USC Annenberg

Norman Lear Center
Media Impact Project

November 2022

Measuring the Impact of Science Journalism

The Pulitzer Center's *Connected Coastlines* Initiative

APPENDICES

Prepared for the Howard Hughes Medical Institute

Report design by Veronica Jauriqui

TABLE OF CONTENTS

APPENDIX A 37
Roundtable and Interview Participants

APPENDIX B 39
List of Analyzed *Connected Coastlines* Stories

APPENDIX C 40
Content Analysis Coding and Reliability

APPENDIX D 45
Twitter Analysis Methods

APPENDIX E 51
Case Studies Key Informants and Methods

APPENDIX A

ROUNDTABLE AND INTERVIEW PARTICIPANTS

ROUNDTABLE PARTICIPANTS

- Dan Barkin, ret. journalist and editor
- Hal Bernton, *Seattle Times*
- Marjee Chmiel, Howard Hughes Medical Institute
- Kathy Dello, North Carlina Director of State Climate Office
- Sammy Fretwell, *The State*
- Dave Hendrickson, *The News & Observer*
- Leilani Rania Ganser (Pulitzer Center)
- Julie Raymond-Yakoubian (Kawerak)
- Rich Stone (HMMI)
- Steve Sapienza (Pulitzer Center)
- Geoff Scott (Center for Oceans and Human Health and Climate Change Interactions)
- Rich Stone (HMMI)
- Maggie Suggs (Appalachian State University)
- Robyn Tomlin (*The News & Observer*)
- Merritt Turetsky (University of Colorado)
- Adam Wagner (*The News & Observer*)

INTERVIEW PARTICIPANTS

Pulitzer-Supported Journalists

- Tony Bartelme, Senior Projects Reporter, *Post & Courier*
- Hal Bernton, Staff Reporter, *Seattle Times*
- Joanna Detz, Co-founder & Publisher, *EcoRI News*
- Sammy Fretwell, Environment Reporter, *The State/McClatchy*
- Frank Graff, Reporter/Producer, PBS North Carolina
- Mark Hibbs, Editor, *Coastal Review*
- Jack Igelman, Reporter, *Carolina Coastal Press*
- Andrew Lewis, Reporter, *New Jersey Coastline*

Other Science Journalists

- Yessenia Funes, Climate Writer, *Atmos*
- Chip Giller, Founder, *Grist*
- Carolyn Gramling, Staff Writer, *Science News*
- Christina Lee Larson, Science & Environment Correspondent, Associated Press
- David Malakoff, Deputy News Editor, *Science Magazine*
- Hillary Rosner, Freelancer, *National Geographic, Wired, Scientific American*
- Ashley Smart, Senior Editor, *Undark*
- John Sutter, Freelancer, CNN

Climate and Science Communication Experts

- Emma Frances Bloomfield, Assistant Professor, Communication Studies, UNLV
- Martin Kaplan, Director, USC Annenberg Norman Lear Center
- Bruce Lewenstein, Professor, Dept. of Communication, Cornell University
- Jeff Nesbit, Executive Director, Climate Nexus
- Claudia Tebaldi, Senior Scientist, JGCRI/Climate Central

APPENDIX B

LIST OF ANALYZED *CONNECTED COASTLINES* STORIES

A total of 142 CC stories were included in the sample for quantitative Twitter analysis. Of these, 137 were included in the content analysis sample (five stories were excluded: three from foreign language publications, one comic book, one duplicate story). Five additional stories (four published in the *New York Times*, one video on YouTube) were excluded from the correlational analysis.

The list of all analyzed stories can be found here.

APPENDIX C

CONTENT ANALYSIS CODING AND RELIABILITY

CODING PROCEDURE

In addition to a NLC project associate, one undergraduate USC student was trained over a two-week period on the coding process and was familiarized with codebook items. The training period included multiple rounds of testing and codebook refinement. In June and July 2022, a total of 137 stories were coded using a Qualtrics survey form for descriptive variables (story title, story URL, publication name/date/region, author name[s] and demographics) and the following variables aligned with the ten best practices:

- **Undefined jargon:** In context of this study, jargon meant any language that is generally used by professional scientists that non-scientists would have difficulty understanding. This item was intended to measure readability and lack of clarity for the average reader. Thus, stories that included defined jargon were not counted, as they were still considered understandable and congruent with best practices regarding readability.

- **Metaphors:** Relevant metaphors communicated a scientific concept by comparing it to an everyday, well-known phenomenon. Analogies were also coded as metaphors (for example, "ocean acidification is like osteoporosis of the sea").

- **Flesch-Kincaid Grade Level (FKGL):** This was computed by entering the story's body text into a readability calculator website. FKGL computations use a story's average sentence and syllable length to indicate how difficult it is to understand that story in terms of years of education necessary for comprehension (usually corresponding with a U.S. grade level). For instance, a story with a 10th grade FKGL score should be generally understandable to anyone with a 10th grade education. Subheadings, image captions, and other extraneous text were removed from body text before FKGL calculation. Stories that were primarily audio-based (i.e. podcasts) were not analyzed for FKGL.

- **Use of infographics or data visualizations**

- **Interview subjects depicted in a photo**

- **Locality:** The location of the story's subject matter was then compared to the publication's location to determine the "locality" of the story. Stories that described events taking place in the same state as the publication were coded as "local" while stories describing events in an adjacent state were coded as "somewhat local" and stories that described events in a faraway locale were coded as "not local."

- **Mentions of health-related problems:** Health problems regarding human beings (that is, not animals or pets) were defined as any physical or mental problem that one would see a medical professional for treatment.

- **Connecting health issues to climate change:** These linkages could be indirect. For instance, a person's bacterial infection could be related to a local outbreak that was indirectly linked to rising sea levels or ocean warming.
- **Personal anecdotes:** An anecdote was defined as a real-life story that happened to the speaker (i.e. author or interview subject), based on the experience or observations of that one person, and relevant to the subject of the story.
- **Individual solutions:** Individual solutions are those that hinge on personal decisions, responsibility, or agency. Individual solutions include recycling, riding a bike instead of driving, planting a tree, and installing solar panels.
- **Systemic solutions:** Systemic solutions are those that are generally not possible at the individual level and require collective action to implement or change (for example, creating a green jobs initiative, building a breakwater, passing a law). In the context of this project, some "resilience" efforts did not count as solutions, as they do not necessarily address the underlying environmental causes of a problem (for example, rebuilding a house that was destroyed in a climate-related hurricane).
- **Headline accuracy:** This was determined by comparing the content of the headline to the content of the story. A small sample of "clickbait" headlines from non-CC stories was used to train the coder on how to identify inaccurate headlines. Sensationalized clickbait headlines are those that inaccurately represent the content of the article in order to attract more readers.
- **Degree of certainty:** Does the story qualify the results of a cited scientific study with any caveats or limitations (for instance, mentioning that the study is a preprint or isn't peer-reviewed)? Stories that simply cited or mentioned a scientific study without including descriptions of uncertainty were not counted under this metric.
- **Competing perspectives:** Competing perspectives are those that offer opposing views on the extent to which something is happening or why something is happening. In the context of climate change reporting, including competing perspectives may mean including a debate between a climate scientist and a climate denier regarding the extent to which anthropogenic climate change is real or causing a certain weather phenomenon.
- **Supporting evidence:** Evidence in favor of a competing perspective can be either qualitative or quantitative. Quantitative information means numbers that refer to the balance of evidence. For example, "30% of scientists believe X, while 70% believe Y." Qualitative information means descriptive language about the balance of evidence. For example, "many scientists believe X, while some others do not."
- **Presents results or indications of progress linked to solution:** For example, did the story provide any qualitative accounts of people changing their behavior, or statistics about reductions in pollution?
- **How-to details of solution implementation:** Such stories provided concrete information about specific steps toward implementing a solution–such as showing up to city council meetings, organizing a protest, or calling a phone number.

42. Offering systemic solutions was the only low or borderline reliability item that was not rare in the sample, and reliability for this item was close to the adequacy threshold.

- **At least one woman author**
- **At least one person of color author**
- **At least one person of color depicted in a photograph**
- **Indigenous knowledge:** Indigenous knowledge refers to novel technologies, ideas, or strategies produced within indigenous communities that make unique contributions towards addressing or understanding some kind of real-world problem.

RELIABILITY

A total of 27 stories (20% of the sample) were double-coded to measure inter-rater reliability. First, a preliminary sample of 13 stories was coded by both the student coder and an MIP project associate to measure the initial reliability. Reliability for these stories was calculated using Cohen's Kappa, with the exception of Flesch-Kincaid Grade Level, which as a continuous variable was assessed using Krippendorf's Alpha. Items with borderline or inadequate reliability were discussed between the student intern and the project associate.

After primary coding of all 137 stories, 14 additional stories were randomly selected from the sample to be double-coded. Some items with inadequate reliability were ultimately excluded from the study. Reliability values for the remaining items are shown in Table 2, along with percent agreement, the proportion of instances in which the student coder and reliability coder agreed on the presence or absence of the item.

- 0.60 or higher: acceptable reliability
- 0.40-0.59: borderline reliability
- Less than 0.40: low reliability
- Feature not present in sample, unable to calculate reliability

Low or borderline reliability was mostly because the feature in question (e.g., use of metaphors, competing perspectives) was rarely present in the sample.[42] In these cases, we can be confident that the feature in question is rare, but the precise frequency of presence should be interpreted with caution. For two items, we were unable to calculate reliability because there was no variability in the sample; the feature was either always absent (undefined jargon) or always present (headline accurately captures story content). In both of these instances, the coders agreed 100% of the time.

Table C1.
Content Analysis Items and Reliability

	Reliability	% Agreement
1. REPLACE SCIENTIFIC JARGON WITH METAPHORS		
Does this story include any undefined scientific jargon?	—	100%
Does this story use any metaphors to communicate the meaning of scientific concepts?	0.29	85%
What is the Flesch-Kincaid Grade Level for this story?	0.90	—
2. USE IMAGES STRATEGICALLY		
Does this story include any infographics or visualizations of data?	1.00	100%
Are there any interview subjects depicted in any photograph?	0.92	96%
3. BRING SCIENCE CLOSE TO HOME		
Is the story local?	1.00	100%
4. CONNECT SCIENCE TO HEALTH OUTCOMES		
Does the story mention any health-related problems regarding human beings?	0.71	100%
Does the story explicitly connect any health issues to climate change?	1.00	100%
5. HUMANIZE COVERAGE WITH PERSONAL STORIES		
Does the story provide any personal anecdote relevant to the subject of the story?	0.76	89%
6. BALANCE PERSONAL STORIES WITH SYSTEMIC CAUSES AND SOLUTIONS		

Does this story offer individual solutions?	0.69	85%
Does this story offer systemic solutions?	0.58	80%
7. AVOID SENSATIONALISM		
Does the headline accurately capture the content of the story?	—	100%
Does this story communicate anything about a scientific study's degree of certainty?	0.00	78%
8. USE WEIGHT-OF-EVIDENCE REPORTING TO COUNTER FALSE BALANCE		
Does the story include any competing perspectives on a problem?	0.24	74%
Does the story provide evidence supporting either perspective?	0.00	96%
9. ADOPT SOLUTIONS JOURNALISM TECHNIQUES		
Does the story present results, or indications of progress, linked to the solution?	0.20	74%
Does the story address the problem solving and how-to details of implementation?	0.41	85%
10. FOSTER DIVERSITY IN NEWSROOMS AND COVERAGE		
Is at least one author a woman?	0.86	94%
Is at least one author a person of color?	0.73	85%
Are there any people of color depicted in any photograph?	0.85	93%
Does this story make any reference to indigenous knowledge?	0.76	93%

APPENDIX D

TWITTER ANALYSIS METHODS

43. Search terms highlighted in gray generated no results. Search terms highlighted in yellow were the most relevant, and were retained in the final filtered dataset.

SEARCH PARAMETERS

In August, 2021, we acquired a Twitter Academic Research account (API v2) to access real time and historical data from Twitter. For each of the two CC stories selected for the deep-dive analysis, we identified a set of search terms, including story title, link to the story, project title, Twitter handles of the outlet and the authors, and key content-related terms.[43]

"From Rust to Resilience" Story

From Rust to Resilience, Climate Change and the Great Lakes Cities, Connected Coastlines, @karilydersen1, @ensiamedia, #RusttoResilience #GreatLakes, #CCNow, #letsbuildabetternormal, #cjs2020, #EnsiaEdge, Ensia, Kari Lydersen, Cuyahoga River, Great Lakes ecological recovery, Great Lakes warming, climate refuge, zebra mussels, quagga mussels, invasive sea lampreys, Ballast water regulations

"Flesh-Eating Bacteria" Story

A Flesh-Eating Bacteria Lurking in the Ocean Is Killing People in the Carolinas, Beyond the Beach, Connected Coastlines, @sfretwell83, @anaarikibandooq, @ Sofialmdo, @newsobserver,* Sammy Fretwell, Ali Raj, Sofia Moutinho, Vibrio, Bacteria vibrio, toxic vibrio, vibrio infections, vibrio-related illness, North Carolina's beach water testing program, Vibrio microbes, vibrio forecasting system, Hurricane Florence, tainted shellfish

We searched the period from **two weeks before to two weeks after** each story's publication date for the above search terms, using the Boolean operator "OR" to combine them. For each tweet that mentioned any of the relevant search terms, we retrieved pre-determined tweet fields and author fields:

Tweet fields:

- Text
- Type of tweet
- Username

44. Data were collected between April 27, 2015 and February 21, 2018, funded by a Canada Foundation for Innovation JELF Grant to Chris Bauch, University of Waterloo. The data were labeled by three reviewers who cross-checked and agreed upon all classifications.

- Created timestamp
- tweet location
- #Likes
- #Comments
- tweet URL
- Embedded media - Video, Photo, GIF

Author fields (for each tweet):

- Author's location
- Follower Count
- Following Count

We retrieved four **types of tweets** to capture various ways in which Twitter users engage in conversations:

- **Original:** Originally tweeted by a user
- **Replied to:** Another or the same user left a comment in a thread, which is a series of connected Tweets from any user
- **Retweeted:** Reshared the original tweet itself
- **Quoted:** Another or the same user retweeted and added extra words

After examining the full dataset, we generated a filtered dataset using the content-specific search terms that were most relevant to each story (as opposed to geographic or outlet-related terms). We used both the full dataset (N = 14,010 tweets) and filtered dataset (N = 2,271 tweets) to analyze the temporal trends in Twitter conversations. The filtered dataset was used for the **thematic**, **sentiment**, and **amplifier analyses**.

SENTIMENT ANALYSIS METHODS

The most commonly available sentiment classification models are pre-trained to score the sentiment of a given text as negative, neutral, or positive. For topics with an overall negative tone, such as climate change, this type of model skews toward negative sentiment ratings due to the nature of the topic rather than the text itself (i.e., Twitter users' attitudes toward climate change). Thus, a USC student programmer developed a new sentiment model to classify each tweet based on a combination of machine learning (ML) and human coding.

Machine Learning

To develop the ML model, we used a training dataset retrieved from the online data science platform Kaggle.[44] The dataset contains 43,943 labeled tweets pertaining to climate change, classified into one of four categories as follows:

- **News:** The message links to factual news about climate change.
- **Pro:** The message supports the belief in man-made climate change.

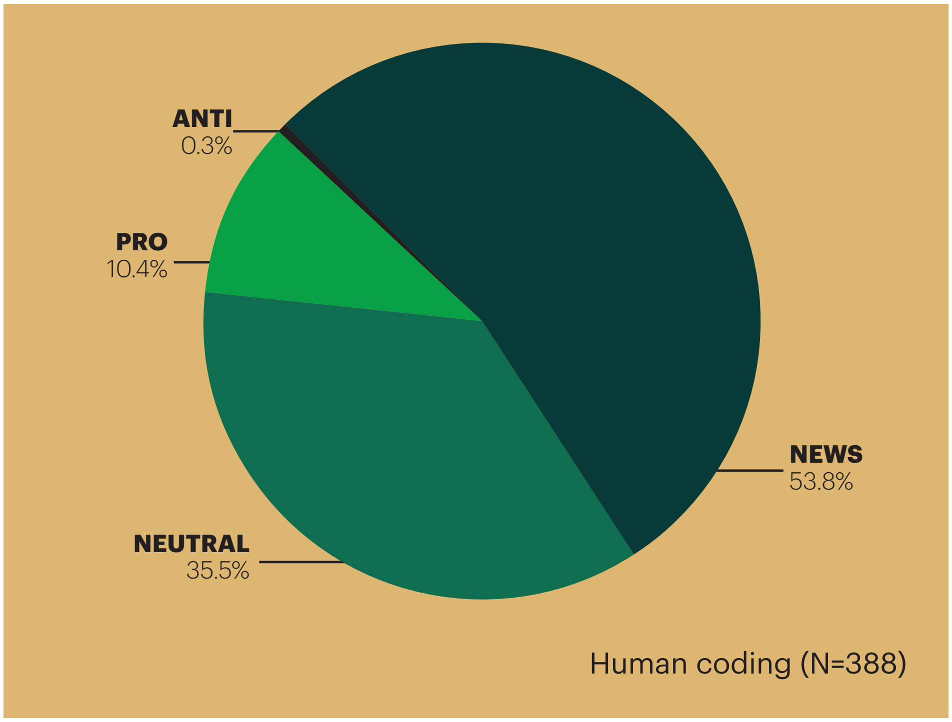
- **Anti:** The message is against the belief in man-made climate change.
- **Neutral:** The message neither supports (Pro) nor refutes (Anti) the belief in man-made climate change and does not link to news (News).

From this training dataset, 75% of tweets were used as training data and the remaining 25% test data to select the most accurate and best-fitting ML algorithm for classifying climate change related data, which was determined to be the Linear SVC model.[45] We applied this model to the complete Kaggle dataset, and then the filtered dataset of the two CC stories.

Human Coding to Refine ML Sentiment Model

To improve the accuracy of the initial ML sentiment model, two human coders analyzed 20% of tweets randomly selected from the filtered sample (N = 399 tweets). Human coders evaluated Twitter texts only, and did not click the link (URL) to tweets available online. Interrater reliability was within an acceptable range (Cohen's kappa = .64).

Figure D1.
Figure D1. Human coders' classification of tweets from the filtered CC sample.[46]

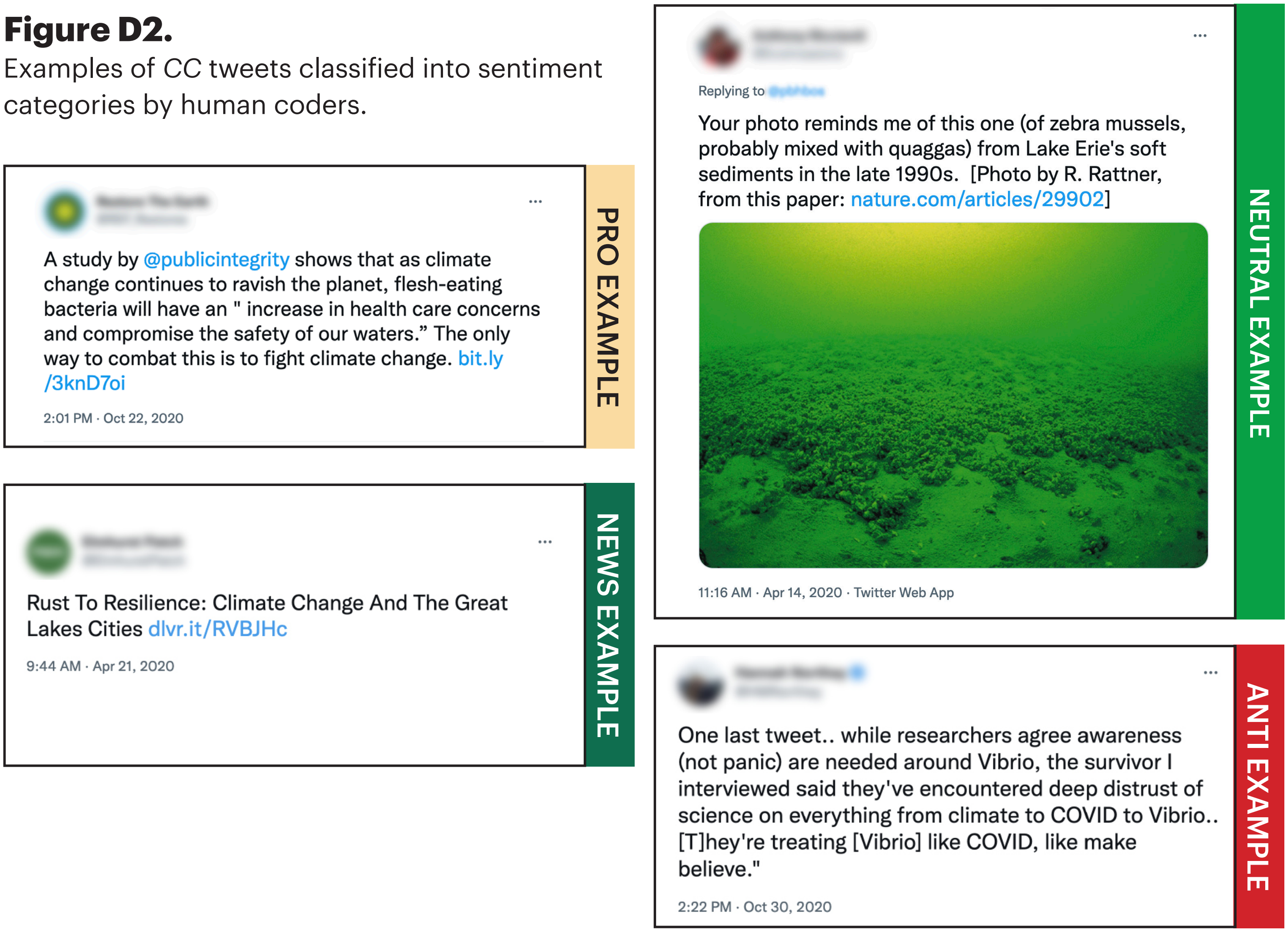


45. Data were fed into different Machine Learning algorithms, including Multinomial Naive Bayes, KNeighborsClassifier, Logistic Regression, Random Forest Classifier, and Linear SVC. Of these models, Linear SVC model achieved the highest accuracy of about 70% when applied to the test data. The Linear SVC model was further fine-tuned by modifying the parameters of the classifier, achieving an accuracy of 73% and improved F1 score (combining "recall" and "precision") for each of the four classifications (labels) used. The closer to 1 the F1 score is, the better the model. (F1 scores: Pro = 0.80, News = 0.75, Anti = 0.58, Neutral = 0.57.)

46. Eleven tweets classified as "unclear" were removed from the analysis.

Figure D2.

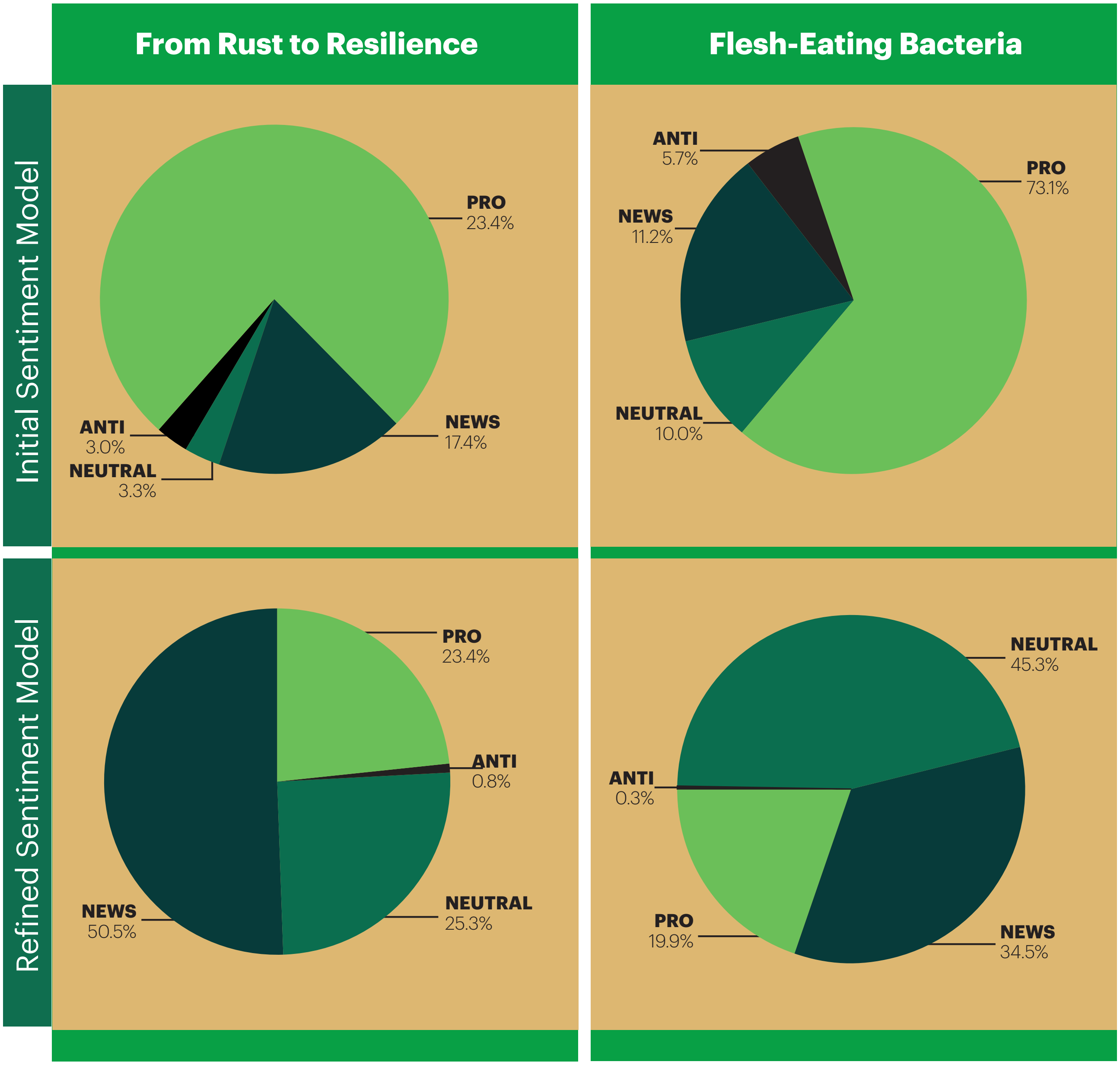
Examples of *CC* tweets classified into sentiment categories by human coders.



We added the human-coded data to the Kaggle training dataset (N = 43,943) and applied the model to the entire dataset (N = 44,460).[47] The sentiment of tweets in the final filtered CC sample (N = 2,271) was reclassified based on the refined model.

47. The sample used for human coding (N=399) combined with the Kaggle training dataset (N=44,061) adds up to 44,460.

Figure D3. Results of sentiment analysis based on the initial model and the refined model.



In the refined sentiment model, the proportion of tweets classified as News increased by about three times, and Neutral by about four times. This is largely because human coders distinguish tweets that simply relay the information about climate change (News) from those that reveal Twitter user's substantive reactions to the belief in man-made climate change (Pro). For example, according to human coders' classifications, tweets that explicitly recognize or imply the impact of climate change on human health/safety, and the role of humans were classified as Pro (10%). Tweets that simply state the title or link of climate change related stories were classified as News (54%). Tweets that add relatively informative or objective comments related to climate change were classified as Neutral (36%). Two tweets were classified as Anti, refuting the belief in man-made climate change (see Figures D2 and D3).

Human Coding to Contextualize Findings

To add further context to the findings from the refined ML sentiment model, two human coders analyzed six additional variables in the randomly-selected subsample (N = 399). Results of the reliability analysis indicate there was acceptable agreement between the coders on all items except for emotional reactions, for which reliability was borderline.

Table D1.
Interrater reliability for human coded items

Variable	Cohen's Kappa
Please label this tweet from the following classes? • **News**: The message links to factual news about climate change. • **Pro:** The message supports the belief in man-made climate change. • **Anti:** The message is against the belief in man-made climate change. • **Neutral:** The message neither supports (Pro) nor refutes (Anti) the belief in man-made climate change and does not link to news (News). • **Unclear**	0.64
Does this tweet discuss climate change? (Yes, No, Unclear)	0.91
If yes, what is the tweet's sentiment about *effects of climate change*? (Positive, Neutral, Negative, Unclear)	0.90
Does this tweet mention any *emotional reactions* [frustration, fear, inspiration, relief, etc] of the text/post author? (Yes, No, Unclear)	0.45
Does this tweet mention any *behavioral reactions* of the text/post author? (Yes, No, Unclear)	0.87
What kind of behavioral reactions of the tweet author are indicated? (Seek information, Engage with others, Call for action, Other: please describe.)	0.58

APPENDIX E

CASE STUDIES KEY INFORMANTS AND METHODS

KEY INFORMANTS

Carolinas story: "Flesh-Eating Bacteria":

- Sammy Fretwell, Author of Story, Journalist
- Dan Barkin, Editor at story publication
- Adam Waxman, Regional Journalism Development Director
- Jack Igelman, Journalist

Great Lakes story: "From Rust to Resilience":

- Kari Lydersen, Author of Story, Journalist
- Mary Hoff, Editor at time of story publication
- Jon Kealing, Director, Institute for Nonprofit News
- Sandy Svoboda, Director, Detroit Public Television

48. Environmental and Natural Resources State Tracking Bill Database. National Conference of State Legislatures. www.ncsl.org/research/environment-and-natural-resources/environment-and-natural-resources-state-bill-tracking-database.aspx

49. Online Business Research & News Media Database. LexisNexis. www.lexisnexis.com/en-us/home.page

50. Muck Rack. muckrack.com

51. CoverageBook. coveragebook.com

52. Backlink Checker. Ahrefs. ahrefs.com/backlink-checker

53. Plagiarism Checker. Dupli Checker. www.duplichecker.com/

54. CrowdTangle. www.crowdtangle.com

55. ATLAS.ti. atlasti.com

METHODS & ANALYSIS

Analytics	Methods	Tools and Data Sources
Regional legislative and policy trends	Relevant legislative database keyword searches in the last ten years, mapped over time and location	Bill text as indexed by the National Conference of State Legislatures[48]
Regional journalistic trends	Unique and relevant keyword searches in news databases, one search 6 months before the first story in our sample and one search 6 months after the last story in our sample	Scan of 15,000 U.S. news outlets indexed by *Lexis-Nexis*[49]
Story-related metrics	• Analyzed data provided by Pulitzer, authors, editors, and directors • Online content reach analysis to document sources that are linking, citing, and pulling content from the stories • Social media impression analysis	• Analytics gathered by Muck Rack[50] and Coverage-Book[51] • Recirculation analysis using Ahrefs[52] and Dupli Checker[53] • Social media analysis with CrowdTangle[54]
Insights and narratives related to story impact	Interviews with 8 key informants connected to the case study articles: authors, editors, directors, experts	Interviews were transcribed and coded for emergent themes and impact narratives using ATLAS.ti[55]
Uptake of identified best practices in the stories	Content analysis of published story, identification of best practices brought up in key informant interviews	Interviews were transcribed and coded for mentions of best practices using ATLAS.ti