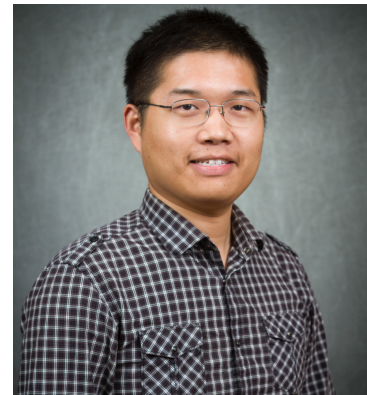# Challenges in Combating Disinformation on Social Media

Huan Liu

Joint work with **Kai Shu**

at IIT Chicago

# 10 Wonderful Examples Of Using Artificial Intelligence (AI) For Good

Bernard Marr Contributor ⓘ
Enterprise Tech

Spot "Fake News"

It's true: AI is the engine that pushes "fake news" out to the masses, but Google, Microsoft, and grassroots effort Fake News Challenge are using AI (machine learning and natural language processing) to assess the truth of articles automatically. Due to the trillions of posts, Facebook must monitor and the impossibility of manually doing it, the company also uses artificial intelligence to find words and patterns that could indicate fake news. Other tools that rely on AI to analyze content include Spike, Snopes, Hoaxy, and more.
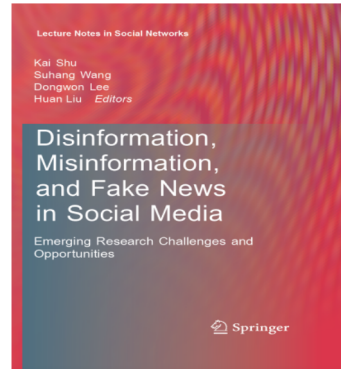
Misinformation and Disinformation
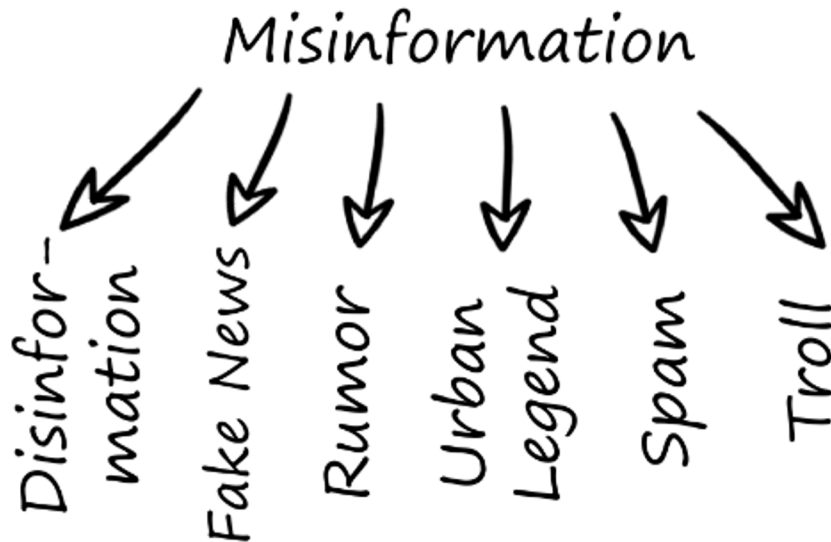
[Misinformation in social media: Definition, manipulation, and detection](#)

L Wu, F Morstatter, KM Carley, H Liu
ACM SIGKDD Explorations Newsletter, 2019, 21 (2), 80-90

Disinformation Challenges

- Why it persists
  - It's not a new problem, but its reincarnation in social media wreaks havoc unprecedentedly
- Why it is difficult to deal with
  "A lie will gallop halfway round the world before the truth has time to pull its breeches on."
- Why it is insufficient with a pure data-driven approach

Fake News Detection on Social Media: A Data Mining Perspective, K Shu, A Sliva, S Wang, J Tang, H Liu, ACM SIGKDD

Rampant Misinformation during COVID-19
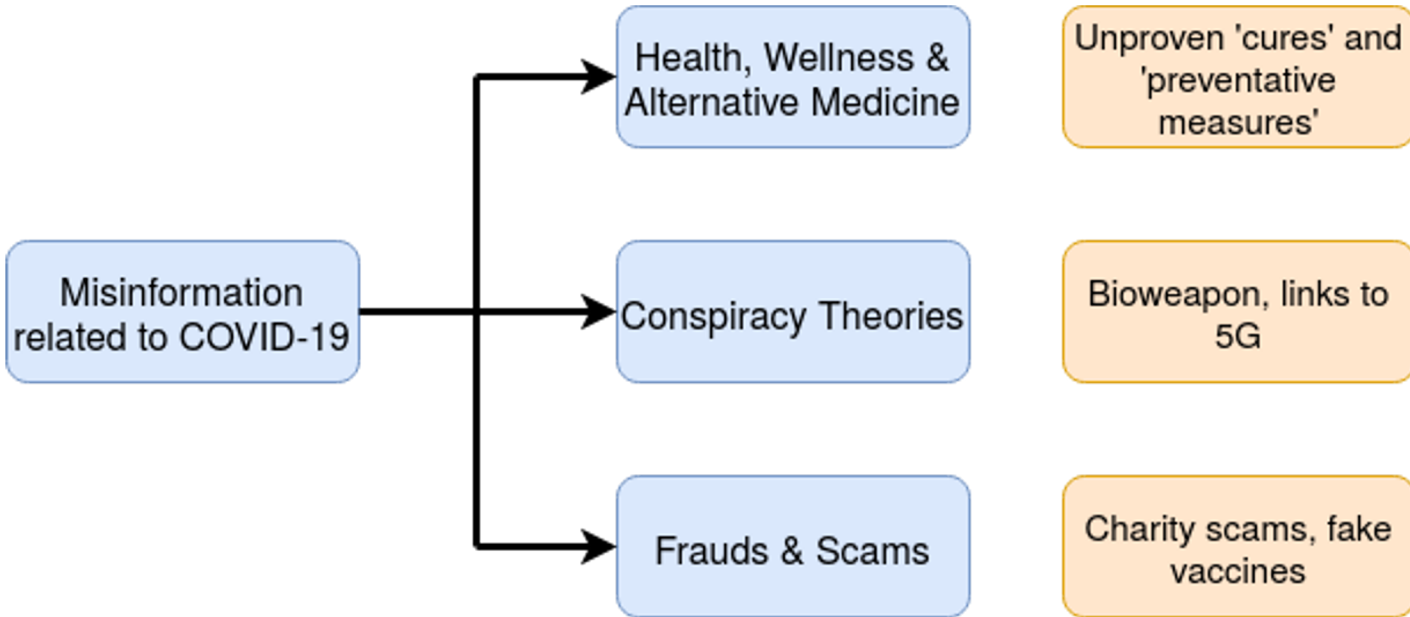
Health Misinformation

1. Often even **official news outlets** transmit health-related misinformation
   https://www.wired.com/story/the-fda-revokes-its-emergency-use-authorization-for-hydroxychloroquine/

2. People **lack adequate health literacy,** hope for miracles, and are vulnerable to health hoaxes

3. In the absence of a cure, people are often **desperate to believe** whatever 'cure' they come across online

Conspiracy Theories

1. Presence of **multiple types of transmission channels** makes the conspiracy information hard to be detected and stopped

2. The **fear and panic** among people also boosts the spread of conspiracy theories

3. Bill Gates who became a <u>top target of conspiracy theories</u> of COVID-19 said: "*even to repeat it would give it credibility*"

Impact of COVID-19 Misinformation

FTC COVID-19 Complaints
January 1, 2020 - May 7, 2020

Category	Count
Fraud	22,109
Other	12,069
Identity Theft	2,816
Do Not Call	1,899
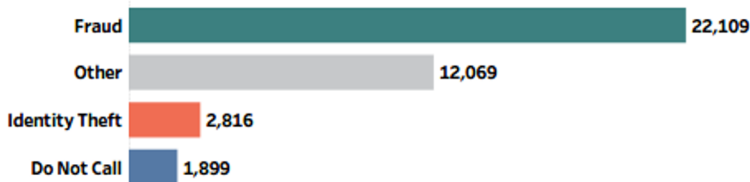
38,675 OVERALL REPORTS

$27.24M TOTAL FRAUD LOSS
*43.1% of fraud complaints report a loss

$503 MEDIAN FRAUD LOSS

NCDC ✓
@NCDCgov

#FactsNotFear@WHO has NOT approved the use of chloroquine for #COVID19 management. Scientists are working hard to confirm the safety of several drugs for this disease.

Please DO NOT engage in self-medication. This will cause harm and can lead to death.#COVID19Nigeria

BUSINESS INSIDER **Business Insider** ✓
@businessinsider

77 cell phone towers have been set on fire so far due to a weird coronavirus 5G conspiracy theory
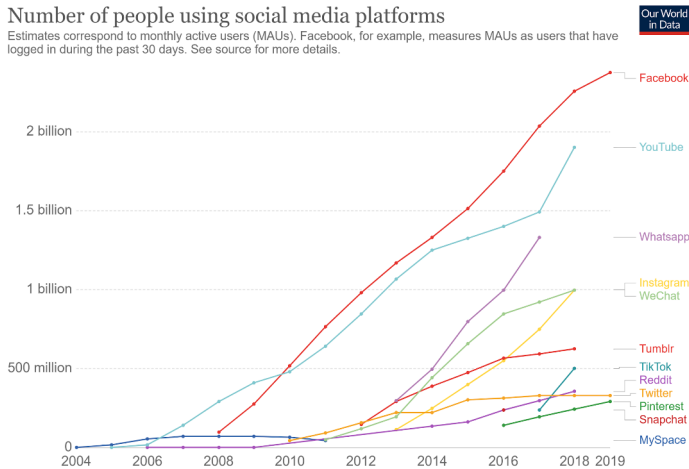

Urgent need to stop the spread of misinformation!

Social Media for Information Sharing

- People are increasingly using social media for information sharing, social networking, etc

- 68% of Americans get news on social media

Number of people using social media platforms
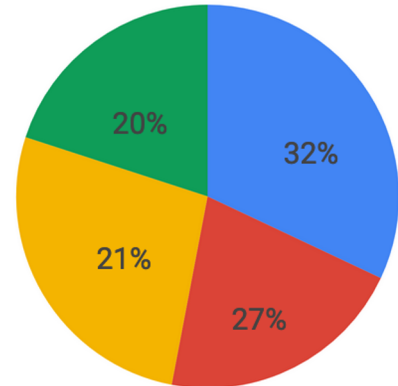Estimates correspond to monthly active users (MAUs). Facebook, for example, measures MAUs as users that have logged in during the past 30 days. See source for more details.

Our World in Data

Source: Statista and TNW (2019)   CC BY

• Never • Hardly Ever • Sometimes • Often

32%
27%
21%
20%

https://www.journalism.org/2018/09/10/news-use-across-social-media-platforms-2018/

Disinformation is Rampant on Social Media

- Disinformation is false information [news or non-news] with an evil intention aiming to mislead the public

- Fake news is news with intentionally false information
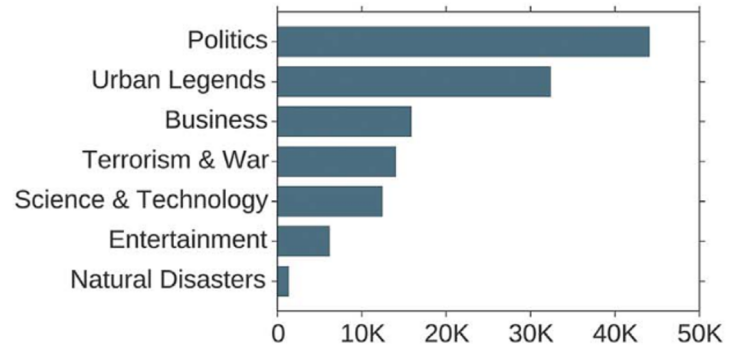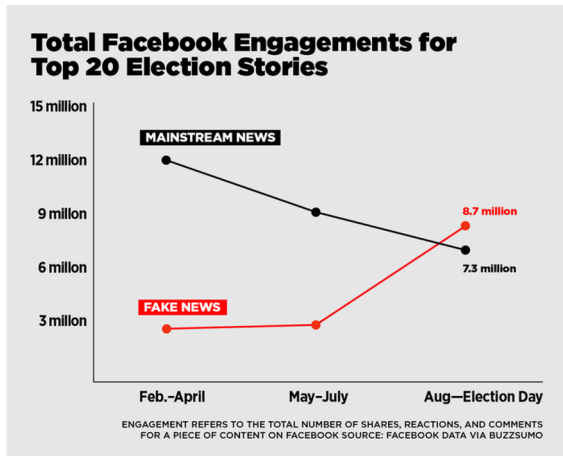
Allcott, Hunt, and Matthew Gentzkow. "Social media and fake news in the 2016 election." *Journal of economic perspectives* 31.2 (2017): 211-36.

Vosoughi, Soroush, Deb Roy, and Sinan Aral. "The spread of true and false news online." *Science* 359.6380 (2018)

Fake News - Disinformation

- Humans are susceptible to fake news
  - Typical accuracy in the range of 55-58%
  - Limited resources (time, information, and expertise)
  - **Confirmation Bias**: individuals tend to believe fake news when it confirms their pre-existing knowledge

- Fake news can have detrimental societal effects
  - Confusing readers
  - Misleading people to false information
  - Changing the way people respond to credible news

Why it is so challenging

- *Real* <u>fake news detection</u> is not just another competition
  - A competition gives a dataset with ground truth and shows who can fare best

- Fake news detection is complex in many dimensions

- We discuss some imperative challenges
  - Data, Detection, Explainability, and Mitigation or Containment

Challenges in Fact-checking

- Requiring annotators with domain expertise
- Labor-intensive and time-consuming

Facebook posts

stated on April 3, 2020 in a Facebook post:

Walmart is adopting a staggered shopping schedule by age group during the COVID-19 pandemic.

Where and how can we find ground truth fast?

By Samantha Putterman
April 17, 2020

No, Walmart hasn't announced a staggered shopping schedule by age group

The alleged Walmart announcement, which contains multiple grammatical errors, reads:

"Due to the COVID-19 pandemic effective immediately Walmart is adopting a staggered shopping schedule as follows. We apologize for the (sic) inconeince: Monday Age 66+, Tuesday 56-65, Wednesday 46-55, Thursday 36-45, Friday 25-35, Saturday 24 and below, Sunday Emergency shopping only."

. . .

https://www.politifact.com/factchecks/2020/apr/17/facebook-posts/no-walmart-hasnt-announced-staggered-shopping-sche/

Data Repository for COVID-19

- Datasets related to COVID-19 with multi-disciplines
  - Spatial-temporal data
  - Social media,
  - Academic articles …

- A *meta-data* repository, MetaCOVID, to encourage data sharing and donation from research community and promote collaborations

https://github.com/bigheiniu/awesome-coronavirus19-dataset
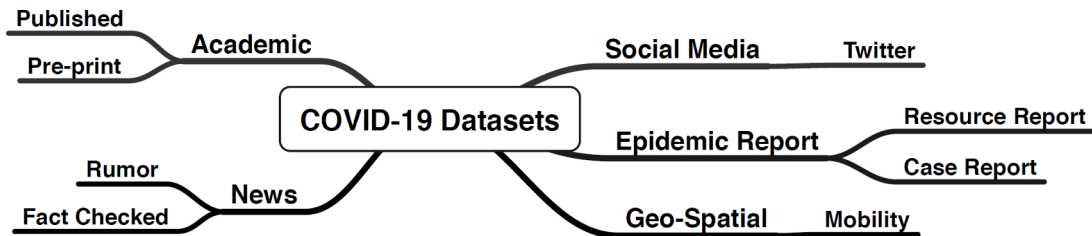
Weak Social Supervision for FN Detection

- ## News Content
  - Intentionally written to mislead people
  - Diverse in terms of topics, styles, and media platforms

- ## Where to find ground truth?

- ## Social Context
  - Social interactions are massive, incomplete, unstructured, and noisy
  - Effective methods are needed to leverage rich social signals

News Content

Auxiliary Information

Social Context
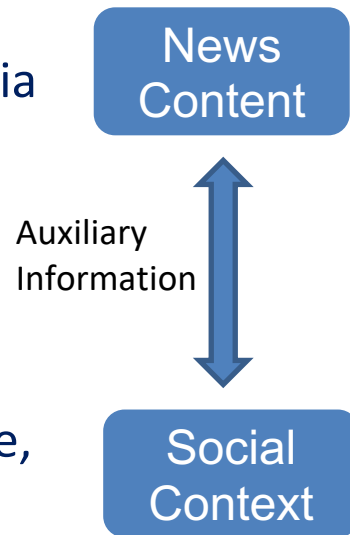
Kai Shu, Ahmed Hassan Awadallah, Susan Dumais, and Huan Liu. ``Detecting Fake News with Weak Social Supervision'', IEEE Intelligent Systems, 2020

Explainable Fake News Detection

- Existing work focuses on ***detecting*** fake news, but cannot ***explain how*** it is detected as fake

- Explanation is important
  - Provides insights and knowledge to domain experts
  - Explainable features from noisy auxiliary information can further improve detection performance
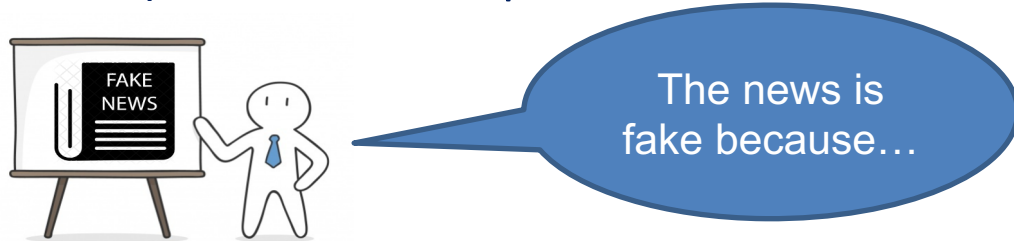
Kai Shu, Limeng Cui, Suhang Wang, Dongwon Lee, and Huan Liu. ``dEFEND: Explainable Fake News Detection'', KDD 2019, August 4-8, 2019. Anchorage, Alaska.

Challenges in Mitigation

- Is mitigation easy? Or easier than detection?

- Accurate detection ≠ successful mitigation

- Mitigation involves *users*
  - It is complex with new dimensions

- Users have their own judgments/opinions
  - Stark contrast in response from the Tulsa rally

- Users exist on social media "alone together"

Are SM Users Stupid?

"Stupidity is knowing the truth, seeing the truth but still believing the lies.
And that is ***more infectious*** than any other disease."

- Richard Feynman

It's more complex than that: *mitigation is about users and humans are imperfect*

We, Users, are Irrational

- Are we susceptible to disinformation?
  - "No, I am immune to disinformation"

- Echo chambers
  - Examine your social networks

- Filter bubbles
  - Look at the news sources you get your daily news

Be open-minded and admit that we're all fallible ☺

A Few More Challenges

- Conversational receptiveness
  - We are so eager to convey our ideas or beliefs
  https://www.scientificamerican.com/article/the-right-way-to-talk-across-divides

- $If \ A \rightarrow B, then \ \neg A \rightarrow \neg B$ (let's pause and take poll)
  - What do you think of the following scenario
  "Black Lives Matter" -> A = Black, B = their lives matter

  An easy to commit, common logic error

  - How can we effectively converse, let alone reason, with wrong logic?

Lessons Learned

- Fake news detection is difficult
  - A moving target with changing topics

- Data is key
  - Impractical to label data at scale and fast

- Early detection is critical
  - Data-driven approaches are limited

- Mitigation is important, but
  - We all have our own preferences ☹

Need for Multidisciplinary Research

- Integrating theories from different disciplines
  - Diversity helps
  - Learning with weak social supervision

- How can build novel *algorithms* help combat disinformation?

- **Kai Shu** is continuing this line of research at IIT, Chicago with his own students ☺

Thank You All

Mining Disinformation
- Data-oriented
  - Intention Estimation
  - Benchmark Datasets
  - Generated Data
- Feature-oriented
  - News Content
  - Social Context
- Model-oriented
  - Semi-Supervised
  - Weakly-supervised
  - Unsupervised
  - Supervised
- Application-oriented
  - Diffusion
  - Mitigation

COVID-19 Datasets
- Academic
  - Published
  - Pre-print
- News
  - Rumor
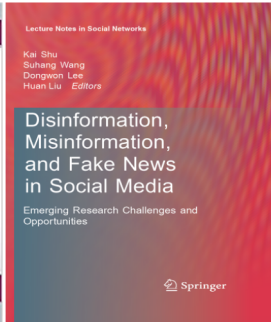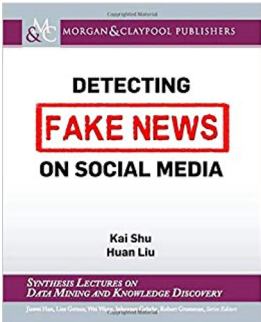 - Fact Checked
- Social Media — Twitter
- Epidemic Report
  - Resource Report
  - Case Report
- Geo-Spatial — Mobility

https://github.com/bigheiniu/awesome-coronavirus19-dataset

FakeNewsNet
https://github.com/KaiDML/FakeNewsNet

Fake News Research
Fundamental Theories, Detection Strategies & Open Problems
KDD2019
wsdm2019

https://www.fake-news-tutorial.com/

MORGAN & CLAYPOOL PUBLISHERS
DETECTING FAKE NEWS ON SOCIAL MEDIA
Kai Shu
Huan Liu
SYNTHESIS LECTURES ON DATA MINING AND KNOWLEDGE DISCOVERY

Lecture Notes in Social Networks
Kai Shu
Suhang Wang
Dongwon Lee
Huan Liu *Editors*
Disinformation, Misinformation, and Fake News in Social Media
Emerging Research Challenges and Opportunities
Springer

KDnuggets™ **Top Stories Last Week**
A Quick Guide to Fake News Detection on Social Media

DATA SKEPTIC — Algorithmatic Detection of Fake News

datanami — AI Squares Off Against Fake News

Psychology Today — What is Fake News? Is Debate Worth the Effort?

2018 SBP-BRiMS
Challenge Problem Winner
Disinformation Challenge
THIS CERTIFICATE IS AWARDED TO
FAKENEWSTRACKER: TOWARDS FAKE NEWS COLLECTION, DETECTION, AND VISUALIZATION
BY
KAI SHU, DEEPAK MAHUDESWARAN AND HUAN LIU

http://dmml.asu.edu/dfn/   https://www.springer.com/gp/book/9783030426989   • • •

FakeNewsTracker
http://blogtrackers.fulton.asu.edu:3000/#/