Fake & Real
in news and method
In collaboration with...

Nir Grinberg

Kenneth Joseph

Lisa Friedland

Briony Swire-Thompson
Does Finland exist?

- Spoof post in reddit
- Vigorous debate on elements of conspiracy
- Some people clearly take it seriously...
Key elements of conspiracy...

- Result of agreement between Russia and Japan
- The landmass that we think of as Finland is really part of the Baltic sea
- Forged GPS and satellite imagery
- “No country in the world could possibly be that good”
Yes, Finland does exist...

But how do you know?
It’s because of a network of informational influences
Truth is a network phenomenon

It’s derived from the people we know and sources we trust

We have diverse understandings of the world that tend to cluster together in the network

Democracy is what happens in the spaces between in the network
Today...

Fake news: the role that social media play
Story within the story is the methodology→
How to derive real insights regarding fakery?
Fake news in 2016 election
Fake news...

• Definition: misinformation with the trappings of “legitimately” produced news but without the underlying organizational processes. (Lazer et al 2018)

• Not: bad news, unfavorable news coverage. Not even errors in news reporting.

• Subgenre of misinformation, particularly pernicious because it undermines legitimate sources of news

• [Note that “real news” is also subject to attack]
THE rush of the day seemed to have ended in the New York office of The Associated Press on the afternoon of June 8, 1921. The wires hummed monotonously with Wall Street's closing prices, baseball scores, the price of pork, to-morrow's weather. In about twenty rushing wireless inquiries to the commanders. The wire men were connecting the news room with North Sydney, Nova Scotia, by direct cable in order to pick up any wireless messages intercepted there. Ten minutes passed. They worked fast. But rumor worked faster.
But the mechanics of spread have changed...
Why Twitter?

• Perhaps the most conducive medium for rapid dissemination of information
• Does not encourage careful scrutiny
• Creates favorable conditions for fake news
• It’s important
• It’s accessible
• It has areas that are dysfunctional...
Method: The story within a story...
How to count in a house of mirrors?
Big data: measurement conundrums

Strength: capturing real behavior (contrast to surveys)
Weaknesses: vast heterogeneity and massive manipulation

The Parable of Google Flu: Traps in Big Data Analysis

David Lazer, Ryan Kennedy, Gary King, Alessandro Vespignani

In February 2013, Google Flu Trends (GFT) made headlines but not for a reason that Google executives or the creators of the flu tracking system would have hoped. Nature reported that GFT was predicting more than double the proportion of doctor visits for influenza-like illness than the algorithm in 2009, and this model has run ever since, with a few changes announced in October 2013 (10, 15).

Although not widely reported until 2013, the new GFT has been persistently overestimating flu prevalence for a much longer time.
Twitter: a scientific conundrum

Regular people
News media
Celebrities
Corporations
Bots, bots, and more bots
etc...
Our objective: build sample of voters

“Regular” people, and most salient to news consumption
Available voter files
Links to some basic demographics
Studying voters on Twitter is not easy!

... exclude all non-voting humans
## Linking Twitter Accounts to Voter Records

### Voter Registrations

<table>
<thead>
<tr>
<th>Name</th>
<th>State</th>
<th>Party</th>
<th>Age</th>
<th>Match to Twitter?</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kenny Joseph</td>
<td>MA</td>
<td>D</td>
<td>52</td>
<td>Yes!</td>
<td></td>
</tr>
<tr>
<td>Nir Grinberg</td>
<td>NY</td>
<td>D</td>
<td>19</td>
<td>No, not unique</td>
<td></td>
</tr>
<tr>
<td>Nir Grinberg</td>
<td>NY</td>
<td>R</td>
<td>25</td>
<td>No, not unique</td>
<td></td>
</tr>
<tr>
<td>Amelia Tzray</td>
<td>CA</td>
<td>...</td>
<td>...</td>
<td>No, not on Twitter</td>
<td></td>
</tr>
</tbody>
</table>

### Results

- **No match, not in voter records**
- **Twitter universe**
- **Matching 2 million registered voters: proportion on Twitter**
  - No matches on twitter
  - >1 match
  - Exactly 1 match...
    - 10% blank location field
    - 2% location not parsable
    - 1% correct U.S. state: 22,861
  - No matches to correct state

**Result:** panel of ~16k U.S. voters on Twitter
Linking Twitter Accounts to Voter Records

Focus on unusual names (no John Smiths)
Use geo info in profiles
Demographics compared to Pew survey of voters on Twitter
We focus on **political news** around the election by:

1. Analyzing tweets with external URLs (not back to Twitter)
2. Created during Aug-Dec 2016
3. Train a classifier for detecting political tweets
Detecting political tweets

1. Start with full text of tweet (incl. handles, hashtags, URLs).
2. Have whitelist of political keywords (111 terms): candidate names and handles, popular hashtags, words about elections and officials. High precision list -- avoid false positives.
4. Train classifier. (New model daily!)
5. Apply classifier to negatives → detects tweets the whitelist missed.

During study period,

- Whitelist matches ~10-20% of tweets per day.
- Classifier detects another 2-3% as political.
Challenge: How much fake news did people see on Twitter?

- Fake news can have continued impact on individuals (through familiarity) even if they don’t recall or tweet about it.

- What counts as “seeing”? an impression? a click?

- How can one estimate exposure on Twitter without having privileged access to Twitter server logs?
Data summary

16,442 panel member accounts

10.1M Political URLs analyzed (non-unique)

648M Potential exposures (32.4M impressions)

Covering Aug. 1st - Dec. 6th 2016
Operationalizing fake news

- Political content coming from sources lacking proper editorial process and factual accountability

- Working definition:
  - Domain-level labeling
  - Start with fake news lists from trusted fact-checking orgs → "black"
  - To expand the set, use Snopes to identify sources (177) with false claims
  - Manually annotate these domains as “red”, “orange” (= fake news), “yellow” or “green” (= non-fake)
Operationalizing fake news

- **Black** (383 sites)
  - Sites from pre-existing lists: purely fictional and deceptive
  - e.g., abcnews.com.co, Denver Guardian

- **Red** (64 sites)
  - Publish major falsehoods with little regard for the truth
  - e.g., InfoWars, Gateway Pundit; “Stunning UFO sighting in Phoenix”

- **Orange** (65 sites)
  - Minor or occasional falsehoods
  - e.g., Daily Caller, Zero Hedge; “Ben Carson is Almost Certainly Suffering from a Mental Illness”
Findings
5.1% of political content in voters’ Timelines was fake news
Fake news is highly concentrated!

Exposures

% of Panel Members

% of All Exposures
Fake news is highly concentrated!

Exposures

Sharing

16 accounts (of 16,442) responsible for 80% of fake news
Who are the people that consume and share so much fake news?
Superspreaders

- Are the accounts human? Are they who we think they are?
- What are they doing?
Superspreaders: Who are they?

Among the 16 accounts sharing 80% of fake news:

- 14 Republicans, 2 Democrats
- 12 women, 4 men
- Ages: 47-77
- States: Florida, Michigan, Maine, Oklahoma (x 2), California (x 2), New York (x 2), Georgia (x 2), Ohio, Virginia, ...
- Professions: flight attendant, accountant, physician, tea party organizer, former Avon representative, marketing communications, retired military (x 2), administrative assistant at state senate, fundamentalist Christian, one-time candidate for local office, ...
Superspreaders: What are they doing?
Superspreaders: What are they doing?

Cyborgs: humans using automation tools
- Apps can auto-retweet sources, quickly share links, queue posts for optimal posting times
- Higher volume than bots OR humans alone?
- Embedded in human social networks
Among the 99%:
Who saw fake news?
Who shared it?
Measuring political affiliation

- Why? We expect behavioral differences among people of the same party (e.g. moderate vs. extreme)

- Our approach: use people’s “news diet” and precinct-level vote share to estimate a continuous [-1, 1] score.

- Even people who don’t share political content on Twitter get a score!
Measuring political affiliation

How well does it work?

- Predicts out-of-sample party registration with 82% AUC

- Our site-level alignment scores correlate really well with previous work:
  - Bakshy et al. 2015, $p = .91$ (.87, .94)
  - Budak et al. 2016, $p = .89$ (0.68,0.96)
Who saw fake news?

Fraction of fake news

- R*: Extreme right
- R: Right
- C: Center
- L: Left
- L*: Extreme left

Political exp. (weekly)

Age

- 20
- 40
- 60
- 80
Who shared fake news?
Is fake news more viral than regular news?
What other sites were consumed by voters who were exposed to fake news?

Where did fake news fit within the broader media ecosystem?
Co-consumption of fake news

- Blue sites were seen more often by registered Democrats
- Red sites were seen more often by registered Republicans
- Filled in nodes are fake news websites
Co-consumption of fake news

Filled in nodes are fake news websites

Red sites were seen more often by registered Republicans

Blue sites were seen more often by registered Democrats

Filled in nodes are fake news websites

Several NBC affiliates

(Several, but not all)
>77% of political news exposures were from one of the 81 sites in this “mainstream media cluster”

People who saw URLs from one fake news site likely saw URLs from multiple fake news sites.
Conclusions

• At its peak, fake news sources accounted for more than one in twelve political URLs in people’s Timelines, and more than one in five political URLs being shared.

• People, not just bots, shared and consumed fake news.

• Exposures and sharing were extremely concentrated:
  • 80% of fake news funneled to just 1.1% of people.
  • 80% of shares came from only 0.1% of people.
Conclusions

• Those most likely to be exposed to and share fake news were those who were conservative leaning and older adult.

• Selective treatment of fake news: left-leaning individuals shared fake news less than regular news, while people on the right shared fake news just like regular news.

• Co-consumption network patterns confirm that those who saw one fake news site were likely to view multiple.
Implications for designing Twitter

• **Interventions for sharing** can be highly efficient: targeting a small number of people can limit the volume of fake news considerably. Also, just limiting posting rates could have major impact.

• **Detection**: patterns in sharing make it potentially easy to detect fake news sources via detection of clusters.

• **Informing consumers**: A small number of sites account for the vast majority of exposures to fake news. Systems can help users make informed decision about content from these sources.