Dr Jessica Kunert and Professor Neil Thurman

Male, metropolitan, and mainstream

Bias in algorithmic news sourcing from Twitter

Ecrea Journalism Studies Conference 2017
23-24 March 2017, Odense, Denmark

For journalists, social media can be…

- A channel for dissemination of own material
- A pointer for news and trends (Broersma/Graham 2013: 448f.)
  ➔ timely, sometimes obscure news
- Source for “ambient” news with the audience as both a receiver and a sender (Hermida 2010)

Problems and challenges (Craig 2014: 106)

- Multitude of accounts
  ➔ Bots: content polluters, valuable content aggregators!? (Kyumin/Eoff/Caverlee 2011; Lokot/Diakopoulos 2015; Larsson/Moe 2015)
- Large frequency of messages
- Speed with which information is spread

➔ "how to keep up?"

Problems and challenges (cont.)

- What's news and what‘s not?
  ➔ lots of clutter (41% on Twitter „pointless babble“ in 2009 – “eating sandwich”; Pear Analytics 2009)
  ➔ wrong/hijacked hashtags
- Verification of information
  ➔ “s/he has a gun!”
- Who to listen to? Elites or nonelites or both (Henderson/Miller 2014)?

News detection with algorithms

"The stories that show in your News Feed are **influenced by your connections and activity** on Facebook. This helps you to see more stories that interest you **from friends you interact with the most**. The **number of comments and likes** a post receives and **what kind of story it is** (ex: photo, video, status update) can also make it more likely to appear in your News Feed."
(Facebook Help Center)

Challenges arising with algorithms choosing content:

- Not everyone sees the same
- Algorithmic accountability
  (Diakopoulos 2014; Mittelstadt 2016)
  ➔ How do the algorithms work?
  ➔ "Why do we see what we see?"
- Privacy issues (who is tracked and how?)

How are the social media posts chosen that journalists see?

How can algorithmic news sourcing help journalists?

Why would journalists even need it?

"With the proliferation of smartphones and social media, it means that there are lot more witnesses to a lot more events. We can't be at everything.
Our tool helps shift some of the burden of witnessing and lets journalists do much more of the high value-added work."

(Reg Chua, Executive editor of data and innovation, Reuters)

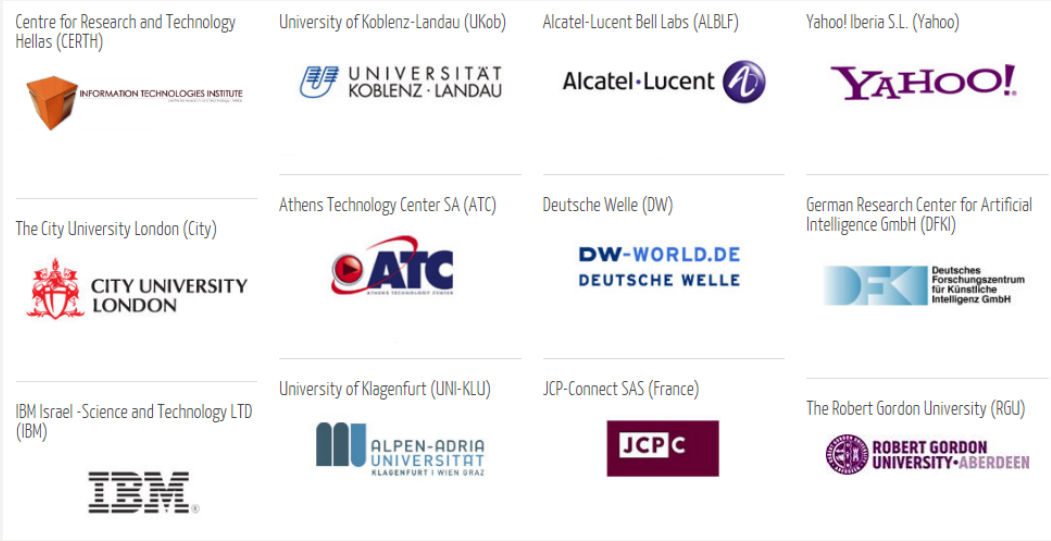
Source: http://www.niemanlab.org/2016/11/reuters-built-its-own-algorithmic-prediction-tool-to-help-it-spot-and-verify-breaking-news-on-twitter/

Why Twitter?

➔ Twitter is open to everyone: corporate, general public…

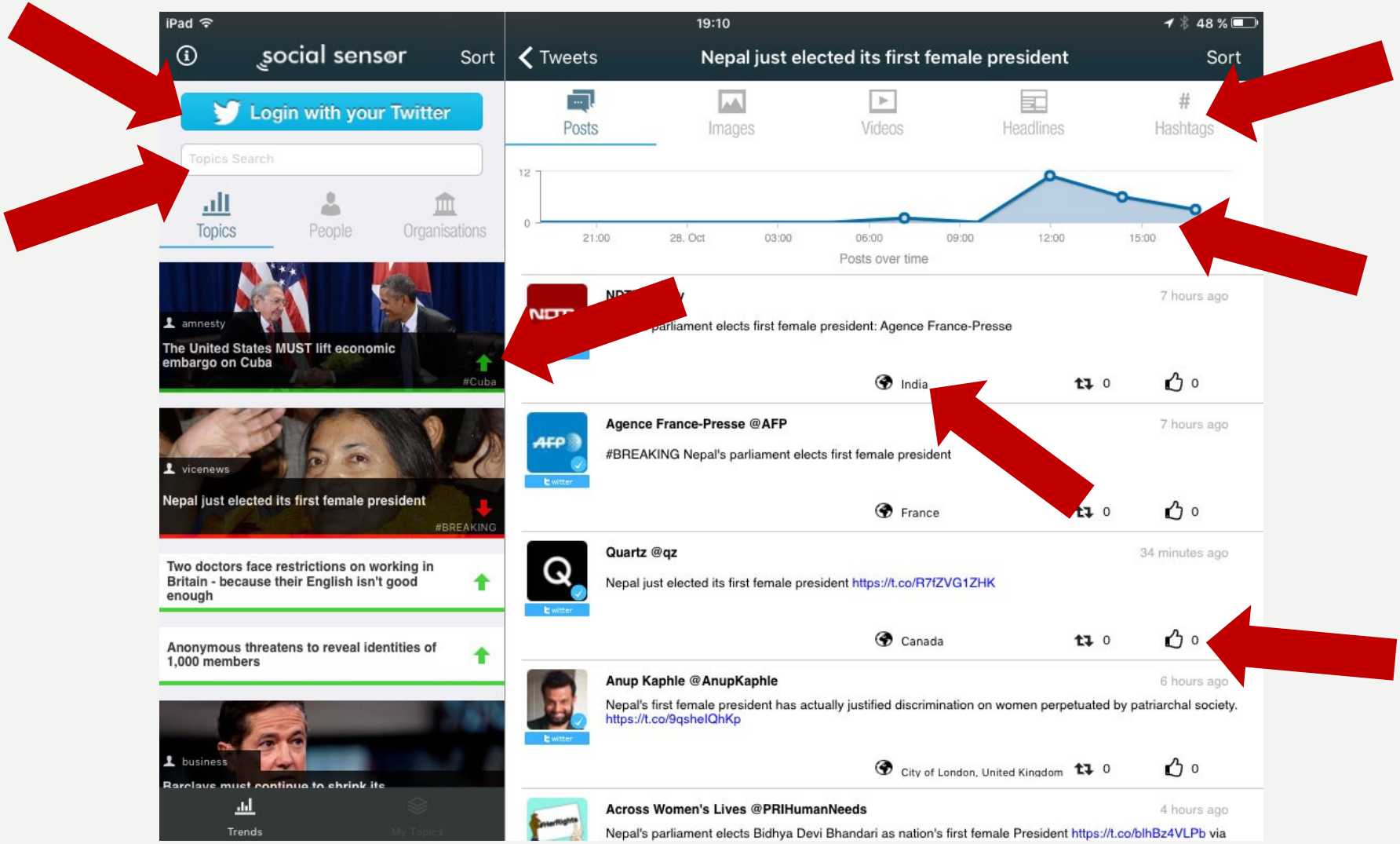
➔ Twitter as an especially timely source for breaking news events

- Interdisciplinary project with researchers, media companies, corporate
- ➔ Goal: application for autonomous detection, clustering, prioritising of news and infotainment from social media



- Works in real time
- Diverse material: text, images, audio and video

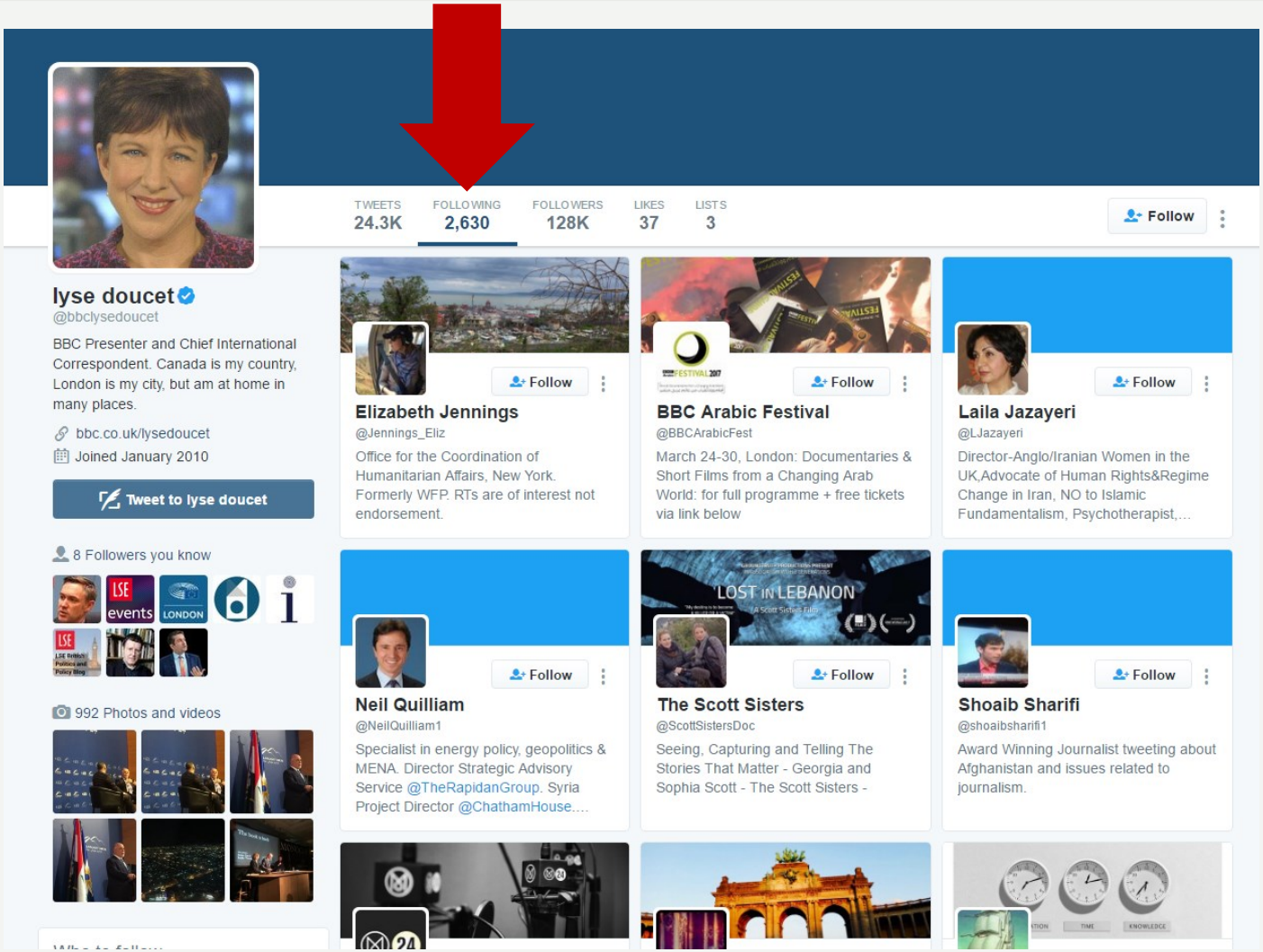
Social Sensor



Twitter: Monitoring accounts via ‚newshounds'

- Possible newshounds are selected (initial seed: 500 journalists, then people they followed)

Social Sensor



Monitoring accounts via ‚newshounds'

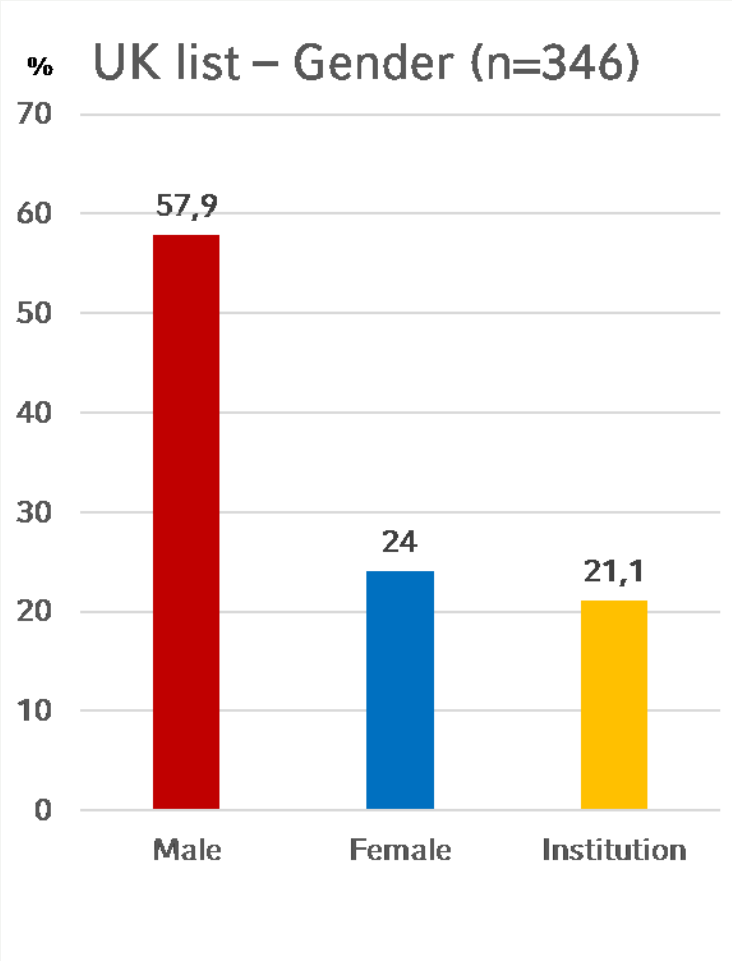
- Possible newshounds are selected (initial seed: 500 journalists, then people they followed)
- Scoring system prioritises newshounds according activity, number of followers, whether they were verified… ➔ different scoring formulas were tried
- Final newshounds database with ca. 6,000 newshounds for news lists (& a celebrity list with 846 for the soft news)

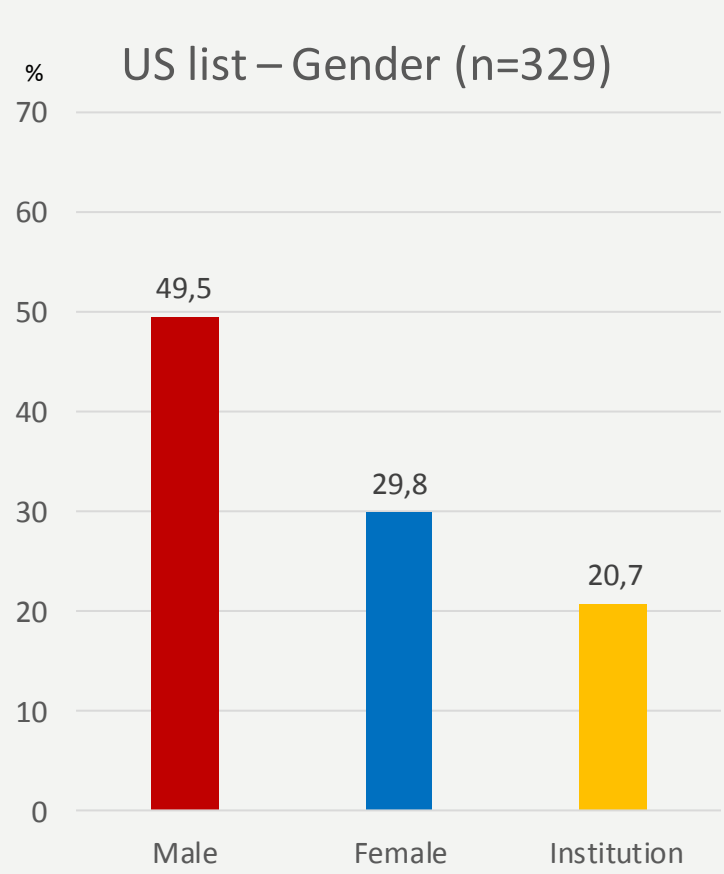
Empircial study of Twitter lists

- **Sample** (random sample, 95 % CI, 5 % MOE)
  - UK news (n=346),
  - US news (n=333)
  - … & Celebrity list (n=260)
    (Moon/Hadley 2014: 300: Twitter is more often used as a source for soft news than for hard news by TV and newspapers)

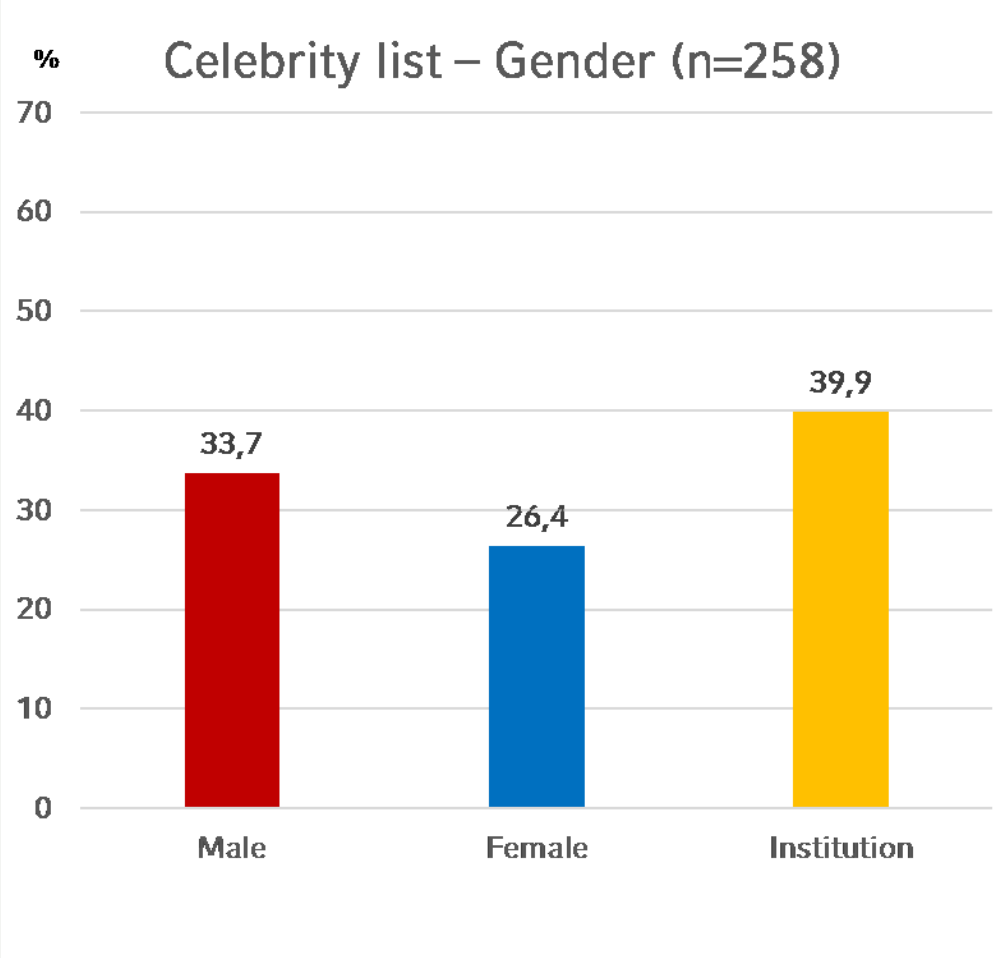
Which characteristics do the newshounds show – and who are thus the sources of Social Sensor?

Results: Male bias

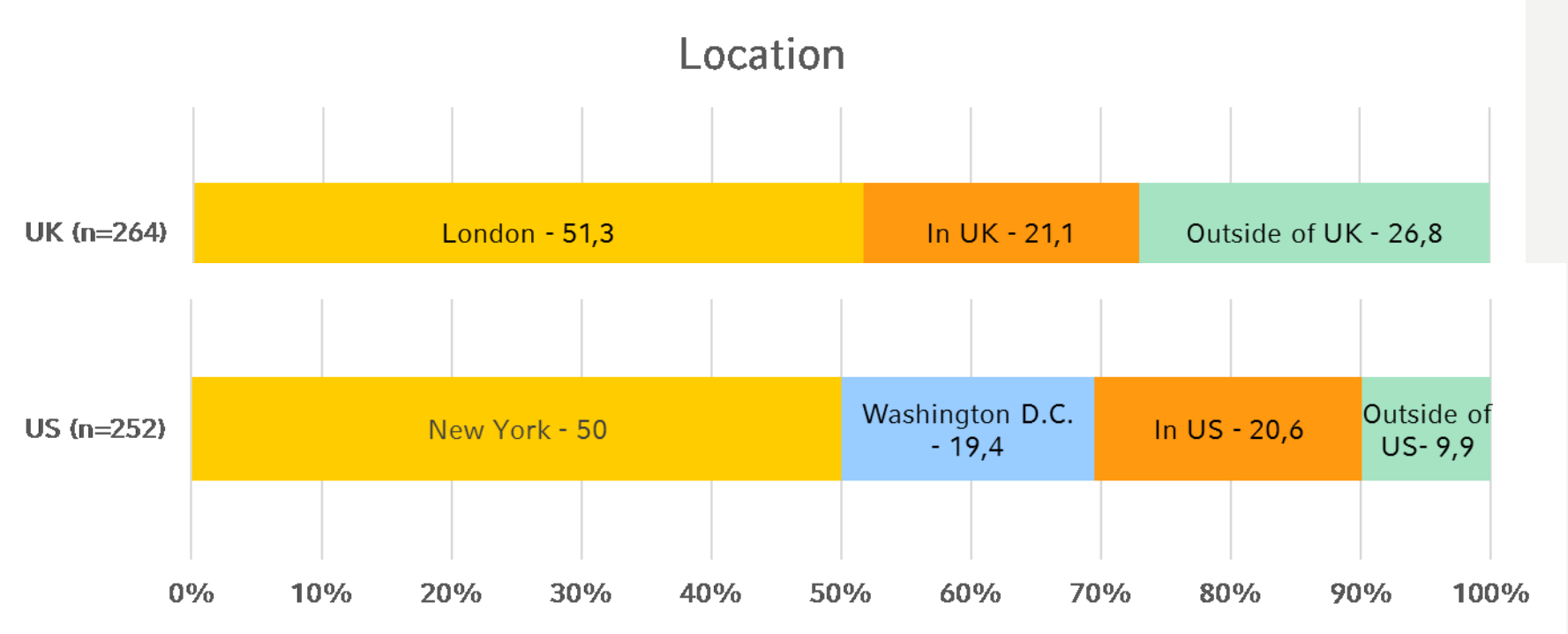




Results: Male bias

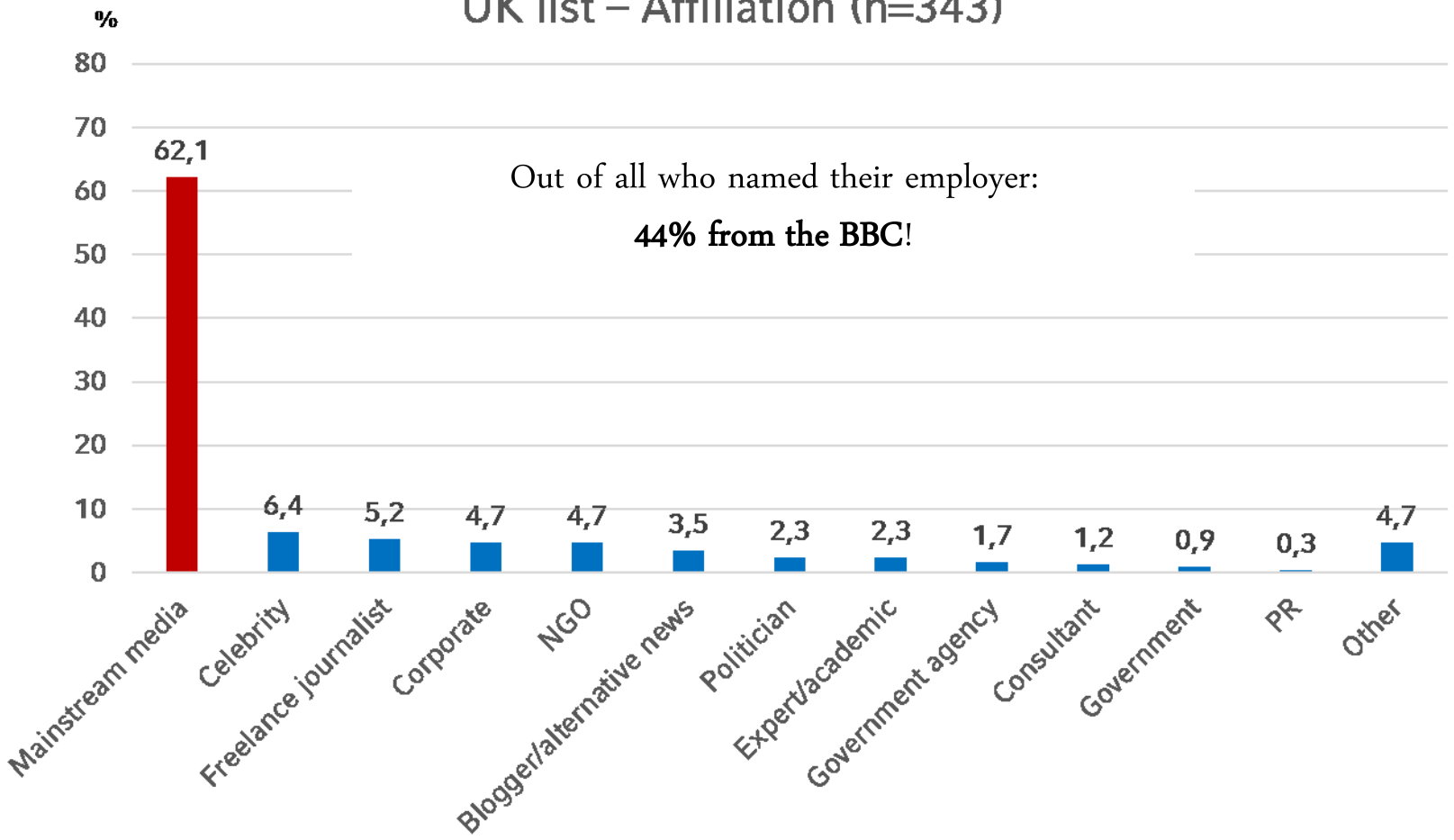


Results: Metropolitan bias



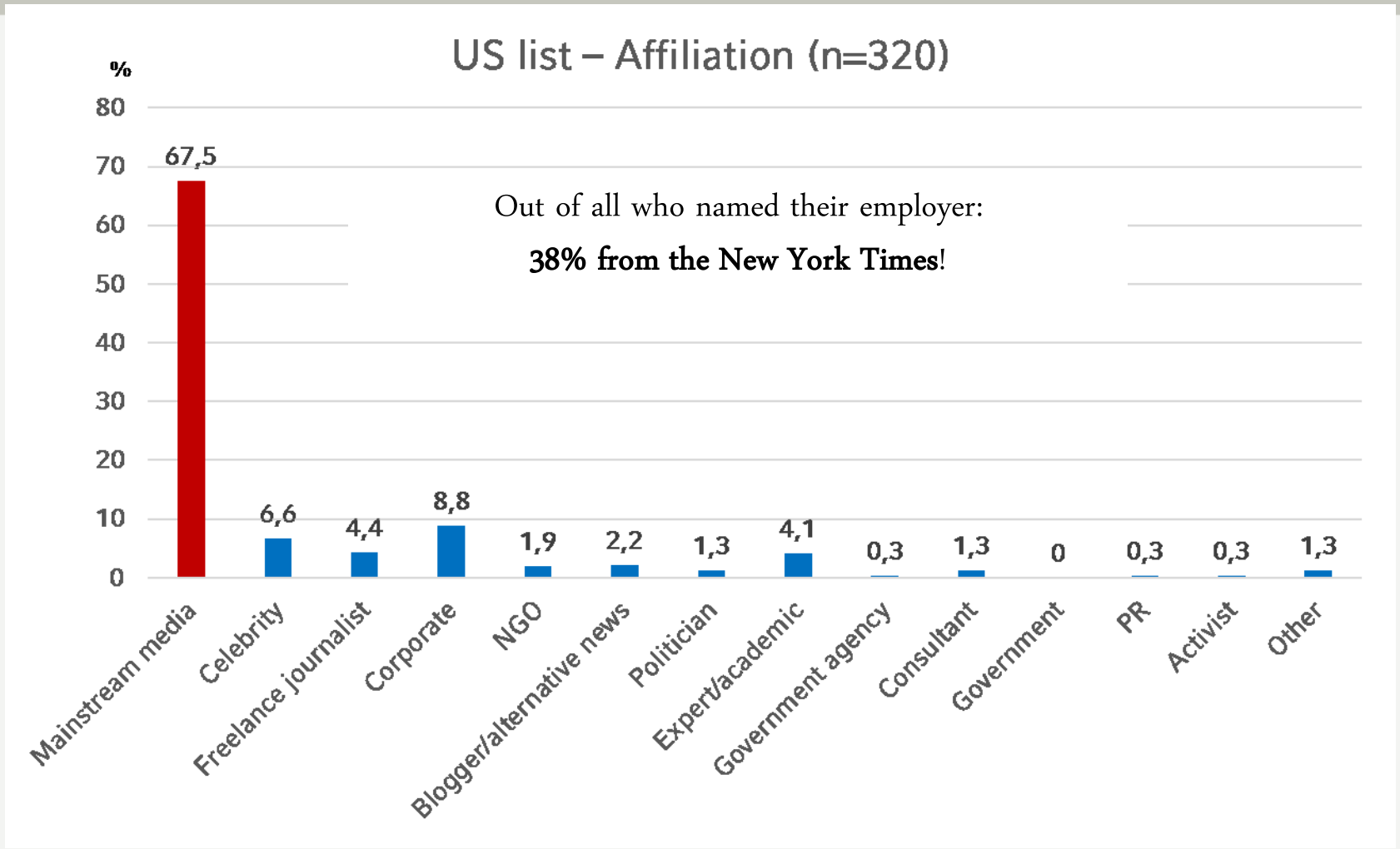
➔ Media cities with the big outlets, but…

Results: Mainstream media bias



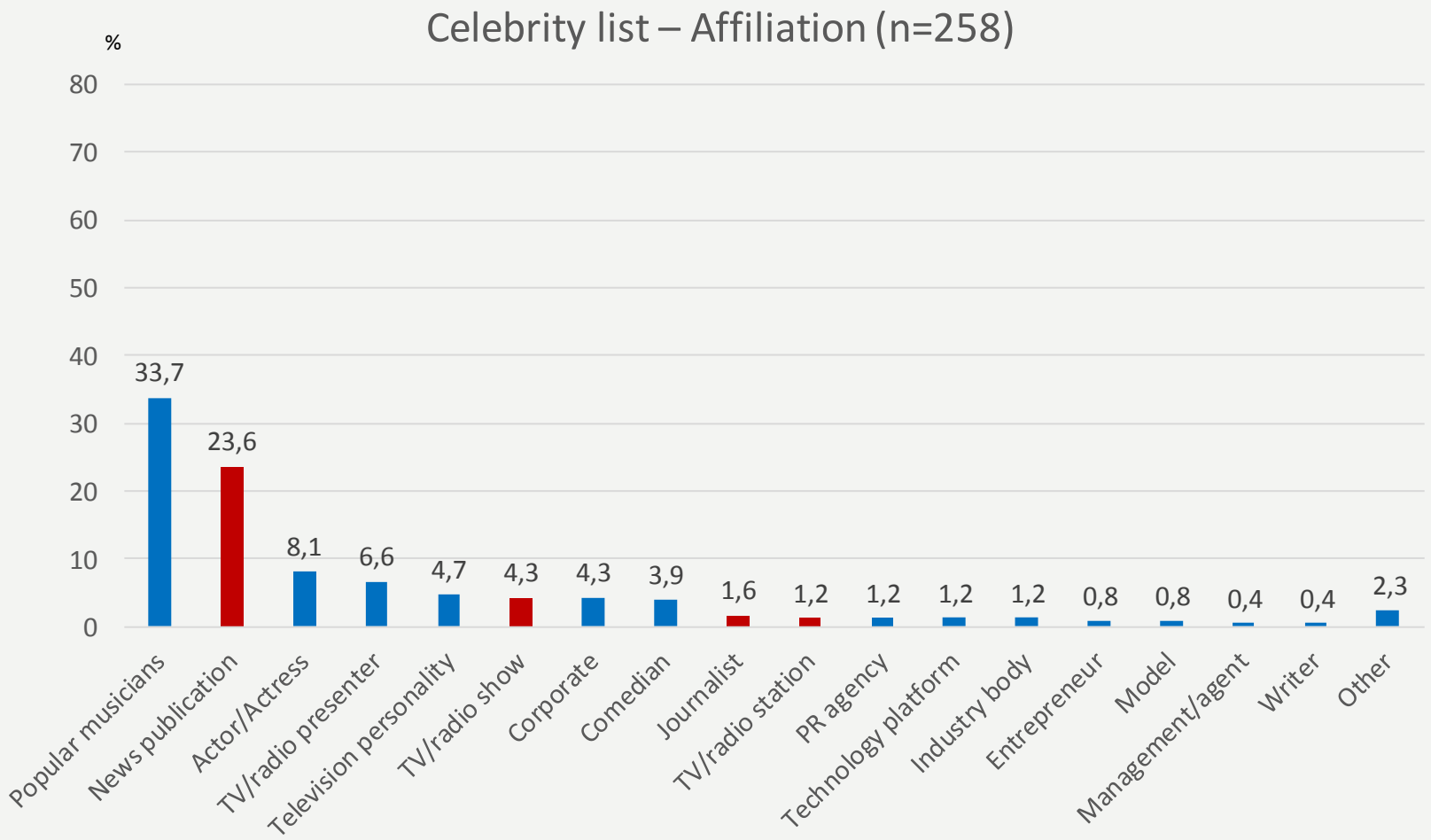
➔ Are the journalists referencing themselves?

Results: Mainstream media bias



➔ Same here!

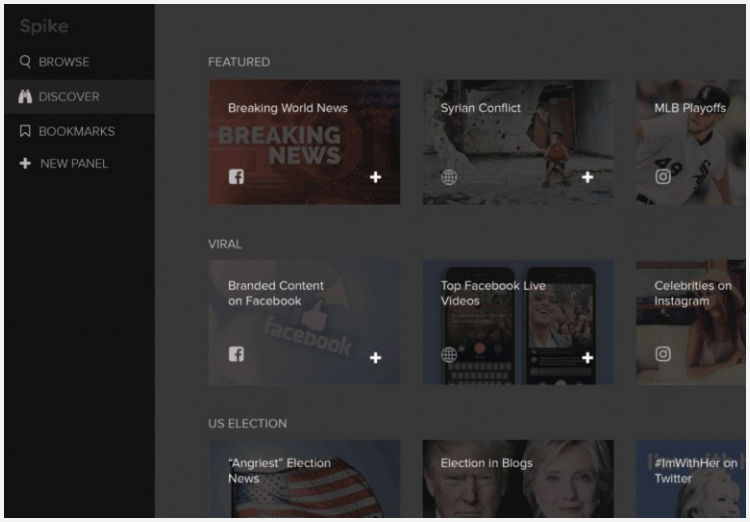
Results: Mainstream media bias

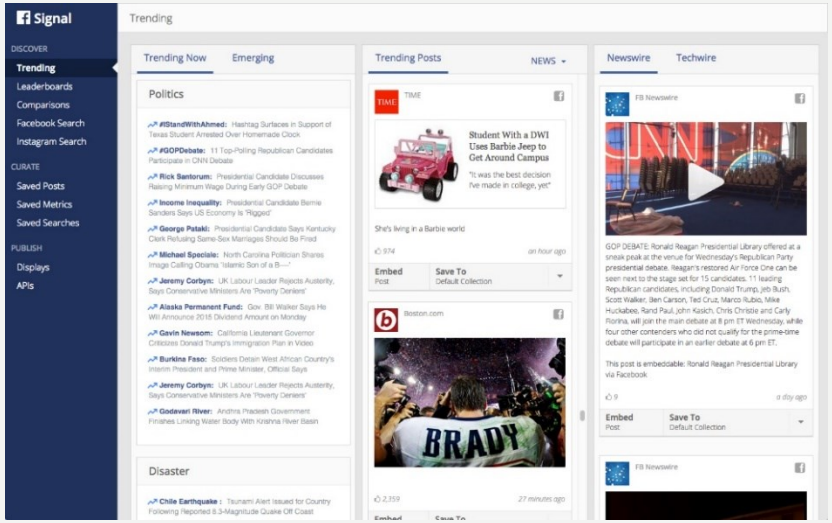


➔ Even here, over 30% are media outlets or journalists!

Other tools for algorithmic news sourcing

- Reuters News Tracer (Twitter)
- NewsWhip Spike (many platforms)
- Dataminr (Twitter)
- Facebook Signal (Facebook & Instagram)
- Bloomberg‘s initiative





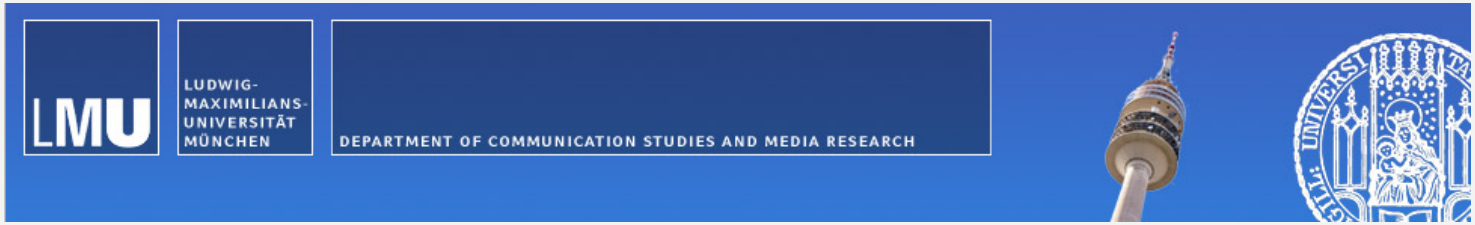
Conclusions

… But: who witnesses whom and how?

- How are the algorithms programmed?
- How to make sure that not the same few people are listened to? What about the “nonelites” (Henderson/Miller 2014)?
- What about an update of the newshound list? (inactive accounts)

Thank you!

Department of Communication Studies and Media Research
Ludwig-Maximilians-Universität München



References

Broersma, Marcel, and Todd Graham. 2013. "Twitter as a news source." *Journalism Practice* 7 (4): 446–64.

Craig, David A. 2014. "The Ethics of Verification in Journalistic Use of Social Media." In *From Twitter to Tahir Square: Ethics in Social Media and New Media Communication. Volume 1*, edited by Bala A. Musa and Jim Willis, 105–18. Santa Barbara: Praeger.

Henderson, Keren, and Andrea Miller. 2014. "Twitter's role in the Modern Newsroom: Circumventing the Gatekeepers and Pounding the Digital Pavement." In *From Twitter to Tahir Square: Ethics in Social Media and New Media Communication. Volume 1*, edited by Bala A. Musa and Jim Willis, 3–19. Santa Barbara: Praeger.

Hermida, Alfred. 2010. "From TV to Twitter: How Ambient News Became Ambient Journalism." Media/Culture Journal 13 (2).

Kyumin, Lee, Brian D. Eoff, and James Caverlee. 2015. *Seven Months with the Devils: A Long-Term Study of Content Polluters on Twitter. Proceedings of the 5th AAAI International Conference on Weblogs and Social Media.* Menlo Park, CA: The AAAI Pess.

Larsson, Anders O., and Hallvard Moe. 2015. "Bots or journalists? News sharing on Twitter." *Communications* 40 (3): 361–70.

Lokot, Tetyana, and Nicholas Diakopoulos. 2016. "News Bots." *Digital Journalism* 4 (6): 682–99.

Mittelstadt, Brent. 2016. "Auditing for Transparency in Content Personalization Systems." *International Journal of Communication* 10: 4991–5002.

Moon, Soo J., and Patrick Hadley. 2014. "Routinizing a New Technology in the Newsroom: Twitter as a News Source in Mainstream Media." *Journal of Broadcasting & Electronic Media* 58 (2): 289–305.

Pear Analytics. 2009. "Twitter Study - August 2009." Accessed March 16, 2017. https://pearanalytics.com/wp-content/uploads/2009/08/Twitter-Study-August-2009.pdf.