www.covidstates.org

THE STATE OF THE NATION:
A 50-STATE COVID-19 SURVEY
REPORT #18: COVID-19 FAKE NEWS ON TWITTER

USA, October 2020

David Lazer, Northeastern University
Damian J. Ruck, Northeastern University
Alexi Quintana, Northeastern University
Sarah Shugars, New York University
Kenneth Joseph, University at Buffalo
Nir Grinberg, Ben-Gurion University
Ryan J. Gallagher, Northeastern University
Luke Horgan, Northeastern University
Adina Gitomer, Northeastern University
Aleszu Bajak, Northeastern University
Matthew A. Baum, Harvard University
Katherine Ognyanova, Rutgers University
Hong Qu, Northeastern University
William R. Hobbs, Cornell University
Stefan McCabe, Northeastern University
Jon Green, Northeastern University
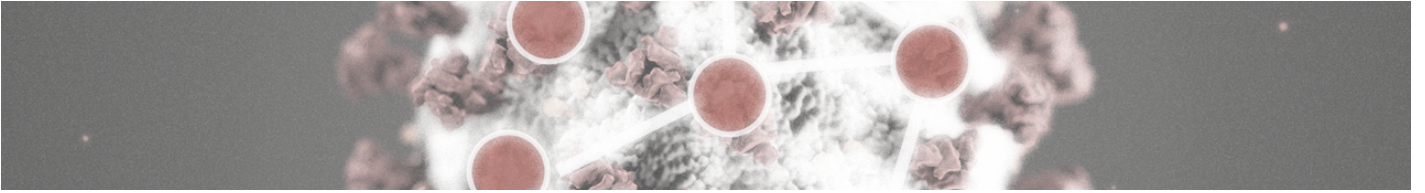
Report of October 21, 2020, v.1

From: The COVID-19 Consortium for Understanding the Public's Policy Preferences Across States

A joint project of:
Northeastern University, Harvard University, Rutgers University, and Northwestern University

Authors: David Lazer (Northeastern University); Damian J. Ruck (Northeastern University); Alexi Quintana (Northeastern University); Sarah Shugars (New York University); Kenneth Joseph (University at Buffalo); Nir Grinberg (Ben-Gurion University); Ryan J. Gallagher (Northeastern University); Luke Horgan (Northeastern University); Adina Gitomer (Northeastern University); Aleszu Bajak (Northeastern University), Matthew A. Baum (Harvard University); Katherine Ognyanova (Rutgers University); Hong Qu (Northeastern University); William R. Hobbs (Cornell University); Stefan McCabe (Northeastern University); and Jon Green (Northeastern University)

Contact information:

For additional information and press requests contact David Lazer at d.lazer@neu.edu
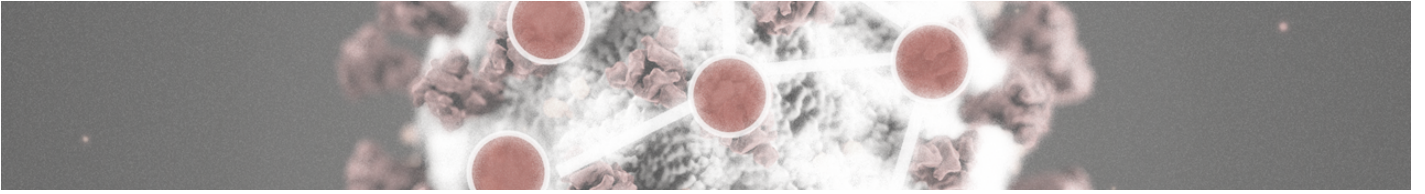
Executive Summary

The COVID-19 pandemic created an unprecedented disruption of everyday life in America that has been exacerbated by the infodemic of COVID-19 misinformation spread on social media.

In this report, we examine the tweets of 1.6 million registered American voters to ask: *who is sharing COVID-19 fake news and what are they sharing*?

We find that older people and Republicans are more likely to share URLs from fake news web domains. In addition, we show that the most shared fake news web domain is The Gateway Pundit, which received far more shares than even the second most popular fake news domain.

Note on methods

Between January 1st and September 30th, 2020, we collected COVID-19 related tweets from a sample of registered voters in America. We examined the content posted by a list of accounts matched to demographic information such as age, race, gender and political party affiliation.[1] Our full panel contains 1.6 million accounts, of which 527,958 tweeted about COVID-19. The total number of COVID-19 tweets is 29,662,169.

We retained only COVID-19 tweets by filtering using a broad list of 974 multi-lingual keywords, phrases and hashtags related to COVID-19. The keyword list contains words directly related to COVID-19 (e.g. coronavirus, COVID-19) and also those related to phenomena that occurred as a result of the virus (e.g. "reopening"). A tweet was included in the sample if at least one item from our list is contained in the tweet text, quoted text, hashtag or any part of the URL string – this does not include the content from the linked web page. For more details on COVID-19 tweet selection, see Gallagher et al 2020.[2]

We then collected the COVID-19 URLs shared by our panel and classified them as originating from a web domain that we classified as either fake, not fake or unknown. Of the 29,662,169 COVID-19 tweets in our panel, 7,062,989 contained at least one URL. The fake news classifications were based on Grinberg et al,[3] which has a three tier classification system for fake news domains: orange, red and black. Here we classify "black" and "red" domains as fake, where "black" domains are "a set of websites taken from preexisting lists of fake news sources," and "red" domains "spread falsehoods that clearly reflect a flawed editorial process." We do not include domains classified as "orange" because the authors "were less certain that the falsehoods stemmed from a systematically flawed process."

Fake classifications were made for web domains, so every URL from a fake news domain is classified as such, regardless of the content of the particular URL. We correct for repeat sharers -- that is, people who share the same URL multiple times – by counting each URL at most one time per user, regardless of how many times they share it. We also remove URLs from platforms, such as YouTube, Facebook, Instagram, etc., so our sample contains mainly news domains.

The findings presented below are the product of a collaboration centered at the [Lazer Lab](#) at the Network Science Institute, Northeastern University, Boston. All analyses were done on aggregate data using data collection protocols reviewed by Northeastern University's Institutional Review Board.

To explore the data yourself, see our [COVID-19 tweets data dashboard](#).

---

[1] https://targetearly.targetsmart.com/modeling.html

[2] Gallagher, R. J., Doroshenko, L., Shugars, S., Lazer, D., & Foucault Welles, B. (2020). *Sustained Online Amplification of COVID-19 Elites in the United States*. http://arxiv.org/abs/2009.07255

[3] Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, *363*(6425), 374–378.

COVID-19 fake news sharing on Twitter

Social media acts as a conduit for fake news websites, where we define fake news as information that mirrors legitimate news in form, but "lacks the news media's editorial norms and processes for ensuring the accuracy and credibility of information."[4] During the 2016 election, for example, many researchers and journalists alike failed to track the weaponization of misinformation on social media, leaving them to retroactively discover which demographics had been most active in sharing fake news after the election had taken place. It is important to understand in real time which parts of the population are sharing fake news on Twitter.

The COVID-19 pandemic has been a once-in-a-generation disruption for Americans. According to the CDC, by October 2020, there were over 8 million cases of COVID-19 and over 200,000 deaths.[5] The consequences for Americans have been wide-ranging, from coping with staggering numbers of illnesses and deaths, to restrictions on freedom of movement, to mass unemployment and economic crisis. There has been a great deal of confusion and misinformation surrounding COVID-19 – a so-called Infodemic[6] – with much of it occurring online. The BBC documented the direct costs of COVID-19 misinformation, which include alcohol and cleaning product poisonings, assault, property damage and heightened racism.[7]

In our panel, between January 1st and September 30th, 2020, the vast majority of shared URLs from COVID-19 tweets are either from known, reputable domains (60%), or domains with unknown quality (38.9%). URLs from domains that publish fake news only comprise 1.1% of the URLs from COVID-19 tweets. However, this is likely an underestimate because if we include web domains rated as "orange" in our fake news classification system, the percentage of shared fake news URLs increases to 1.8%.

---

[4] Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096.

[5] https://covid.cdc.gov/covid-data-tracker/#cases_totalcases, accessed on October 8th 2020

[6] Lancet Infectious Diseases Editorial Board. (2020). The COVID-19 infodemic. *The Lancet Infectious Diseases*, *20*(8), 875.

[7] https://www.bbc.co.uk/news/stories-52731624

Age and political identity differences in fake news sharing

We first look at age differences in COVID-19 sharing by splitting our sample into four age bands: 18-29, 30-49, 50-64 and 65 and over. **Older registered voters share the most URLs per capita on the topic of COVID-19 (see Figure 1)**; this is true for all political orientations. For example, political independents aged 18-29 shared 5.7 URLs per capita, those aged 30-49 shared 12.4 URLs per capita, those aged 50-64 shared 27.7 URLs per capita and those aged over 65 shared 54.1 URLs per capita.

In addition, **older voters also share the most URLs from fake news domains (see Figure 2);** again, this is true across political orientations**.** For instance, political independents aged 18-29 shared 0.01 fake news URLs per capita, those aged 30-49 shared 0.09 per capita, those aged 50-64 shared 0.48 per capita, and those over 65 shared 1.98 per capita.

Figure 1 - shared COVID-19 URLs per capita, by party and age

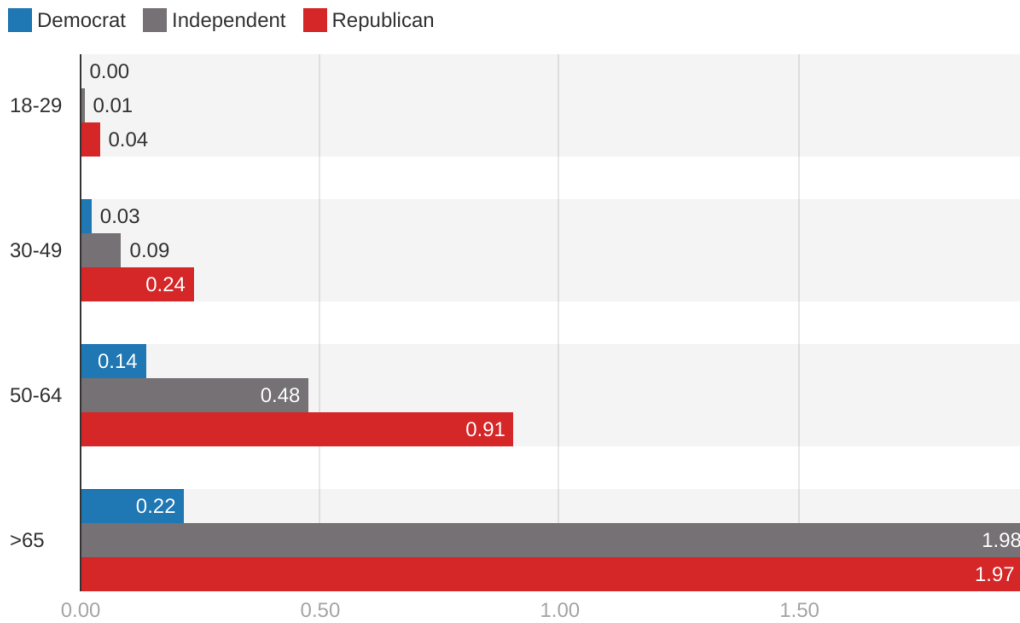Per capita number of URLs shared in COVID-19 tweets, corrected for users who repeatedly share the same URL
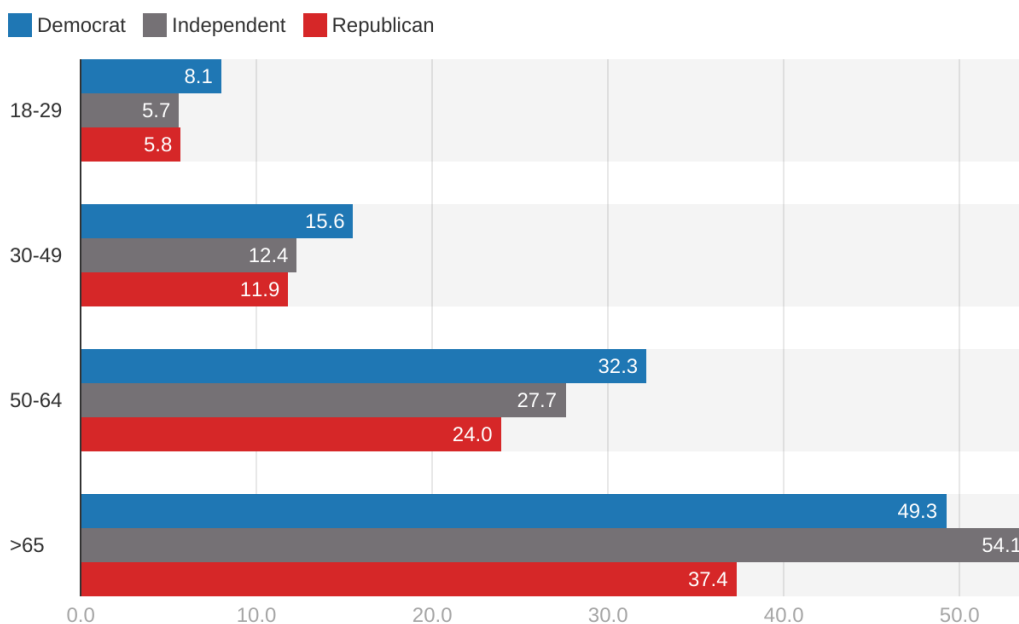
Total number of COVID-19 tweets with URLs = 7,062,989, Time period 01/01/2020-09/30/2020
Source: Lazer Lab at the Network Science Institute, www.covidstates.org • Created with Datawrapper

We next look at the percentage of COVID-19 URLs shared by each age group that come from fake news domains. **These results further suggest that older voters** – **across all political orientations** – **share more URLs from fake news domains than younger voters (see Figure 3).** For political independents, 0.2% of all URLs shared by those aged 18-29 come from fake domains; for those aged 30-49, it is 0.7%; for those aged 50-64, it is 1.7%; for those over 65, it is 3.7%.

Figure 2 - per capita number of shared URLs from fake news domains, by party and age

Per capita number of URLs from web domains classified as fake news shared in COVID-19 tweets
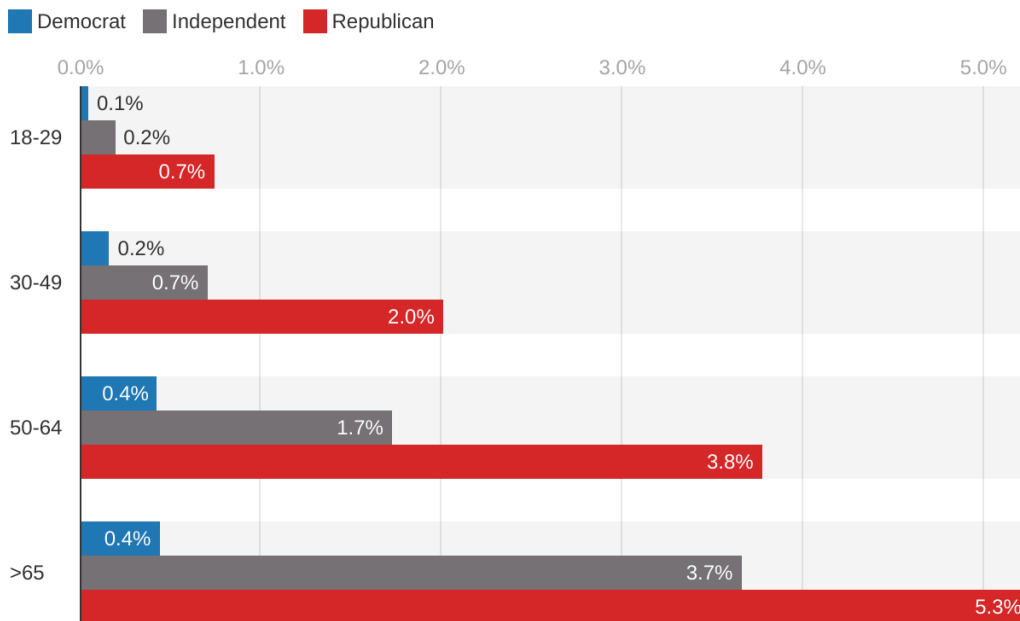

■ Democrat ■ Independent ■ Republican

18-29
- 0.00
- 0.01
- 0.04

30-49
- 0.03
- 0.09
- 0.24

50-64
- 0.14
- 0.48
- 0.91

>65
- 0.22
- 1.98
- 1.97

Total number of COVID-19 tweets with URLs = 7,062,989, Time period 01/01/2020-09/30/2020
Source: Lazer Lab at the Network Science Institute, www.covidstates.org • Created with Datawrapper

Figure 3 - percentage of shared URLs from fake news domains, by party and age

Percent of COVID-19 tweets containing URLs from web domains classified as fake news

■ Democrat ■ Independent ■ Republican

18-29
- 0.1%
- 0.2%
- 0.7%

30-49
- 0.2%
- 0.7%
- 2.0%

**50-64**
- 0.4%
- 1.7%
- 3.8%

>65
- 0.4%
- 3.7%
- 5.3%

Total number of COVID-19 tweets with URLs = 7,062,989, Time period 01/01/2020-09/30/2020
Source: Lazer Lab at the Network Science Institute, www.covidstates.org • Created with Datawrapper

It is interesting to contrast this with [survey findings](#) from the COVID-19 Consortium for Understanding the Public's Policy Preferences Across States,[8] which showed **that younger people, regardless of political orientation, are more likely to believe one of eleven pieces of COVID-19 misinformation when compared to older people (see Figure 4)**. Democrats between the ages of 18 and 29 believe, on average, 1.92 pieces of COVID-19 misinformation, while Democrats over 65 believe only 0.87 pieces. Similarly, Republicans between the ages of 18 and 29 believe 2.26 pieces of COVID-19 misinformation, whereas Republicans over 65 only believe 1.21 pieces.

Figure 4 - belief in misinformation by party and age

Average number of COVID-19 misinformation claims (out of eleven) believed by survey participants
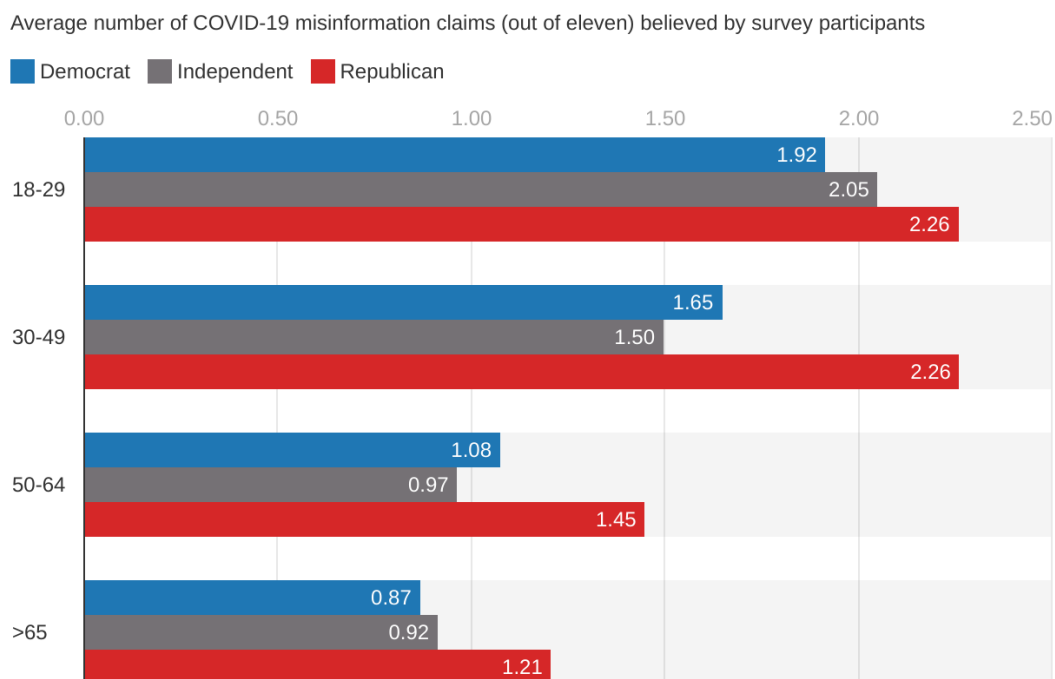
National Sample, N=21,196, Time Period: 08/07/2020-08/28/2020

Source: The COVID-19 Consortium for Understanding the Public's Policy Preferences Across States (A joint project of: Northeastern University, Harvard University, Rutgers University and Northwestern University) www.covidstates.org
 • Created with Datawrapper

These findings suggests that while older people share more misinformation, they are actually more informed. This seeming paradox of mismatched belief and sharing patterns with respect to age may indicate that Twitter users have different belief patterns than the general population, or may point to an age-related mismatch between believing and sharing on Twitter. The driving force(s) behind this phenomenon will be the subject of future research.

---

[8] Baum, M. A., Ognyanova, K., Chwe, H., Quintana, A., Perlis, R. H., Lazer, D., Druckman, J., Santillana, M.,Lin, J., Della Volpe, J., Simonson, M., & Green, J. (2020). *The State of the Nation: A 50-State COVID-19 Survey Report #14: Misinformation and Vaccine Acceptance USA*. www.covidstates.org

For political groups, there are small differences in the per capita number of shared URLs of all kinds in COVID-19 tweets, with **Democrats generally sharing more COVID-19 URLs than Republicans across all age groups (see Figure 1)**. However**, Republicans share far more URLs from fake news domains per capita (Figure 2),** which is true across all age groups except for over-65-year-olds, for whom Independents share slightly more fake news URLs per capita.

We next look at the proportion of total URLs shared by each political group that come from fake news domains. We find that **registered Republicans share proportionally more URLs from fake news domains than registered Democrats and Independents (see Figure 3)**, which is true for all age groups**.** For those aged over 65**,** about 0.4% of the URLs shared by Democrats are from fake news domains; the corresponding numbers for Independents and Republicans are 3.7% and 5.3%, respectively.

It is notable that in 2020 registered Republicans and older people are more likely to share URLs from fake news domains on the topic of COVID-19, as the same demographics were also more likely to do so in general during the 2016 presidential election.[9] This time, however, we have identified these demographics ahead of the election. This comes with the caveat that it is actually younger people who are more likely to believe COVID-19 misinformation, even though they are less likely to share it on Twitter.

Most-shared fake news domain

We examine the URLs from our sample's COVID-19 tweets to quantify which fake news domains are shared most often. The **fake news domain with the most shares is the Gateway Pundit**, which had 7.6 shares per 1000 COVID-19 URLs in March, and has now increased to 12.1 shares per 1000 COVID-19 URLs in September**. Since March, the Gateway Pundit has received an order of magnitude more shares than the second most shared fake news domain (see Figure 5)**.

The popularity of the Gateway Pundit is striking even when taking into account all shared web domains, including those that are reputable or of unknown quality. **In August and September respectively, the Gateway Pundit was ranked the 4th and 6th most shared domain (see Figure 6)**. In August, the only domains with more shares were the New York Times, the Washington Post and CNN, all of which are mainstream media.

---

[9]Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, *363*(6425), 374–378.

Figure 5 - Shares per 1000 URLs, top 5 fake news domains

Monthly URL shares per 1,000 for the top 5 most shared fake news domains
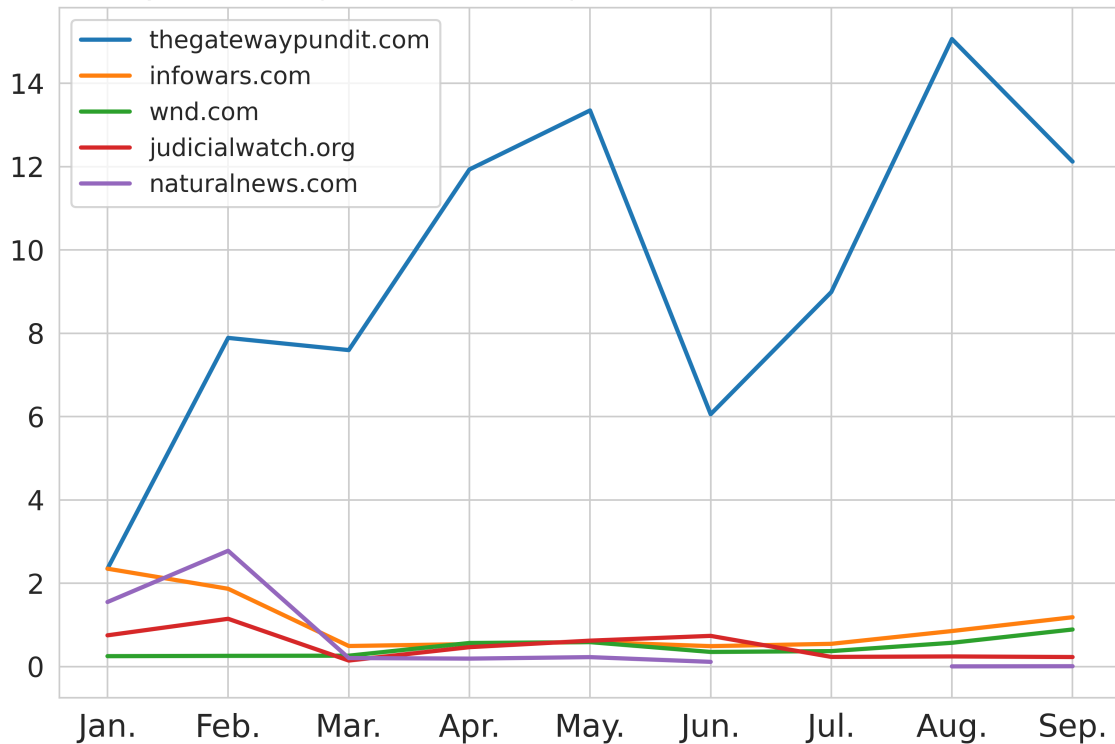
Total number of COVID-19 tweets with URLS = 7,062,989, TIme period 01/01/2020-09/30/2020
Source: Lazer Lab at the Network Science Institute, www.covidstates.org

Finally, it seems that the Gateway Pundit has increased in popularity during the COVID-19 pandemic, with *Similar Web* showing visits at 15.6 million in April, and 26.2 million in September.[10] Moreover, information from *Alexa.com* shows that Twitter is the most likely website to direct visitors to the Gateway Pundit.[11]

These data show that many registered voters in our panel are sharing COVID-19 stories from the Gateway Pundit. We might ask: what does this domain do differently compared to other non-credible domains? To what extent does its demonstrated popularity extend beyond the topic of COVID-19?

---

[10] https://www.similarweb.com/website/thegatewaypundit.com/

[11] https://www.alexa.com/siteinfo/thegatewaypundit.com#section_traffic shows that 10.9% of visitors to thegatewaypundit.com had just visited twitter.com. This is true for the 60 days preceding October 8th, 2020.

Figure 6 - Shares per 1000 URLs, overall top 50 domains

Shares per 1000 URLs for the Gateway Pundit, compared to the other top 50 most popular domains (top 3 are labeled). The red circles contain the monthly popularity ranking of the Gateway Pundit.
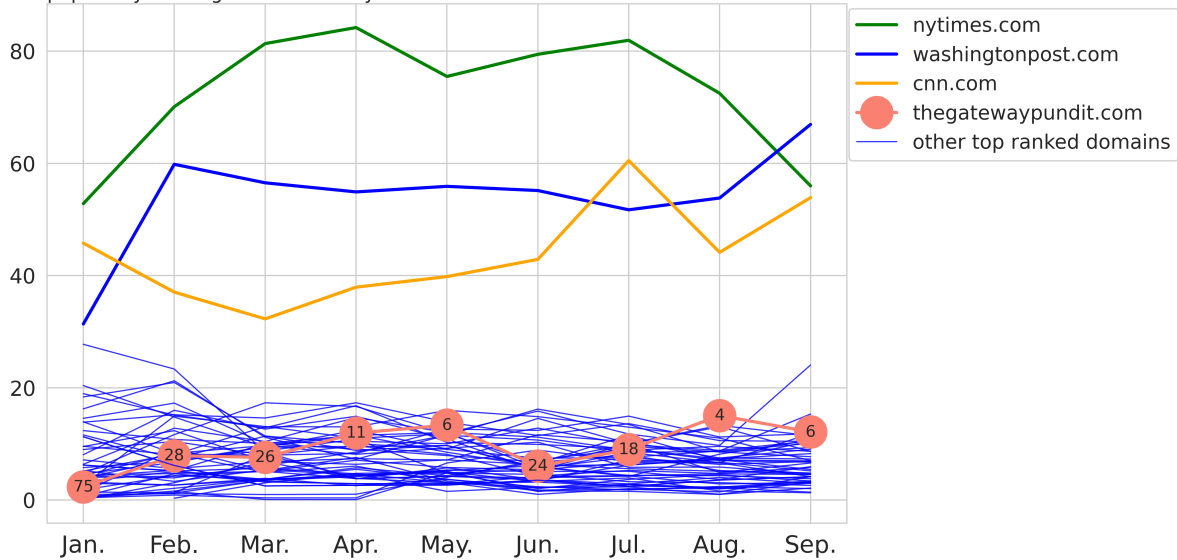
Total number of COVID-19 tweets with URLS = 7,062,989, TIme period 01/01/2020-09/30/2020
Source: Lazer Lab at the Network Science Institute, www.covidstates.org

Conclusion

COVID-19 has been a massive disruption in 2020, exacerbated by the spread of misinformation on social media. In this report we ask *who is sharing COVID-19 fake news on Twitter, and which URLs are being shared?* We show that older people are more likely to share URLs from fake news domains, despite being less likely to believe COVID-19 misinformation. We also show that Republicans are more likely to share fake news than Democrats and Independents. Finally, the fake news domain that has been shared the most in COVID-19 tweets is the Gateway Pundit, which, throughout the COVID-19 pandemic, has received an order of magnitude more shares than even the second most popular fake news domain.