Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Social Anatomy of Climate Change Denial in the United States

Jianxun Yang
Nanjing University
Dimitrios Gounaridis
University of Michigan-Ann Arbor https://orcid.org/0000-0003-3145-7284
Miaomiao Liu
Nanjing University
Joshua Newell (jpnewell@umich.edu)
jpnewell@umich.edu

Article

Keywords:

Posted Date: December 7th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2163106/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

The Social Anatomy of Climate Change Denial in the United States

Abstract

Using Twitter data, this study evaluates and maps climate change denialism across the United States. We estimate that 14.8% of Americans do not believe in climate change. This denialism highest in the central and southern U.S. However, it also persists in clusters within states where belief in climate change is high. Political affiliation was the strongest determinant, followed by level of education, COVID-19 vaccination rates, carbon intensity of the regional economy, and income. A coordinated social media network in the Twittersphere uses periodic events, such as cold weather and climate conferences, to sow disbelief about climate change and science in general. Donald Trump was the most influential, followed by conservative media outlets, and right-wing activists. As a form of knowledge vulnerability, this denialism renders communities unprepared to take steps to increase resilience. We recommend that social media companies flag accounts that spread climate misinformation and initiate targeted educational campaigns.

Main Text

Climate change denialism persists in the United States, with estimates ranging from 12% to 26% of the U.S. population.[1,2] It is more pronounced in some states and regions.[3] Reasons for this denialism are multifaceted: Political affiliation and ideology, income, education, and exposure to extreme weather events are important factors.[4-6] Denialism is more prevalent where local economies are highly dependent on fossil fuels,[7] in rural communities, and in populations where mistrust in science is pronounced.[8,9] Social media reaches millions of users, providing a key mechanism for influencers to spread misinformation and contributing to the persistent segmentation of populations.[10] The ability of social media to influence and cement attitudes was apparent in the response to the vaccines for COVID-19.[11]

Understanding how and why climate change opinion varies geographically and documenting it at an actionable scale is crucial for the success of communication campaigns, outreach, and other interventions.[12,13] Most estimates of the extent and geographic configuration of climate change denialism rely primarily on national surveys, with the Yale Climate Opinion Survey the only dataset that provides estimates at the state and county levels for the entire U.S.[3] These survey efforts, however, are time-intensive and expensive. The Yale Survey combines data from more than 2,500 national surveys and uses multinomial regression modeling to downscale estimates to subnational levels. Independent representative surveys conducted in states and metropolitan areas validate the predictions from the Yale Survey models.[3]

Mining social media data (for example, Facebook, YouTube, and Twitter) is a tantalizing alternative to survey-based approaches.[14,15] Twitter is a social media platform with an extensive data repository. By adjusting for the skew toward certain demographic groups in Twitter users, data from this platform is useful for estimating public views on an array of topics, such as politics, social issues, and COVID-19 vaccination rates.[16,17] Data from Twitter has also been used in predictive modeling of election outcomes.[18] Account holders can misuse Twitter to oppose scientific knowledge and spread misinformation.[19]

This study harnessed Twitter data to (i) estimate the prevalence of climate change denialism at the state and county level, (ii) identify characteristics of climate change deniers, (iii) understand how social media promulgates climate change denialism including the key influencers, and (iv) determine how world events are leveraged to promulgate climate change attitudes

To answer these questions, we used a Deep Learning text recognition model to classify 7.4 million geocoded tweets, collected between September 2017 to May 2019, containing keywords related to climate change posted by 1.3 million unique users in the U.S (see Online Methods). We classified these tweets about climate change into ‘for’ (belief) and ‘against’ (denial). Our analysis resulted in a profile of climate change deniers at the county level, insight into the networks of social media figures influential in promoting climate change denial, and knowledge of how these influencers use current events to foster this denial.

After confirming the validity of using Twitter data instead of survey data to capture public opinion on climate change at policy-relevant geographical scales, we found that denialism clusters in particular regions (and counties) of the country and amongst certain socio-demographic groups. Our analysis showed that politicians, media figures, and conservative activists promulgate misinformation in the Twittersphere and that denialists and climate change believers formed mostly separate Twitter communities, creating echo chambers. Such information provides a basis for developing strategies to counter this knowledge vulnerability and reduce the spread of mis- or disinformation by targeting the communities most at risk for failing to take steps to increase resilience to the effects of climate change.

Results

Where in the U.S. is climate change denial prevalent?

We found that 14.8% of Americans deny that climate change is real (**Fig. 1A**), a percentage consistent with previous national studies (**Fig. S4**). Using geolocation information, we determined that denialism was highest in the Central part of the U.S. and in the South, with more than 20% of the populations of OK, MS, AL, and ND consisting of deniers. Along the West and East Coasts and New England, belief in climate change was highest. However, climate change denial varies substantially within states, often clustering in geographic swaths across multiple counties (**Fig. 1B**). For example, in Shasta County, California climate change denial was as high as 52%; yet overall less than 12% of the population of California were climate change deniers. Similarly, the average percentage of deniers was 21% in Texas, but the county-level ranged from 13% in Travis County to 67% in Hockley County.

To validate these results, we compared them to the Yale Climate Opinion Surveys at the national, state, and county levels (**Fig. S5**). The mean absolute difference between the two models was 3.0 percentage points (S.D. = 2.7) with the Twitter data yielding a higher percentage of deniers (**Fig. S5A**). Compared to the Yale Survey, our model showed higher proportions of deniers in Southern states (for example, MS, AL, TN, and TX). However, state-level and county-level percentages of believers and deniers were highly correlated between the two datasets ($p < 0.001$) (**Fig. S5B – E**).

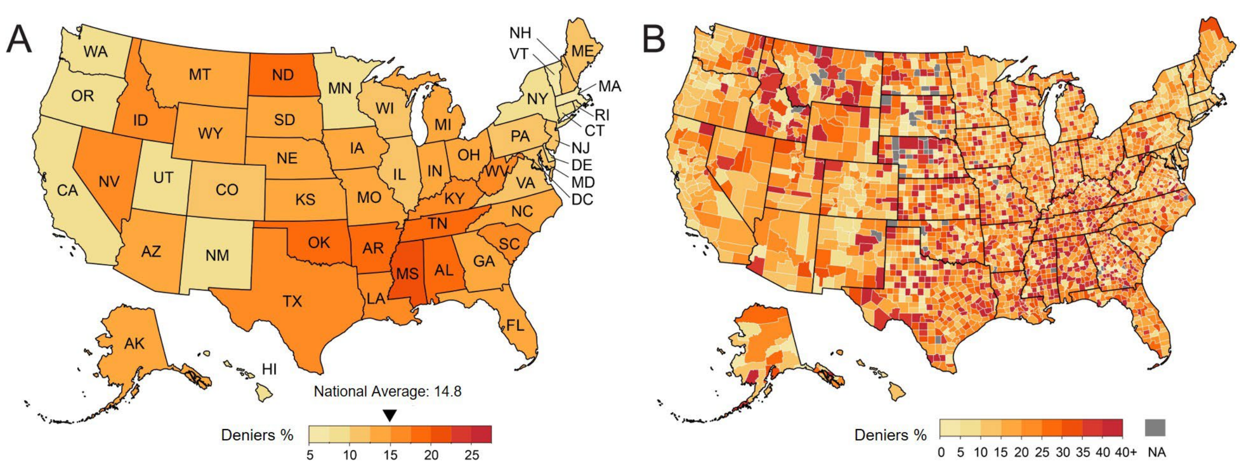

Fig. 1. Climate change denialism in the United States, by state (A) and county (B).

What type of people are climate change deniers?

We performed bivariate correlation analysis with data from multiple publicly available sources (see Online Methods) to characterize climate change deniers (**Table 1**). We evaluated the following characteristics of populations in those regions that were associated with the Twitter profiles for a positive or negative correlation with climate change denial: Political affiliation, race or ethnicity, median income, college education, COVID-19 vaccination rate (proxy for belief in science in general), carbon-intensive economies reliant on fossil fuels, rural or urban county, and local weather patterns (Table 1). At both the county and state levels, populations with a high percentage of Republican voters had the strongest correlation with climate change deniers. Carbon dependency of the economy was also significantly high at the state level. The strongest negative correlations at both state and county levels were educational attainment and COVID-19 vaccination rates. Integrating these data into a weighted least squares regression model, we defined a profile of a "typical" climate change denier (Table 2). The profile had the following characteristics: Republican without a college degree and without COVID-19 vaccination living in an area with a high average annual temperature, such as southern states.

Table 1. State- and county-level weighted Pearson correlations. Total number of tweets per county and per state were used as the universal weights in the model.

	State level		County level	
	Correlation	p value	Correlation	p value
Political Affiliation (Republican)	0.86	<0.001	0.63	<0.001
Education (Population % with a College Degree)	-0.79	<0.001	-0.49	<0.001
COVID Vaccination Rate	-0.77	<0.001	-0.48	<0.001
Carbon Intensity of Economy	0.75	<0.001	/	/
Median Income	-0.73	<0.001	-0.33	<0.001
Urbanization Rate	/	/	0.30	<0.001
Race - Asian	-0.42	0.002	-0.32	<0.001
Weather - Mean Temperature	0.46	<0.001	0.25	<0.001
Race - White	0.27	0.338	0.22	<0.001
Weather - Extreme Natural Hazards	-0.27	0.051	-0.13	<0.001
Race - Black	0.046	0.002	-0.12	<0.001
Weather - Temperature Anomalies	-0.13	0.391	-0.02	0.210

Table 2. Results of the weighted least squares regression model fitted at the county level (N = 1960). Notes: Total number of tweets per county was used as the universal weights in the model. Counties with less than 50 tweets were excluded. Variance Inflation Factor (VIF) < 5 indicate low multicollinearity of the multiple regression variables used in the model. C.I.: Confidence Intervals of regression coefficients.

	Coefficient	VIF	C.I. 2.5%	C.I. 97.5%
(Intercept)	1384.31		1187.59	1581.02
Political Affiliation (Republican)	0.16***	2.33	0.14	0.18
COVID-19 Vaccination rate	-0.09***	1.73	-0.12	-0.07
Education - College Degree	-0.06***	3.63	-0.10	-0.03
Hazard Risk Score	0.87	1.80	-0.17	1.90
Mean temperature (2010-2020)	14.70***	1.16	10.30	19.10
Median income	0.01	2.57	0.00	0.03
Adjusted R-squared: 0.47	***p < 0.001			

To gain additional insight into the geographical relationship between denialism and political affiliation at the county level, we used the bivariate LISA (Local Indicators of Spatial Association) model[20] to identify which counties with high rates of denialism or belief are spatially associated with high rates of Republican or Democratic voters. Clusters of deniers that coincide with high rates of Republican voters were spatially contiguous and covered large swaths of the interior West (Idaho, Montana, Wyoming), Central (Nebraska, Kansas, Oklahoma, Texas), and Appalachia regions (West Virginia, Tennessee) of the U.S. (**Fig. 2**). These findings are consistent with our regression modeling and bivariate correlations: These regions tended to have high rates of carbon dependency of the economy, low vaccination rates, and large rural populations. Conversely, clusters of believers and high rates of Democratic voters were most prevalent along the coasts (California, Washington), the New England Region, the Great Lakes, and the Southwest (Arizona), close to populous metropolitan areas and technological hubs.

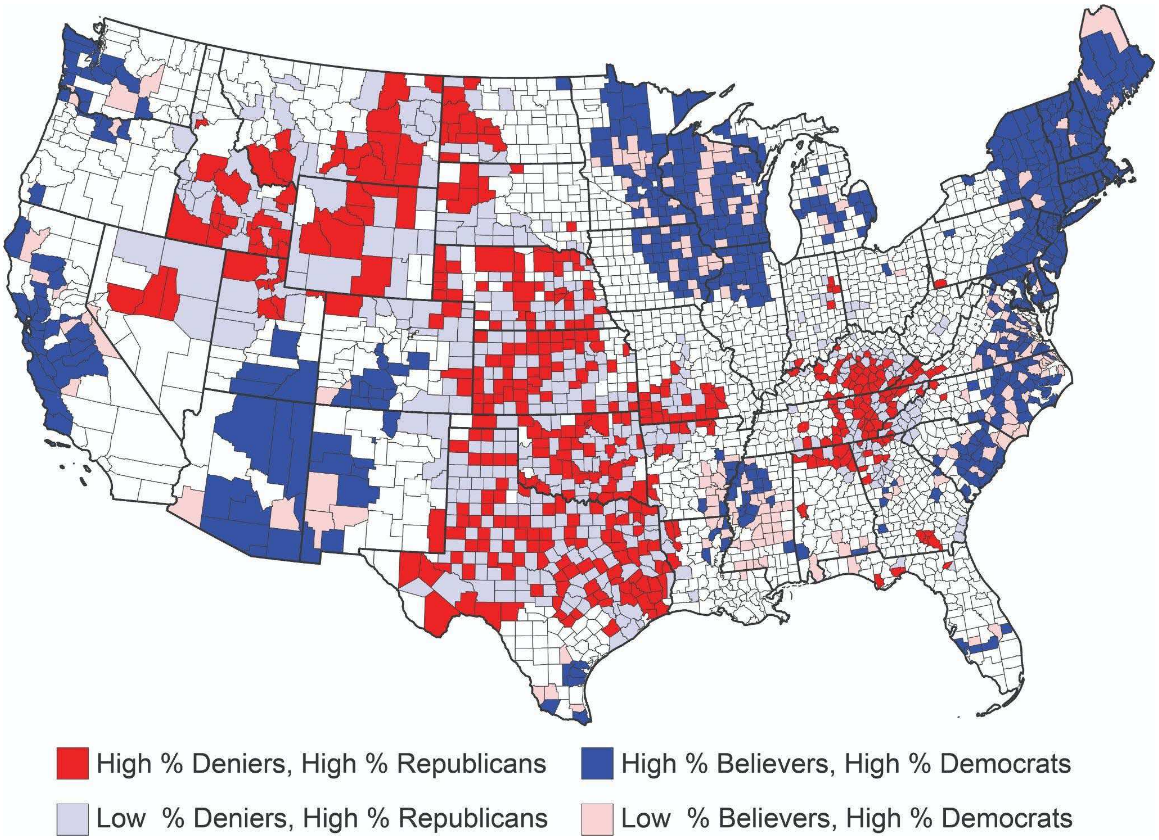

Fig. 2. Spatial clusters of climate change denialism and belief in relation to political affiliation.

Who are climate change influencers in the Twittersphere?

To delineate how polarized opinion forms in the Twittersphere, we constructed Twitter networks (based on the 1200 most retweeted users in the sample), analyzed how users interact, and identified key influencers (**Fig. 3**). To identify closely linked users assumed to share similar views, we evaluated co-retweeting, in which a single user retweets two or more other users.[21] Two distinct communities emerged, a denier and a believer community (**Fig. 3A**). The community of climate change believers (blue nodes) is larger, with 1029 users and ~224,000 co-retweets, giving it a broader reach and influence on Twitter than the denier community (red nodes), which its 171 users and ~15,000 co-retweets. The proportion of deniers among the top 1200 influential users (14.3%) aligned with the national percentage of climate change deniers identified in our model (14.8%).

Both believers and deniers mostly shared information and interacted within their own community. Users from the two communities were rarely co-retweeted, as illustrated by the distance between the cluster of nodes for each community and the low number of edges connecting the two communities. Among ~ 230,000 co-retweets, only 4083 (< 0.02%) were between users having opposite views on climate change. This low percentage of co-retweets of contrasting views highlights an *echo-chamber effect*. We found that a few nodes bridge the gap between the two communities, notably conservative news outlets such as *Fox News* (@FoxNews) and the *Washington Examiner* (@dcexaminer).

To identify the most influential users, we calculated the eigenvector centrality value per Twitter user. A high score means that a user is co-retweeted with many other users who also have high scores. Among climate change deniers, former U.S. President Donald Trump (@realDonaldTrump) had the biggest influence (**Fig. 3B**). Three groups of influential deniers were heavily co-retweeted with President Trump: (i) conservative media outlets that regularly broadcast contrarian views on climate change, including alt-right news and blogs such as *The Daily Wire*, *Daily Caller*, *Breitbart* and *thebradfordfile*; mis/disinformation websites that publish misleading and false claims about climate change, include *TownHall Media* and the *Climate Depot;* (iii) right-wing producers, political commentators, and activists. Collectively, in concert with former President Trump and close colleagues, these three groups formed an organized and coordinated social media network, enabling climate change denialism to amplify and expand.

In contrast, the larger blue community was more diffuse. Politicians dominated the most influential users (**Fig. 3C**). Of the top 30 influential believers, 15 accounts belong to figures of the Democratic Party, such as Alexandria Ocasio-Cortez (@*AOC*), Bernie Sanders *(@SenSanders*), and Kamala Harris (@*KamalaHarris*) (**Table S1**). Eight of the top 30 nodes were popular media outlets, or websites, such as *CNN, NBC, ABC, The Hill, The Washington Post, and New York Times*. Other influential nodes included popular science communicators and entertainers advocating scientific consensus.

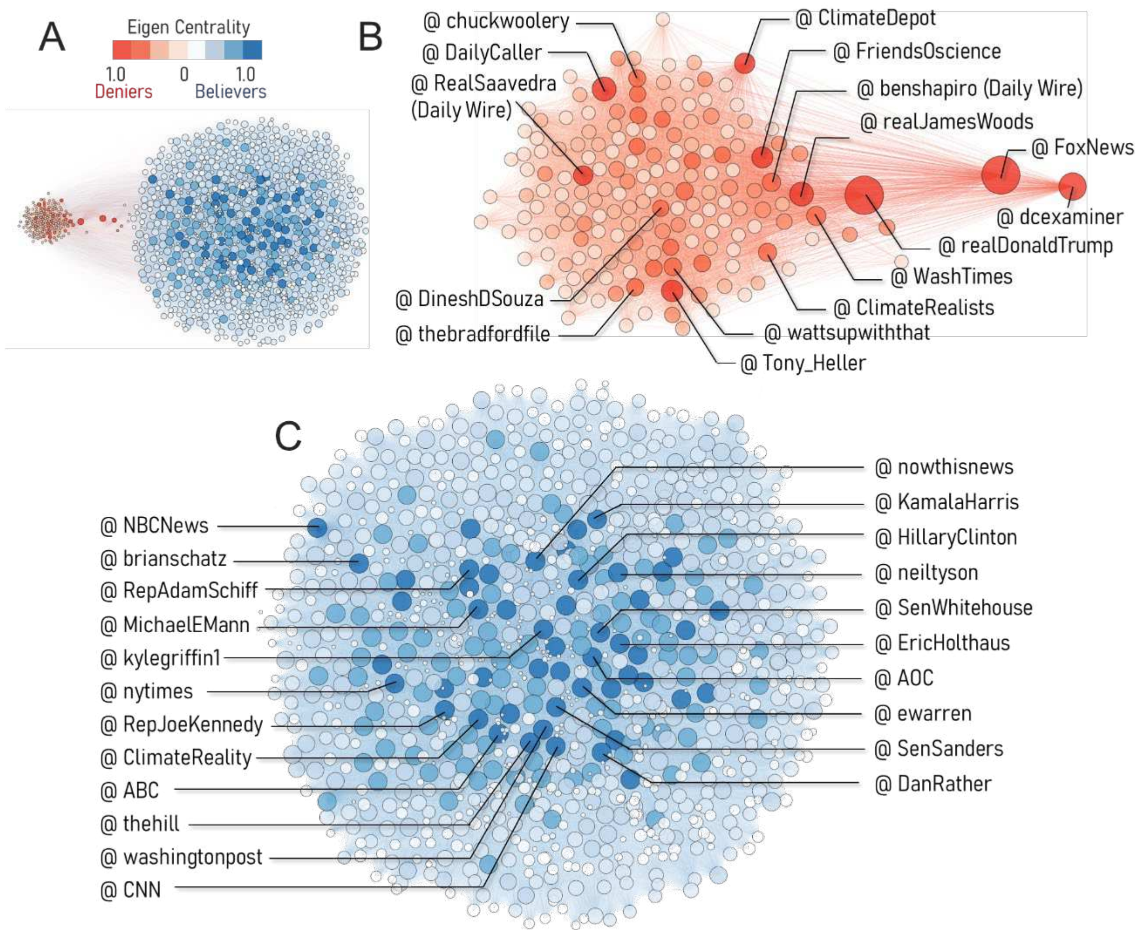

Fig. 3. Influencers detected in climate change co-retweeted networks. (A) Co-retweeted networks formed by the 1200 most retweeted users in the US. The nodes represent unique accounts; the edges represent co-retweeted relationships. The size of nodes and the shade of the nodes' color are proportional to their influence, as measured by eigenvector centrality scores. The high density of edges within the communities makes many individual edges not resolvable. The top influencers in the community of climate change deniers (B) and believers (C) are labeled with the usernames. In panel Band C, edges to users in the other community are not displayed.

How does tweeting and topic use related to climate change vary over time?

To investigate the dynamics of tweeting activity for both communities and to understand how each perceives and responds to real-world events, we performed topic modeling and time series analysis of tweet volume. Such an analysis revealed how each group reacts selectively and opportunistically to the 17 events that occurred during the period of data collection (November 2017 – May 2019).

Consistent with the larger size of the believer community, this community had a consistent pattern of climate change tweet activity throughout the sampling period (**Fig. 4A)**. In contrast, the denier community had lower activity overall. However, both communities had periods of high activity with spikes that exceeded the average pattern. The number of these high spikes was lower for the denier community. By manually identifying events that potentially triggered these large spikes, we found that deniers and believers do not always respond to the same events. Only 6 events triggered higher than average tweet volume by the denier community (**Table S3**): Three were related to extreme cold weather events, two were related to United Nations activities about climate change —the Intergovernmental Panel on Climate Change (IPCC) and the meeting of the United Nations Framework Convention on Climate Change (UNFCCC COP24), and the last was an attack on climate change deniers by Bill Nye in an HBO broadcast. Intriguingly, two of the highest spikes by the believer community occurred with events associated with President Trump that sparked high activity in the denier community, suggesting that these communities tried to influence or counter each other.

To gain further insight into whether the groups attempted to counter each other, we classified tweets of believers and deniers for these 17 events based on the five climate change narratives (**Fig. 4B**) proposed by Cook (2019)[22]. Overall, the major narrative in the believer community was "There is still time to adapt," representing 42% of the total tweets). In contrast, deniers focused tweeting activity on the message "Climate change is not real," as indicated by 48% of the tweets falling into this category.

Although weather events were associated with spikes in tweets from both communities, events viewed as abnormal weather caused by climate change [the California Wildfires (event 9) and Hurricane Florence (event 10)] triggered a high volume of tweets among believers and events viewed as colder-than-expected weather [a snowstorm in Texas (event 3) and a blizzard in the Mid-Atlantic and New England regions (event 6)] triggered a surge in tweets amongst deniers. Both of the colder-than-expected weather events provided an opportunity for the deniers to espouse that climate change is not real (64% of total tweets for both events), to delegitimize scientific consensus (12% of total) and to reaffirm the claim that the changing climate is a normal geologic process and foment doubt that human activities are a source of this change (13% of total).

Consistent with an attempt to counter each other's messages, the December 2017 tweet by Donald Trump casting doubt on global warming due to a blizzard (event 5) triggered the believer community to issue tweets emphasizing that climate change is unequivocal (32% of total) and that there is clear scientific consensus (35% of total). A common refrain among deniers was that climate change is a conspiracy theory or hoax (59% of total) and a shadowy attempt to dupe the public into bearing the costs of decarbonization, while generating enormous wealth for Blue 'elites' (9% of total). These tweets were heavily re-tweeted by conservative media (e.g., @DailyCaller), right-wing activists (e.g., @chuckwoolery), and mis/disinformation sites (e.g., @wattsupwiththat) (**Table S1**).

Conflicting messages were also common in response to UNFCCC COP24 (event 13 consistent with an attempt to influence opinion. Believers overwhelmingly advocated for timely collective action or promoted campaigns showcasing impacts of and solutions to climate change (50% of the total). Deniers focused on conspiracy theories (climate change is not real, 46%) or the Democratic party agenda of impractical solutions (26%).

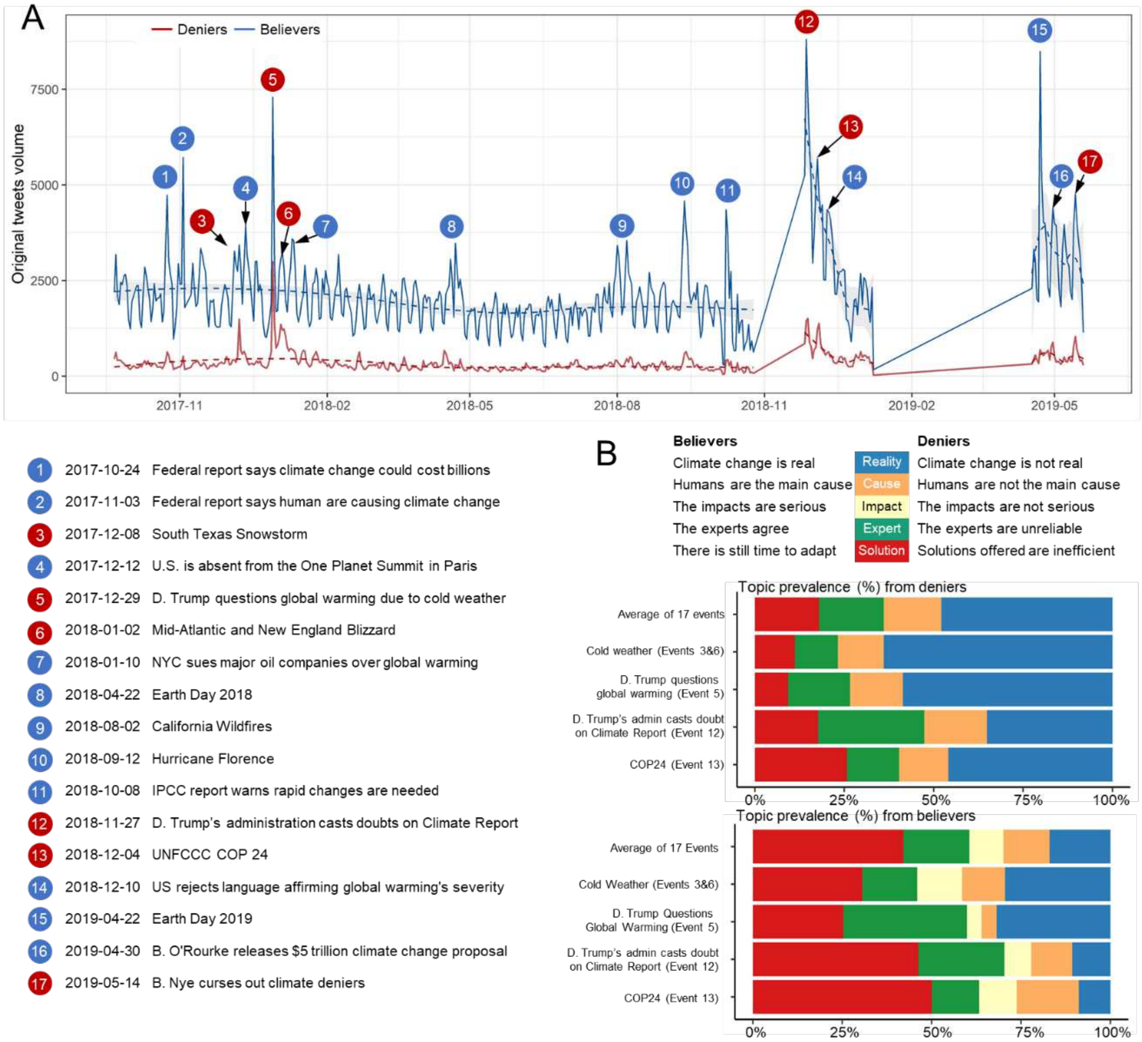

Fig. 4. Events that drive tweet volume among deniers and believers and topic prevalence for typical events. *(A)* Original tweet volume per day and locally weighted regression lines are plotted over time for both climate change deniers and believers. Events that sparked online discussions are labeled alongside the tweets volume spikes numerically and detailed in lower left. Red bubbles denote the events that a large group of deniers are actively involved with (>1000 original tweets). The gap in November 2018 and between January and April 2019 was due to discontinued data collection. *(B)*Topic prevalence for typical major events[22]: Events 3 & 6 represent extreme cold weather events; event 5 represents top denier influencer Donald Trump tweeting about cold weather and doubts global warming; event 12 represents top denier influencer Donald Trump refuting the validity of climate change report; and event 13, a United Nations climate change conference (COP24), represents an event that engaged both deniers and believers.

Discussion

Using data from Twitter, we delineated a comprehensive anatomy of climate change denialism in the U.S. at the state and county levels. We identified geographic clusters of climate change denial in Republican counties, especially rural ones, and among residents do not have a college education. This provides critical knowledge for targeting populations that would benefit especially from targeted efforts to expand awareness of the risks associated with climate and strategies to increase local resilience.

The strong correlation between denialism and low COVID-19 vaccination rates indicated a broad skepticism of science generally amongst the climate change deniers, which corresponds to resistance to science-based public policies such as shelter-in-place COVID-19 mandates[23] or mask usage.[24] This finding indicates that communities with high prevalence of climate change deniers are at risk for discounting other science-based health or safety recommendations.

We acknowledge limitations associated with the model and the bivariate analysis and took steps to address them. We minimized the effect of low population density by normalizing our input data by county population and using a weighted approach using the total count of tweets as weights. To minimize the effect of inaccurately interpreting tweets as for or against climate change, for example due to sarcastic or ambivalent language, we calculated a confidence for each prediction and removed those with low confidence. Additional details are in the Online Methods.

Classifying tweets based on the Cook's five categories[22] enables identification of commonly deployed rhetorical strategies deployed to promote climate misinformation and in science denialism more broadly.[25] In our 7.3 million tweet sample, these techniques included *fake experts,* who have possess little to no expertise about the underlying science but nonetheless convey messages that cast doubt. They serve as a *credible messenger* in which someone shares the same moral values and uses language consistent with existing beliefs.[26] One such example is the tweet by the Trump administration casting doubts on the Climate Report, which was retweeted heavily by supporters. Then there are *logical fallacies*, such as a Trump tweet questioning global warming because of an unusual cold weather event that went viral.[27] Other common strategies include *impossible expectations* as well as *cherry picking* to attack climate change science and scientists.

Combatting misinformation requires effective refutation strategies.[28] Deploying such strategies on Twitter, however, is challenging as denier and believer communities are isolated from each other, leading to echo chambers[19]. Only 0.02% of the co-retweets about climate change were between users having opposing views. Consequently, this leads one to conclude that believers have limited ability in reaching deniers through Twitter. One strategy is to label denialism tweets as misinformation. However, some evidence suggests that this can strengthen opposition rather than change attitudes.[29]

Another option is to suspend or ban accounts that disseminate misinformation or dangerous information. For example, Twitter banned Donald Trump from using Twitter because of tweets maintaining election fraud and supporting the January 6 capital riots.[30] Twitter also banned accounts for spreading COVID-19 misinformation and calling for violence against media.[31] To date, climate change denialism does not appear to trigger account bans or suspension on Twitter but this should be seriously considered. As with COVID-19, climate change is a humanitarian crisis that will affect millions, albeit at a more elongated temporal scale.

Communities face increasing risks related to climate change, such as flooding, wildfire, heat stress, and sea-level rise. The scientific community is starting identify especially vulnerable communities and regions.[32] Climate change denialism is also a risk, in the form of knowledge vulnerability. Those who discount climate change as a natural rather than human-induced process tend to underestimate their current (and future) risk to it. This renders them less likely to take necessary steps to mitigate and adapt to it.

Online Methods

Opinion data

As primary data, we used an open access dataset created by George Washington University that is available from the GWU Libraries Dataverse.[33] This dataset was created using the Twitter Stream API and contains ~40 million tweets related to climate change and global warming. It covers a two-year period from September 2017 to May 2019. We initially retrieved ~27.3 million raw tweets based on tweet IDs. The ~30% loss of tweets was due to deleted or inactive accounts since 2019.

To extract tweets located in the U.S., we developed a rule based on the geo-attributes in the raw data. We extracted the self-reported location information in an account profile. A large proportion of users (> 73%) provided the location information in our dataset. To standardize the addresses and improve the geocoding process, we first transformed all the user locations to lower case and removed the URL links, emojis, punctuation marks, and other non-ASCII characters. Next, we extracted all the unique user locations (~ 640,000 "clean" addresses) and standardized all the U.S. state and city abbreviations. As a final step, we manually inspected and removed national level and obviously fake user locations.

After the preprocessing, we used the Nominatim API server to geocode user locations based on the OpenStreetMap database.[34] We removed locations outside the U.S., and classified addresses within the U.S. into two levels: 1) county level with tweets from users reporting their local address, city, or county; 2) state level with tweets from users reporting only the state. In the state-level tweets, we also added the aggregated county-level tweets. We then rejoined these unique U.S. addresses and the corresponding geographical coordinates to the original datasets by spatial level. The geocoding yielded ~1.3 million unique users and ~5.2 million county-level tweets and ~7.4 million state-level tweets, from which ~2.2 million tweets had state-level only information. To reduce the incidence of non-human accounts in our sample, we removed users who tweeted more than 20 times a day. **Fig. S1-S2** presents the data spatial distribution and representativeness analysis.

Tweet classification

To identify climate change opinions on Twitter, we built a tweet classifier based on the Transformer, a deep learning model in the field of natural language processing.[35] We parameterized the model to classify tweets as either believing in the existence of climate change (predicted as 'for') or denying that climate change is real (predicted as 'against'). Instead of training a model de novo, the Transformer uses language models pre-trained on large text corpora in an unsupervised manner and then uses user-labeled training samples to fine-tune the model for specific natural language tasks. Our classifier was built upon OpenAI GPT-2, a large transformer-based language model pre-trained on a database of ~8 million web pages.[36] Previous studies found that the GPT-2 model performs well in classifying short text from social media.[37]

We built a training dataset of manually labeled tweets to fine tune the pre-trained GPT-2 model. Labeled samples were randomly extracted only from the 1.4 million original tweets, excluding re tweets and quotes. Each tweet was reviewed independently by two members of the research team and labeled as either ‘against’ or ‘for’ climate change.

We labeled training tweets as ‘for’ or ‘against’ climate change if they had one of the following viewpoints listed in Table SX. This labeling resulted in a balanced sample of 6,500 tweets (3300 ‘for’ tweets and 3200 ‘against’ tweets) that we used as a training set for the model. Tweets with ambiguous messages, sarcastic language or tweets that were irrelevant to climate change were discarded from the training dataset.

Table SX. Classification of tweets used for training the model as ‘for’ or ‘against’ climate change.

‘For’ (Belief): N = 3300 tweets	‘Against’ (Denial): N = 3200 tweets
Climate change concern: The user believes climate change is real and worries about its negative consequences. *Advocate for action:* The user calls for collective actions and supports any adaptation and mitigation policies. *Scientific consensus:* The user advocates for the scientific evidence on climate change and recognizes the role of greenhouse gas emissions caused by human activities.	*Trend denialism:* The user shows disbelief that the earth is warming and climate change is happening. *Attribution denialism:* The user believes climate change is happening, but it is a natural, unpreventable process and anthropogenic greenhous gases are not the dominant driver. *Impact denialism:* The user believes climate change will not have significant negative impacts on the environment and humanity. *Evidence denialism:* The user doubts there is trustworthy scientific consensus on climate change.

Our model was built upon the *Huggingface* Transformers[38] library and implemented in *PyTorch*.[39] To increase the model’s predictive accuracy, we fine-tuned the parameters that resulted in an optimum learning rate at 1e-5, with dropouts at 0.1. Tweets with sarcastic, ambiguous or irrelevant messages were evaluated with the model, but the predictions based on these tweets tended to be invalid or random. To overcome this limitation, we used the *Softmax* function embodied in *PyTorch*, which calculated the prediction confidence for every individual tweet. Based on this score, we removed predictions with low confidence (CI < 0.75). The final classification was performed on the complete set of 7.4 million tweets from the collection period. We then aggregated tweets at the county and state levels and calculated percentages of ‘against’ tweets and ‘for’ tweets as proxies of deniers and believers.

To evaluate the model’s performance, we performed a series of validation tests. We manually labeled an independent validation dataset to test model accuracy. To ensure the validation dataset was balanced across the two categories and was spatially representative, we randomly extracted 30 unique original tweets from each state. Our fine-tuned model achieved an overall accuracy of 0.91 and F1 score of 0.90 (**Fig. S3).**

Our model predictions were compared with US-wide estimates of climate change opinion based on representative surveys, showing that our model provided a percentage for U.S. climate change deniers within the range of those determined from the surveys (**Fig. S4**). To validate our results at the sub-national level, we referred to the Yale Climate Opinion Surveys. The Yale Climate Opinion Surveys use a downscaling statistical model based on national survey data and are the only surveys that provide climate change opinion estimates at the state and county levels. We compared these data with our model results at both state and county levels by calculating Pearson correlation. To normalize the data, we weighted the variables per population of each state and county (US Census 2018).

Correlation analysis

To examine what drives climate change opinion, we performed a series of correlation analyses. Studies have shown that climate change opinion is mainly driven by political affiliation, socio-demographics, local microclimate, and personal experience with extreme weather events.[40] We examined variables that are among the top drivers of climate change opinion: political affiliation, COVID vaccination rate (proxy for belief in science in general), urbanization rate, education, income, race, carbon intensity of economy, natural hazard risk, and temperature anomaly.

We used the percentage of 'against' and 'for' tweets to reflect the prevalence of deniers and believers across the U.S. at the county and state levels. For political affiliation, we acquired 20 years (2000-2020) of county-level U.S. Presidential election returns from the MIT Election Data and Science Lab (https://electionlab.mit.edu/data). We calculated the average percentage of Democrats and Republicans per state and county, weighted by the county population. For science skepticism, we used the county-level COVID-19 vaccination rates as a proxy, using data from the CDC (https://www.cdc.gov/coronavirus/2019-ncov/vaccines/distributing/reporting-counties.html). For educational attainment, race, and income, we used data from the US Census Bureau's 2020 American Community Survey, which provides estimates of average characteristics from 2016 through 2020 at the state and county levels. Specifically, we used the number of people who have at least a Bachelor's college degree, number of people per race, and the median family income. For county-level natural hazard risk, we used the National Risk Index developed by FEMA (https://www.fema.gov/flood-maps/products-tools/national-risk-index). An overall risk score was calculated for each county, measuring the expected annual loss due to 18 types of natural hazards. To calculate temperature anomaly, we acquired historic 30-year annual mean temperature (1981-2010) and the mean for recent years (2015-2019) from the PRISM climate group (https://prism.oregonstate.edu/). County-level temperature anomaly was then obtained by calculating the standard deviation between annual mean temperature of recent years and the 30-year averages. To investigate the association between state-level carbon dependency of economy and climate change opinion, we used energy-related carbon emissions per gross domestic product (GDP) for each state from the Energy Information Administration (https://www.eia.gov/environment/emissions/state/). The unit of carbon intensity is the metric tons of energy-related carbon dioxide per million dollars of GDP. A six-level urban-rural classification at the county level was from the National Center for Health Statistics data systems (https://www.cdc.gov/nchs/data_access/urban_rural.htm).

To account for variations in population across counties and states, we normalized all data expressed as counts. We adjusted the total county population as: PopulationAdj = Total population / 10,000. Then, we normalized each variable by population by dividing the counts of people for each variable by the adjusted population: Normalized Variable = Variable count / PopulationAdj. Based on the normalized data, we calculated bivariate weighted Pearson correlations between climate change opinion and each of these variables using the total count of tweets per county as the weight. The same data were used as predictors for the regression model. We used the weighted ordinary least squares for the total count of tweets per county as the universal weight.

To identify spatial clusters of climate change denialism or belief at the county level in relation to political affiliation (Republican or Democrat), we applied the bivariate Local Indicators of Spatial Association (LISA).[20] We applied the second order Queen contiguity weights at the county level and ran the models with 999 permutations and significance at $p < 0.05$. This approach was carried out in the open-source software *Geoda*.[41]

Co-retweeted network analysis

We constructed a co-retweeted network to delineate interactions and identify the most influential Twitter users from both sides. Co-retweeting is defined as the act of a single user retweeting two or more other users. We used these events to create undirected weighted edges between the co-retweeted accounts. The more users retweet two other users, the more weight the edge gains. Accordingly, we assumed that the more co-retweets two accounts receive, the more likely their views are related. The co-retweeted network represents engaged communities with similar opinions.

To construct the co-retweeted network, we first calculated the total sum of retweets as a measure of overall influence for each user account in our 7.2 million tweets dataset. We selected the 1200 most retweeted accounts for further processing, along with all the users who have retweeted them. We then constructed the retweet matrix A where the rows represent the 1200 top accounts, and the columns represent the rest of user accounts. Elements in matrix A are binary: A value of 1 means that the public account has retweeted the corresponding top influential account and 0 means the public account has not retweeted the top influential account. We then multiplied matrix A with its transposed matrix A^T and transformed it into the co-retweeted square matrix B. Matrix B has 1,200 rows and columns that represent the influential accounts. The upper and lower diagonal cells of matrix B contain the total number of times that two influential accounts are co-retweeted. We exported all the unique pairs of influential accounts and their co-retweets as the edge table for further network analysis.

Our co-retweeted network was visualized in *Gephi*, using the Force Atlas algorithm[42], which clusters nodes based on their connections. The distance between two nodes was weighted by the number of co-retweets. We then applied the Louvain community-detection algorithm [43] and separated the nodes as two communities based on modularity scores. To detect opinion leaders in each community, we calculated the Eigen centrality values for each node based on the *igraph* package in R.[44] The number of co-retweets for each node was set as the weight. To facilitate visualization, we extracted the top 30 influencers from each community (**Table S1** for deniers and **Table S2** for believers). The eigenvalues are scaled to a maximum score of one.

Time-series analysis and topic modeling

To examine the dynamics of tweeting activity regarding climate change, we identified 17 major climate change-related events that happened during September 2017 to May 2019 and analyzed the tweet volume of both deniers and believers during this period. To delineate the major climate change-related topics discussed, to understand how the prevalence of each topic evolved over time, and to explore how each group perceived the event, we employed the Latent Dirichlet Allocation (LDA) algorithm[45] to automatically extract the main topics. We specified the number of topics before training the model. We devised a five-category classification scheme following Cook's (2019)[22] categories of misinformation: a) climate change is/is not real; b) humans are/are not the main cause; c) the impacts are/are not serious; d) the experts agree/are unreliable; e) there is still time to adapt/solutions offered are inefficient.

The model was implemented in Python's gensim package along with the Java-based package Mallet to accelerate data processing.[46] We ran topic modeling separately for tweets classified as from 'believers' or 'deniers.' We preprocessed the original ~7.2 million tweets, keeping original tweets and excluding retweets with the same text. We removed all the @mentions, hashtags, punctuation marks, and changed all characters to lower case. From keywords, we removed "climate change" and "global warming" because these words occurred too frequently and would dominate as distinct topics. After this pre-processing, we tokenized every tweet and created bigrams and trigrams because some words often occurred together as phrases. We reduced words to their common word stem and dropped duplicates to ensure the text corpora analyzed by the model was clean and distinct.

Study limitations

Our modeling has some limitations. In rural areas with low population densities, the sample sizes are relatively small, so uncertainty is higher than with more densely populated areas. This is a recognized limitation of Twitter and even more pronounced in countries where use of social media is limited. To minimize this effect, we normalized our input data by county population and employed a weighted approach using the total count of tweets as weights both for the calculation of bivariate relationships and for the regression models (see Methods). Second, our classification scheme labeled tweets as either believing or denying climate change. National surveys indicate a cohort of people (5-15%) who remain neutral or may not have a particular opinion on the topic. We used climate change related keywords in our binary classification that indicated a clear position (for or against) on the issue. Classifying these tweets can be challenging as a portion of our sample uses sarcastic or ambivalent language that is virtually impossible for the model to distinguish. To address this, we calculated confidence for each prediction (see Methods) to filter out those with low confidence (CI < 0.75) that are closer to being random.

References

1 MacInnis, B. & Krosnick, J. A. Climate insights 2020: Surveying American public opinion on climate change and the environment. (2020).

2 Leiserowitz, A., Roser-Renouf, C., Marlon, J. & Maibach, E. Global Warming's Six Americas: a review and recommendations for climate change communication. *Current Opinion in Behavioral Sciences* **42**, 97-103, (2021).

3 Howe, P. D., Mildenberger, M., Marlon, J. R. & Leiserowitz, A. Geographic variation in opinions on climate change at state and local scales in the USA. *Nature Climate Change* **5**, 596-603, (2015).

4 Hornsey, M. J., Harris, E. A., Bain, P. G. & Fielding, K. S. Meta-analyses of the determinants and outcomes of belief in climate change. *Nature Climate Change* **6**, 622-626, (2016).

5 Hornsey, M. J., Harris, E. A. & Fielding, K. S. Relationships among conspiratorial beliefs, conservatism and climate scepticism across nations. *Nature Climate Change* **8**, 614-620, (2018).

6 McCright, A. M. & Dunlap, R. E. Cool dudes: The denial of climate change among conservative white males in the United States. *Global Environmental Change-Human and Policy Dimensions* **21**, 1163-1172, (2011).

7 Knight, K. W. Does fossil fuel dependence influence public awareness and perception of climate change? A cross-national investigation. *International Journal of Sociology* **48**, 295-313, (2018).

8 Long, E. F., Chen, M. K. & Rohla, R. Political storms: Emergent partisan skepticism of hurricane risks. *Science Advances* **6**, (2020).

9 Weckroth, M. & Ala-Mantila, S. Socioeconomic geography of climate change views in Europe. *Global Environmental Change-Human and Policy Dimensions* **72**, 102453, (2022).

10 Pennycook, G. *et al.* Shifting attention to accuracy can reduce misinformation online. *Nature* **592**, 590-595, (2021).

11 Johnson, N. F. *et al.* The online competition between pro- and anti-vaccination views. *Nature* **582**, 230-233, (2020).

12 Goldberg, M. H., Gustafson, A., Rosenthal, S. A. & Leiserowitz, A. Shifting Republican views on climate change through targeted advertising. *Nature Climate Change* **11**, 573-577, (2021).

13 Zhang, B. *et al.* Experimental effects of climate messages vary geographically. *Nature Climate Change* **8**, 370-374, (2018).

14 Stieglitz, S. & Dang-Xuan, L. Social media and political communication: a social media analytics framework. *Social network analysis and mining* **3**, 1277-1291, (2013).

15 Kirilenko, A. P. & Stepchenkova, S. O. Public microblogging on climate change: One year of Twitter worldwide. *Global Environmental Change-Human and Policy Dimensions* **26**, 171-182, (2014).

16 Jaidka, K. *et al.* Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 10165-10171, (2020).

17 Grossman, G., Kim, S., Rexer, J. M. & Thirumurthy, H. Political partisanship influences behavioral responses to governors' recommendations for COVID-19 prevention in the United States. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 24144-24153, (2020).

18 Bovet, A., Morone, F. & Makse, H. A. Validation of Twitter opinion trends with national polling aggregates: Hillary Clinton vs Donald Trump. *Scientific Reports* **8**, (2018).

19 Pennycook, G. & Rand, D. G. Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 2521-2526, (2019).

20 Anselin, L. Local Indicators of Spatial Association—LISA. *Geographical Analysis* **27**, 93-115, (1995).

21 Finn, S., Mustafaraj, E. & Metaxas, P. T. The co-retweeted network and its applications for measuring the perceived political polarization. (2014).

22 Cook, J. in *Handbook of research on deception, fake news, and misinformation online. Advances in media, entertainment, and the arts (AMEA) book series.* 281-306 (Information Science Reference/IGI Global, 2019).

23 Brzezinski, A., Kecht, V., Van Dijcke, D. & Wright, A. L. Science skepticism reduced compliance with COVID-19 shelter-in-place policies in the United States. *Nature Human Behaviour* **5**, 1519-1527, (2021).

24 Merkley, E. & Loewen, P. J. Anti-intellectualism and the mass public's response to the COVID-19 pandemic. *Nature Human Behaviour* **5**, 706-715, (2021).

25 Ceccarelli, L. Manufactured scientific controversy: Science, rhetoric, and public debate. *Rhetoric and Public Affairs* **14**, 195-228, (2011).

26 Corner, A. *et al.* How do young people engage with climate change? The role of knowledge, values, message framing, and trusted communicators. *WIREs Climate Change* **6**, 523-534, (2015).

27 Farmer, G. T. & Cook, J. in *Climate Change Science: A Modern Synthesis: Volume 1 - The Physical Climate* (eds G. Thomas Farmer & John Cook) 445-466 (Springer Netherlands, 2013).

28 Schmid, P. & Betsch, C. Effective strategies for rebutting science denialism in public discussions. *Nature Human Behaviour* **3**, 931-939, (2019).

29 Christenson, D. P., Kreps, S. E. & Kriner, D. L. Contemporary Presidency: Going Public in an Era of Social Media: Tweets, Corrections, and Public Opinion. *Presidential Studies Quarterly* **51**, 151-165, (2021).

30 Tollefson, J. Tracking QAnon: how Trump turned conspiracy-theory research upside down. *Nature* **590**, 192-194, (2021).

31 Bak-Coleman, J. B. *et al.* Combining interventions to reduce the spread of viral misinformation. *Nature Human Behaviour*, (2022).

32 Rentschler, J., Salhab, M. & Jafino, B. A. Flood exposure and poverty in 188 countries. *Nature Communications* **13**, 3527, (2022).

33 Littman, J. & Wrubel, L. Climate Change Tweets Ids. *Harvard Dataverse*, (2019).

34 OpenStreetMap. <https://www.openstreetmap.org/ > (2020).

35 Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. 4171-4186, (2019).

36 Radford, A. *et al.* Language models are unsupervised multitask learners. *OpenAI blog* **1**, 9, (2019).

37 Fagni, T., Falchi, F., Gambini, M., Martella, A. & Tesconi, M. TweepFake: About detecting deepfake tweets. *PLOS ONE* **16**, e0251415, (2021).

38 Wolf, T. *et al.* 38-45 (Association for Computational Linguistics).

39 Paszke, A. *et al.* Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems* **32**, (2019).

40 Howe, P. D., Marlon, J. R., Mildenberger, M. & Shield, B. S. How will climate change shape climate opinion? *Environmental Research Letters* **14**, 113001, (2019).

41 Anselin, L., Syabri, I. & Kho, Y. in *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications* (eds Manfred M. Fischer & Arthur Getis) 73-89 (Springer Berlin Heidelberg, 2010).

42 Jacomy, M., Venturini, T., Heymann, S. & Bastian, M. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi Software. *PLOS ONE* **9**, e98679, (2014).

43 Meo, P. D., Ferrara, E., Fiumara, G. & Provetti, A. Generalized Louvain method for community detection in large networks. *2011 11th International Conference on Intelligent Systems Design and Applications*, 88-93, (2011).

44 Csardi, G. & Nepusz, T. The igraph software package for complex network research. *InterJournal, complex systems* **1695**, 1-9, (2006).

45 Blei, D. M., Ng, A. Y. & Jordan, M. I. Latent dirichlet allocation. *Journal of Machine Learning Research* **3**, 993-1022, (2003).

46 McCallum, A. K. *MALLET: A Machine Learning for Language Toolkit*, <http://mallet.cs.umass.edu> (2002).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TwitterandClimateChangeSupplementaryMaterialfinal.pdf
- R.pdf