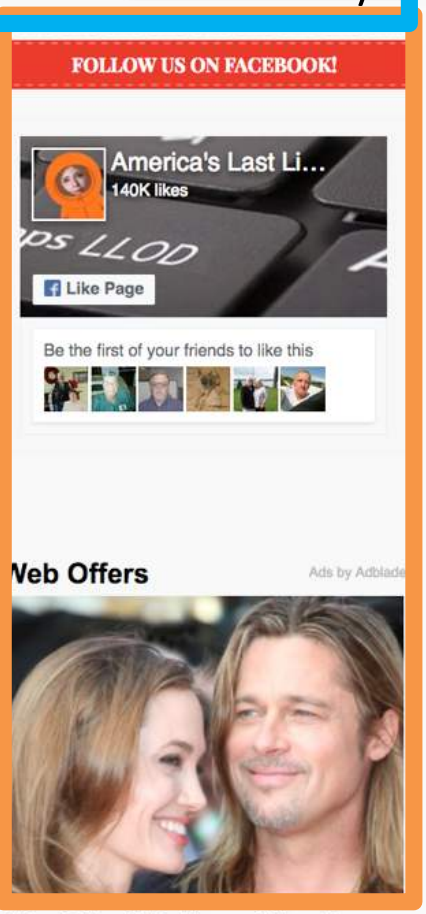
HOME › LOOSE MEAT HOAGIES › Latest Dead Body Tied To The Clintons Turns Up Valuable Evidence

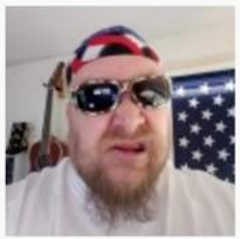
Latest Dead Body Tied To The Clintons Turns Up Valuable Evidence

December 20, 2017 Flagg Eagleton Loose Meat Hoagies, Potater Special

Another body found in a shallow grave in Arkansas has been tied to the Clintons. Marcia Ann Bosley, who was the head housekeeper at the Governor's Mansion before she was mysteriously let go in 1985, has been positively identified after spending more than 3 decades in the ground.

Bosley, who had no family when she was let go, was assumed to have left the area for other opportunities. She was never reported missing. According to the Hamilton County Coroner:





About Flagg Eagleton 36 Articles

Flagg Eagleton is the son of an American potato farmer and a patriot. After spending 4 years in the Navy and 7 on welfare picking himself up by the bootstraps, Flagg finally got his HVAC certificate and is hard at work keeping the mobile homes of Tallahassee at a comfy 83 degrees.

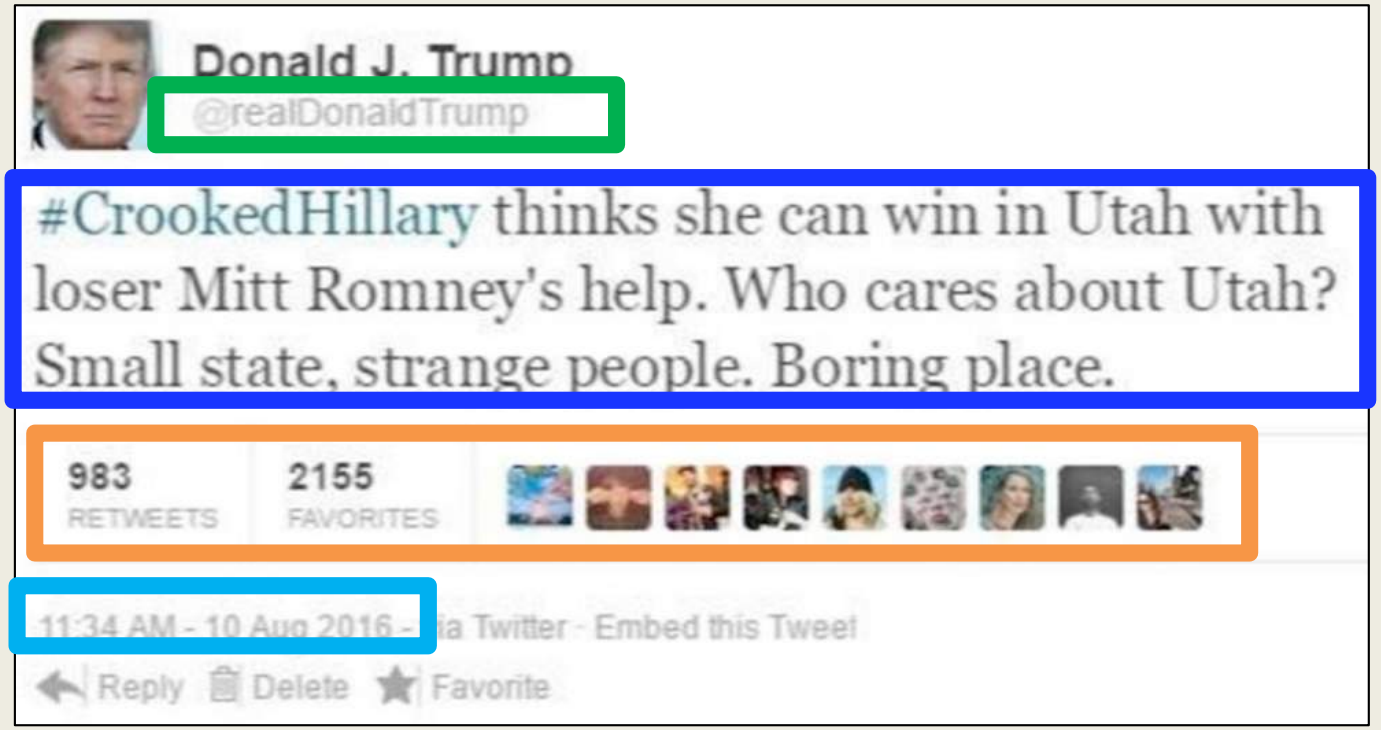
Content: Text

Content: Image

Content: Ad

Context

Features: Fake News Example



Content

User

Network

Context

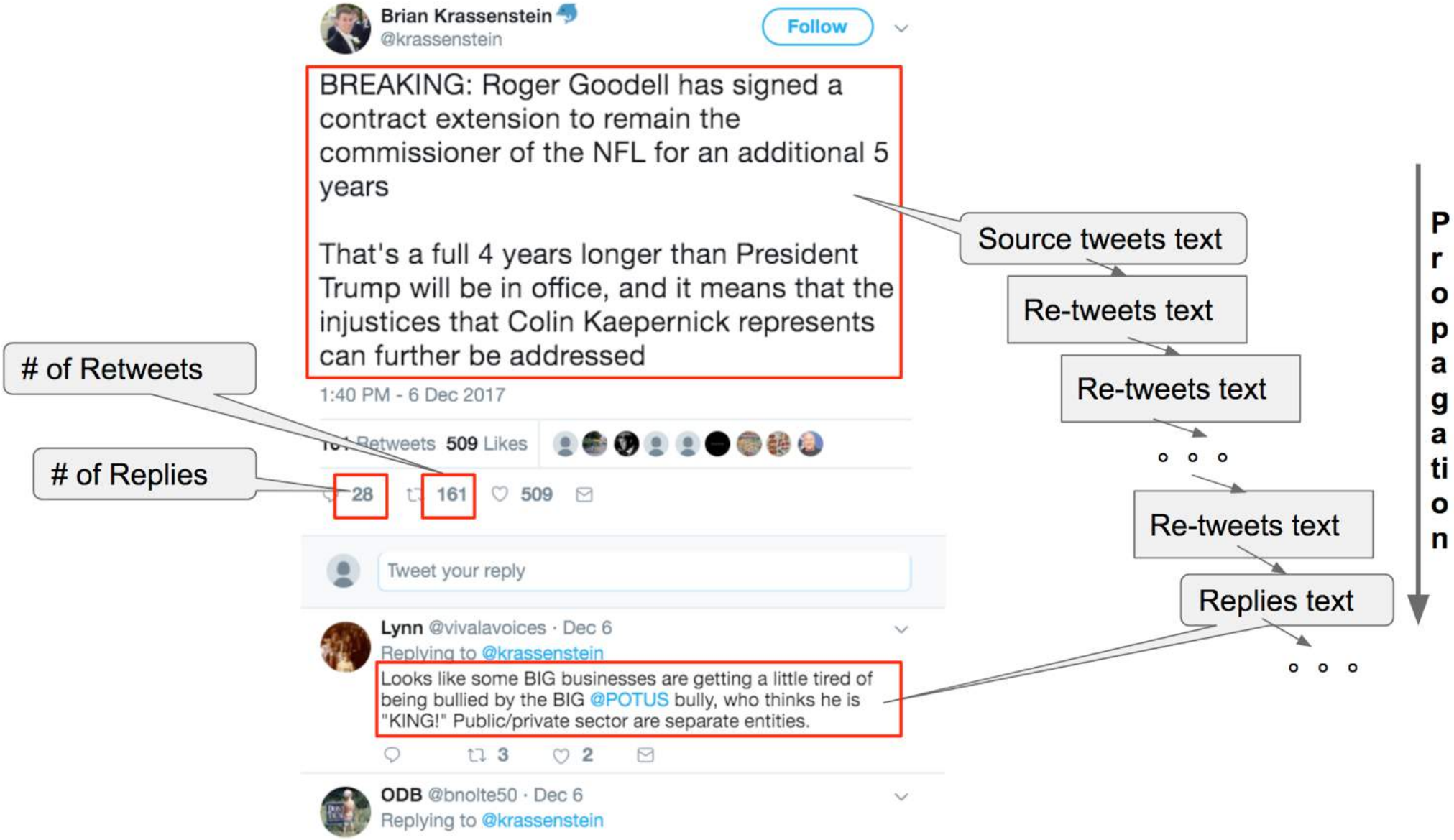
Features: Content

	Fake News	Real News
Lexical	"you", "should"	"today"
Syntactic	Present tense	Past tense
Message quality	• Spelling mistakes • Not in AP style • No last names when citing	• Goes through proof • Reading and editing process • Adheres to AP style
Pedigree	Originated in an obscure site or social media post	Originated from a well-known site
Quotes	• Unverified sources • Absence of quotes	• Verified quotes (real people and quotes) • Always has quotes, if not it is a commentary
Photos	• Altered pixel structure • Picture is out of context with the event	• Normal pixel structure • Pictures will correspond to the event at hand

Features: Network Node

Originality = original tweets vs. retweeted

Features: Network Path



[Qazvinian et al., EMNLP 2011]

- Two RQs:
  - Q1: Detect rumor tweets
  - Q2: Detect rumor spreaders

(non-rumor) "As Obama bows to Muslim leaders Americans are less safe not only at home but also overseas. Note: The terror alert in Europe... "

(rumor) "RT @johnnyA99 Ann Coulter Tells Larry King Why People Think Obama Is A Muslim http://bit.ly/9rs6pa #Hussein via @NewsBusters #tcot .."

(confirm) "RT @moronwatch: Obama's a Muslim. Or if he's not, he sure looks like one #whyimvotingre-publican." ⇦ Spreader

(deny) "Barack Obama is a Christian man who had a Christian wedding with 2 kids baptised in Jesus name. Tea Party clowns call that muslim #p2 #gop"

⇖ Refuter

[Qazvinian et al., EMNLP 2011]

- Dataset
  - 10K tweets manually annotated by two judges
    - 3.6K non-rumor, 6.8K rumor (2.9K spreader)
  - Kappa = 0.954 for Q1 / 0.853 for Q2

Q1

(0) "McCain Divorces Palin over her 'untruths and out right lies' in the book written for her. McCain's team says Palin is a petty liar and phony"

(1) "Sarah and Todd Palin to divorce, according to local Alaska paper. http://ow.ly/iNxF"

⇦ Rumor Tweet

Q2

(12) "Sarah Palin Divorce Rumor Debunked on Facebook http://ff.im/62Evd"

⇦ Refuter

(11) "Todd and Sarah Palin to divorce http://bit.ly/15StNc"

⇦ Spreader

[Qazvinian et al., EMNLP 2011]

- Features
  - Content: lexical, part-of-speech patterns, URLs, hashtags
  - Network: *x* retweets a message *t* from *y*
    - *t* is more likely to be a rumor if: (1) *y* has a history of posting rumors, or (2) *x* has posted/re-tweeted rumors in the past
- Learning
  - Learn +/- Bayes classifiers
  - Q1: F1=0.897
  - Q2: F1=0.925

$$\frac{P(\theta_i^+|t)}{P(\theta_i^-|t)} = \frac{P(\theta_i^+)}{P(\theta_i^-)} \frac{P(t|\theta_i^+)}{P(t|\theta_i^-)}$$

[Zubiaga et al., SocInfo 2017]

- RQ: Detect rumor tweets
- Dataset
  - For an event *e*, collect subsequent re-tweets *e*
  - 5.8K manually annotated by journalists (1.9K rumor)
    - Tweet is rumor if unconfirmed "at the time of posting"
- Features
  - Content: word2vec, part-of-speech tag, symbols (?,!,.)
  - Social: # of tweets, # of follows, age
  - Context: how others reacted to tweet
    - Aggregating rumourous and non-rumourous posts preceding the tweet being classifed will improve detection

[Zubiaga et al., SocInfo 2017]

- Context
  - How others reacted to tweet in a thread
  - Capture the thread as a linear chain
  - Use Conditional Random Field (CRF)
- Result

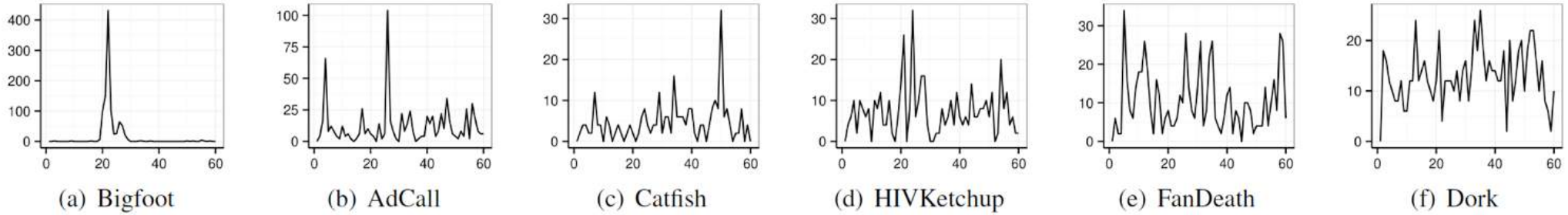
	Content			Social			Cont + Social		
Classifier	P	R	F1	P	R	F1	P	R	F1
SVM	0.355	0.445	0.395	0.337	**0.524**	**0.410**	0.337	0.483	0.397
Random Forest	0.271	0.087	0.131	0.343	0.433	0.382	0.275	0.099	0.145
Naive Bayes	0.309	**0.723**	0.433	0.294	0.010	0.020	0.310	**0.723**	0.434
MaxEnt	0.329	0.425	0.371	0.336	0.476	0.394	0.338	0.442	0.383
CRF	**0.683**	0.545	**0.606**	**0.462**	0.268	0.339	**0.667**	0.556	**0.607**

[Kwon et al., ICDM 2013]

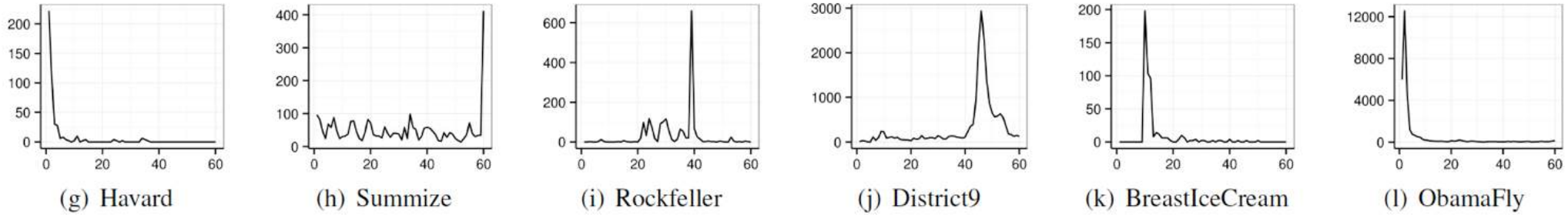
- RQ: Detect rumor tweets
- Dataset
  - 102 topics: 47 rumors and 55 non-rumors
  - Manually annotated by 4 judges (ICC=0.992)
- Features
  - Temporal
  - Structural
  - Linguistic: Linguistic Inquiry and Word Count (LIWC)

[Kwon et al., ICDM 2013]

- Temporal
  - Rumors tend to have multiple and periodic spikes
  - Non-rumors have a single prominent spike



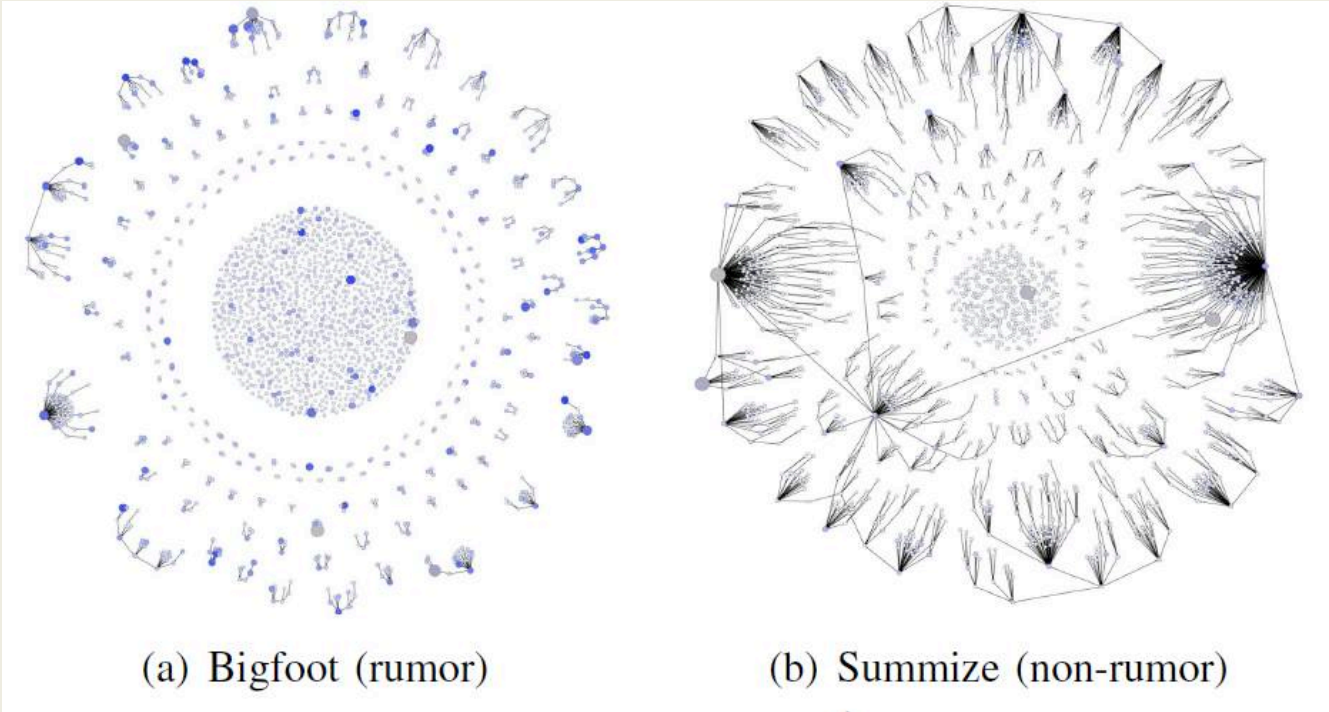
Rumor



Non-Rumor

[Kwon et al., ICDM 2013]

- Structural
  - A topic is diffused from A to B, iff: (1) B follows A on Twitter and (2) B posted about a given topic by mentioning the appropriate keywords only after A did so

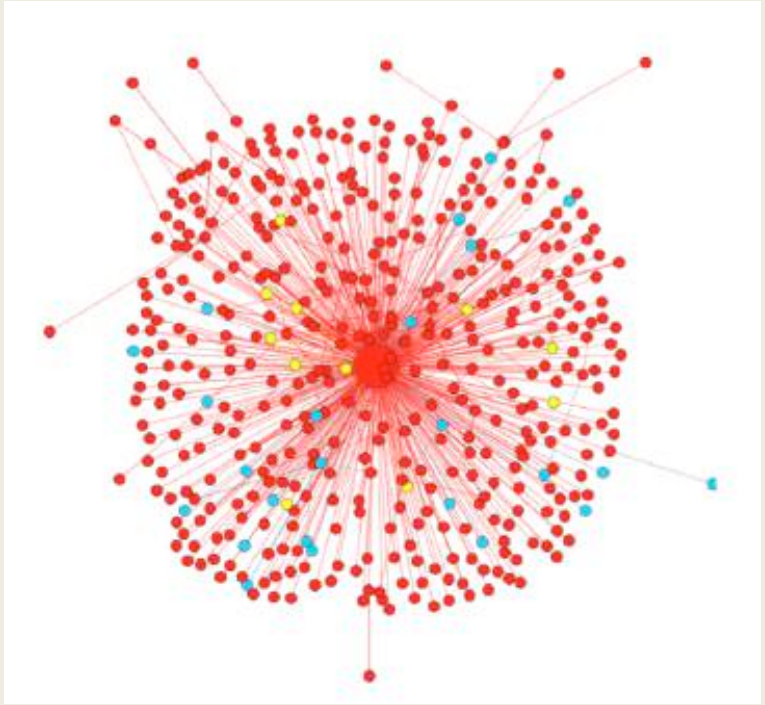


(a) Bigfoot (rumor)

(b) Summize (non-rumor)

[Zhao et al., WWW 2015]

- RQ: Detect/rank rumor tweet clusters
- Insight: When there is a rumor, a few posts raise questions → signature text phrases for skepticism
  - Eg, "Is this true?", "Really?", "What?"

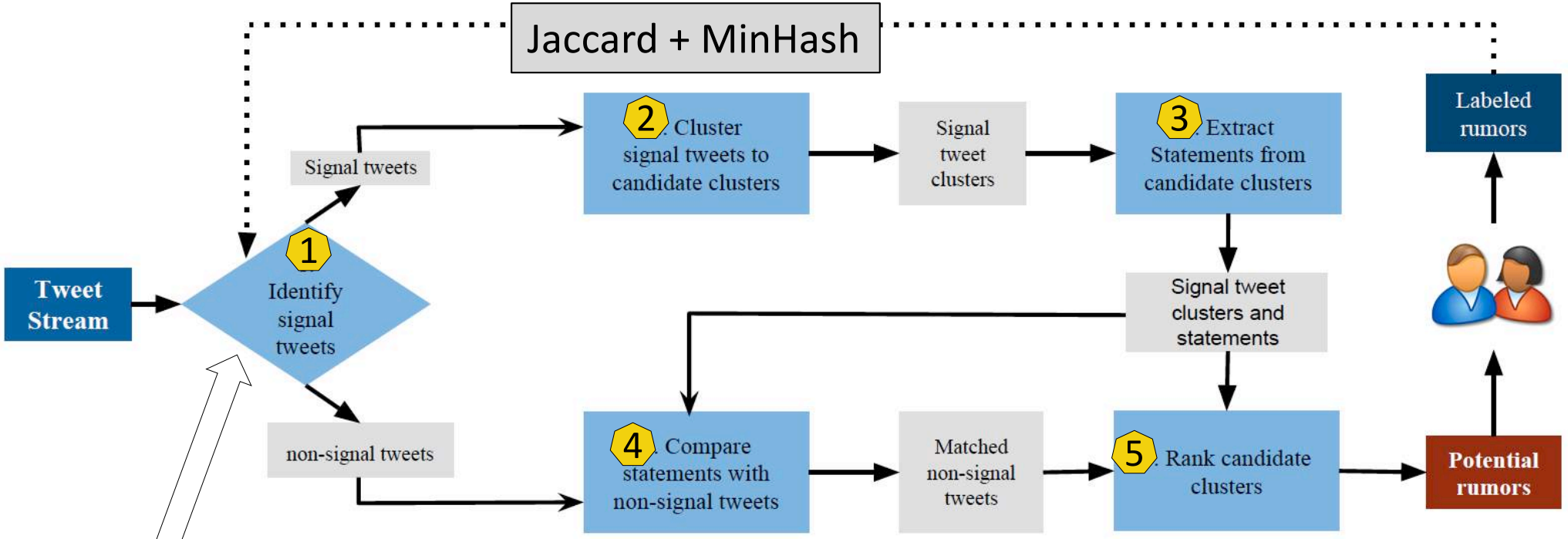


60 seconds after a hacked twitter account sent out the White House rumor, there were already sufficient enquiry tweets (blue nodes)

[Zhao et al., WWW 2015]

- Method

What??!! Sharks on highway
Really?? @AP: Sharks on highway



is (that | this | it) true
wh[a]*t[?!][?1]*
(real? | really ? | unconfirmed)
(rumor | debunk)
(that | this | it) is not true

Rank "not" by popularity
Rank using features (eg, % of signal tweets, entropy ratio, tweet length, retweets, URLs, hashtags)

[Zhao et al., WWW 2015]

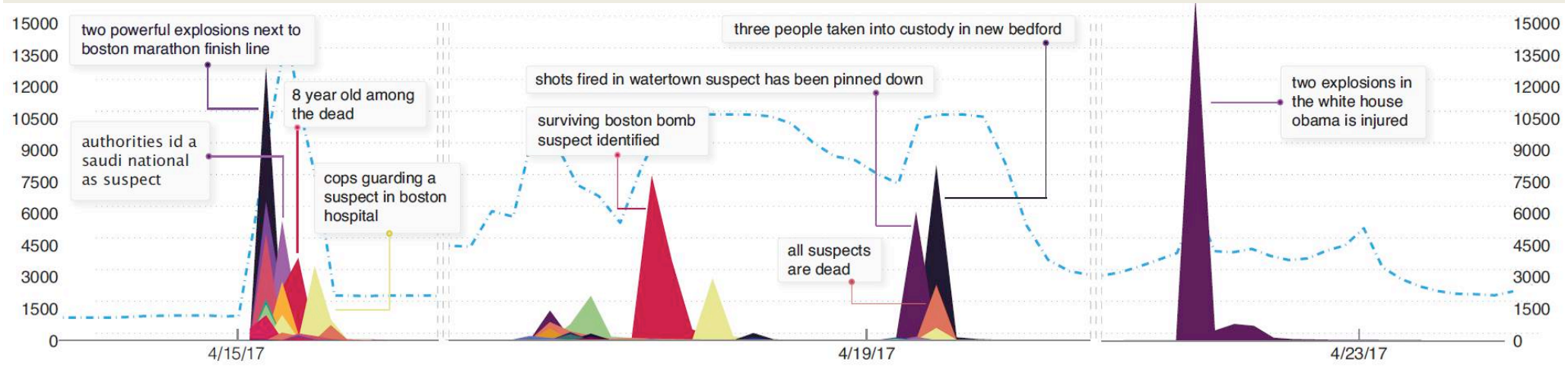
- Dataset
  - Boston Bombing: 30M tweets
  - Random Month: 1.2B tweets
  - Manually annotated by 2 judges (Kappa=0.76)
- Result

(rumor | debunk)
(that | this | it) is not true

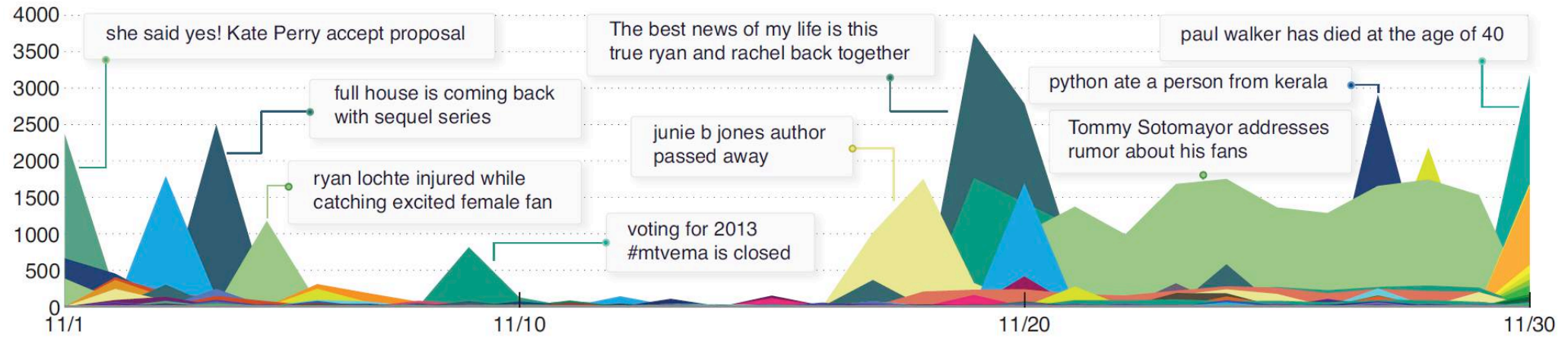
⇦

Method	Data Set	Candidates Detected	Real Rumors	Precision
Trending Topics	BOSTON	720	71	0.099
Hashtag Tracking	BOSTON	720	35	0.049
Corrections only	BOSTON	109	52	0.466
Enquiries+ Corrections	BOSTON	194	**110**	**0.521**
Corrections only	GARDENHOSE	312	87	**0.279**
Enquiries+ Corrections	GARDENHOSE	350	**92**	0.263

[Zhao et al., WWW 2015]



Rumors about Boston Bombing

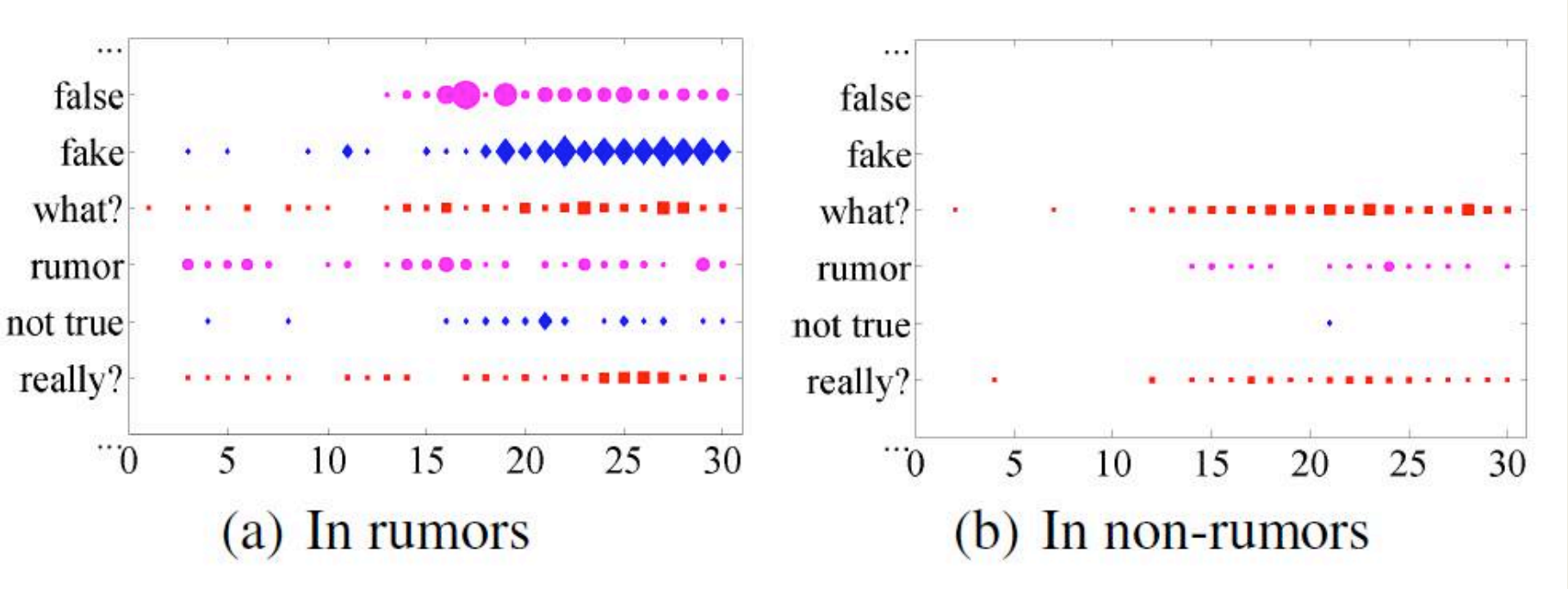


Rumors during Nov. 2013

[Ma et al., IJCAI 2016]

- RQ: Detect rumor tweet clusters (early)
- Insight:
  - Sequential nature of text streams in SN → RNN suitable
  - Connections between units in an RNN form a direct cycle and create an internal state of the network (LeCun et al., 2015)
  - Can capture the dynamic temporal signals characteristic of rumor diffusion

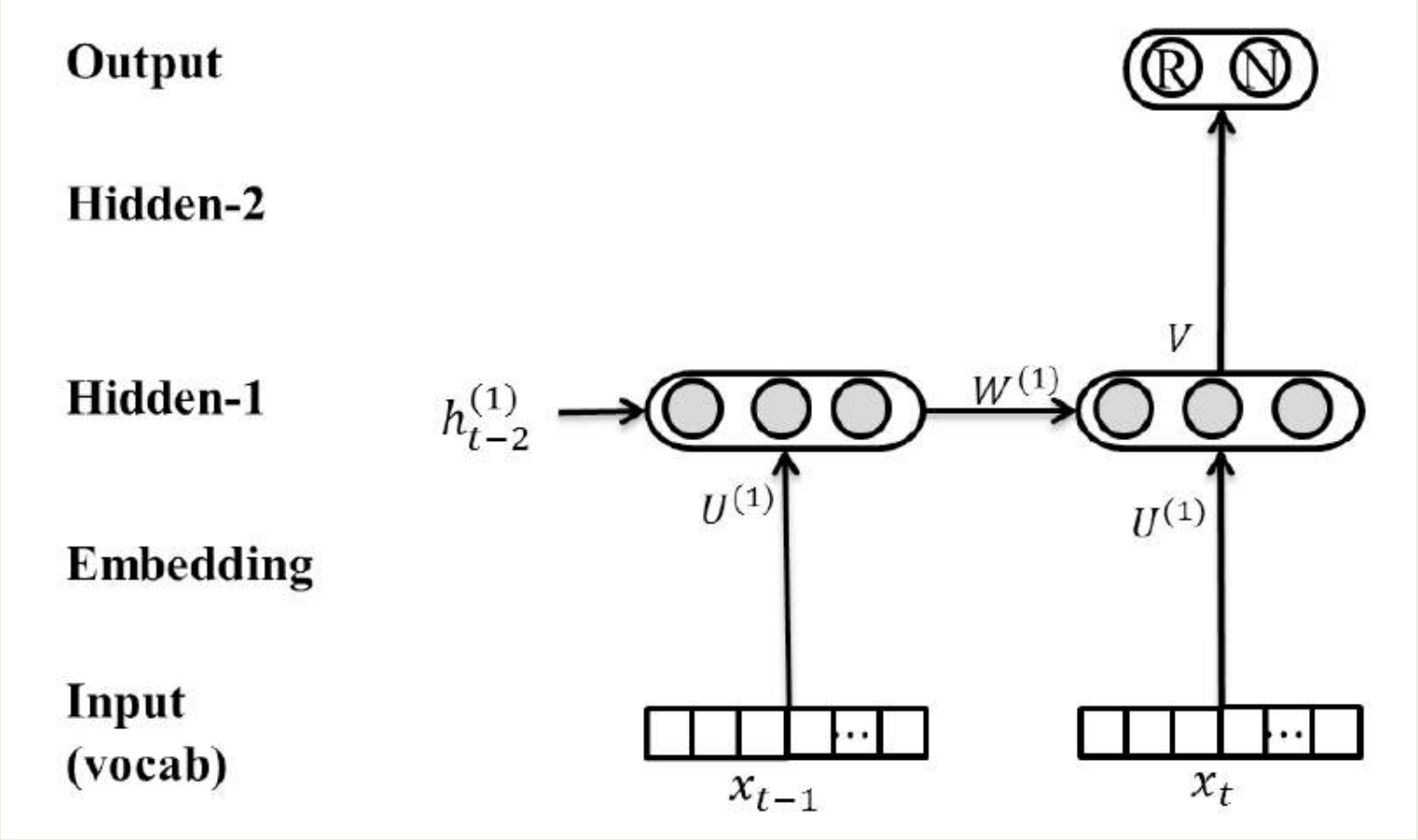
[Ma et al., IJCAI 2016]



- Continuous stream of posts/comments/replies
  - Continuous variable-length time series
- Leads to both the temporal and textual representations

[Ma et al., IJCAI 2016]

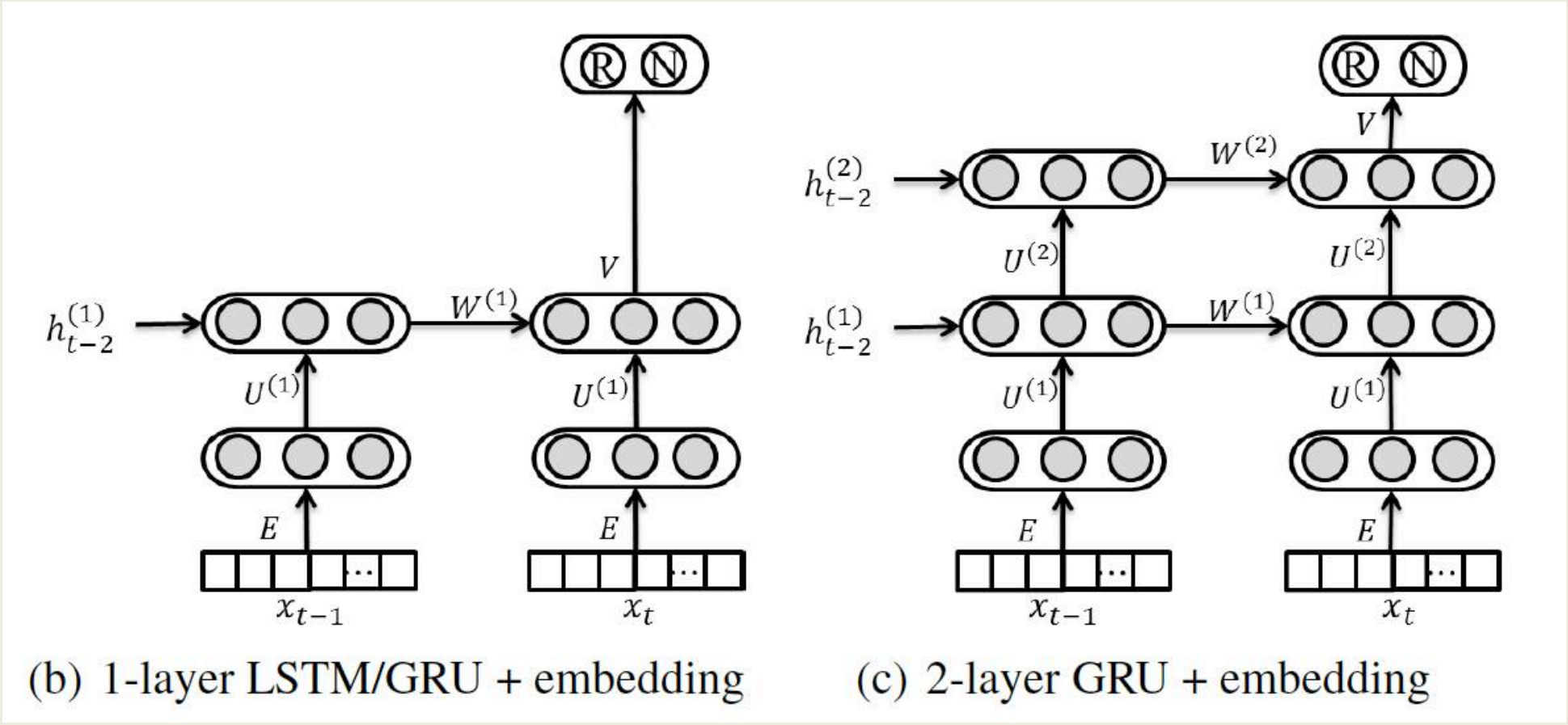
- Basic tanh-RNN



Top-K *tf-idf* term
(eg, K=5,000)

[Ma et al., IJCAI 2016]

- Improvement



(b) 1-layer LSTM/GRU + embedding (c) 2-layer GRU + embedding

[Ma et al., IJCAI 2016]

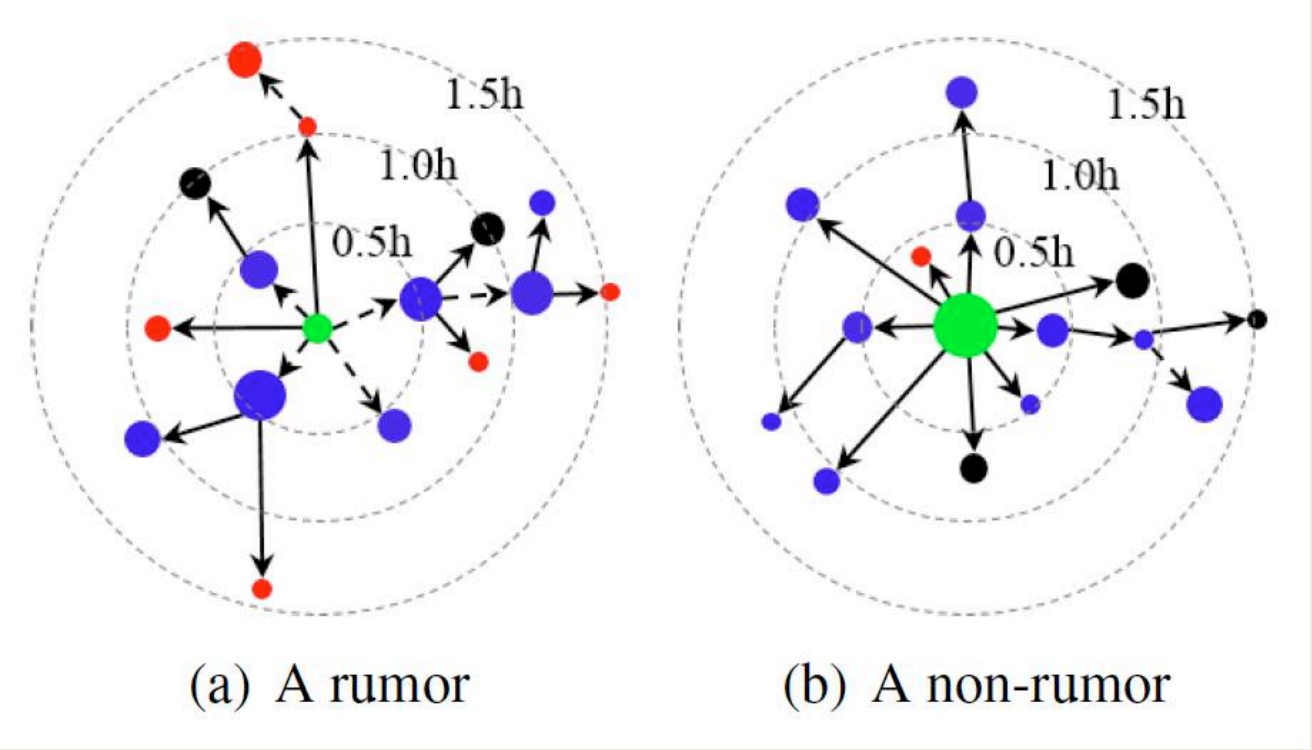
- Dataset
  - Twitter: 992 events (498 rumors, 494 non-rumors)
    - Verified via snopes.com
  - Weibo: 4,664 events (2,313 rumors, 2,351 non-rumors)
- Result

(a) Twitter dataset

Method	Class	Accuracy	Precision	Recall	F_1
DT-Rank	R	0.644	0.638	0.675	0.656
	N		0.652	0.613	0.632
SVM-RBF	R	0.722	0.856	0.526	0.651
	N		0.663	**0.914**	0.769
DTC	R	0.731	0.724	0.757	0.740
	N		0.739	0.704	0.721
RFC	R	0.772	0.717	0.908	0.801
	N		0.870	0.634	0.734
SVM-TS	R	0.808	0.735	**0.963**	0.834
	N		**0.947**	0.652	0.772
tanh-RNN	R	0.827	0.847	0.833	0.840
	N		0.804	0.820	0.812
LSTM-1	R	0.855	0.855	0.883	0.869
	N		0.854	0.820	0.837
GRU-1	R	0.864	**0.857**	0.900	0.878
	N		0.872	0.820	0.845
GRU-2	R	**0.881**	0.851	0.950	**0.898**
	N		0.930	0.800	**0.860**

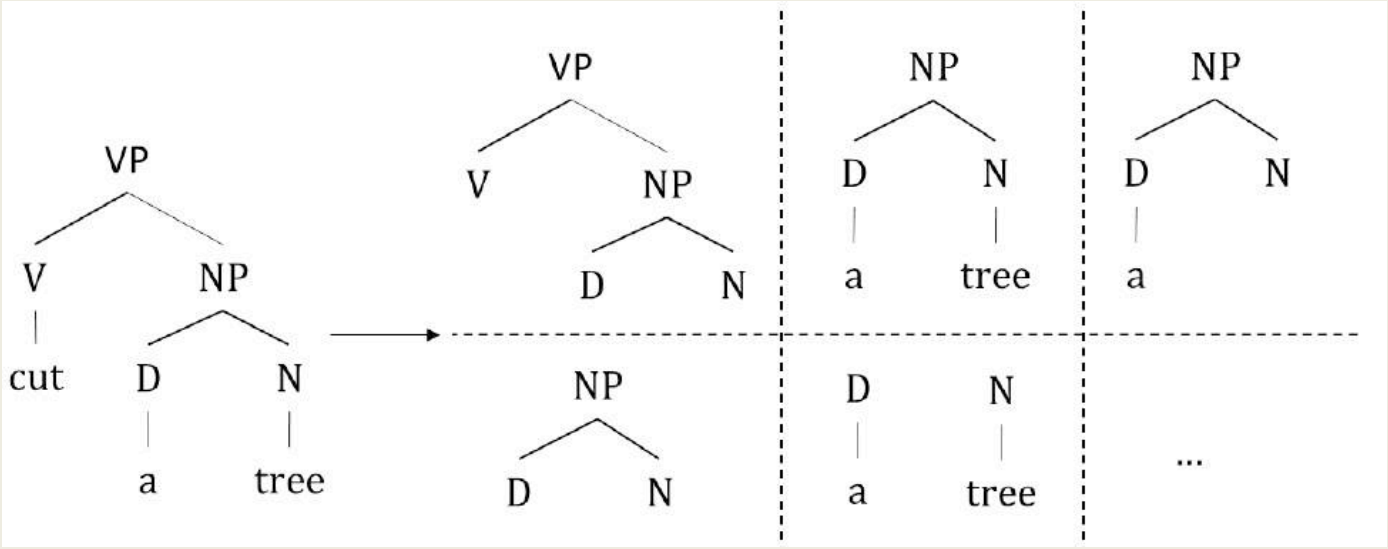
[Ma et al., ACL 2017]

- RQ: Detect rumor events
- Insight: Capture the substructure/content of propagation as tree

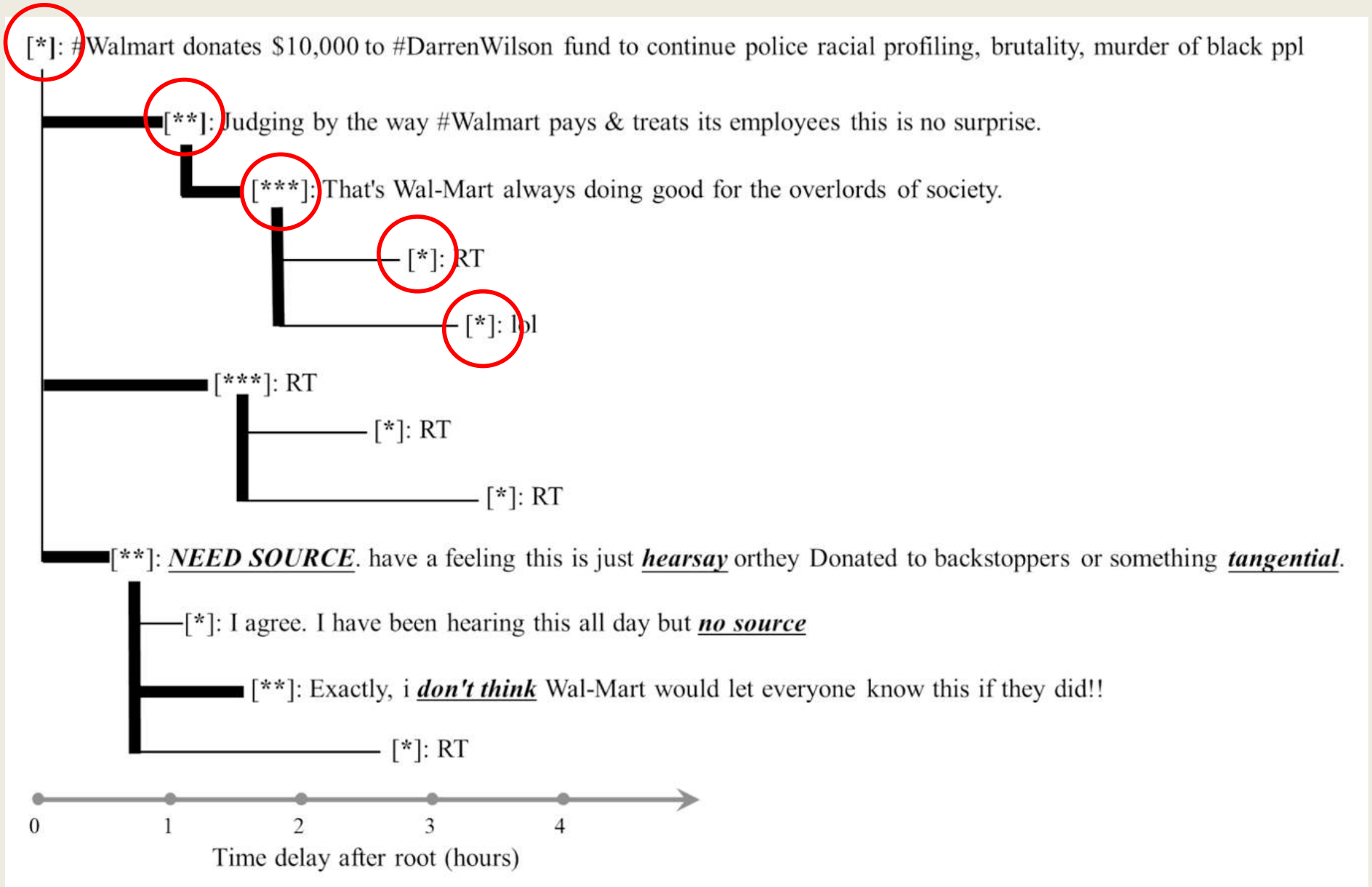


[Ma et al., ACL 2017]

- Tree-kernel
  - From NLP
  - To compute the syntactic and semantic similarity between two natural language sentences by implicitly counting # of common subtrees between their corresponding parse trees



[Ma et al., ACL 2017]



[Ma et al., ACL 2017]

- Extension
  - Parse tree node is represented by enumerable nominal value (e.g., part-of-speech tag)
  - Propagation tree node is given as a vector of continuous numerical values representing the basic properties of the node
  - Similarity of subtree must be “softer”

$$f(v_i, v_j) = e^{-t}\left(\alpha\mathcal{E}(u_i, u_j) + (1-\alpha)\mathcal{J}(c_i, c_j)\right)$$

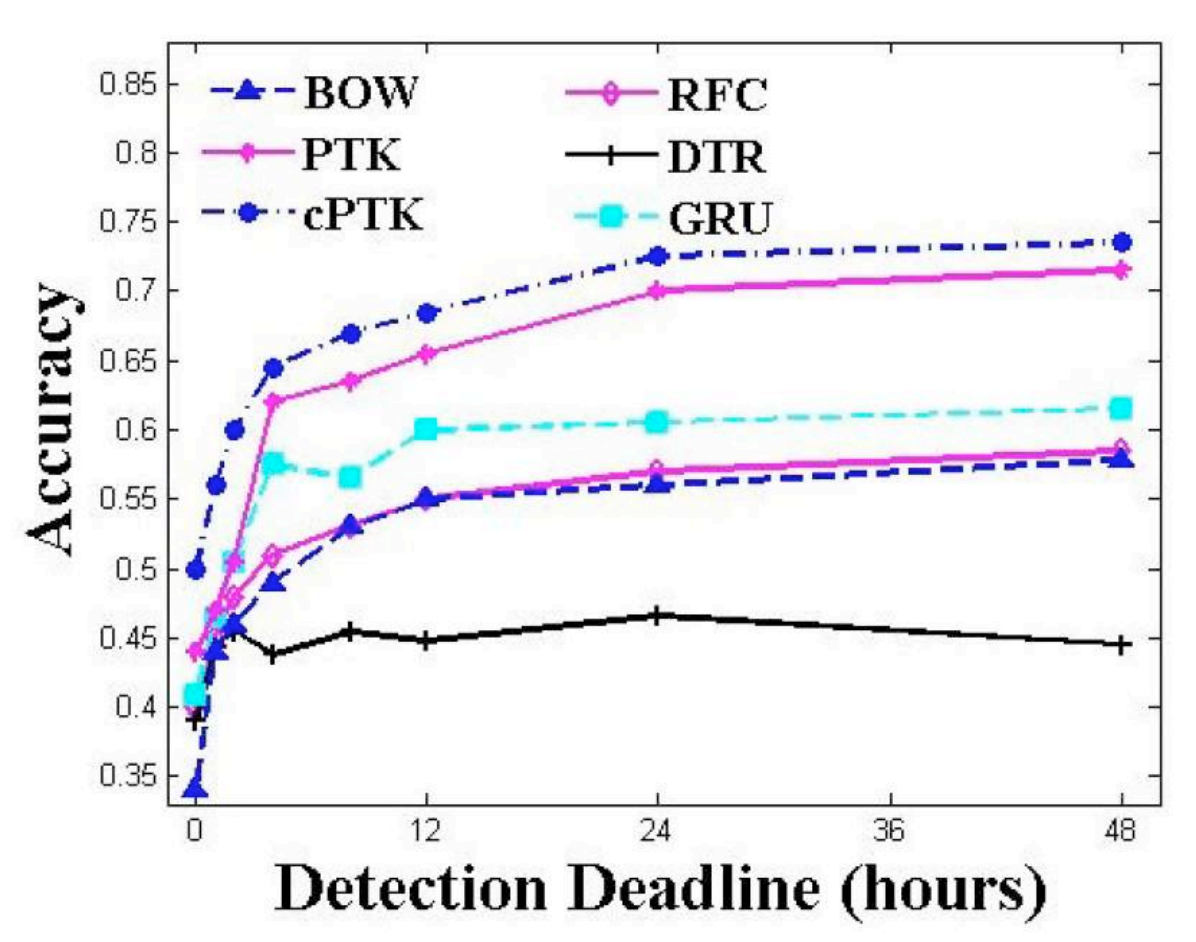
Time lag

User feature

Content feature

[Ma et al., ACL 2017]

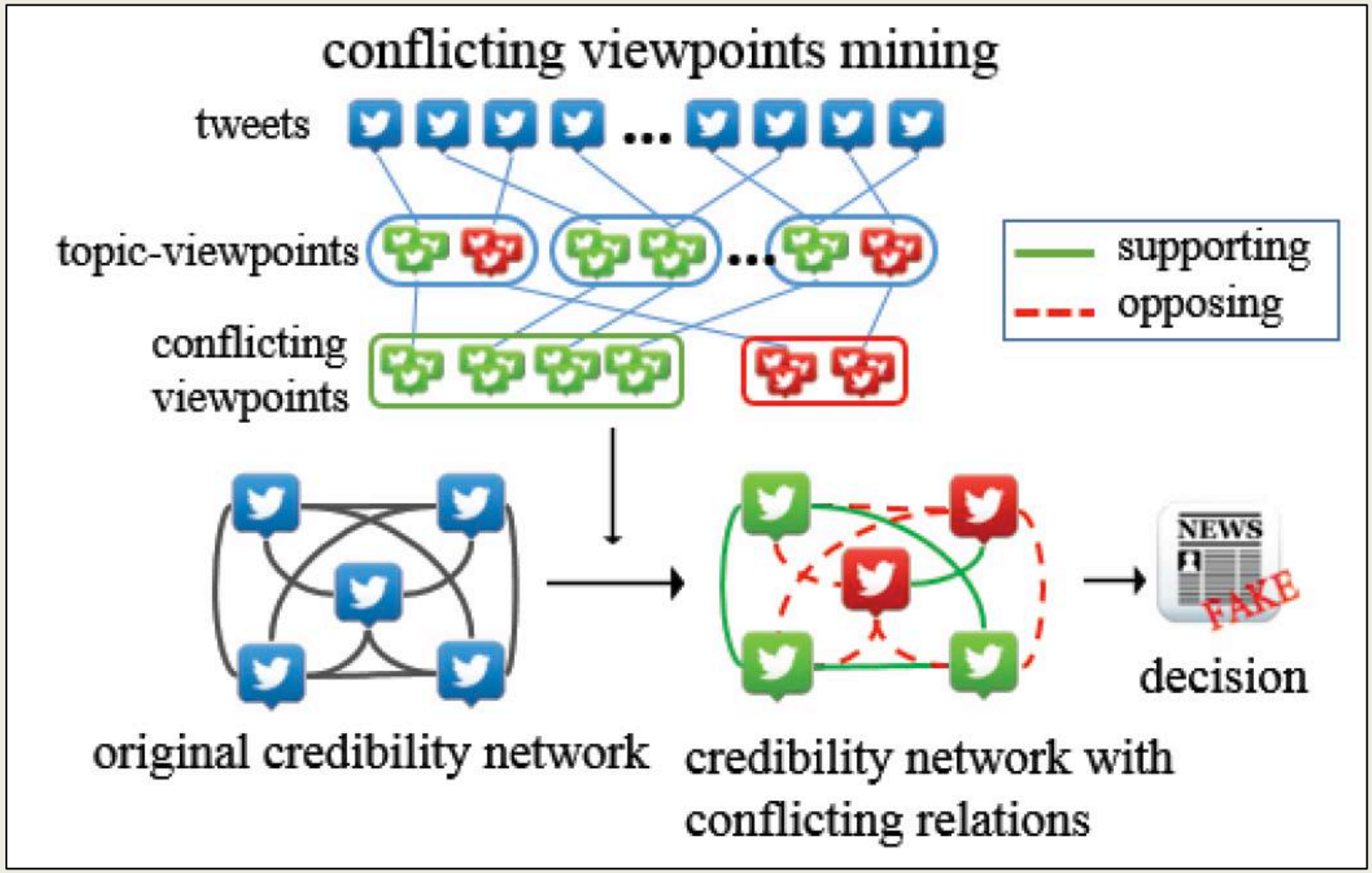
- Dataset: same as [Ma et al., IJCAI 2016]
- Result



⇦ Tree kernel learning

[Jin et al., AAAI 2016]

- Insight: fake news has conflicting viewpoints



What is Lacking Now?

- Lack of datasets
  - Labeled, Community-Agreed, Benchmark
  - Uncertainty in comparison of methods

- What is the “after” the detection?
  - How to convince users to believe in the detection?
  - By fact-checkers or machines?
  - How to present the detection results?

South Park slams Facebook for selling fake news

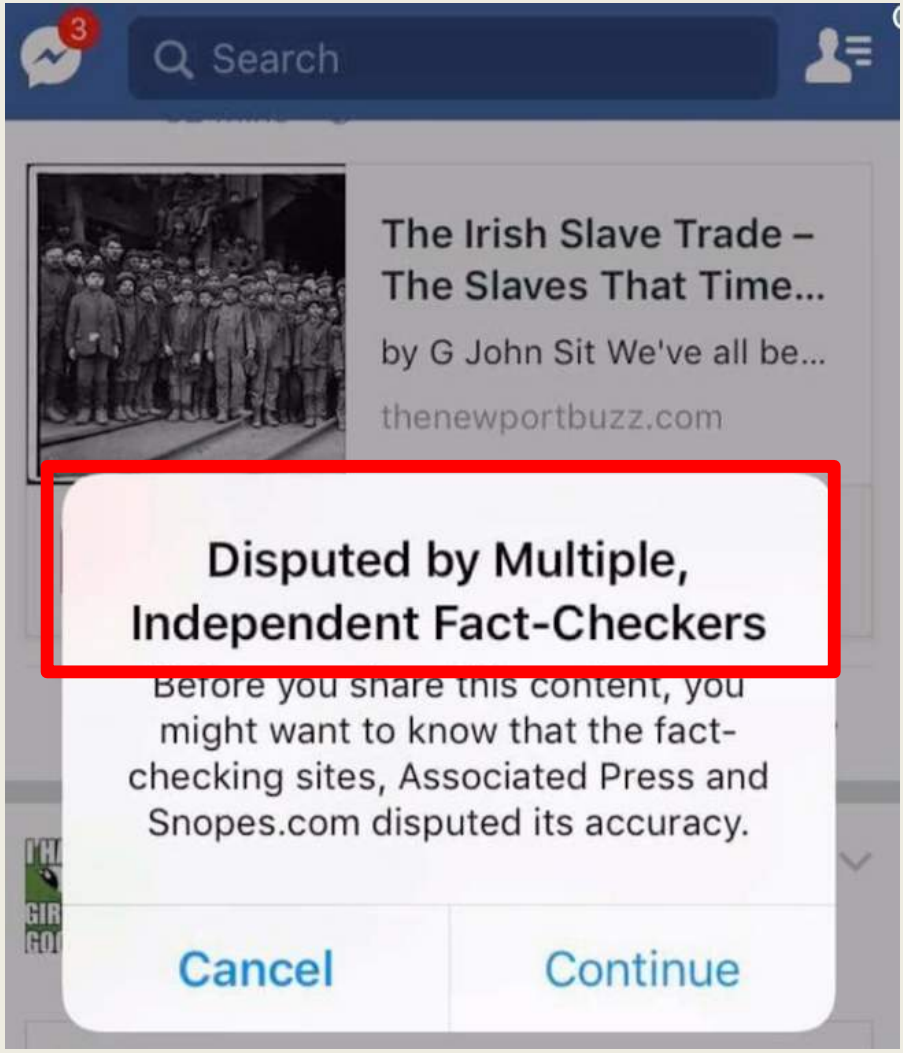
Posted 16 hours ago by Josh Constine (@joshconstine)

Next Story

"I make money from Facebook for my fake content in order to pay Facebook to promote my fake stories," said Professor Chaos in one of the most brutal and succinct criticisms of the social network to date. The latest episode of South Park pulled no punches in its take-down of the Facebook fake news scandal. It poses Mark Zuckerberg as an indecipherable bully

Eg, How to Present the Detection Result?

The warning message that appears when some Facebook users try to post a fake news article. Photograph: Facebook



Eg, How to Present the Detection Result?

- Yale study [Pennycook & Rand, 2017]
  - Tagging fake news by a label "disputed by 3rd party" ➔ little effect
  - "Implied truth effect"
- Dartmouth study [Blari et al., 2017]
  - Portion of users rating fake news as legit ➔ 29% in unlabeled to 19% with "disputed by 3rd party" label
- Better labeling?
  - Showing more evidences: eg, "disputed by 3rd party" & "agreed by x% readers"?

Eg, How to P

January 19, 2018

News Feed FYI: Helping Ensure News on Facebook Is From Trusted Sources

1. "Do you recognize the following websites?" (Yes, No)
2. "How much do you trust each of these domains?" (Entirely, A lot, Somewhat, Barely, Not at all)

Derive the ratio of those who trust the source to those who are familiar with it

News Feed that are designed to help bring
e meaningful connections on Facebook. As a
cluding news, video and posts from brands.

e we're making this year: to make sure the
news people see, while less overall, is high quality.

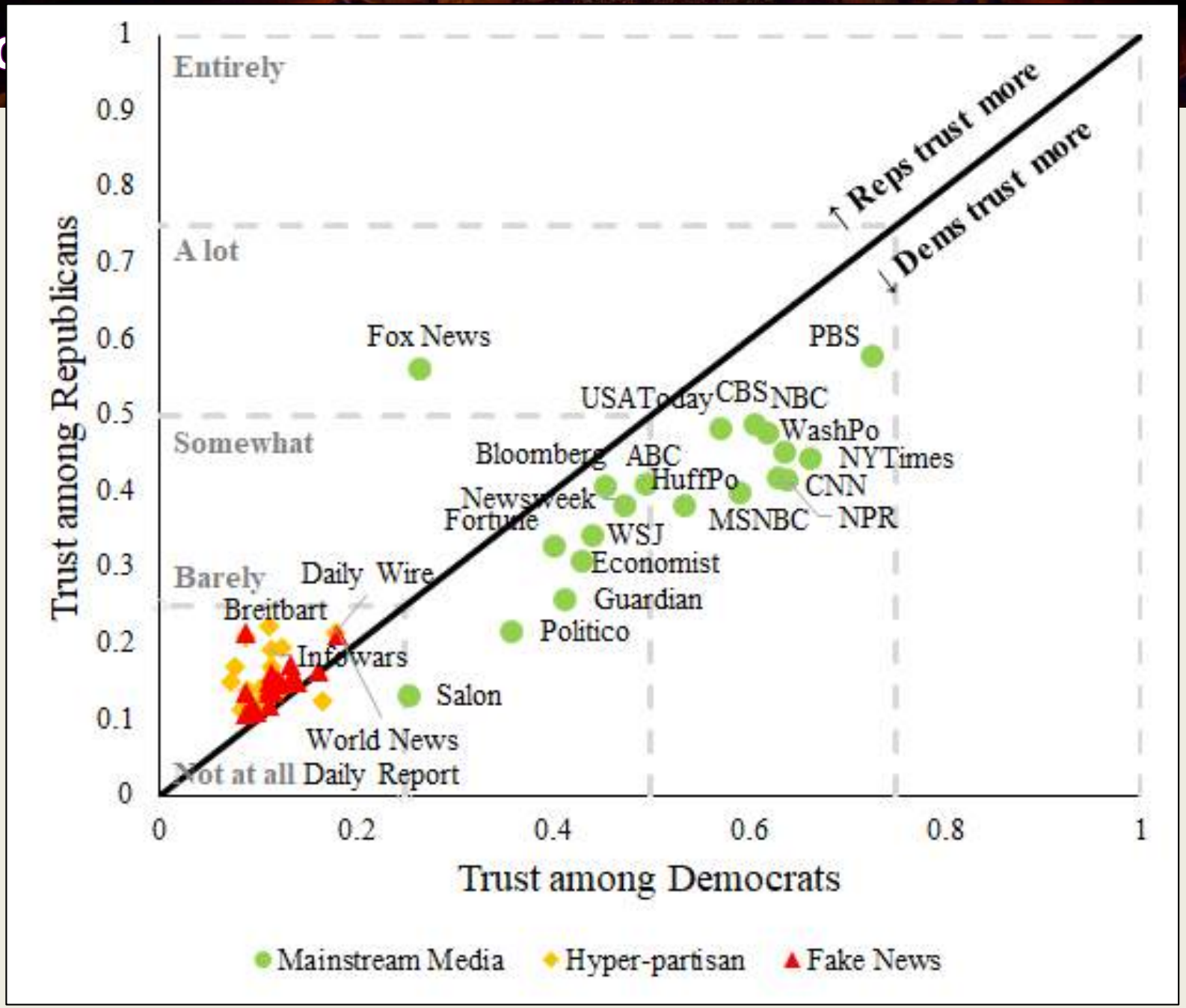
Mark Zuckerberg
about 2 weeks ago

time we all spend on Facebook is

urage meaningful social interactions with family and friends over passive consumption. As a result, you'll see less public content, including news, video, and posts from brands. After this change, we expect news to make up roughly 4% of News Feed -- down from roughly 5% today. This is a big change, but news will always be a critical wa... See More

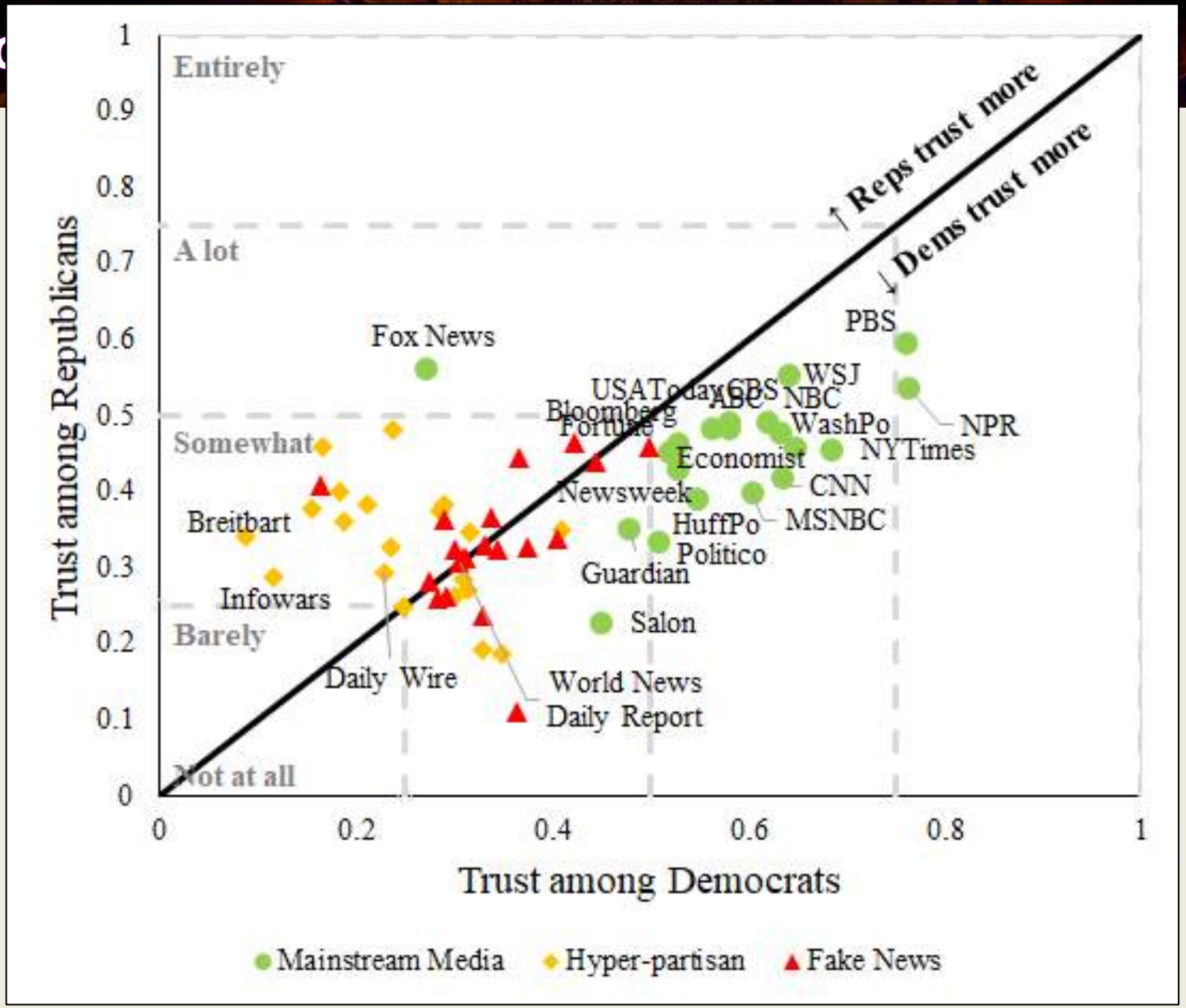
140K 16K 10K

Eg, Ho



[Pennycock and Rand, 2018], preliminary release

Eg, Ho



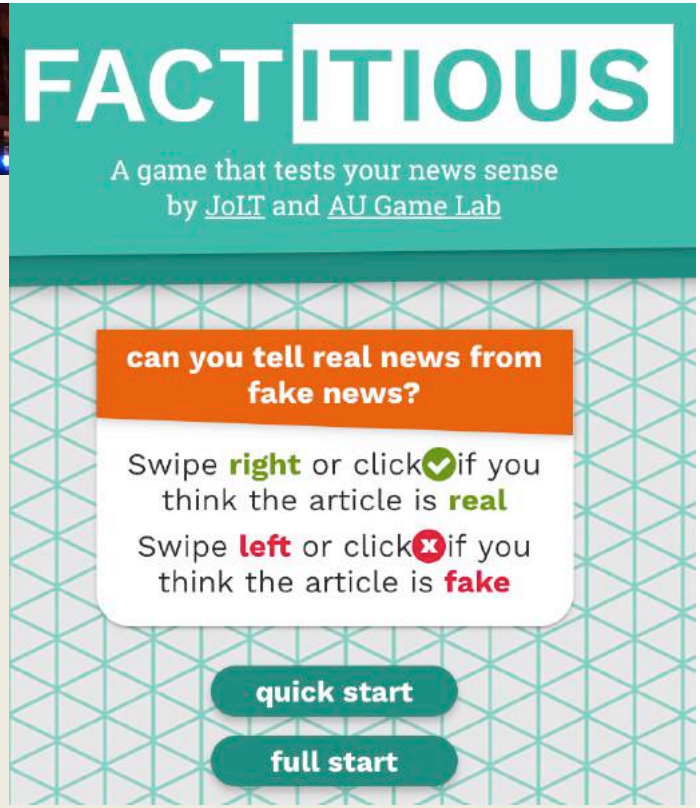
When unfamiliar ratings removed (as FB planned)

Fake News Datasets

- ASU FakeNewsNet: https://github.com/KaiDMML/FakeNewsNet
- PHEME rumor: https://goo.gl/5JXAXt
- NewsVerification: https://goo.gl/jDZJpS
- BuzzFeedNews: https://goo.gl/TUnw2K
- LIAR: https://goo.gl/wkNYhv
- BS Detector: https://github.com/bs-detector/bs-detector
- FakeNewsSites: https://www.kaggle.com/mrisdal/fake-news/data

Fake News Games

- http://factitious.augamestudio.com/
- https://www.fakeittomakeitgame.com/
- https://playfakenews.com/



- http://fakenews.game/
- https://boardgamegeek.com/boardgame/235085/fake-news-or-not

Reference

- [Blair et al., 2017] Real Solutions for Fake News? Measuring the Effectiveness of General Warnings and Fact-Check Banners in Reducing Belief in False Stories on Social Media, https://www.dartmouth.edu/~nyhan/fake-news-solutions.pdf
- [Ciampaglia et al., 2018] Research Challenges of Digital Misinformation: Toward a Trustworthy Web, AI Magazine
- [Jin et al., 2016] News Verification by Exploiting Conflicting Social Viewpoints in Microblogs, AAAI
- [Kwon et al., 2013] Prominent Features of Rumor Propagation in Online Social Media, ICDM
- [Ma et al., 2017] Detect Rumors in Microblog Posts Using Propagation Structure via Kernel Learning, ACL
- [Ma et al., 2016] Detecting Rumors from Microblogs with Recurrent Neural Networks, IJCAI
- [Pennycook and Rand, 2017] Who Falls for Fake News? The Roles of Analytic Thinking, Motivated Reasoning, Political Ideology, and Bullshit Receptivity, SSRN

Reference

- [Pennycook and Rand, 2018] Crowdsourcing judgments of news source quality, (preliminary release)
- [Qazvinian et al., 2011] Rumor has it: Identifying Misinformation in Microblogs, EMNLP
- [Shu et al., 2017] Fake News Detection on Social Media: A Data Mining Perspective, KDD Exploration
- [Wu and Liu, 2018] Tracing Fake-News Footprints: Characterizing Social Media Messages by How They Propagate, WSDM
- [Wu et al., 2017] Mining Misinformation in Social Media: Understanding Its Rampant Spread, Harm, and Intervention, ICDM
- [Zhao et al., 2015] Enquiring Minds: Early Detection of Rumors in Social Media from Enquiry Posts, WWW
- [Zubiaga et al., 2017] Exploiting Context for Rumour Detection in Social Media, SocInfo
- [Zubiaga et al., 2018] Detection and Resolution of Rumours in Social Media: A Survey, CSUR