to writing. H.R.D. contributed the lava flow models and contributed to writing. K.R.A. produced the DEM, cross-sectional figure, and the supplementary caldera collapse movie and contributed to writing. P.A.N. and C.K. analyzed gas data and contributed to writing. R.L.L. and C.G. contributed petrology analysis. P.L. processed the ALOS-2 interferograms. J.B. and M.H.Z. compiled GIS data and produced the lava flow map. J.L. Babb and J.L. Ball contributed to media organization and writing. A.K.D. processed the DEMs and collected, processed, and managed the UAS data and contributed to writing. I.A.J. and A.M. collected and processed the geodetic data and contributed to writing. M. Pa and C.P. collected and analyzed geologic data. W.A.T. contributed to seismic analysis and conceptualized the manuscript. M.B., P.D., and J.C.C. contributed to seismic analysis. All other authors contributed to the eruption response or writing of the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data access information and additional methodologies are provided in the manuscript or supplementary material and in these references: geologic data (*16*, *39*), seismic data (*40*, *41*), geodetic data (*42*, *43*), and digital elevation data (*44*, *45*).

SUPPLEMENTARY MATERIALS
www.sciencemag.org/content/363/6425/367/suppl/DC1
Materials and Methods
Figs. S1 and S2
References (*47*–*57*)
Movies S1 and S2

11 October 2018; accepted 3 December 2018
Published online 11 December 2018
10.1126/science.aav7046

POLITICAL SCIENCE

Fake news on Twitter during the 2016 U.S. presidential election

Nir Grinberg[1,2]*, Kenneth Joseph[3]*, Lisa Friedland[1]*, Briony Swire-Thompson[1,2], David Lazer[1,2]†

The spread of fake news on social media became a public concern in the United States after the 2016 presidential election. We examined exposure to and sharing of fake news by registered voters on Twitter and found that engagement with fake news sources was extremely concentrated. Only 1% of individuals accounted for 80% of fake news source exposures, and 0.1% accounted for nearly 80% of fake news sources shared. Individuals most likely to engage with fake news sources were conservative leaning, older, and highly engaged with political news. A cluster of fake news sources shared overlapping audiences on the extreme right, but for people across the political spectrum, most political news exposure still came from mainstream media outlets.

In 1925, *Harper's Magazine* published an article titled "Fake news and the public," decrying the ways in which emerging technologies had made it increasingly difficult to separate rumor from fact (*1*). Nearly a century later, fake news has again found its way into the limelight, particularly with regard to the veracity of information on social media and its impact on voters in the 2016 U.S. presidential election. At the heart of these concerns is the notion that a well-functioning democracy depends on its citizens being factually informed (*2*). To understand the scope and scale of misinformation today and most effectively curtail it going forward, we need to examine how ordinary citizens experience misinformation on social media platforms.

To this end, we leveraged a panel of Twitter accounts linked to public voter registration records to study how Americans on Twitter interacted with fake news during the 2016 election season. Of primary interest are three simple but largely unanswered questions: (i) How many stories from fake news sources did individuals see and share on social media? (ii) What were the characteristics of those who engaged with these sources? (iii) How did these individuals interact with the broader political news ecosystem? Initial reports were alarming, showing that the most popular fake news stories in the last 3 months of the presidential campaign generated more shares, reactions, and comments on Facebook than the top real news stories (*3*). However, we do not yet know the scope of the phenomenon, in part because of the difficulty of reliably measuring human behavior from social media data (*4*). Existing studies of fake news on social media have described its spread within platforms (*5*, *6*) and highlighted the disproportionate role played by automated accounts (*7*), but they have been unable to make inferences about the experiences of ordinary citizens.

Outside of social media, fake news has been examined among U.S. voters via surveys and web browsing data (*8*, *9*). These methods suggest that the average American adult saw and remembered one or perhaps several fake news stories about the 2016 election (*8*), that 27% of people visited a fake news source in the final weeks before the election, and that visits to these sources constituted only 2.6% of hard news site visits (*9*). They also show a persistent trend of conservatives consuming more fake news content, with 60% of fake news source visits coming from the most conservative 10% of Americans (*9*). However, because social media platforms have been implicated as a key vector for the transmission of fake news (*8*, *9*), it is critical to study what people saw and shared directly on social media.

Finally, social media data also provide a lens for understanding viewership patterns. Previous studies of the online media ecosystem have found evidence of insulated clusters of far-right content (*10*), rabbit holes of conspiratorial content (*11*), and tight clusters of geographically dispersed content (*12*). We wish to understand how fake news sources were positioned within this ecosystem. In particular, if people who saw content from fake news sources were isolated from mainstream content, they may have been at greater risk of adopting misinformed beliefs.

Data and definitions

Fake news sources

We follow Lazer *et al.* (*13*), who defined fake news outlets as those that have the trappings of legitimately produced news but "lack the news media's editorial norms and processes for ensuring the accuracy and credibility of information." The attribution of "fakeness" is thus not at the level of the story but at that of the publisher [similar to (*9*)].

We distinguished among three classes of fake news sources to allow comparisons of different operational definitions of fake news. The three classes correspond to differences in methods of generating lists of sources as well as perceived differences in the sites' likelihoods of publishing misinformation. We labeled as "black" a set of websites taken from preexisting lists of fake news sources constructed by fact-checkers, journalists, and academics (*8*, *9*) who identified sites that published almost exclusively fabricated stories [see supplementary materials (SM) section S.5 for details]. To measure fake news more comprehensively, we labeled additional websites as "red" or "orange" via a manual annotation process of sites identified by Snopes.com as sources of questionable claims. Sites with a red label (e.g., Infowars.com) spread falsehoods that clearly reflected a flawed editorial process, and sites with an orange label represented cases where annotators were less certain that the falsehoods stemmed from a systematically flawed process. There were 171 black, 64 red, and 65 orange fake news sources appearing at least once in our data.

Voters on Twitter

To focus on the experiences of real people on Twitter, we linked a sample of U.S. voter registration records to Twitter accounts to form a panel (see SM S.1). We collected tweets sent by the 16,442 accounts in our panel that were active during the 2016 election season (1 August to 6 December 2016) and obtained lists of their followers and followees (accounts they followed).

[1]Network Science Institute, Northeastern University, Boston, MA, USA. [2]Institute for Quantitative Social Science, Harvard University, Cambridge, MA, USA. [3]Department of Computer Science and Engineering, University at Buffalo, SUNY, Buffalo, NY, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: d.lazer@northeastern.edu

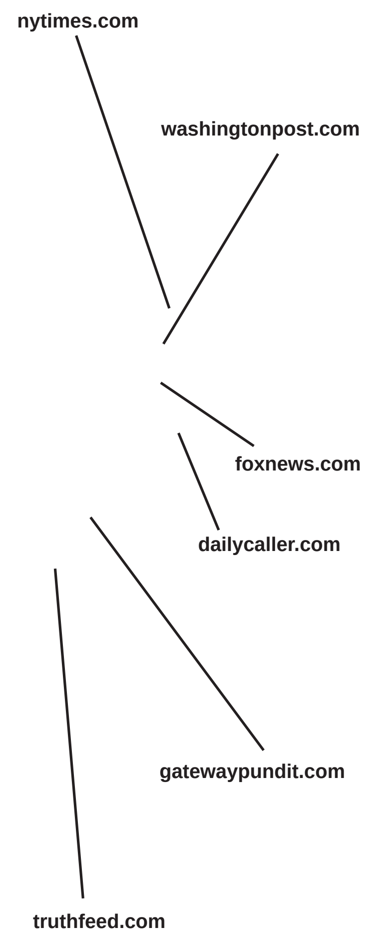

Fig. 5. Coexposure network. Each node is a political news, blog, or fact-checking website. Edges link pairs of sites where an unusually high number of (nonoutlier) panel members were exposed to content from both sites, controlling for the popularity of each site. Filled nodes represent fake news sources. Node colors indicate groups (1, green; 2, orange; 3, purple; 4, gray) identified via an ensemble of clustering algorithms. Sites with the highest exposures are sized slightly larger. See fig. S10 for node labels.

although our sample roughly reflects the demographics of registered voters on Twitter, it might systematically differ from that population in other ways.

Our findings suggest immediate points of leverage to reduce the spread of misinformation. Social media platforms could discourage users from following or sharing content from the handful of established fake news sources that are most pervasive. They could also adopt policies that disincentivize frequent posting, which would be effective against flooding techniques (25) while affecting only a small number of accounts. For example, platforms could algorithmically demote content from frequent posters or prioritize users who have not posted that day. For illustrative purposes, a simulation of capping political URLs at 20 per day resulted in a reduction of 32% of content from fake news sources while affecting only 1% of content posted by nonsupersharers. (SM S.15). Finally, because fake news sources have shared audiences, platforms could establish closer partnerships with fact-checking organizations to proactively watch top producers of misinformation and examine content from new sites that emerge in the vicinity of fake news sources in a coexposure network. Such interventions do raise the question of what roles platforms should play in constraining the information people consume. Nonetheless, the proposed interventions could contribute to delivering corrective information to affected populations, increase the effectiveness of corrections, foster equal balance of voice and attention on social media, and more broadly enhance the resiliency of information systems to misinformation campaigns during key moments of the democratic process.

REFERENCES AND NOTES

1. E. McKernon, Harper's Magazine 151, 528–536 (1925).
2. J. H. Kuklinski, P. J. Quirk, J. Jerit, D. Schwieder, R. F. Rich, J. Polit. 62, 790–816 (2000).
3. C. Silverman, "This analysis shows how viral fake election news stories outperformed real news on Facebook" (2016); www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook.
4. D. Ruths, J. Pfeffer, Science 346, 1063–1064 (2014).
5. S. Vosoughi, D. Roy, S. Aral, Science 359, 1146–1151 (2018).
6. M. Del Vicario et al., Proc. Natl. Acad. Sci. U.S.A. 113, 554–559 (2016).
7. C. Shao et al., Nat. Commun. 9, 4787 (2018).
8. H. Allcott, M. Gentzkow, J. Econ. Perspect. 31, 211–236 (2017).
9. A. Guess, B. Nyhan, J. Reifler, "Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign" (2018); www.dartmouth.edu/~nyhan/fake-news-2016.pdf.
10. Y. Benkler, R. Faris, H. Roberts, Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics (Oxford Univ. Press, 2018).
11. K. Starbird, in Proceedings of the 11th International AAAI Conference on Web and Social Media (AAAI, 2017), pp. 230–239.
12. A. L. Schmidt et al., Proc. Natl. Acad. Sci. U.S.A. 114, 3035–3039 (2017).
13. D. M. J. Lazer et al., Science 359, 1094–1096 (2018).
14. S. Greenwood, A. Perrin, M. Duggan, "Social media update 2016" (Pew Research Center, 2016).
15. Z. Chu, S. Gianvecchio, H. Wang, S. Jajodia, IEEE Trans. Depend. Secure Comput. 9, 811–824 (2012).
16. L. X. Wang, A. Ramachandran, A. Chaintreau, in Workshops of the 10th International AAAI Conference on Web and Social Media (AAAI, 2016), pp. 108–113.
17. P. N. Howard, B. Kollanyi, S. Bradshaw, L.-M. Neudert, "Social media, news and political information during the US election: Was polarizing content concentrated in swing states?" Data Memo 2017.8 (Oxford Project on Computational Propaganda, 2017).
18. S. Bhatt, S. Joglekar, S. Bano, N. Sastry, in Proceedings of the 27th International World Wide Web Conference (ACM, 2018), pp. 545–554.
19. Z. Kunda, Psychol. Bull. 108, 480–498 (1990).
20. N. Dianati, Phys. Rev. E 93, 012304 (2016).
21. E. Mustafaraj, S. Finn, C. Whitlock, P. T. Metaxas, in Proceedings of the 3rd International Conference on Social Computing (IEEE, 2011), pp. 103–110.
22. C. R. Sunstein, #Republic: Divided Democracy in the Age of Social Media (Princeton Univ. Press, 2018).
23. B. Swire, U. K. H. Ecker, S. Lewandowsky, J. Exp. Psychol. Learn. Mem. Cogn. 43, 1948–1961 (2017).
24. Z. Tufekci, in Proceedings of the 8th International AAAI Conference on Weblogs and Social Media (AAAI, 2014), pp. 505–514.
25. M. E. Roberts, Censored: Distraction and Diversion Inside China's Great Firewall (Princeton Univ. Press, 2018).
26. N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, D. Lazer, Public Replication Package for Fake news on Twitter. Zenodo (2019); https://doi.org/10.5281/zenodo.2483311.
27. N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, D. Lazer, Protected Replication data for Fake news on Twitter. Zenodo (2019); https://doi.org/10.5281/zenodo.2485428.

ACKNOWLEDGMENTS

We thank L. Adamic, Y. Benkler, S. McCabe, and the three anonymous reviewers for thoughtful feedback on the manuscript and TargetSmart for access to voter data. The research was approved by Northeastern University's Institutional Review Board. All opinions expressed in this article are those of the authors alone. Funding: D.L. acknowledges support by the ESRC ES N012283/1 and ARO W911NF-12-1-0556. Author contributions: D.L. conceived of the study. N.G., K.J., and L.F. collected and processed data, carried out statistical modeling, and produced visualizations. N.G., K.J., L.F., and B.S.-T. performed literature review and annotated data. All authors devised analyses and wrote and revised the paper. Competing interests: The authors declare no competing interests. Data and materials availability: Aggregate data and code from this study are freely available at Zenodo (26). Deidentified individual-level data are also available at Zenodo (27) upon signing a usage agreement stating that: (i) you shall not attempt to identify, reidentify, or otherwise deanonymize the dataset and (ii) you shall not further share, distribute, publish, or otherwise disseminate the dataset without Northeastern University's prior written approval.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/363/6425/374/suppl/DC1
Supplementary Text
Figs. S1 to S14
Tables S1 to S7
References (28–68)

23 May 2018; accepted 2 January 2019
10.1126/science.aau2706