CESifo WORKING PAPERS

10429
2023

May 2023

Analyzing Climate Change Policy Narratives with the Character-Role Narrative Framework

Kai Gehring, Matteo Grigoletto

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10429*

Analyzing Climate Change Policy Narratives with the Character-Role Narrative Framework

Abstract

Understanding behavioral aspects of collective decision-making is an important challenge for economics, and narratives are a crucial group-based mechanism that influences human decision-making. This paper introduces the Character-Role Narrative Framework as a tool to systematically analyze narratives, and applies it to study US climate change policy on Twitter over the 2010-2021 period. We build on the idea of the so-called drama triangle that suggests, within the context of a topic, the essence of a narrative is captured by its characters in one of three essential roles: hero, villain, and victim. We show how this intuitive framework can be easily integrated into an empirical pipeline and scaled up to large text corpora using supervised machine learning. In our application to US climate change policy narratives, we find strong changes in the frequency of simple and complex character-role narratives over time. Using contagiousness, popularity, and sparking conversation as three distinct dimensions of virality, we show that narratives that are simple, feature human characters and emphasize villains tend to be more viral. Focusing on Donald Trump as an example of a populist leader, we demonstrate that populism is linked to a higher share of such simple, human, and villain-focused narratives.

JEL-Codes: C800, D720, H100, P160, Q540.

Keywords: narrative economics, text-as-data, machine learning, large language models, climate change, virality, populism.

Kai Gehring
Wyss Academy at the
University of Bern / Switzerland
mail@kai-gehring.net

Matteo Grigoletto
Wyss Academy at the
University of Bern / Switzerland
matteogrigoletto@gmail.com

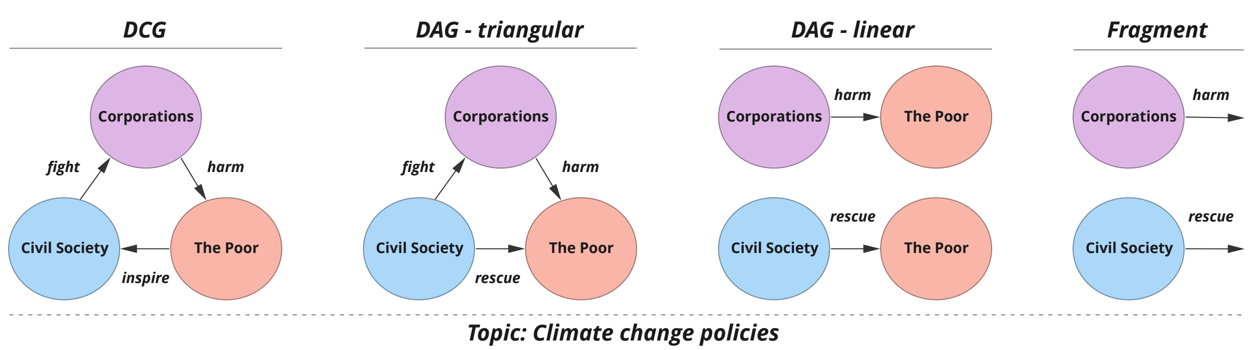
May 11, 2023
We thank Lina Goetze and Lorenz Gschwent for excellent research assistance, all remaining mistakes are ours. We are grateful for helpful suggestions by Toke Aidt, Elliott Ash, Vera Eichenauer, Joachim Voth and Paul Schaudt, as well as comments from seminars at Uppsala University, University of Bolzano, ETH Zurich and the University of Bern, as well as participants and discussants at the EPCS 2023 conference, Bolzano BBQ 2022 workshop, the 31th Silvaplana Workshop on Political Economy, the 4th Monash-Warwick-Zurich Text-as-Data Workshop, the 6th Economics of Media Bias Workshop, and the NLP for Climate Politics workshop.

1 Introduction

Narratives are one key mechanism that influences human decisions, highlighted in popular science books (Harari, 2014; Shiller, 2017) but also increasingly in theoretical (Bénabou *et al.*, 2020) and empirical studies (e.g., Bursztyn *et al.*, 2023b; Esposito *et al.*, 2023) in economics. In this paper, we build on insights from several disciplines to develop a novel framework – the **Character-Role Narrative Framework** – and a pipeline to measure narratives using supervised machine learning. Complementing recent survey-based papers analyzing support for climate change policies as a key policy question (Andre *et al.*, 2022; Dechezleprêtre *et al.*, 2022), we then apply our framework to measure the frequency and virality of climate change policy narratives on the social media platform Twitter.

Shiller (2017) highlights the importance of narratives for economics, initially relying on anecdotal examples and a broad definition of narratives as viral stories. Eliaz and Spiegler (2020) and Andre *et al.* (2023) take an important step forward to analyze narratives more systematically as causal or temporal sequences. The quote by climate activist Greta Thunberg and the accompanying illustration demonstrate how complex narrative structures can be represented as causal sequences using Directed Cyclic Graphs (DCGs) or Directed Acyclic Graphs (DAGs).

*"**Global greenhouse emissions** are still on the rise, oil production is soaring and energy companies are making sky-high profits while countless people struggle to pay their bills. [...] A critical mass of people – especially younger people – are demanding change and will no longer tolerate the procrastination, denial and complacency that created this state of emergency."* Greta Thunberg, *The New Statesman*, 19th Oct. 2022



These prior approaches have, however, two important limitations. First, coding the full causal graph structure is done manually and cannot easily be scaled-up for large corpora of text. Second, in much of human conversation as well as in newspapers, social media or speeches, narratives are simple and reflect only a fragment of those complex causal structures.

The core idea of our **Character-Role Narrative Framework** is that in the context of a specific topic – here climate change policies (*global greenhouse emissions*) – identifying its characters and their roles is sufficient to capture the essence of most narratives. The nodes of the graphs represent the characters: corporations (*energy companies*), the poor (*countless people*), and civil society (*younger people, activists*). Building on the so-called drama triangle (Karpman, 1968), characters reflect one of three essential roles: hero, villain, and victim. Defining narratives based on these character-roles is an intuitive approach that can be efficiently scaled-up to capture both simple and complex narratives.

Any entity can constitute a character, and it is helpful to broadly distinguish two categories: human and instrument characters. Human characters comprise entities composed of people, such as corporations, governments and political movements. Instrument characters comprise more abstract entities such as policies, laws, or technologies. Our **Character-Role Narrative Framework** thus defines narratives as follows:

> *"Complex narratives frame a topic by providing a temporal or causal sequence between two or more characters in three essential roles: hero, villain or victim; simple narratives are fragments of those sequences that feature only a topic and one character-role."*

In addition to this definition, we propose a pipeline to implement the framework empirically, that is easily adaptable to almost any context. Because the drama triangle is intuitive for humans, it can easily be used to annotate a training data set. This is then employed in supervised machine learning, adapting language models to the specific prediction task. Measuring the topic and character-roles does often, but not always, allow to retrieve the underlying causal structure. However, even in those cases and for simple narratives, measuring character-roles still provides valuable information. Why? Narratives matter because they can influence preferences and actions about aspects such as the role of government in the economy, institutions like central banks or specific policies, and depicting those as characters in specific roles is a key technique how narratives exert that influence.

We apply our framework and pipeline to analyze climate change policy narratives on social media in the US. Coping with climate change is arguably one of the most urgent challenges humanity is facing, but also a "wicked problem". Due to the time-lag between policies and their effect, and the strong re-distributional impact, discussions around climate change policies are often controversial and divisive. Survey evidence suggests that there is a lot of variation within and across countries in support of climate change policies, correlated with individual socio-economic as well as more psychological factors (Dechezleprêtre *et al.*, 2022). We move beyond survey evidence by studying the frequency and virality of climate change policy narratives on Twitter, which has developed into a major political arena for testing or positioning narratives in most Western countries (Halberstam and Knight, 2016; Acemoglu *et al.*, 2018; Macaulay and Song, 2022). We focus on the United States, which is a major CO2-emitter and a politically decisive country in the fight against climate change, but also deeply divided on the appropriate climate change policies.

To explore the possibilities of our framework, we use a large set of 12 human (e.g., government, corporations) and instrument (e.g., green technology, emission pricing) characters and their roles (hero, villain, or victim) in the context of climate change policy. In a first step, we use a set of keywords related to climate change policies (Oehl *et al.*, 2017) to download more than two million English-language tweets through the Twitter API from every Saturday over the 2010-2021 period. We then recruit three human coders through the Amazon Mechanical Turk platform and provide them with a detailed codebook to create a training dataset comprising over 10,000 tweets. We predict 18 character-roles that appear with a minimum frequency in the training data, with a high prediction performance. Hence, our framework can be implemented well with a reasonable amount of human-annotated text even for a large set of characters.

Examining our first set of descriptive results, we identify substantial differences and significant changes in the frequency of specific character-role narratives from 2010 to 2021. To illustrate a character-role narrative, consider Government-Villain, where the US government is portrayed as a villain in climate change policies, or Corporation-Hero, in which corporations are depicted as heroes taking positive actions against climate change. We document a marked increase in Government-Villain narratives, while for corporations villain narratives decrease and hero narratives almost triple. At the same time, narratives depicting The Poor –in both the US and globally – as the main victims of climate change policies grow substantially. As for instrument characters, Green Tech-Hero narratives, initially quite prevalent, experience a decline of over 50%, while Nuclear Tech-Hero narratives see a more than fourfold increase. These findings offer valuable insights into the evolving character-role narratives that shape the US climate change discourse.

Our second analytical contribution examines the virality of climate change policy narratives on social media. By working with Twitter data, we examine virality across three dimensions: contagiousness (retweets), the ability to spark a conversation (replies), and popularity (likes). We observe that factors influencing virality share similarities between contagiousness and popularity, while factors affecting conversation differ significantly. Across dimensions, hero narratives about civil society (e.g., Civil Society-Hero) and villain narratives about the US government or the BRICS (emerging economies, particularly China and India) stand out as viral. Our approach also allows us to identify specific character-role combinations associated with high virality for each dimension, with complex narratives featuring civil society as the hero opposing a Government-Villain as a prominent example.

Building on this descriptive analysis of specific character-roles, we then explore seven main determinants of virality more systematically using a Poisson regression framework. We find that, in terms of contagiousness and popularity, human characters outperform instrumental ones; however, no significant difference emerges concerning their ability to spark conversation. Generally, villain narratives demonstrate higher virality than hero or victim narratives, with the exception of instrument characters like Green Tech or Emission Pricing. In general, simple narratives featuring villain roles and human characters, exhibit a higher tendency for virality.

With these results on virality in mind we then investigate the impact of populism on the frequency of specific climate change policy narratives, comparing the Trump presidency as an illustrative example with both Obama's and Biden's term. Populism is often associated with the simplification and polarization of complex issues, potentially obstructing nuanced climate policy discussions. We use a regression spanning the entire period that accounts for a time trend and other controls as well as an event study specification focusing on the six months before and after the changes in presidency. In all specifications, we uncover a compelling link between the Trump years and a significant rise of viral aspects, such as human characters, more villain-focused and simpler narratives. The stable and significant differences between Trump and Obama, followed by a reversal under Biden, suggests this relationship could be causal. These findings contribute to a broader understanding of how populism may shape public narratives and the policy discourse in general.

We highlight our contributions to different strands of literature in section 2. Then, section 3 introduces the framework while section 4 outlines the implementation pipeline. Section 5 describes the data and section 6 evaluates the model performance. Finally, we present our results on climate change policy narratives in section 7, and conclude in section 8.

2 Contributions and Links to Literature

This paper contributes to various strands of literature. First, we contribute to the growing "narrative economics" literature (Shiller, 2017; Bénabou *et al.*, 2020; Eliaz and Spiegler, 2020), that emphasizes the critical role of narratives in shaping economic phenomena and decision-making processes. Despite the increasing interest, there is currently no coherent and systematic definition or measurement of narratives in this context. To address this gap, we propose the **Character-Role Narrative Framework**. Drawing on insights from various disciplines ranging from political science (Terry, 1997; Verweij *et al.*, 2006; Jones and McBeth, 2010; Jiangli, 2020), sociology (Brittain, 2006; Polletta *et al.*, 2011; Merry, 2016; O'Brien, 2018), communications studies (Anker, 2005; Gomez-Zara *et al.*, 2018) to literature (Fog *et al.*, 2010; Puckett, 2016), our framework provides a clear and specific approach for understanding and measuring narratives. Unlike narrower definitions that focus solely on causal or temporal chains, our approach encompasses a wider range of narratives. Additionally, unlike broader definitions that may lack specificity, our framework provides a clear and systematic way to identify and classify narratives. By introducing the **Character-Role Narrative Framework**, this paper offers a promising tool for researchers to more effectively understand and measure narratives in various domains.

Second, in addition to the conceptual advancements, we present and assess an efficient methodology to implement the **Character-Role Narrative Framework** in empirical applications. We build upon the foundational work of Gentzkow *et al.* (2019) and O'Neill *et al.* (2021), who provide comprehensive overviews of text-as-data approaches for economists, as well as specific narrative analysis packages by Shahaf *et al.* (2013) and Ash *et al.* (2023). These packages can serve as important inputs for the initial phase of our methodology, facilitating topic and character selection. Despite the extensive number of character-roles, our model achieves exceptional performance by employing a large language model (transformer) together with the ML algorithm XGBoost. The ongoing developments in computational linguistics will further enhance character-role predictions. Our methodology is accompanied by comprehensive instructions and adaptable code, fostering the replicability of our method.

Third, our unified framework that incorporates the most pertinent insights from various disciplines can help to bridge the gap between narrative analysis approaches in diverse economic fields. Recent empirical research investigates the impact of narratives across diverse topics, advancing our understanding of their role in political economy and macroeconomic contexts. This body of work spans subjects such as racism (Esposito *et al.*, 2023), diverging narratives during the COVID-pandemic (Bursztyn *et al.*, 2023b), and monetary policy perception and inflation (Andre *et al.*, 2023; Macaulay and Song, 2022).[1] Overall, there is an increasing interest in macroeconomics and finance to understand how collective behavioral phenomena like animal spirits (Akerlof and Shiller, 2010) can drive economic decisions. The **Character-Role Narrative Framework** can be adapted to study these types of narratives, for instance whether central

[1] While not specifically framing it as a narrative, Cagé *et al.* (2023) relate to the concept of heroes and villains by studying the impact of French general Petain's perception as a combat hero on political views and behavior during WWII.

banks and their instruments are perceived as heroes or villains in the fight against inflation, or as victims of uncooperative governments.

Fourth, by examining the frequency and virality of specific character-role narratives in real social media interactions, our results provide many additional insights and more nuances to an emerging and quickly growing literature on the political economy of climate change. Dechezleprêtre *et al.* (2022) demonstrate that, in a global survey of 20 countries, support for climate change policies hinges more on an individual's perception of the policy's effectiveness on their well-being and political orientation, rather than on their knowledge or understanding of climate change facts. Andre *et al.* (2022) highlight the role of social norms, moral values and the perceived behavior of others in determining the willingness to financially support climate change policy. Levi (2021) demonstrate that individual responsibility and political trust are strong predictors of carbon tax policy support in their study for 23 European countries.

Fifth, our approach and results also contribute and link to important strands of literature in other disciplines that we cannot fully acknowledge here. Berger and Milkman (2012) shows that New York Times articles which lead to high-arousal and activate emotions tend to be more viral. Our study provides more nuances to better understand the virality of online content, suggesting that human characters and villain roles are more successful in raising arousal and activating emotions. In political science, Bernauer and McGrath (2016) examine communication strategies for garnering public support for climate change, Keohane (2015) focuses on global political challenges posed by climate change, and Huber *et al.* (2020) use a conjoint experiment and a survey to study public support for climate change policies. Bernauer (2013) provide an overview of climate change research in political science.

Finally, our work contributes to a large literature in media economics (see Zhuravskaya *et al.* (2020) for a comprehensive review), by providing a novel framework and pipeline to analyze narratives in media. While initial seminal studies focus on newspapers (Gentzkow and Shapiro, 2010) more recent work investigates extensively the internet (Cagé *et al.*, 2020; Campante *et al.*, 2018) and social media (Acemoglu *et al.*, 2018; Enikolopov *et al.*, 2020; Müller and Schwarz, 2023). In contrast to previous findings that emphasize the limited persuasive power of facts compared to narratives (Bursztyn *et al.*, 2023a) and their inability to change policy conclusions from right-wing fake news Barrera *et al.* (2020), our approach facilitates a more nuanced exploration of narrative dynamics. By building on examples such as Campante *et al.* (2023), which demonstrates how politicians construct narratives by connecting unrelated events, our work paves the way for future research to uncover the underlying mechanisms driving the impact of narratives on public opinion and policy discourse.

3 Defining Narratives with the Character-Role Narrative Framework

3.1 Relation to Prior Approaches

Since starting to work on narratives, economists have successfully developed formal theoretical models and applied quasi-experimental approaches to establish causality in specific contexts. Nonetheless, the lack of a common framework that can be easily adapted and scaled up limits progress, makes communication unnecessarily hard, and yields results that are difficult to compare. This section begins by carving out the differences and commonalities in prior approaches – broadly distinguishing between broad and narrow narrative definitions – to then show how our framework builds on and augments these approaches.

Broad definitions include the seminal contributions by Shiller (2017, 2020) who defines a narrative rather loosely as "[...]a simple story or easily expressed explanation of events that many people want to bring up in conversation" (Shiller, 2017, p. 968). In Shiller (2017)'s initial work, virality was also regarded as a defining feature of a narrative. Many empirical papers investigating the impact of narratives provide similarly broad definitions. For example, Esposito *et al.* (2023) describe narratives as the "framing of memories and selective recollection of facts[...]" (p.2). While the broad scope allows those definitions to encompass almost any narrative, such an approach makes it hard to compare empirical results and cannot form the basis for measuring narratives more systematically.

In contrast, narrow definitions that view narratives only as temporal or causal sequences are analytically very useful, but can be too narrow to capture narratives comprehensively. Nonetheless, modeling narratives as directed graphs of sequences between different entities (Andre *et al.*, 2023) or similarly as mapping of "actions to consequences" (Eliaz and Spiegler, 2020, p.2) is an important step. As we show with the example in the introduction, directed acyclic graphs (DAGs) help to incorporate and graphically order variables according to the underlying causal or temporal model. Ash *et al.* (2023) develop a package that allows deriving all possible subject-verb-object sequences in selected texts, which can be seen as a subset of narratives that fit a specific DAG structure. We argue that DAG-based causal sequences constitute an important subset of all narratives, but miss some important ones. First, some complex narratives with a more circular reasoning are better represented as cyclic instead of acyclic graphs. Second, statements like "corporations are greedy" or "immigrants cannot be trusted" have no clear sequence, but represent common narratives.

The **Character-Role Narrative Framework** aims to reconcile the directed graph approach with those simple narratives by viewing them as fragments of the underlying grander narratives. Our framework aims to be broad enough to incorporate a larger share of narratives that are empirically important, but much more precise than initial broad definitions. Moreover, instead of making virality an ex-ante condition, our framework aims to allows analyzing which aspects of a narrative contribute to its virality. The next section describes our framework using narrative examples from various domains.

3.2 The Character-Role Narrative Framework

The **Character-Role Narrative Framework** identifies the topic and the characters in three fundamental roles – hero, villain, and victim – as the essence of a narrative. While other character schemes may broaden the range of roles, the drama triangle (Karpman, 1968) is consistently recognized as the central character set in human storytelling across various disciplines. Other character-roles tend to be peripheral and gain relevance in the story only through their connection and relation to the three main roles. While social identities and roles evolve constantly over time (Bergstrand and Jasper, 2018), the drama triangle remains relatively unchanged since the biblical story about the hero David fighting the giant Goliath to save his victims. Bergstrand and Jasper (2018) show empirically that most character traits can be clustered and assigned to those three roles, which seem to resemble the way that humans intuitively comprehend, interpret and memorize stories.

While using the drama triangle simplifies the content of a narrative, it goes much beyond mirroring positive and negative sentiments. The ability of the hero-villain-victim roles to capture the essence of a narrative is reflected in prior work in disciplines ranging from political science (Terry, 1997; Verweij *et al.*, 2006; Jones and McBeth, 2010; Jiangli, 2020), sociology (Brittain, 2006; Polletta *et al.*, 2011; Merry, 2016; O'Brien, 2018), communication studies (Anker, 2005; Gomez-Zara *et al.*, 2018) to literature (Fog *et al.*, 2010; Puckett, 2016). Villains often set a story into motion with their actions (Fog *et al.*, 2010). They can be characterized by their evil nature or their villainy emerges more passively, as they fail to act to pursue good. The hero, often embodying moral values, pursues an important goal or strives to save the usually helpless victim from the villain (Bergstrand and Jasper, 2018). Both hero and villain roles would not be fully captured by simple sentiment dictionaries, and victims do not fit into a simple one-dimensional positive-negative scheme.

Table 1 shows how our framework can be applied to narratives across different topics and domains, as well as its benefits and limitations. For each panel, it displays the topic, an example narrative, how this would be represented as a causal or temporal sequence (DCG or DAG), and the **Character-Role Narrative Framework** encoding. In each panel, we display a complex grander narrative and a possible simple narrative that relates to this grander narrative. We provide examples about immigration, monetary policy, and climate change policies, and use each to illustrate important aspects of the framework.

Consider the first panel in the table. Within the topic immigration, an example of a popular, grander narrative is 'The government lets in too many lazy immigrants, who only exploit our welfare state.' This narrative could be well represented as a unidirectional DAG. In the **Character-Role Narrative Framework**, we would identify government and immigrants both as villains, and the welfare state or the taxpayers financing it as the victims. In this example, the causal structure of the DAG could easily be inferred from the character-roles.

In many texts or conversations, people might express only a fragment of this complex narrative, for example 'Immigrants are lazy'. In this case, the only character-role would be Immigrants-Villain. How much we can infer from this character-role depends, ultimately, on the clarity and scope of the topic of interest. For instance, if the topic is 'Crime' or 'Welfare State', the character-role interpretation becomes clear. In the former, Immigrants-Villain refers to narratives on the impact that immigration has on criminality, while in the latter, narratives on the detrimental effect of immigrants on the Welfare State.

In the second panel, we use a very prominent narrative about the role of cryptocurrencies. Some version of 'Look at inflation, reckless central banks harm the poor. Bitcoin fixes this' was extremely viral on social media and beyond. As the illustration shows, this narrative would need to be represented in a DCG, rather than in a DAG. The **Character-Role Narrative Framework** would identify central banks as villains, the poor as victims and Bitcoin as the hero. In this example, it might not be possible to retrieve the exact causal structure from those character-roles as they might relate to more than one possible grander narrative. How problematic is that? Considering that framing characters in a certain role is a main channel how narratives influences preferences and decisions, this information is valuable and often even sufficient to understand the intent and possible effect of the narrative.

Table 1: *Examples of Narratives*

	Complex Narrative temporal/causal sequence	**Simple Narrative** sequence fragment
Topic: Immigration		
DAG / DCG / Fragment	Government → Immigrants → Welfare State	Immigrants →
Character-Role Narrative Framework	Topic: Immigration Character-roles: Government-Villain, Immigrants-Villain, Welfare State-Victim	Topic: Immigration Character-roles: Immigrants-Villain
Example of narrative	*The government lets in too many lazy immigrants, who only exploit our welfare state.*	*Immigrants are lazy.*
Topic: Inflation		
DAG / DCG / Fragment	Central banks → Inflation → The Poor ← Bitcoin	Bitcoin →
Character-Role Narrative Framework	Topic: Inflation Character-roles: Central Bank-Villain, The Poor-Victim, Bitcoin-Hero	Topic: Inflation Character-roles: Bitcoin-Hero
Example of narrative	*Look at inflation, reckless central bank policy harms the poor. Bitcoin fixes this.*	*Bitcoin fixes inflation problems.*
Topic: Climate Change Policy (i.)		
DAG / DCG / Fragment	BRICS → CO2 → The Poor	BRICS →
Character-Role Narrative Framework	Topic: Climate change Character-roles: BRICS-Villain, The Poor-Victim	Topic: Climate change Character-roles: BRICS-Villain
Example of narrative	*Why should the US working class bear the costs of climate change policies, which are useless because China irresponsibly increases its CO2 emissions so much anyway.*	*China irresponsibly increases its CO2 emissions.*
Topic: Climate Change Policy (ii.)		
DAG / DCG / Fragment	Corporations → Green Tech → Climate change	Corporations →
Character-Role Narrative Framework	Topic: Climate change policy Character-role: Corporations-Hero, Green Tech-Hero	Topic: Climate change policy Character-role: Corporations-Hero
Example of narrative	*If we want to stop climate change, we need innovative corporations to scale up renewable energy.*	*To fight climate change, we need innovative corporations.*

Notes: The table offers examples of narratives from three distinct topics. For each example, we provide the grander, complex narrative as well as a potential fragment that originates from it, forming a simple narrative. Additionally, we demonstrate how the **Character-Role Narrative Framework** would encode the sequence/fragment that captures the narrative.

The third and fourth panels show examples of climate policy narratives that are either very frequent or viral in our empirical application. The third narrative portrays BRICS (especially China and India) as villains exacerbating global warming, while the poor in the US are victims who suffer from climate change and/or costly climate policies. Our framework captures this complex narrative through BRICS-Villain and Poor-Victim character-roles. This narrative is often used to argue against ambitious climate change policies in the US, shifting blame to a scapegoat. Some variations depict BRICS actions directly harming the poor, while others emphasize that BRICS emissions outweigh the benefits of US policies burdening the poor. While condensing this into just two character-roles looses some of the nuance, they reveal this intent and the essence of those narratives well.

The fourth narrative is very frequent within the US climate change discourse, which could for example be expressed as 'If we want to stop climate change, we need innovative corporations to scale up renewable energy.' For this complex narrative, our framework would capture the character-roles Corporations-Hero and Green Tech-Hero. A related simple narrative that is also quite frequent could for example be 'To fight climate change, we need innovative corporations', featuring only Corporations-Hero. In both cases, we can infer from the topic and character-role that the purpose of the narrative is to convince readers that corporations are an important part of solving climate change problems. If a text contains both Corporation-Hero and Green Tech-Hero, this reveals the grander narrative structure: corporations are the heroes because they will foster the development and scaling of renewable energy.

In conclusion, the **Character-Role Narrative Framework** captures the essence of narratives in an intuitive way and can be adapted to any topic or domain. In many cases, we can infer at least the most important underlying causal or temporal sequences from the character-roles, but not in all. In particular simple narratives that involve only one character-role could theoretically be linked to more than one grander narrative. Still, we are convinced that measuring the topic and the three character-roles is a good compromise. Even when the full causal sequence cannot be retrieved, it provides useful information because the framing of characters in those roles is directly related to the mechanism how narratives can shape preferences and public opinion. While using a set of more than three roles could capture further nuances of a narrative, it would make it much harder to implement the framework empirically.

4 Measuring Narratives

4.1 Methodological Choices

Textual analysis methods can be broadly categorized into dictionary methods and machine learning approaches. Dictionary methods are quite straightforward, focusing on the frequency or share of certain keywords within the text. This approach can be particularly effective when predefined dictionaries exist for capturing the sentiment of specific emotions. However, the simple word-counting mechanism lacks the capacity to account for context and complex sentence structures. To tackle these limitations, newer approaches attempt to combine dictionaries with Natural Language Processing (NLP) tools, enabling a more nuanced understanding of the role of specific words within sentences and their interdependencies Gehring *et al.* (2022).

On the other hand, unsupervised machine learning approaches, such as topic models (Hansen *et al.*, 2018), employ algorithms to cluster similar text elements. This method requires users to set specific parameter values, which subsequently yield potential topics. The users then need to distil these potential topics into a smaller, manageable set, and interpret and label each topic based on the words it contains. Both dictionary methods and unsupervised learning demand a degree of manual intervention and the filtering of input text to reduce computational load. Even more recent tools, like RELATIO (Ash *et al.*, 2023), still require users to condense and aggregate a large set of potential causal sequences and entities into a smaller, more relevant set of narratives. This emphasizes that regardless of the degree of automation, human judgment remains a crucial part of text analysis.

Among the various advancements in Natural Language Processing (NLP), the development of Large Language Models (LLMs), often referred to as Transformer models, is particularly notable. These models are initially trained using unsupervised learning techniques on a vast amount of text data, learning language structure, grammar, and context. However, their performance can be substantially enhanced by additional supervised learning, or fine-tuning, for specific tasks or domains. This involves creating a training data set for a specific topic, providing the model with detailed information about the desired prediction, and tailoring the model to the nuances of the task at hand. This two-step process leverages the broad language understanding of the LLM from unsupervised learning while adapting it to specific tasks through supervised learning.

One of the reasons why this approach is significant is that the interpretation of narratives, characters, and roles is culture-specific. People from different countries, such as Kenya, Australia, or the United States, may interpret a given text differently, even if they share the same official language. This doesn't refute the existence of universal *motifs* (Michalopoulos and Xue, 2021). Still, it highlights the importance of cultural context in understanding narratives, suggesting that annotators from the same country or culture should create the training dataset.

In the forthcoming subsection, we outline our approach for a pipeline that utilizes a popular, open-source Transformer model. This choice ensures replicability and broad accessibility of our approach. As LLMs continue to evolve, their predictive performance will improve, which, when combined with supervised learning, allows researchers to leverage a relatively small amount of training data for complex predictions. Our proposed supervised machine learning pipeline, detailed in the next section, represents a practical, efficient, and replicable methodology for the systematic measurement of narratives in text data.

4.2 Pipeline

This section outlines our 5-step pipeline, an instruction for (i.) selecting and defining the topic of interest, (ii.) identifying a source and extracting data, (iii.) identifying relevant characters and drafting a codebook for annotation, (iv.) annotating a training set of data and (v.) training and using the prediction model. The model combines the strengths of novel large language models - Transformer models - with a supervised machine learning algorithm XGBoost. The guideline is kept concise here. We provide detailed information on how to replicate or adapt our pipeline in Appendix A.

1. Selecting and defining the topic of interest

A well-defined topic is a prerequisite for a fruitful narrative analysis. The clearer the topic, the more straightforward the identification of relevant characters and the exploration of the research question. We propose two main strategies for topic selection and definition:

- For researchers initiating their inquiry from the beginning, we recommend immersing in the relevant scientific literature. Qualitative research often provides nuanced insights, aiding in the identification of key themes and issues related to the topic.
- For those with access to an existing text corpus, leveraging Natural Language Processing (NLP) methods proves to be efficient. Topic modeling helps isolate the main discussion clusters in the text, providing a preliminary understanding of the discourse structure. More sophisticated tools like RELATIO, for instance, assist in identifying common relationships among entities within the text corpus.

In our application, we engage with the discourse on the political economy of climate change. After reviewing the relevant literature, we identify two main streams of discussion: one concerning the scientific evidence for climate change, and the other debating the appropriate policy responses. Given our focus on the political economy aspect, we concentrate on climate change policies, consciously steering clear of the debate over the reality and predictability of climate change.

2. Identifying data sources and extracting data

After selecting the topic, the next step consists in gathering data. The most common sources of text data in economics are digitized newspapers and social media. However, researchers can also consider resources like transcribed TV, radio, and YouTube broadcasts, or open-ended responses from surveys for more unique insights. Although our pipeline is compatible with any text type, we recommend annotating data and training models separately for different sources due to their structural variations.

For this study, our focus is on narratives about climate change policies in the United States, collected from the social media platform Twitter over the period 2010-2021. We choose the US due to the significant role Twitter plays in shaping and disseminating narratives in this country. The data collection process involves querying the Twitter historical APIv2 with a set of keywords adapted from Oehl *et al.* (2017). This results in two datasets: one for model training and a larger one for narrative detection. After data cleaning, our analysis dataset comprises 3,279,730 tweets, with 26% of them originating from the US. subsection A.2 provides details about data extraction.

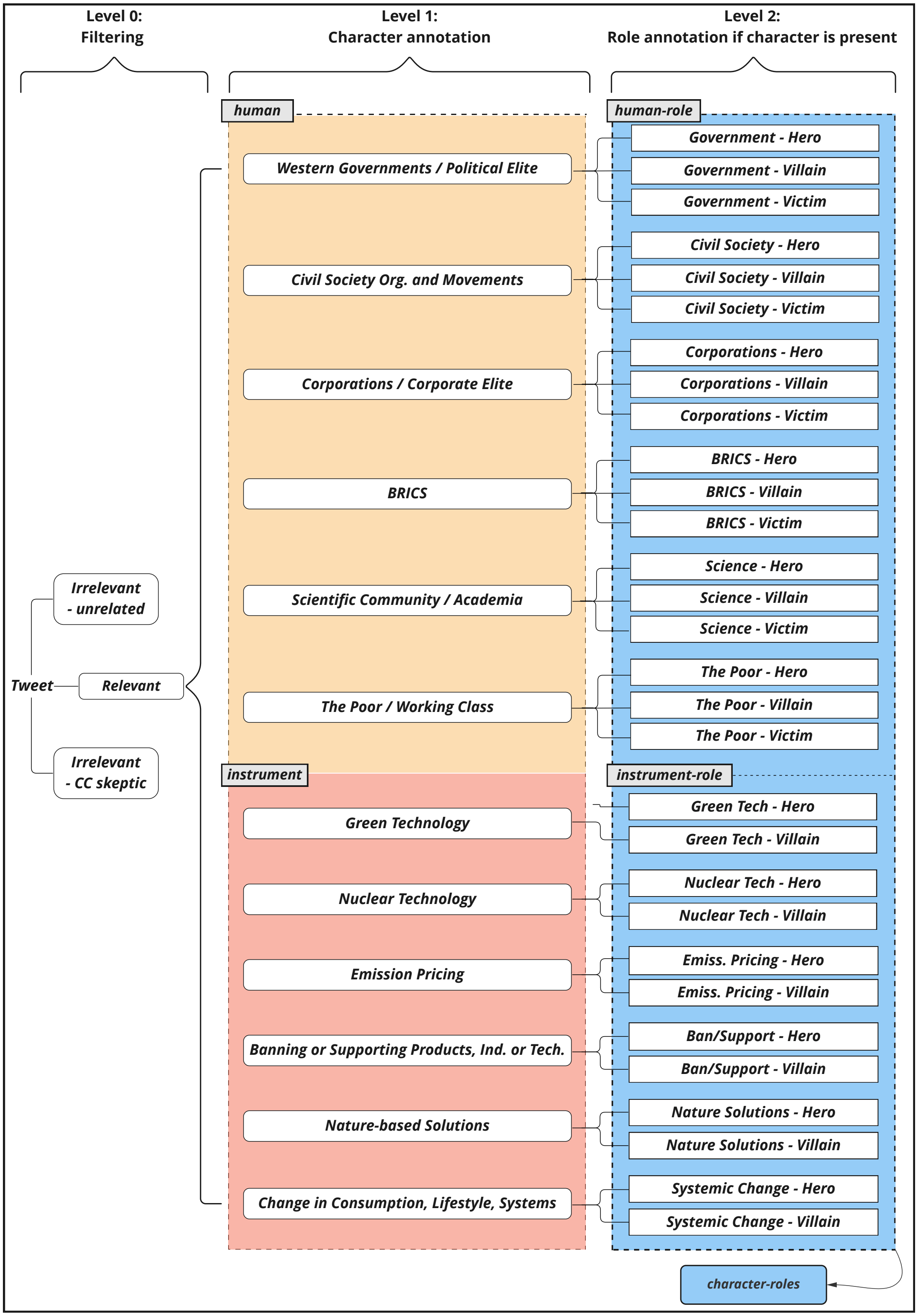
3. Identifying characters and preparing the codebook

The third step is a critical juncture in the pipeline, focusing on the identification of relevant characters within the topic and the development of a codebook for annotation. A codebook is a comprehensive guide outlining how data should be interpreted and categorized for the purpose of analysis.

The selection of characters is largely influenced by the researcher's interests and the research question. However, it is essential to balance the desire to explore a broad array of characters with the practicality of model training. A larger and more complex set of characters requires a larger dataset for training and may increase computational costs. Furthermore, defining precise and narrowly scoped characters could improve prediction performance as broad or vague categories can be challenging for the model to learn.

Characters can be identified through various methods, including literature review, topic modeling, and entity recognition tools. The codebook guides the annotators through the process, which we recommend to structure into three main steps: assessing relevance, identifying characters, and assigning a role. Figure 1 illustrates those steps and shows all human and instrument characters.

Figure 1: *Annotation Codebook: Schematic Representation*



Notes: The figure shows a schematic representation of the codebook used for the manual annotation process. In level 0 each tweet is classified as *irrelevant* or *relevant*. For those labeled as *relevant*, coders assess the presence of a character in level 1. In level 2, if a character was detected, coders further assess whether the character is depicted as *hero*, *villain* or *victim*. We report the codebook in its full length in Appendix A.

In our study, we curated the initial list of human and instrument characters from both the academic literature on climate change as well as climate change policy discussions in the media. The lists were further refined using entity recognition and word frequency, and validated through iterative feedback from external reviewers. The resulting codebook and annotation process are outlined in Appendix A.

4. Annotating of the training dataset

The fourth step involves annotating a training dataset. This process can be performed by the researchers themselves or by external coders, and can involve one or multiple coders. The number of coders can influence the quality of the training dataset. While a single coder may provide a consistent dataset, leading to high performance, the model might risk reflecting the coder's biases. Conversely, multiple coders can yield a more diverse, but potentially noisier, dataset that may result in a less biased model albeit with slightly lower performance.

In our study, we employ external coders from Amazon Mechanical Turk. After a selection process, the top three coders were chosen from over 60 interested individuals. Coders annotated tweets in a three-level process depicted in Figure 1: Level 0 is filtering out irrelevant tweets, Level 1 is detecting the presence of human and instrument characters, and Level 2 assigning the roles. Our final training dataset consists of 10,230 tweets, partly randomly drawn, partly oversampled using specific keywords to increase representation of underrepresented characters (see subsection A.4).

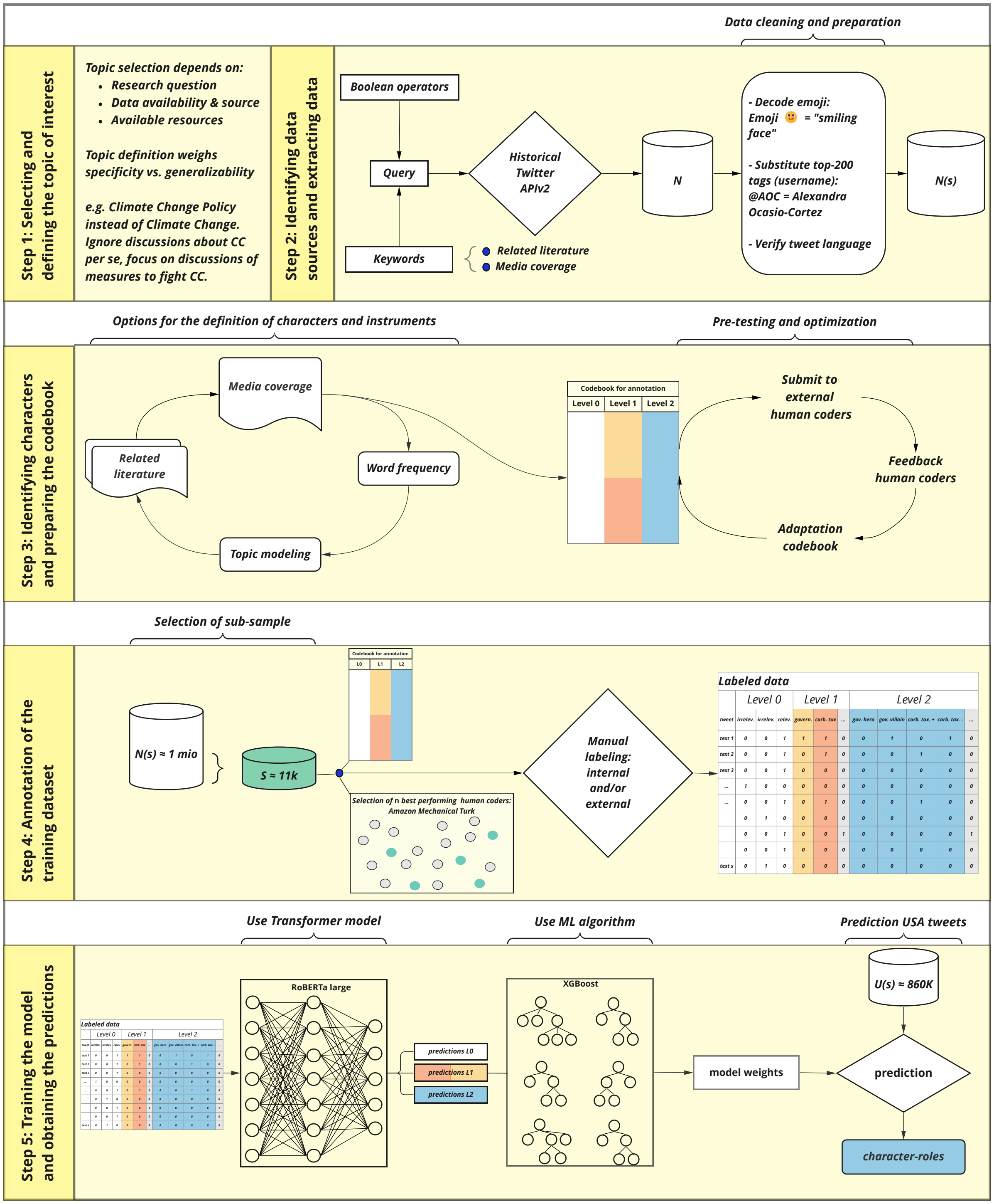
5. Training the model and obtaining predictions

The final step involves training the model and generating predictions. The objective is to predict the presence of specific character-role combinations in the input text, such as a tweet, paragraph, or article.

In this study, we use a Transformer model (see Vaswani *et al.*, 2017), RoBERTa, and the ensemble algorithm XGBoost. The integration of these models allows us to take advantage of the advanced natural language processing capabilities of RoBERTa and the scalability and robustness of XGBoost's decision tree-based gradient boosting algorithm (Chen and Guestrin, 2016). RoBERTa was chosen for its balance of performance and accessibility; for a more detailed discussion on why RoBERTa was chosen, please see the 'Methodological Choices' section. XGBoost was chosen due to its versatility in handling different types and distributions of data, its robustness to outliers, and its ability to handle missing values.

Our model is applied to the subset of US-originating tweets, totaling 859,738. The flexibility of our pipeline allows for future adaptation with newer models as they become available. Figure 2 provides a visual representation of our 5-step pipeline.

Figure 2: *Pipeline of Work: Schematic Representation*



Notes: The figure presents a schematic visualization of the pipeline followed by our analysis. The pipeline can be adapted to measure narratives in different settings. The process comprises five main steps. First, the selection of the topic of interest and the formulation of the research question. Second, the extraction of text data from the source of interest. Third, the identification of the relevant characters and the formulation of the annotation codebook. Fourth, the annotation of a training dataset, via external or internal coders. Fifth, the training and application of the prediction model. The latter is a combination of the Transformer model and the XGBoost algorithm. Figure A.1 provides more information on inputs, outputs and accessibility of our code.

5 Data

5.1 Twitter Data

We perform our analysis on a set containing a total of 3,359,627 tweets, collected from each Saturday between 2010 and 2021 (details in appendix A.2). We exclude non-English tweets and retweets while querying the Twitter API, resulting in a set of exclusively original tweets in English. We conduct ex-post checks to drop potential duplicates, decreasing the number to a total of 3,279,730 non-duplicated tweets.

Our focus is on the conversation about climate change policies within the United States. Thus, we narrow down our analysis set to only those tweets that we can locate in the US. This final data set consists of 859,738 tweets. We choose the US as our focus due to its position as world's second-largest emitter of CO2 (Sohail *et al.*, 2022). The discussion about climate change has always been particularly controversial in the country, with Twitter serving as an important arena of exchange. Tweets do not inherently come with geographical information and we use two approaches to locate them. First, we exploit the information that users often provide in their profile about their location. Second, we rely on the use of geo-localization tags that a minority of the tweets include in their meta-data. In fact, Twitter allows users to "tag" their tweets to a specific location, assigning a coordinate box. More details can be found in appendix A.

Table 2 provides descriptive statistics for the analysis tweets originating from the US. The average number of words in the tweets of our sample is 23, excluding in this hashtags and mentions. The average tweet in our sample is liked 11 times and re-tweeted 2.2 times. There is a lot of variation in the sample, suggesting that those differences are an important feature of a tweet.[2]

Table 2: *Descriptive Statistics: Tweets from the United States*

	Mean	**Median**	**St. Dev.**	**Min.**	**Max.**
Virality					
No. of Retweets (Contagiousness)	2.2	0	244	0	194,217
No. of Replies (Conversation)	.78	0	52	0	30,887
No. of Likes (Popularity)	11	0	1,266	0	896,759
Control Variables					
No. of Words	23	20	13	0	79
No. of Hashtags	.41	0	1.1	0	28
No. of Mentions	1.7	1	5	0	51
No. of Followers	10,203	625	356,243	0	133,245,480
No. of Following	2,497	749	13,217	0	4,066,970
No. of Tweets	49,975	16,118	114,533	1	3,671,810
No. of Observations	859,738				

Notes: The table reports descriptive statistics for the tweets extracted via the Twitter APIv2, that we can geo-localize at least at the US-country level. For all the variables indicated we provide information on average value, median, standard deviation, minimum and maximum values across the sample. The number of words per tweet is calculated disregarding hashtags and mentions. Table C.2 reports the same statistics for the full extracted dataset.

[2] We provide descriptive statistics for the full sample of extracted tweets in table C.2. Figure C.3 shows the spatial distribution of tweets in US. We provide information on the localization of tweets according to precision in Figure C.2.

5.2 Annotated Data

Before annotation, we decided to only select character-roles that appear in at least 100 annotated tweets. We make an exception only for the character-role Nuclear Tech-Villain, due to its high performance despite the lower number of annotated tweets. We randomly draw the initial annotation dataset from the sample of tweets used to train the model. Researchers can over-sample particular character-roles of interest, even if they are less frequent, by utilizing pre-selected keywords. For some character-roles, we use specific keywords to pre-select additional tweets, as explained in Appendix A. More annotated tweets generally improve the performance, but the number depends on the researcher's budget.

Table 3 provides a summary of the frequency with which character-roles are present in the annotated set of tweets. Panel A and B report the frequencies of each character-role element respectively for human and instrument characters. The fourth column of panel (a) also shows how often a character appears in a more neutral way without taking on one of the three roles. Overall one downside of our character selection is that only one character, The Poor, appears sufficiently often as a victim. A good choice, retrospectively, would have been to separating the poor in the US and people in poor countries that are also affected by climate change.

Table 3: *Frequency of Character-Roles in Annotated Tweets*

(a) *Human-role*

	Hero	Villain	Victim	None	Total
Government	430	1,372	42	649	2,493
Civil Society	358	111	23	126	618
Corporations	200	390	18	146	754
BRICS	125	231	30	121	507
Science	195	30	13	178	416
The Poor	98	63	358	114	633

(b) *Instrument-role*

	Hero	Villain	Total
Green Technology	798	123	921
Nuclear Technology	207	76	283
Emission Pricing	225	189	414
Ban/Support	269	42	311
Nature Solutions	139	13	152
Systemic Change	800	121	921

Notes: The table shows the absolute frequencies of character-roles in the annotated data. Panel (a) displays how often characters of the *human* type were annotated. Column *None* of Panel (a) collects information for cases where the character is present in the tweet but it is not depicted in one of the three roles. Panel (b) displays how often characters of the *instrument* type were annotated. When identified, these characters are always depicted either as *hero* or *villain*.

6 Model Performance

6.1 Prediction by Character-Role

In machine learning, model performance is typically evaluated using a number of key indicators. For our study, we primarily use the F1-score, a standard measure that combines precision and recall, two critical indicators in evaluating the performance of classification models. Precision, denoted as p, is the proportion of true positives (tp) to the sum of true positives and false positives (fp):

$$p = \frac{tp}{tp + fp}$$

Recall, denoted as r, is the proportion of true positives to the sum of true positives and false negatives (fn):

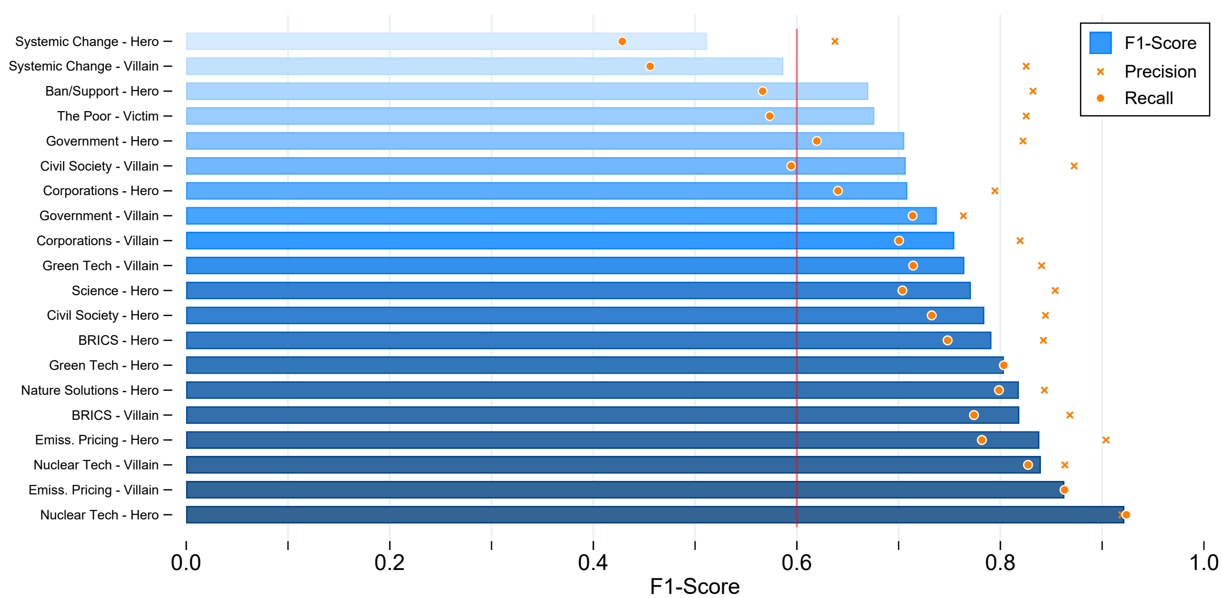
$$r = \frac{tp}{tp + fn}$$

The F1-score, then, is the harmonic mean of precision and recall. It is a balanced measure that gives equal weight to both precision and recall, making it particularly suitable for our study where both false positives and false negatives are of concern.

Many studies in economics and social sciences often report accuracy, which is a more limited measure of performance. Therefore, in our work, we compare our results with those of similar studies, such as Salminen *et al.* (2020), who explore a range of models predicting hate speech in tweets with F1-scores ranging from 0.6 to 0.91. Given the complexity of our task, predicting multiple character-roles simultaneously, we set a minimum target F1-score of 0.6.

Our model performs remarkably well, with F1-scores varying between 0.51 and 0.92 across different character-roles (Figure 3). Sixteen out of twenty character-roles have an F1-score of 0.7 or higher, and seven have a score of 0.8 or higher. Precision is consistently higher than recall, suggesting that the model excels at identifying instances of a character-role while avoiding false positives. For our main analysis we use the 18 character-roles that have a performance higher than 0.6, dropping the Systemic Change character. Before continuing our analysis, the next section will provide an overview of factors that can or could affect the performance of the model

Figure 3: *Model Performance by Character-Role*



Notes: The figure shows the performance of our prediction model for all character-roles of our analysis. The bars indicate the F1-Score, the harmonic mean of two indicators, precision and recall. Precision is defined as $p = \frac{tp}{tp+fp}$, where tp stands for true positive while fp for false positive. Recall is defined as $r = \frac{tp}{tp+fn}$, where tp stands for true positive while fn stands for false negative. In Figure F.4 we include the model performance for the categories *Denier* and *Irrelevant* from level 0 of the prediction process.

6.2 What Affects Performance?

The predictive performance of the model depends on several factors, including the complexity and interpretability of character-roles, the number of annotated tweets, and the level of agreement between coders.

The complexity of character-roles influences the model's performance. For instance, the character-role Nuclear Technology is less complex with a limited set of representative terms, making it easier to predict. Hence, for Nuclear-Villain 76 training tweets are sufficient for a very good performance. In contrast, the character-role Systemic Change, being broad and complex, is more challenging to predict due to varied interpretations among coders.

Our analysis reveals that there is not a strong positive correlation between the number of annotated tweets and the model's performance (Figure F.5a). This may seem counter-intuitive as one might expect more training data to enhance the model's performance. However, this underscores the fact that the quality of annotations and the clarity of the character-roles may matter more than the sheer number of annotations. Thus, the number of annotated tweets should not be the sole focus; the narrative's complexity and the coders' agreement on interpretation also significantly influence the model's performance.

Indeed, our findings show a strong positive relationship between the agreement among coders and model performance (Figure F.5b). This emphasizes the importance of clearly defined character-roles and a good codebook to ensure consistent interpretation among coders. Despite the negative impact of coder disagreement, as seen in the 'Systemic Change' example, we also find that having more coders annotate tweets for the same character-role narratives can be beneficial on average (Figure F.5c). This suggests that the diversity of perspectives can enhance the model's performance, even if it also introduces some level of disagreement.

In light of these findings, we recommend a minimum of 100 annotated tweets per character-role, while also taking into consideration the narrative's complexity and the agreement among coders. Rigorous pre-testing, careful selection, and clear definition of character-roles are also key to achieving optimal performance. However, there exists a trade-off between having multiple coders per character-role and the associated costs. Therefore, it's important to balance the benefits of diverse interpretations with the additional resources required.

In essence, our model's performance hinges on the balance between the number of annotated tweets, character-role clarity, and coder diversity. A good start is having around 100 annotated tweets per character-role, but the interpretability and complexity of these roles are equally crucial. Multiple coders can enhance performance, even if there is disagreements among them. However, this requires additional resources. Therefore, focusing on clear, unambiguous character-roles and thorough pre-testing can substantially boost the model's predictive capabilities.

7 Analytical Results

7.1 Narrative Frequency and Structure

The first research question that we address is the frequency with which different character-roles appear, and how their share changes over time. To assess this, we use our model to predict narratives in our sample of 859,738 English-language, climate-related tweets originating from the United States. Overall, our time period ranges from January 2010 to September 2021, covering a considerable period with many economic and political changes, including decisive climate meetings.

The share and changes in climate policy narratives are captured in Figure 4. The figure illustrates the frequency of appearance of different character-roles in tweets, but it is not yet distinguishing particular combinations of different character-roles in the same tweet, which we explore below. As of 2021, the most common character-roles on average are, by far, Green-Tech-Hero (28.1%) and Government-Villain (26.1%). Other common elements are Government-Hero (9.1%), Corporations-Hero (6.3%), Corporations-Villain (5.6%), Emission Pricing-Hero (3.9%) and The Poor-Victim (3.4%). All other elements have shares lower than 3%.

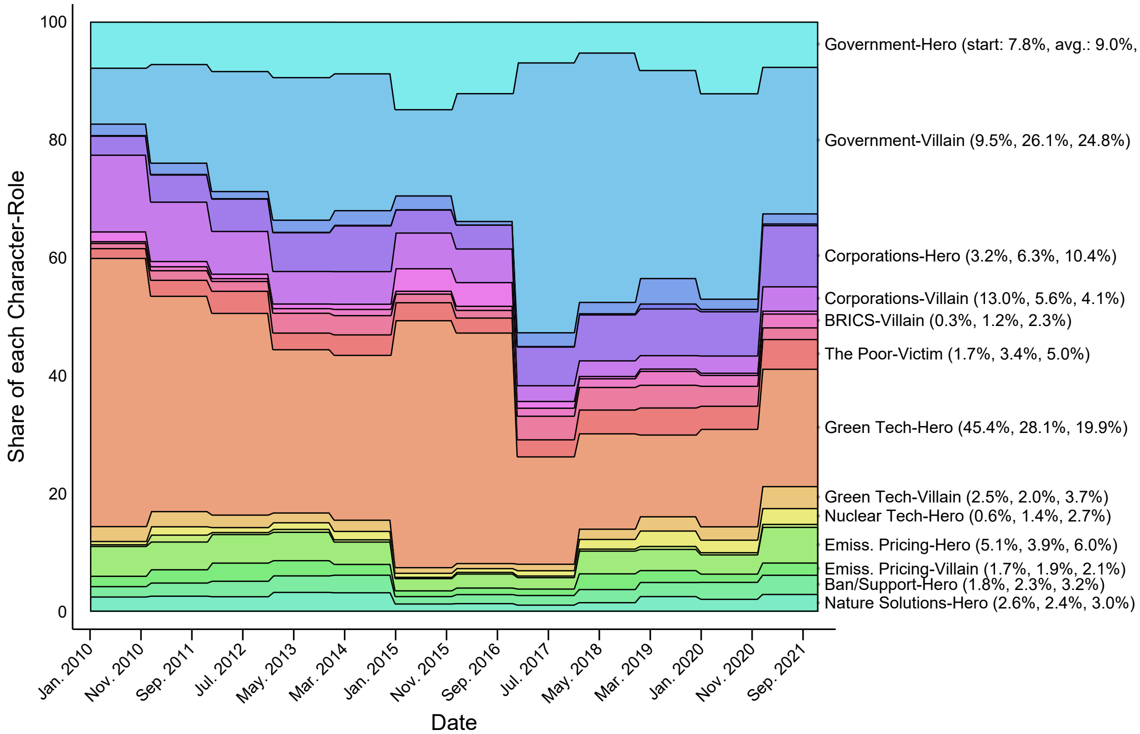
There are major changes over time, however, marking drastic shifts in the discussion about climate change policies. Government-Villain narratives started at just 9.5% in the 2010s. The rise of this character-role could signal both increasing impatience by those who want more ambitious policies, but also increasing anger by those who already find current interventions as too much of a burden or interference in free markets. It goes along with an increase of The Poor-Victim, the narrative of common people suffering from climate change and bearing the worst consequences of the policies addressing it.

Another drastic change concerns the decline of the Green Tech-Hero narrative, from 45.4% to 19.9%. This is not a sign of generally declining optimism about technology, as Nuclear Tech-Hero increased from 0.6% to 2.7%. Hero narratives about Emission Pricing and Nature Based Solutions also increased in time, but less dramatically. The strongest increase, besides Nuclear Tech narratives, is in hero narratives about banning or subsidizing specific products or industries.

Taken as a whole, these changes suggest a growing frustration with government inaction or even support for fossil industries and other contributors to climate change.[3] Instead of relying on government to be the hero, we see a marked increase in the Corporations-Hero narrative from 3.2% to 10.4%, while Corporation-Villain narratives decline from 12.8% to 4.1%. Science as the hero also more than doubles, at smaller levels, whereas Civil Society-Hero stagnates at low levels. Hence, as of 2021, the narratives that are increasing in frequency are centered around hero roles for corporations, science and nuclear technology, and villain role for the government.

[3] Note that we are not capturing the most recent climate bills in the US, which might reinstate more positive narratives about the government's role.

Figure 4: *Frequency of Character-Roles over Time*



Notes: The figure shows the frequency of each character-role over time. We include only character-role for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. The shares of character-roles are cumulative (i.e. sum up to 100) and are computed on a yearly basis. We label the most common character-roles. In parentheses, each label indicates the share of the corresponding character-role in the first period, the average share across all periods, and the share in the last period. Unlabeled character-roles are the following: Civil Society-Hero, Civil Society-Villain, BRICS-Hero, Science-Hero and Nuclear Tech-Villain. Figure C.4b shows the yearly version weighted by retweets to account for the virality of narratives. Figure C.5a shows the share computed on a six-month basis. Figure C.5b shows the six-month version weighted by retweets.

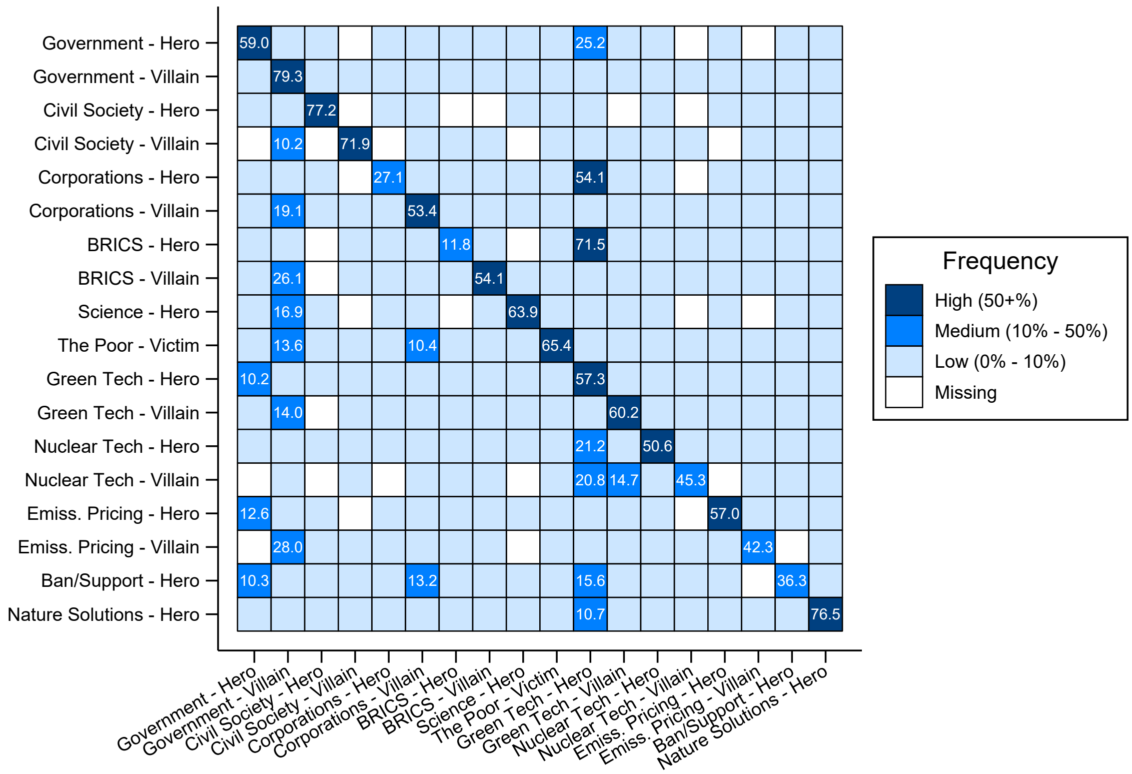
While the previous Figure shows the overall frequency with which a specific character-role appears, Figure 5 depicts all possible combinations of that with other character-roles. The Figure shows all combinations in a matrix form, with the diagonal showing occurrences of the character-role in simple narratives, hence without any other character-roles. The matrix should be read by row, with each row representing one character-role relative frequency, summing up to 100%. One interpretation is that the lower the share displayed on the diagonal, the more complex discussions involving that character-role.[4]

For 16 out of 18, the simple narrative form is the most common way in which that character-role appears. The exceptions from this pattern are Corporations-Hero and BRICS-Hero, which appear most often not alone, respectively 27.1% and 11.8%, but paired with Green Tech-Hero, respectively 54.1% and 71.5%. Narratives that appear more than 75% of the times alone are Government-Villain, Civil Society-Hero and Nature Solutions-Hero. There are also combinations that never occur together, like Science-Hero

[4] Some tweets comprise combinations of three or more character-roles hence they enter the matrix for each pairwise combination. These tweets represent a minority, as shown in Figure C.6. About 20% of all climate policy narratives contain two different elements, but only less than 3% feature three or more.

and Emission Pricing-Villain. Some other patterns about common complex narrative combinations are visible. Complex narratives including Green Tech-Hero comprise a large share of the complex narratives present in the tweets. Although to a lesser extent, a similar observation applies to Government-Villain.

Figure 5: *Frequency of Narratives*



Notes: The figure shows how frequently each character-role occurs with another one (or by itself). We include only character-roles for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. Each square of the matrix is divided by the respective row-wise total. The diagonale depicts how often each character-role appears in a simple narrative, i.e. when it is the only character-role detected in the tweet. Tweets containing three or more character-roles enter the matrix more than once, e.g. a tweet containing Government-Hero, Corporations-Villain and Ban/Support-Hero will enter the matrix for each pairwise combination of the character-roles that it includes. Tweets with more than two character-roles represent less than 3% of the data as shown in Figure C.6.

There are several other noteworthy character-role combinations emerging in Figure 5. A large share of Government-Hero narratives link the element to Green Tech-Hero. For Emission Pricing, a large share of villain narratives are linked to Government-Villain narratives, and a large share of hero narratives are linked with Government-Hero narratives. This is in line with previous research (Fisher *et al.*, 2013; Kousser and Tranter, 2018) and it exemplifies how polarized the discussion about pricing emissions to internalize externalities is, closely linked to perceptions about the role of government. Ban/Support-Hero appears often together with Corporations-Villain and Green Tech-Hero. Proponents of ban or support narratives seem to most commonly link those with narratives that villainous companies do not act if not being forced to. Alternatively, the latter combination concerns most calls to ban fossil fuels or subsidies to help the breakthrough of renewable energy.

Finally, the discussion around Nuclear vs Green Tech (renewable energies) is complex and interesting. A high share of Nuclear Tech-Hero narratives, around 21.2%, also feature the Green Tech-Hero character-role. Those tweets often argue that technology will be a key element to fight climate change, and that it is necessary to overcome ideology as both nuclear and renewables are needed to save humanity from the climate collapse. The contrast to that are the 20.8% combining Nuclear Tech-Villain with a Green Tech-Hero narrative. This usually portray nuclear as an obstacle to the faster adoption of the necessary renewable green technology.

Although Twitter is dominated by short and relatively simple narratives, likely due to the platform's word limit, it is important to note that complex narratives play also a significant role in the discourse on climate policy. Our approach is effective in capturing both simple, slogan-like narratives and more intricate narratives featuring multiple character-roles. Although our supervised machine learning method may sacrifice some precision in capturing the exact causal links and direction of the narrative conveyed in the text, it enables the analysis of large amounts of data and allows for a nuanced depiction of the discussion.

7.2 Narrative Virality

7.2.1 Virality of Simple and Complex Character-Role Narratives

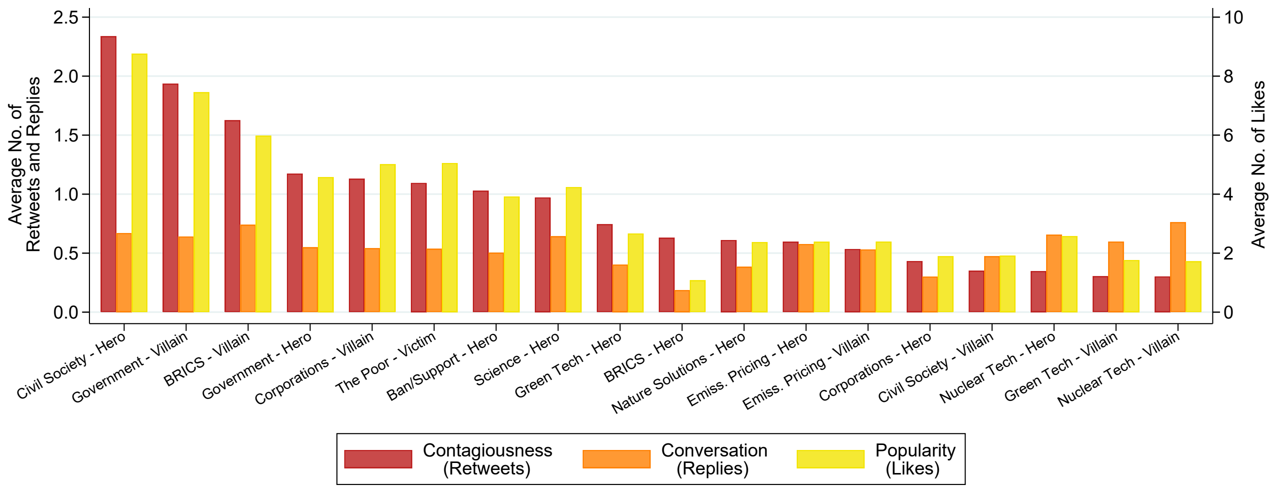
So far we have focused solely on the supply and composition of narratives by Twitter users. Nevertheless, it is important to note that, in contrast to more traditional media, social media is characterized by its social nature. Scholars such as Bénabou *et al.* (2020) and Shiller (2017) have emphasized that one essential aspect of narratives is their contagiousness. Our approach and use of Twitter data allow us to study this contagiousness by analyzing the retweet ratios of individual tweets featuring one or multiple character-role combinations.

Another important feature of social media is the ability to directly interact with other users and their content. By examining the number of replies that a tweet receives, we can determine which types of character-role narratives are more likely to spark conversations. Additionally, we investigate the popularity of narratives by analyzing the number of likes a tweet receives. Given the investment of time and effort required to create content on social media, users seek rewards for their contributions. The number of likes a piece of content receives is a useful indicator of its popularity, and serves as a measure of the return on a user's investment in Twitter activity.

Figure 6 displays the average virality of climate policy related tweets containing a specific character-role, according to the three dimensions of virality we outline above: contagiousness measured in retweets, conversation measured in replies and popularity measured in likes. We find that distinguishing between these dimensions adds important nuance to the investigation into narratives. This is exemplified by the remarkable differences in ranking character-roles according to specific virality measure. Contagiousness and popularity are more closely related, while the character-roles that spark conversation differ significantly from the other two dimensions.

The three most contagious (retweets) and popular (likes) character-roles are Civil Society-Hero, Government-Villain, and BRICS-Villain. While Government-Villain is also a very frequent character-role, the other two are less often used in tweets (see Figure 4). Government-Hero, The Poor-Victim, Ban/Support-Hero and Science-Hero share a similar degree of contagiousness and popularity. BRICS-Hero stands out for being quite contagious in that it generates a lot of retweets, but not very popular in terms of generating likes. Regarding their ability to generate conversation, or at least engagement in terms of replies, there are some remarkable differences. In particular, Nuclear Technology and Emission Pricing spark a lot of debate, relative to their low contagiousness and popularity.

Figure 6: *Virality by Character-Role*

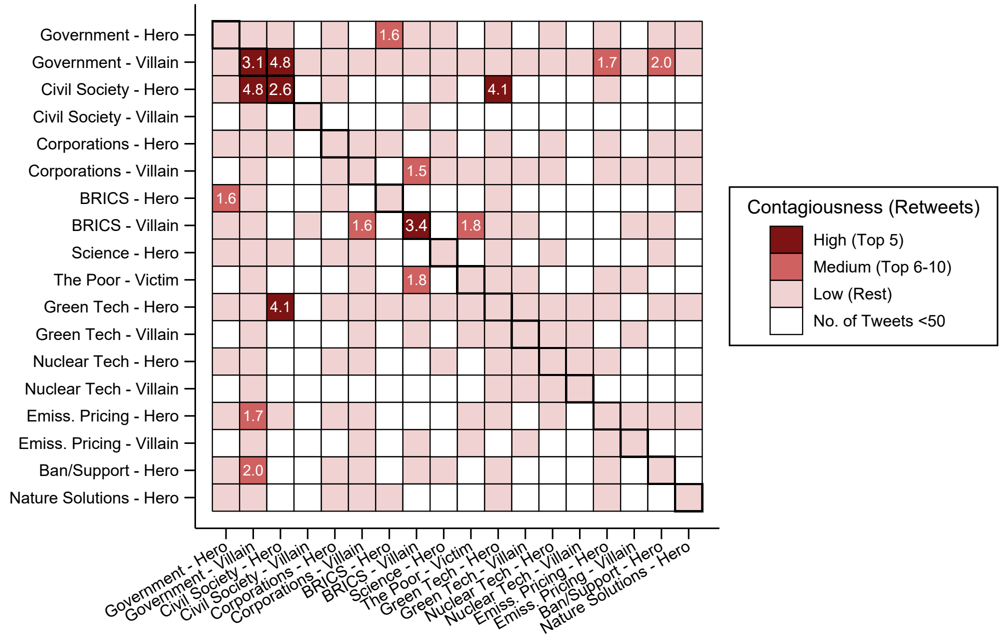


Notes: The figure illustrates the average rates of retweeting (contagiousness), replying (conversation), and liking (popularity) for different character-roles. We include only character-roles for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. Thes rates are calculated by dividing the total sum of retweets, replies, and likes of tweets that contain a specific character-role by the total count of those tweets. Retweets and replies are shown on the left y-axis, likes on the right. In Figure C.7 we show the distribution tweets over the virality indicators.

As before, we now fully exploit the possibilities of our framework, and analyze virality for all combinations of two or more character-roles. We do so for contagiousness in Figure 7, conversation in Figure 8, and popularity in Figure 9. Squares of these matrices report, respectively, the average number of retweets, replies and likes obtained by the specific character-role combination exemplified by the matrix entry. Simple narratives are shown on the diagonal, complex narrative combinations in the other entries of the matrices. The matrices are symmetrical but shown entirely for ease of reading.

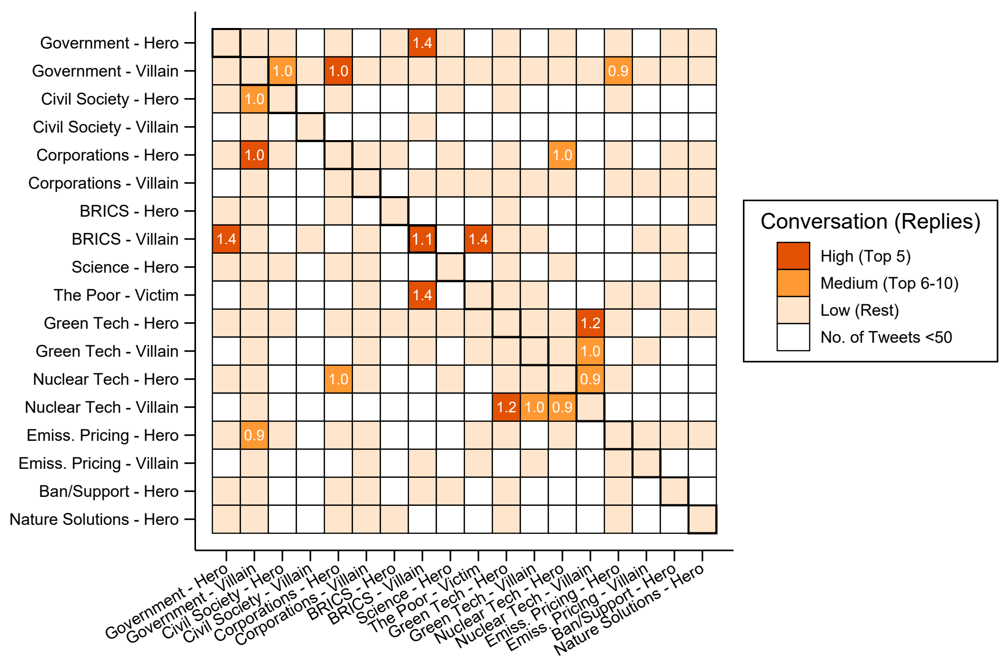
Some complex narratives are among the most viral, with similarities and differences across virality dimensions. Contagiousness and popularity show similar patterns with complex narratives combining Government-Villain and Civil Society-Hero being the most viral. These tweets often depict the government as unable or unwilling to address the issue of climate change seriously. Civil society, often represented by grassroot movements and activists, is depicted as the hero using political activism to force the government to act. Narratives depicting Civil Society-Hero + Green Tech-Hero are also in the top 5 for both dimensions and they highlight the hero roles of bottom-up initiatives by local governments, NGOs or ordinary citizens and of renewable energy for a more sustainable, clean future.

Figure 7: *Virality: Contagiousness Measured in Retweets*



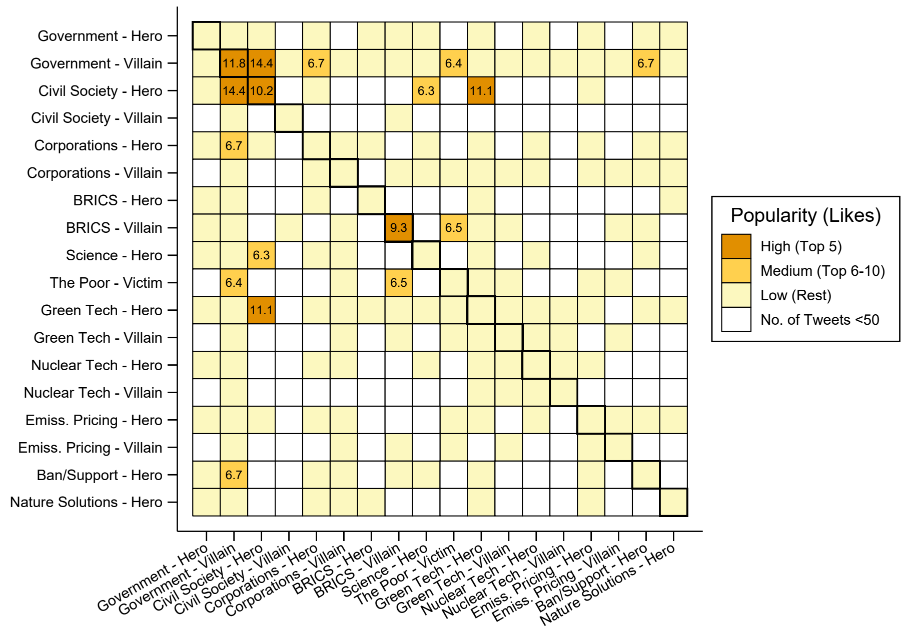
Notes: The figure shows the contagiousness of tweets measured as retweet rate. We include only character-roles for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. We include only character-role combinations that appear in 50 or more tweets in our analysis period. The matrix is symmetric but shown entirely for ease of reading. The diagonale displays the contagiousness of each character-role when appearing in a simple narrative. The intervals of the three levels indicated in the legend are the following: low, below 1.55 retweets; medium, from 1.55 to 2.05; high, above 2.05 retweets. Figure C.8 shows the same information considering a threshold of 25 tweets to include the character-role combinations in the matrix.

Figure 8: *Virality: Conversation Measured in Replies*



Notes: The figure shows the conversation sparked by tweets measured as reply rate. We include only character-roles for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. We include only character-role combinations that appear in 50 or more tweets in our analysis period. The matrix is symmetric but shown entirely for ease of reading. The diagonale displays the conversation sparked by each character-role when appearing in a simple narrative. The intervals of the three levels indicated in the legend are the following: low, below 0.85 replies; medium, from 0.855 to 1.04; high, above 1.02 replies. Figure C.9 shows the same information considering a threshold of 25 tweets to include the character-role combinations in the matrix.

Figure 9: *Virality: Popularity Measured in Likes*



Notes: The figure shows the popularity of tweets measured as like rate. We include only character-roles for which the model performs with an F1-score above .6 and we use geolocated tweets that are identified at least at the US-country level. We only include character-role combinations that appear in 50 or more tweets in our analysis period. The matrix is symmetric but shown entirely for ease of reading. The diagonal displays the popularity of each character-role when appearing in a simple narrative. The intervals of the three levels indicated in the legend are the following: low, below 6 likes; medium, from 6 to 8; high, above 8 likes. Figure C.10 shows the same information considering a threshold of 25 tweets to include the character-role combinations in the matrix.

Two more narratives in the top 5 for both contagiousness and popularity are Ban/Support-Hero + Government-Villain and BRICS-Villain + The Poor-Victim. The former are usually narratives about a government that is unwilling or unable to strictly regulate pollution, turning to inefficient simplistic measures such as bans on certain industries. Those narratives capture the image of crony capitalism and revolving doors between government regulators and the industry that should be broken up. BRICS-Villain + The Poor-Victim tweets transport a narrative of China and India being villains as the main emitters of CO2, while the average American is bearing most of the costs of climate change policies. Some of those narratives add the nuance that US climate change policy is a costly but pointless endeavor, because any reduction in the US would be trumped by increases in the BRICS states.

Narratives depicting BRICS-Villain + The Poor-Victim have also high potential of initiating conversations. Additionally, other narratives stand out, including Government-Hero + BRICS-Villain, Nuclear Tech-Villain + Green Tech-Hero and Corporations-Hero + Government-Villain. Government-Hero + BRICS-Villain narratives typically portray the US as the engaged hero in the fight against climate change, while the BRICS as the villains that keep on increasing their emissions. Nuclear Tech-Villain + Green Tech-Hero narratives often assert the dangers of nuclear technology, while green technology is the hero that will deliver increasingly cheap and clean energy. Corporations-Hero + Government-Villain narratives contrast an incapable and bureaucratic governments with dynamic, innovative corporations and the power of competition in free markets. The virality of this narrative reflects the reversal of hero and villain roles in the fight against climate change between government and corporations over time (see Figure 4).

7.2.2 Regression Analysis About Determinants of Virality

Beyond more descriptive results, we also use a regression framework to analyze the determinants of virality more generally. We use the following Poisson pseudo-maximum likelihood regression equation:

$$E[Y_{i,s,t}|D_{i,s,t}] = exp[\alpha + \beta D_{i,s,t} + \theta X_{i,s,t} + \gamma_t + \delta_s + \epsilon_{i,s,t}] \quad \forall i \in C \tag{1}$$

Where $Y_{i,s,t}$ refers to the count of one of our three measures of virality (retweets, replies or likes) for tweet i, originating from state s in year t. Each tweet i in the model is such that $i \in C$, where C is the sample of tweets located in the US with precision at least at the state level. α is a constant term, while γ_t and δ_s refer respectively to the year and state fixed effects. $D_{i,s,t}$ refers to a dummy variable reflecting the comparison of interest. E.g. when comparing villain and hero narratives it equals 1 if a villain narrative is present in the tweet and 0 for hero narratives. We exclude tweets that feature neither of the two, as well as all tweets that feature both simultaneously. $X_{i,s,t}$ collects the number of hashtags, mentions and words in the tweet as well as the number of followers, following and tweets ever produced by the user. We include standard errors clustered at the state level, represented by $\epsilon_{i,s,t}$ in the equation. Figure 10 shows the results as a coefficient plot.

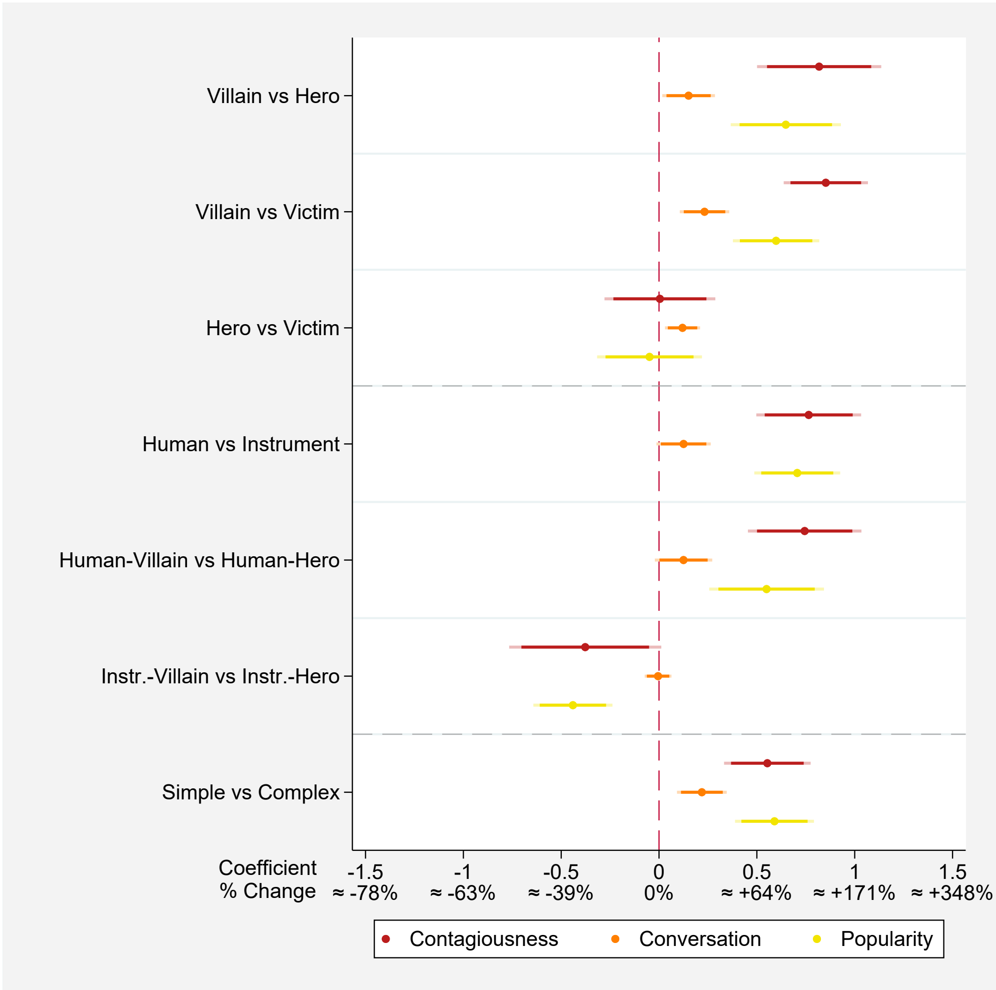
We contribute to the ongoing discussion on the determinants of virality in the media examining seven main comparisons of interest. First, we find that narratives containing a villain are generally more viral than those with a hero or those with a victim. Compared to hero (victim), villain narratives are more contagious by ≈127% (≈134%), spark more conversation by ≈16% (≈26%), and are ≈91% (≈82%) times more popular. Previous research shows that controversy sparks discussion on social media only when very moderate (Chen and Berger, 2013). In the specificity of our context, content depicting narratives about villains seems to have high virality return. When comparing hero and victim narratives we find no significant differences in terms of contagiousness and popularity, but an increase of ≈13% for conversations. In line with Al-Rawi (2019), it seems that users engage in this more costly interaction - relative to sharing or liking a tweet - when they want to discuss the heroes (or solutions) of climate change policy[5].

Second, when comparing human vs instrumental characters, the fourth comparison in Figure 10, we find that narratives containing human characters are more viral in all three dimensions. Interesting differences emerge when we check more specifically if the difference between the virality of hero and villain roles differs between human and instrumental characters, in the fifth and sixth comparison. Villain narratives tend to be more viral for human characters, while for instrumental characters it tends to be the opposite. This seems to be a novel insight, for the later narratives tend to be more successful in term of virality if they emphasize the positive hero-role of the instrument. The differences are strongest for popularity, and the effects for conversation are not statistically significant at the 95% level.

Finally, we test if on average, complex narratives with two or more character-roles are more viral

[5] All coefficient transformations are rounded to the closest unit

Figure 10: *The Determinants of Virality*



Notes: The figure shows the coefficients of Poisson pseudo-maximum likelihood regressions testing the determinants of the virality of narratives expressed as the count of retweets (Contagiousness), replies (Conversation) and likes (Popularity). We include only character-roles for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. The dependent variable is a tweet-level count of retweets/replies/likes, the independent variable a dummy that equals to 1 if a particular narrative is present in the tweet. If the model tests e.g., the contagiousness of villain vs. hero, it includes a dummy that takes the value 1 when the tweet contains a villain narrative, 0 when it contains a hero. The dummy is null when the tweet contains neither one or both simultaneously. The x-axis reports labels for the coefficient values and the correspondent percentage change computed as follows: $\approx e^{\beta} - 1$. All models include state and week-of-the-year FE, a linear weekly time trend and controls for the number of hashtags, mentions and words in the tweet as well as the number of followers, following and tweets ever produced by a user. Standard errors are clustered at the state level. Table D.1 reports the coefficient values.

than simple narratives with only one. While the prior evidence in Figure 7 to Figure 9 shows that the top 5 most viral narratives are often complex combinations, on average simplicity wins over complexity. A simple narrative is 74% more contagious, sparks 24% more conversation and increases likes by 80%. Together with the prior figures this clearly highlights that certain complex narratives can be very powerful, but that on average complexity is a hurdle for virality[6].

[6] Figure E.1 shows tests of further determinants of virality. Figure E.2 shows the same tests performed with linear regressions. Figure E.3 shows that the results hold when dropping tweets originating by potential Twitter bots.

7.3 The Effect of Populism on Climate Policy Narratives

Studying the relationship between populism and narratives is relevant because the strategic use of narratives seems to be a key tool of populists. Previous research links populism in general to decreasing trust in institutions and science, and specifically to rising skepticism about climate change and related policies (Huber *et al.*, 2022). In the US, the climate change policy discourse was always polarized, but it seems overall plausible that Donald Trump and his particularly populist communication style had a significant impact. Climate change policies, together with immigration and China, were a key topic he used both on social media and more generally in his campaign.[7] Trump is often referred to as a prototypical example of a populist politician (Inglehart and Norris, 2016), and many other politicians like Bolsonaro, Le Pen, Salvini or Orban were eager to copy from his playbook. Hence, studying the effect of the Trump presidency on climate change policy narratives in the US is one case with many particularities, but can potentially yield interesting insights for many other settings.

To study the effect of Trump and populism on the frequency of character-role narratives, we begin by using the following regression:

$$y_{i,s,t} = \alpha + \beta_1 T^{Obama} + \beta_2 T^{Biden} + \tau + \gamma X_{i,s,t} + \omega_i + \delta_s + \sigma_w \ + \epsilon_{i,s,t} \tag{2}$$

Where $y_{i,s,t}$ is a dummy variables indicating whether the tweet from user i in state s in year t contains the narrative of interest, e.g. featuring a villain. T^{Obama} and T^{Biden} are dummy variables taking on the value 1 during the years of the respective administration, and zero otherwise. Hence, we compare both to Trump's term in office. For simplicity, we multiply the β's by -1, so that they each reflect the effect of Trump relative to the two other presidents.

The main challenge for identification is obviously that in such a comparison over time a large range of aspects differ between the terms in office, leading to many omitted and unobserved variables that could bias this effect. τ is a linear time trend, which captures longer term trends that correlate with a larger share of Trump years and could thus bias the effect. For instance, it is plausible that there are long-term trends in polarization that are also reflected in social media narratives.

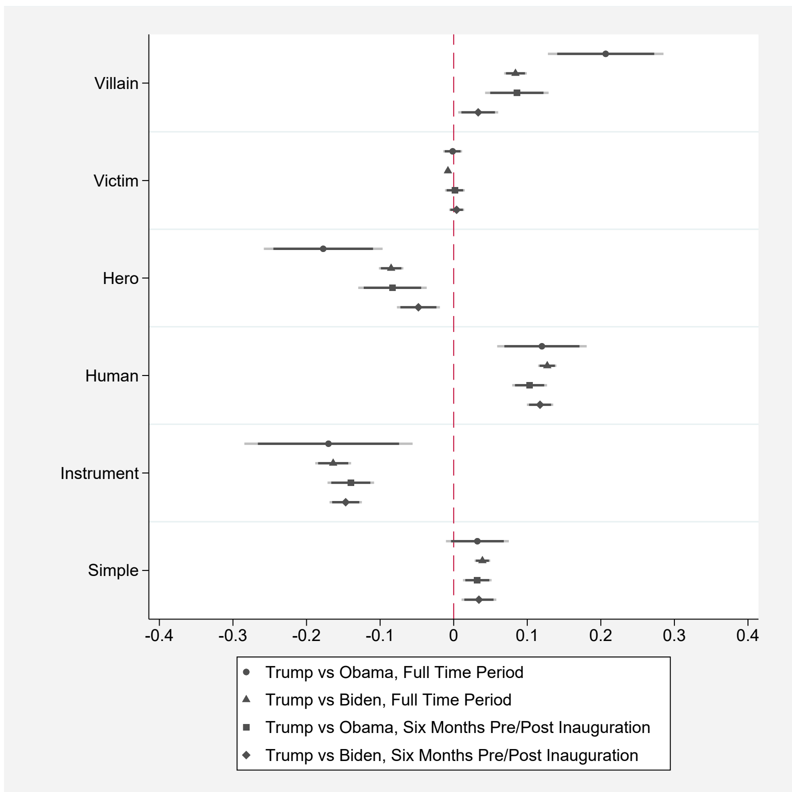
Changes in narratives during the Trump years might also correlate with general changes in the nature of tweets – like Twitter allowing longer Tweets since 2017 – which we can at least partly capture by controlling for the number of hashtags, mentions and words at the tweet level. Regarding the users, everything we can measure in this paper reflects narratives on Twitter for the respective user base and time. While we cannot ultimately determine if a possible Trump influence involves altering the composition of users in Twitter, we can adjust for potential effects on narrative composition and tweet frequency related to climate change policies. We do this at the user level by controlling for the number of followers, other users followed and tweets ever produced by a user $X_{i,s,t}$. Moreover, σ_w is a fixed effect for each calendar week to capture any seasonality, and δ_s are state fixed effects that account for any time-invariant omitted variables at the state level. Standard errors $\epsilon_{i,s,t}$ are clustered at the state-level.

Considering this specification, we thus ask: Is a particular character-role narrative more or less likely

[7] Exiting the Paris agreement was a key promise, which he fulfilled on November 4, 2020.

to appear during the Trump years relative to the years of his Democrat predecessor and successor? We present the coefficients - relative to Obama (symbol: ●) and to Biden (symbol: ▲) - in Figure 11. Some key differences stand out. First, during the Trump years, climate change policy narratives focus significantly more on villains than on heroes, both for human and instrument characters. Second, narratives feature significantly more human and fewer instrument-characters. Third, the share of simple narratives featuring just one character-role is higher during the Trump years. This last coefficient is borderline insignificant compared to Obama, and highly significant compared to Biden. Those results suggest that populism has poisoned the discussion about climate change policies in the US by pushing simple narratives that are more focused on human characters and on blaming a villain rather than highlighting solutions.[8]

Figure 11: *The Trump Effect*



Notes: The figure shows the impact of being under the Trump administration relative to the Obama and Biden administrations on the frequency of particular character-role narrative categories, e.g. villain narratives. The first two coefficients are from the same regression covering full period 2010-2021 and reflect a dummy variable comparing the Trump period with Obama (symbol ●) and with Biden (symbol ▲). The third and fourth coefficient are from separate regressions focusing on an event window 6 months before and after the transition between Trump and Obama (symbol ■) and between Trump and Biden (symbol ⧫). All regressions include state FE, a linear weekly time trend and controls for the number of hashtags, mentions and words in the tweet, as well as for the number of followers, number of other user followed, and overall number of tweets issued by a user. The full period regressions also controls for seasonality using week-of-the-year FE. In all models SE are clustered at the state level. Table D.2 reports the full specifications.

All those results are conditional on the controls and the linear time trend, but there are of course

[8] Figure E.4 shows the *Trump Effect* on different outcomes. Figures E.5 and E.6 show the effect on each character-role of our analysis. Figure E.7 shows the results using logistic regression. Table E.3 and Figure E.8 show that the results are not affected by respectively taking into account natural disasters and potential Twitter bots.

many remaining doubts about causality. A first reassuring observation is that we find almost the same results comparing Trump to the earlier Obama years and the later Biden years. Together with the time trend, this alleviates many of the concerns that we might just be picking up some long-term trends. However, within the limits of our data and set-up, we can go one step further and zoom in more on the change in office between the presidents. Figure 12 restricts the time window to the six months before and after the respective transition, and focuses on the ratios between villain vs. hero, human vs. instrument characters, and simple vs. complex narratives.

We use a time window of six months before and after each presidential transition because there is no clear theoretical prior if we should expect an immediate swift in narratives. The six months period is a compromise that allows incorporating enough week-level observations, while reducing the risk of capturing the impact of other unrelated events. Figure 12 visually confirms the results from the prior regressions over the full period. The figures indicate that there is a strong change in all ratios linked to the transitions. Looking at the horizontal lines depicting the average ratio over the six month periods, we observe both an increase from Obama to Trump as well as a decrease from Trump to Biden.

We then measure the effects in those more restricted time windows using the following specification:

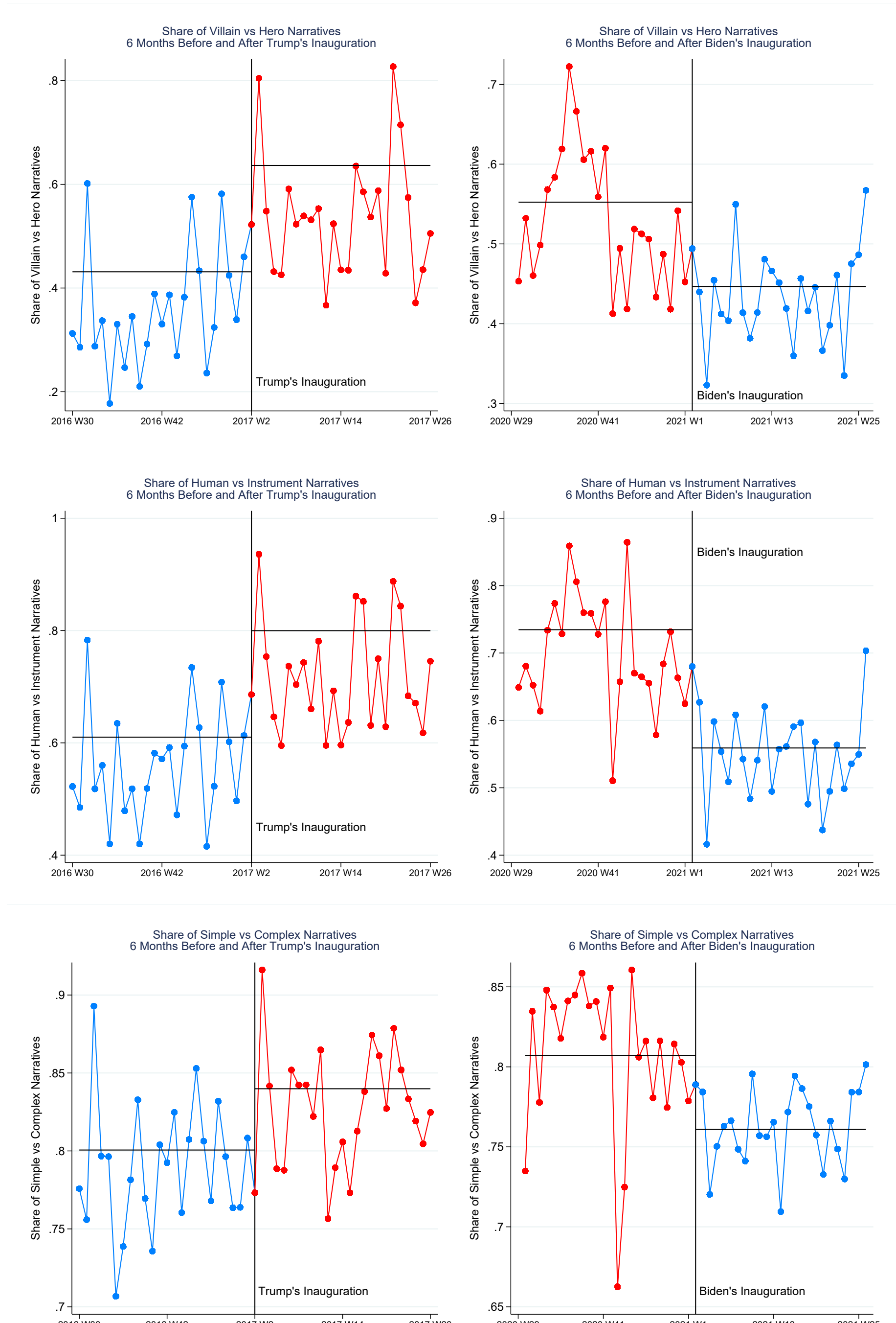
$$y_{i,s,t} = \alpha + \beta_{3,4} T^{Trump} + \tau + \gamma X_{i,s,t} + \delta_s \; + \epsilon_{i,s,t} \quad \forall t \in I^{Trump\,vs\,Obama/Trump\,vs\,Biden} \tag{3}$$

where I corresponds to the 6 months before and after the presidential transition. We use the same outcomes, and each regression measures the effect of a dummy variable T^{Trump} that takes on the value 1 during the Trump months, and 0 otherwise. We still control for state fixed effects δ_s and the linear time trend τ, the same tweet and user-level controls, and cluster standard errors $\epsilon_{i,s,t}$ at the state-level. Given the shorter period, the week-of-the-year FE is not included. We display the coefficients of the Trump vs. Obama regression (β_3, symbol: ■) and Trump vs. Biden regression (β_4, symbol: ♦) in Figure 11.

In line with the graphical impression from Figure 12, the results are similar in magnitude and significance to the full period specification for all outcomes. One difference is that the increase in simple narratives is now highly significant in both specifications. The idea that populism is linked to spreading more simple narratives is also reflected in Eliaz and Spiegler (2020). Overall, we find it very reassuring that the *Trump Effect* is so stable and can be detected in different specifications, and that we find it both compared to Obama and Biden. It is not obvious that switching back from a populist to a less populist president necessarily leads to such a reversal. Hence, the fact that we observe this reversal is both interesting and suggests that the impact of populism might, at least in some regards, be transitory and can be overcome.

Some caveats apply to this analysis. First, as we emphasize throughout the paper, our findings only reflect what is going on in Twitter, not necessarily in the population. Changes in the partisan composition of Twitter users could affect this result. However, we would expect that Trump coming into office led to more republican users, and him leaving – and being banned subsequently – to less republican users. The fact that we find the effects in both comparisons alleviates this concern to some extent. Moreover, instead of the actual transition in power, it could certainly be interesting for future research to explore more specific events like the publication of the election results or presidential debates, or the effect of

Figure 12: *The Trump Effect around Elections*



Notes: The figure provides insights into the appearance of a *Trump Effect* on narrative frequency after Trump's election, and disappearance after the election of Biden. Graphs on the left side display results for the six months around the election of Trump after Obama, while on the right we display results for the six months after the election of Biden. The graphs in the first row show the share of tweets containing a villain narrative out of all tweets containing a villain or a hero narrative, excluding those containing both or neither one. In the second row we show the same for human vs instrument narratives, in the third row for simple vs complex narratives.

specific tweets. Such an approach would require a deeper exploration of these events and daily data while in this paper we explore the overall effect of populism at a weekly level.

8 Concluding Remarks

The role of narratives in political and economic decision making processes is the subject of an increasing amount of economic research. However, the lack of a coherent definition of narratives and a structured way to measure them reliably in large datasets is a barrier to progress. Instead of developing yet another new approach, we analyze how prior approaches in economics (Ash *et al.*, 2023; Andre *et al.*, 2023; Eliaz and Spiegler, 2020), but also how disciplines like political science, sociology, communication studies and literature have been analyzing narratives. We draw on those insights and develop the **Character-Role Narrative Framework**, which defines narratives based on a topic in combination with the characters that are depicted in one of three archetypal roles: hero, villain or victim. It is a simple, yet powerful approach that can be applied in almost any setting to measure narratives.

We provide examples of applying the framework in various domains, and an empirical pipeline that allows researchers to easily adapt and implement the framework in their own work. By using the power of large language models combined with a limited training data set for fine-tuning the models, our pipeline is a very efficient approach to measuring narratives in large text data sets. We demonstrate the applicability of the framework by analyzing narratives around climate change policies using a large-scale Twitter dataset from U.S. users over the 2010-2021 period.

Our analysis reveals significant shifts in climate change narratives over time. Governments are increasingly portrayed as villains, corporations as heroes, and the hero role of renewable energies has seen a drastic decrease. Our framework provides valuable insights into the virality of different climate change narratives, as measured by retweets, replies, and likes. On average, narratives depicting villains, those involving human characters, and simpler narratives are more viral. These insights could inform the design of effective communication strategies around climate change policies. Interestingly, these viral narrative types became more prevalent during the populist Trump presidency, which appeared to shift the narrative focus from potential solutions to climate change towards the villain role of human characters. While we cannot perfectly establish causality, our various strategies suggest that this relationship could be causal.

Our work complements other research on the political economy of climate change policies that is based on survey data (Andre *et al.*, 2022; Dechezleprêtre *et al.*, 2022). Our data capture the discourse on social media, specifically Twitter, which may not be representative of the broader U.S. population's views on climate change. Additionally, endogenous changes in Twitter users over time cannot be perfectly accounted for. Nevertheless, with a large user base in the U.S., Twitter plays a significant role in shaping political and societal discourse.

In conclusion, understanding the nuances and dynamics of climate change narratives can significantly contribute to shaping effective climate policies. By offering a scalable and flexible tool to measure and analyze narratives, the **Character-Role Narrative Framework** opens up exciting possibilities for future research and policy-making. Our insights on climate change policy narratives and their virality provide the basis for further research into the complexity of reaching a political consensus and implementing such policies.

Bibliography

Acemoglu, D., Hassan, T.A. and Tahoun, A. (2018). 'The Power of the Street: Evidence from Egypt's Arab Spring', *The Review of Financial Studies*, vol. 31(1), pp. 1–42.

Akerlof, G.A. and Shiller, R.J. (2010). *Animal Spirits: How Human Psychology Drives the Economy, and why it Matters for Global Capitalism*, Princeton University Press.

Al-Rawi, A. (2019). 'Viral News on Social Media', *Digital Journalism*, vol. 7(1), pp. 63–79.

Andre, P., Boneva, T., Chopra, F. and Falk, A. (2022). 'Misperceived Social Norms and Willingness to Act Against Climate Change', ECONtribute Discussion Paper No. 10.

Andre, P., Haaland, I., Roth, C. and Wohlfart, J. (2023). 'Narratives About the Macroeconomy', *Review of Economic Studies, forthcoming.*

Anker, E. (2005). 'Villains, Victims and Heroes: Melodrama, Media, and September 11', *Journal of Communication*, vol. 55(1), pp. 22–37.

Asayama, S. and Ishii, A. (2017). 'Selling Stories of Techno-Optimism? The Role of Narratives on Discursive Construction of Carbon Capture and Storage in the Japanese Media', *Energy Research & Social Science*, vol. 31, pp. 50–59.

Ash, E., Gauthier, G. and Widmer, P. (2023). 'Text Semantics Capture Political and Economic Narratives', *Political Analysis.*

Barrera, O., Guriev, S., Henry, E. and Zhuravskaya, E. (2020). 'Facts, Alternative Facts, and Fact Checking in Times of Post-Truth Politics', *Journal of Public Economics*, vol. 182, p. 104123.

Baylis, P. (2020). 'Temperature and Temperament: Evidence from Twitter', *Journal of Public Economics*, vol. 184, p. 104161.

Beach, B. and Hanlon, W.W. (2023). 'Historical newspaper data: A researcher's guide and toolkit', *Explorations of Economic History, forthcoming.*

Bénabou, R., Falk, A. and Tirole, J. (2020). 'Narratives, Imperatives, and Moral Reasoning', NBER Working Paper No. w24798.

Benites-Lazaro, L.L., Giatti, L.L. and Giarolla, A. (2018). 'Topic Modeling Method for Analyzing Social Actor Discourses on Climate Change, Energy and Food Security', *Energy Research & Social Science*, vol. 45, pp. 318–330.

Berger, J. and Milkman, K.L. (2012). 'What Makes Online Content Viral?', *Journal of Marketing Research*, vol. 49(2), pp. 192–205.

Bergstrand, K. and Jasper, J.M. (2018). 'Villains, Victims, and Heroes in Character Theory and Affect Control Theory', *Social Psychology Quarterly*, vol. 81(3), pp. 228–247.

Bernauer, T. (2013). 'Climate Change Politics', *Annual review of political science*, vol. 16, pp. 421–448.

Bernauer, T. and McGrath, L.F. (2016). 'Simple Reframing Unlikely to Boost Public Support for Climate Policy', *Nature Climate Change*, vol. 6(7), pp. 680–683.

Brittain, M. (2006). 'Benevolent Invaders, Heroic Victims and Depraved Villains: White Femininity in Media Coverage of the Invasion of Iraq', *En) Gendering the War on Terror: War Stories and Camouflaged Politics*, pp. 73–96.

Bursztyn, L., Egorov, G., Haaland, I.K., Rao, A. and Roth, C. (2023a). 'Justifying Dissent', *Quarterly Journal of Economics, forthcoming.*

Bursztyn, L., Rao, A., Roth, C. and Yanagizawa-Drott, D. (2023b). 'Opinions as Facts', *The Review of Economic Studies, forthcoming.*

Cagé, J., Dagorret, A., Grosjean, P.A. and Jha, S. (2023). 'Heroes and Villains: The Effects of Heroism on Autocratic Values and Nazi Collaboration in France', *The American Economic Review, forthcoming.*

Cagé, J., Hervé, N. and Viaud, M.L. (2020). 'The Production of Information in an Online World', *The Review of Economic Studies*, vol. 87(5), pp. 2126–2164.

Campante, F., Durante, R. and Sobbrio, F. (2018). 'Politics 2.0: The Multifaceted Effect of Broadband

Internet on Political Participation', *Journal of the European Economic Association*, vol. 16(4), pp. 1094–1136.

Campante, F.R., Depetris-Chauvin, E. and Durante, R. (2023). 'The Virus of Fear: The Political Impact of Ebola in the US', *American Economic Journal: Applied Economics, forthcoming.*

Chen, T. and Guestrin, C. (2016). 'XGBoost: A Scalable Tree Boosting System', in (*Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining),*p. 785–794, KDD '16, New York, NY, USA: Association for Computing Machinery.

Chen, Z. and Berger, J. (2013). 'When, Why, and How Controversy Causes Conversation', *Journal of Consumer Research*, vol. 40(3), pp. 580–593.

Chu, Z., Gianvecchio, S., Wang, H. and Jajodia, S. (2012). 'Detecting Automation of Twitter Accounts: Are You a Human, Bot, or Cyborg?', doi:10.1109/TDSC.2012.75.

Crow, D.A. and Lawlor, A. (2016). 'Media in the Policy Process: Using Framing and Narratives to Understand Policy Influences', *Review of Policy Research*, vol. 33(5), pp. 472–491.

Dechezleprêtre, A., Fabre, A., Kruse, T., Planterose, B., Chico, A.S. and Stantcheva, S. (2022). 'Fighting Climate Change: International Attitudes Toward Climate Policies', NBER Working Paper No. 30265.

Eliaz, K. and Spiegler, R. (2020). 'A Model of Competing Narratives', *American Economic Review*, vol. 110(12), pp. 3786–3816.

Enikolopov, R., Makarin, A. and Petrova, M. (2020). 'Social Media and Protest Participation: Evidence from Russia', *Econometrica*, vol. 88(4), pp. 1479–1514.

Esposito, E., Rotesi, T., Saia, A. and Thoenig, M. (2023). 'Reconciliation Narratives: The birth of a Nation after the US Civil War', *American Economic Review forthcoming.*

Feng, Z., Guo, D., Tang, D., Duan, N., Feng, X., Gong, M., Shou, L., Qin, B., Liu, T., Jiang, D. and Zhou, M. (2020). 'CodeBERT: A Pre-Trained Model for Programming and Natural Languages', arXiv Preprint arXiv:2002.08155v4.

Fisher, D.R., Waggle, J. and Leifeld, P. (2013). 'Where Does Political Polarization Come From? Locating Polarization Within the US Climate Change Debate', *American Behavioral Scientist*, vol. 57(1), pp. 70–92.

Fløttum, K. and Gjerstad, Ø. (2017). 'Narratives in Climate Change Discourse', *Wiley Interdisciplinary Reviews: Climate Change*, vol. 8(1), p. e429.

Fog, K., Budtz, C., Munch, P. and Blanchette, S. (2010). 'The Four Elements of Storytelling', in (*Storytelling),*pp. 31–46, Springer.

Ganzevoort, R.R. (2017). 'The Drama Triangle of Religion and Violence', in (*Religion and Violence),*pp. 17–29, Springer.

Gehring, K., Harm Adema, J.A. and Poutvaara, P. (2022). 'Immigrant Narratives', CESifo Working Paper No. 10026.

Gentzkow, M., Kelly, B. and Taddy, M. (2019). 'Text as Data', *Journal of Economic Literature*, vol. 57(3), pp. 535–74.

Gentzkow, M. and Shapiro, J.M. (2010). 'What Drives Media Slant? Evidence from US Daily Newspapers', *Econometrica*, vol. 78(1), pp. 35–71.

Gomez-Zara, D., Boon, M. and Birnbaum, L. (2018). 'Who is the Hero, the Villain, and the Victim? Detection of Roles in News Articles Using Natural Language Techniques', in (*23rd International Conference on Intelligent User Interfaces),*pp. 311–315.

Halberstam, Y. and Knight, B. (2016). 'Homophily, Group Size, and the Diffusion of Political Information in Social Networks: Evidence from Twitter', *Journal of Public Economics*, vol. 143, pp. 73–88.

Hansen, S., McMahon, M. and Prat, A. (2018). 'Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach', *The Quarterly Journal of Economics*, vol. 133(2), pp. 801–870.

Harari, Y.N. (2014). *Sapiens: A Brief History of Humankind*, Random House.

Hermwille, L. and Sanderink, L. (2019). 'Make Fossil Fuels Great Again? The Paris Agreement, Trump, and the US Fossil Fuel Industry', *Global Environmental Politics*, vol. 19(4), pp. 45–62.

Huber, R.A., Greussing, E. and Eberl, J.M. (2022). 'From Populism to Climate Scepticism: The Role of Institutional Trust and Attitudes Towards Science', *Environmental Politics*, vol. 31(7), pp. 1115–1138.

Huber, R.A., Wicki, M.L. and Bernauer, T. (2020). 'Public Support for Environmental Policy Depends on Beliefs Concerning Effectiveness, Intrusiveness, and Fairness', *Environmental politics*, vol. 29(4), pp. 649–673.

Hyvärinen, M. (2010). 'Revisiting the Narrative Turns', *Life Writing*, vol. 7(1), pp. 69–82.

Inglehart, R.F. and Norris, P. (2016). 'Trump, Brexit, and the Rise of Populism: Economic Have-Nots and Cultural Backlash', HKS Working Paper No. RWP16-026.

Jiangli, S. (2020). 'European & American Think Tanks and the Reality of US-China Trade War: An NPF Application', *Journal of Social and Political Sciences*, vol. 3(2).

Johnson, T.F. and Greenwell, M.P. (2022). 'Are Companies Using Twitter to Greenwash and Hide Bad Environmental Performance?', *Energy, Ecology and Environment*, pp. 1–14.

Jones, M.D. and McBeth, M.K. (2010). 'A Narrative Policy Framework: Clear Enough to Be Wrong?', *Policy Studies Journal*, vol. 38(2), pp. 329–353.

Joshi, M., Chen, D., Liu, Y., Weld, D., Zettlemoyer, L. and Levy, O. (2020). 'SpanBERT: Improving Pre-training by Representing and Predicting Spans', vol. 8, pp. 64–77.

Karpman, S. (1968). 'Fairy Tales and Script Drama Analysis', *Transactional Analysis Bulletin*, vol. 7(26), pp. 39–43.

Keohane, R.O. (2015). 'The Global Politics of Climate Change: Challenge for Political Science', *PS: Political Science & Politics*, vol. 48(1), pp. 19–26.

Kirilenko, A.P. and Stepchenkova, S.O. (2014). 'Public Microblogging on Climate Change: One Year of Twitter Worldwide', *Global Environmental Change*, vol. 26, pp. 171–182.

Kousser, T. and Tranter, B. (2018). 'The Influence of Political Leaders on Climate Change Attitudes', *Global Environmental Change*, vol. 50, pp. 100–109.

Kruse, T.L. and Kruse, J.L. (1994). 'Suzuki and the Drama Triangle', *Transactional Analysis Journal*, vol. 24(3), pp. 211–215.

Levi, S. (2021). 'Why Hate Carbon Taxes? Machine Learning Evidence on the Roles of Personal Responsibility, Trust, Revenue Recycling, and Other Factors Across 23 European Countries', *Energy Research & Social Science*, vol. 73, p. 101883.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. and Stoyanov, V. (2019). 'RoBERTa: A Robustly Optimized BERT Pretraining Approach', arXiv Preprint arXiv:1907.11692v1.

Macaulay, A. and Song, W. (2022). 'Narrative-Driven Fluctuations in Sentiment: Evidence Linking Traditional and Social Media', MPRA Paper No. 113620.

McBeth, M.K. and Pearsall, C.A. (2021). 'Using the Narrative Policy Framework in the Teaching of Introduction to Politics', *Journal of Political Science Education*, vol. 17(sup1), pp. 880–896.

McMahon, D. (2005). 'The Drama Triangle', *The Skilled Facilitator Fieldbook*, p. 421.

Merry, M.K. (2016). 'Constructing Policy Narratives in 140 Characters or Less: The Case of Gun Policy Organizations', *Policy Studies Journal*, vol. 44(4), pp. 373–395.

Michalopoulos, S. and Xue, M.M. (2021). 'Folklore', *The Quarterly Journal of Economics*, vol. 136(4), pp. 1993–2046.

Müller, K. and Schwarz, C. (2023). 'From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment', *American Economic Journal: Applied Economics, forthcoming.*

Nguyen, D.Q., Vu, T. and Nguyen, A.T. (2020). 'BERTweet: A Pre-trained Language Model for English Tweets', arXiv Preprint arXiv:2005.10200v2.

Nordensvard, J. and Ketola, M. (2021). 'Populism as an Act of Storytelling: Analyzing the Climate Change Narratives of Donald Trump and Greta Thunberg as Populist Truth-Tellers', *Environmental Politics*, pp. 1–22.

Oehl, B., Schaffer, L.M. and Bernauer, T. (2017). 'How to Measure Public Demand for Policies When

There is no Appropriate Survey Data?', *Journal of Public Policy*, vol. 37(2), pp. 173–204.

O'Neill, L., Anantharama, N., Buntine, W. and Angus, S.D. (2021). 'Quantitative Discourse Analysis at Scale-AI, NLP and the Transformer Revolution', SoDa Laboratories Working Paper Series No. 2021-12.

O'Brien, E. (2018). *Challenging the Human Trafficking Narrative: Victims, Villains and Heroes*, Routledge.

Pearce, W., Holmberg, K., Hellsten, I. and Nerlich, B. (2014). 'Climate Change on Twitter: Topics, Communities and Conversations about the 2013 IPCC Working Group 1 Report', *PloS One*, vol. 9(4), p. e94785.

Polletta, F., Chen, P.C.B., Gardner, B.G. and Motes, A. (2011). 'The Sociology of Storytelling', *Annual Review of Sociology*, vol. 37(1), pp. 109–130.

Puckett, K. (2016). *Narrative Theory*, Cambridge University Press.

Salminen, J., Hopf, M., Chowdhury, S.A., Jung, S.g., Almerekhi, H. and Jansen, B.J. (2020). 'Developing an Online Hate Classifier for Multiple Social Media Platforms', *Human-centric Computing and Information Sciences*, vol. 10(1), pp. 1–34.

Shahaf, D., Yang, J., Suen, C., Jacobs, J., Wang, H. and Leskovec, J. (2013). 'Information Cartography: Creating Zoomable, Large-Scale Maps of Information', in (*Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining),*pp. 1097–1105.

Shiller, R.J. (2017). 'Narrative Economics', *American Economic Review*, vol. 107(4), pp. 967–1004.

Shiller, R.J. (2020). *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*, Princeton University Press.

Shmelev, I. (2015). 'Beyond the Drama Triangle: The Overcoming Self', *Psychology. Journal of Higher School of Economics*, vol. 12(2), pp. 133–149.

Sohail, A., Du, J., Abbasi, B.N. and Ahmed, Z. (2022). 'The Nonlinearity and Nonlinear Convergence of CO2 Emissions: Evidence from Top 20 Highest Emitting Countries', *Environmental Science and Pollution Research*, pp. 1–17.

Tabassum, F., Mubarak, S., Liu, L. and Du, J.T. (2023). 'How Many Features Do We Need to Identify Bots on Twitter?', in (I. Sserwanga, A. Goulding, H. Moulaison-Sandy, J. T. Du, A. L. Soares, V. Hessami and R. D. Frank, eds.), *Information for a Better World: Normality, Virtuality, Physicality, Inclusivity*, pp. 312–327, Cham: Springer Nature Switzerland, ISBN 978-3-031-28035-1.

Terry, L.D. (1997). 'Public Administration and the Theater Metaphor: The Public Administrator as Villain, Hero, and Innocent Victim', *Public Administration Review*, pp. 53–61.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I. (2017). 'Attention is All You Need', *Advances in Neural Information Processing Systems*, vol. 30.

Verweij, M., Douglas, M., Ellis, R., Engel, C., Hendriks, F., Lohmann, S., Ney, S., Rayner, S. and Thompson, M. (2006). 'Clumsy Solutions for a Complex World: The Case of Climate Change', *Public Administration*, vol. 84(4), pp. 817–843.

Zhuravskaya, E., Petrova, M. and Enikolopov, R. (2020). 'Political Effects of the Internet and Social Media', *Annual Review of Economics*, vol. 12, pp. 415–438.

Analyzing Climate Change Policy Narratives with the Character-Role Narrative Framework

Kai Gehring* and Matteo Grigoletto**

Appendix

*University of Bern, Wyss Academy at the University of Bern, CESifo, e-mail: *mail@kai-gehring.net*
**University of Bern, Wyss Academy at the University of Bern, e-mail: *matteogrigoletto@gmail.com*

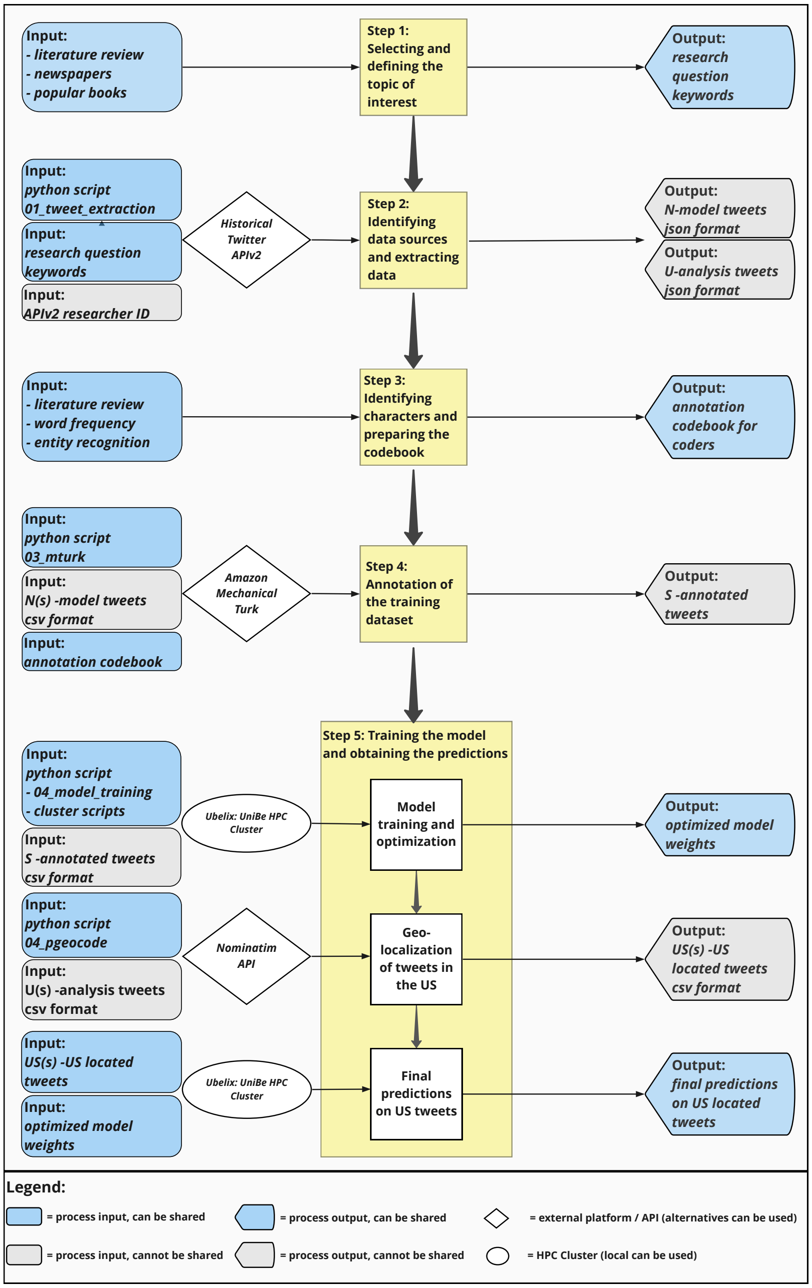
Appendix

Table of Contents

A Replication Codebook

This appendix provides additional details on our pipeline. The aim is to facilitate its replicability and assist researchers in using our methodology for similar projects. While some details may overlap with those mentioned in the paper, this appendix provides complementing information. Figure A.1 integrates the pipeline scheme shown in Figure 2 in the paper, providing more insights on the inputs and outputs of our steps.

Figure A.1: *Process Inputs and Outputs: Schematic Representation*



Notes: The figure presents a schematic visualization of the inputs and outputs of our pipeline. It provides additional information complementing what is shown in Figure 2. We indicate inputs and outputs that we can share to replicate the analysis in azure, these can be adapted to perform different analyses using our pipeline of work. We indicate inputs and outputs that we cannot share for proprietary reasons in light gray.

A.1 Topic Selection

A narrative requires a topic. The first step of our process consists in selecting a well-defined topic. This enhances model training and allows for a clear interpretation of character-roles. Poorly defined topics can make it challenging to interpret character-roles accurately. For instance, portraying civil society as hero, can have completely different meanings in the topics of war and climate change activism.

How to find and define the topic? This depends on the specific research question, but we have two main recommendations. First, in case you are starting from scratch and cannot already access a corpus of text data, begin with the relevant literature. Often, qualitative studies can provide important insights into the structure of narratives. Second, if data are already available, augment the literature review with exploratory analysis. Both manual inspection and exploratory NLP methods - such as word fish or any type of unstructured topic modeling (e.g. Latent Dirichlet Allocation) - can provide an important indication of what subset of the text is similar enough to be potentially summarized as one topic. A more structured tool like RELATIO (Ash *et al.*, 2023) provides information on the most common subject-action-object relation in the text, guiding the researcher in identifying the most relevant topics.

For our application, we first checked the vast literature on climate change discourse. It appears clear that there are two distinct discussions about the topic. One concerns the scientific evidence about the issue. The other embrace the scientific evidence and focuses on which policies should be adopted to address the problem. Our interest is to understand the political economy of climate change and hence we focused on the second discussion. Among the many papers, Oehl *et al.* (2017) provide important insights into the study of climate change policy and propose a new way to measure policy demand. We adopt the keywords proposed by the authors to access the Twitter API, slightly modifying their query to our needs, as explained in the next section.

A.2 Extracting Data

The next step after selecting a topic and defining its scope is deciding on a data source and time frame to start extracting data. It is important to consider the potential differences in narratives structure across sources, like books, newspapers and social media. The most common sources of text data in economics are digitized newspapers (Gehring *et al.*, 2022; Beach and Hanlon, 2023) and social media (Cagé *et al.*, 2020). However, a much larger body of text including transcribed TV, radio and YouTube broadcasts or open-ended survey questions provide potentially interesting material. Our framework and pipeline can be applied to any type of text, but given structural differences it seems advantageous for performance to annotate data and train models separately for most applications.

For our study, we focus on English-language tweets from the United States, posted on the social media platform Twitter over the 2010-2021 period. At the time we write, the historical Twitter APIv2 provides researchers with access to any tweet posted and not deleted since 2006. We focus on the US as a country in which Twitter plays a particularly prominent role for policy discussions, and select 2010 as a start year when Twitter had become a mainstream medium of exchange. The sample of US citizens active on Twitter is of course still not representative, but as a medium provides a particularly interesting chance to observe the creation and spread of narratives over time and across space.

Keywords and query

We indicate here the keywords and rules used for the query of Twitter historical APIv2. Consider the following conditions:

1. The tweet includes at least one of the following terms: 'climate change', 'global warming', 'renewable energy', 'energy policy', 'emission', 'certificate trading', 'green certificate', 'white certificate', 'combined heat', 'power solution', 'energy solution', 'CO2', 'energy efficiency', 'energy saving', 'solar power', 'solar energy', 'wind power', 'wind energy', '*renewable energies*', '*energy policies*', '*ipcc*', '*green growth*', '*green-growth*', '*green wash*', '*green-wash*', '*climate strike*', '*climate action*', '*strike 4 climate*', '*strike for climate*'.
2. The tweet includes at least one of the following terms: 'climate', 'global warming', 'greenhouse' AND at least one of the following terms: 'refining', 'feed-in', 'cogeneration', 'extraction', 'exploitation', 'geotherm', 'hydro', 'agriculture', 'waste management', 'forest', 'wood', '*problem*', '*issue*', '*effect*', '*gas*', '*degrowth*', '*de-growth*', '*fridaysforfuture*', '*fridays4future*', '*scientistsforfuture*', '*scientists4future*'.
3. The tweet includes at least one of the following terms: 'climatechange', 'globalwarming', 'renewableenergy', 'renewableenergies', 'energypolicy', 'energypolicies', 'greencertificate', 'whitecertificate', 'combinedheat', 'powersolution', 'energysolution', 'energyefficiency', 'energysaving', 'solarpower', 'solarenergy', 'windpower', 'windenergy', 'greengrowth', 'greenwash', 'climatestrike', 'climateaction', 'strike4climate', 'strikeforclimate.
4. The tweet includes term 'carbon' AND Tweet does NOT include any of the following terms: 'bicycle', 'bike', 'copy', 'fiber', 'rims', 'altered', 'fork', 'frame', 'dating', 'tacos'.

A tweet is part of our sample if any of the above conditions applies. In addition, a tweet is part of our sample if its text also satisfies all of the following:

a. The tweet does not contain an URL address.

b. The tweet's content is in English language.

c. The tweet is not a retweet.

The following are the changes adopted in deviation from keywords and rules proposed by Oehl *et al.* (2017), the paper of reference for us to define our query:

1. In Oehl *et al.* (2017) any keyword needs to appear in combination with at least one among: 'climate', 'global warming', 'greenhouse'. We do not adopt the condition as baseline but we use it for those words that refer to climate change in a more lose way (see condition No. 2 above).
2. We use some terms that are not present in Oehl *et al.* (2017): 'ipcc', 'climate change', 'energy policies', 'renewable energies', 'green growth', 'green-growth', 'green wash', 'green-wash', 'climate strike', 'climate action', 'strike 4 climate', 'strike for climate', 'problem', 'issue', 'effect', 'gas', 'degrowth', 'de-growth', 'fridaysforfutrue', 'fridays4future', 'scientistsforfuture', 'scientists4future' (see words in *italics* in conditions 1 and 2).
3. We use all multi-word expressions in condition 1 (e.g. 'energy policies') also as hashtags (see condition 3).

4. We use an exclusion restriction tailored towards tweets, because we realized there was a consistent pattern of false positive cases with the word 'carbon' (see condition 4).

Extraction dates

To build the model we use tweets extracted from a set of randomly selected days in the period 2010-2021. We use the calendar option of the online random number generator random.org to randomly select a day within each month of this time period. The extraction was done on 7^{th} and 8^{th} February 2022. The selected days are the following:

2010-01-28, 2010-02-14, 2010-03-13, 2010-04-30, 2010-05-25, 2010-06-30, 2010-07-07, 2010-08-04, 2010-09-14, 2010-10-02, 2010-11-06, 2010-12-14, 2011-01-14, 2011-02-09, 2011-03-01, 2011-04-10, 2011-05-13, 2011-06-19, 2011-07-22, 2011-08-15, 2011-09-08, 2011-10-23, 2011-11-21, 2011-12-17, 2012-01-31, 2012-02-27, 2012-03-26, 2012-04-04, 2012-05-26, 2012-06-18, 2012-07-10, 2012-08-18, 2012-09-20, 2012-10-22, 2012-11-01, 2012-12-03, 2013-01-15, 2013-02-12, 2013-03-27, 2013-04-25, 2013-05-05, 2013-06-18, 2013-07-19, 2013-08-08, 2013-09-25, 2013-10-11, 2013-11-06, 2013-12-01, 2014-01-24, 2014-02-13, 2014-03-04, 2014-04-30, 2014-05-16, 2014-06-23, 2014-07-12, 2014-08-21, 2014-09-26, 2014-10-24, 2014-11-05, 2014-12-06, 2015-01-26, 2015-02-21, 2015-03-20, 2015-04-24, 2015-05-06, 2015-06-09, 2015-07-23, 2015-08-20, 2015-09-15, 2015-10-15, 2015-11-11, 2015-12-21, 2016-01-11, 2016-02-05, 2016-03-22, 2016-04-02, 2016-05-01, 2016-06-19, 2016-07-01, 2016-08-31, 2016-09-09, 2016-10-13, 2016-11-14, 2016-12-22, 2017-01-01, 2017-02-12, 2017-03-25, 2017-04-04, 2017-05-07, 2017-06-05, 2017-07-11, 2017-08-27, 2017-09-14, 2017-10-21, 2017-11-09, 2017-12-21, 2018-01-09, 2018-02-09, 2018-03-30, 2018-04-06, 2018-05-08, 2018-06-05, 2018-07-06, 2018-08-14, 2018-09-16, 2018-10-22, 2018-11-12, 2018-12-15, 2019-01-04, 2019-02-14, 2019-03-15, 2019-04-19, 2019-05-17, 2019-06-21, 2019-07-22, 2019-08-30, 2019-09-19, 2019-10-01, 2019-11-01, 2019-12-01, 2020-01-12, 2020-02-12, 2020-03-22, 2020-04-16, 2020-05-08, 2020-06-22, 2020-07-17, 2020-08-17, 2020-09-26, 2020-10-08, 2020-11-07, 2020-12-18, 2021-01-23, 2021-02-25, 2021-03-20, 2021-04-05, 2021-05-23, 2021-06-12, 2021-07-11, 2021-08-30, 2021-09-25, 2021-10-10, 2021-11-11, 2021-12-30.

To perform the analysis and hence apply the prediction model, we use a larger sample of tweets extracted along the same period of analysis 2010-2021. We collect tweets from every Saturday of every week between January 2010 and December 2021. The extraction was done between 4^{th} and 7^{th} December 2022.

Data managing

We compute a number of steps to clean and organize data after the extraction. We describe these steps in the following points:

1. Despite setting API's filter, some non-English tweets were captured and we had to clean them using langdetect, a python port of the language-detection library in Java. At the time of writing, langdetect is also available as an extension in spaCy.

2. The text of a single tweet might satisfy more than one condition of our query, hence representing a potential duplicate in the extracted data. Each tweet is associated to a uniquely identifying ID that we use to drop potential duplicates.

3. Exclusively for the tweets used to train the model we also drop exact text duplicates. This is done to avoid exact copies in among the tweets used to create our annotated data.

4. Before manual labeling, we clean the tweets of emojis and any other unicode objects that have a UTF-8 code larger than three bits. This is done in order to fulfill the conditions Amazon sets for text representation on the Mechanical Turk platform. We also replace line-breaks in the text with simple spaces.

At the moment of the extraction the dataset of tweets used for the model training process comprised 1,097,424 tweets. After the cleaning and wrangling it consists of 1,070,702. At the moment of the extraction the dataset of tweets used to for the analysis comprised 3,359,627 tweets. After the cleaning and managing it consists of 3,279,730.

A.3 Identifying Relevant Characters

This step represents a key passage of our pipeline. The researcher should aim to obtain two main outcomes at the end of this step. The first is a well-defined list of the characters relevant for her research. The second is a clear and concise codebook that should guide the annotation process, independently from the process happening internally or externally.

There are many different methods the researcher can use to identify the characters of interest, and these methods can be coupled with each other. We propose a short list of recommended methods:

- *Popular sources*: Explore newspapers and popular sources - such as television entertainment programs and podcasts - dealing with the topic of your interest. These might help to understand where the public discussion stands.
- *Word cloud*: Explore your data with simple word cloud methods, it might help to pinpoint the most recurrent terms in your corpus.
- *Word fishing*: Explore your data by searching for words you think exemplify the characters of your interest. If e.g. you believe unions might be an important character, you might want to see how many times the term appears in your data.
- *Related literature*: Explore the related literature and see what are the characters identified as relevant.
- *Entity recognition*: Explore your data with entity recognition methods that allow you to find and extrapolate all 'entities that exist' (think of people, places, organizations and so on). Even better if already embedded in more structured tools like RELATIO (Ash *et al.*, 2023).

Once the characters are defined the researcher should develop a codebook using concise and straightforward language. We recommend structuring the coding instructions, and thus the annotation process, in a step-by-step manner. For each text snippet (tweet, article, paragraph, etc.) the coder should: 1) Evaluate the relevance of the text to the topic of interest, 2) determine the presence of characters, and 3) if characters are present, assess the role of the drama triangle they assume. Additionally, we advise to obtain feedback from external reviewers to iteratively improve the codebook's clarity.

For our work, we incorporate different methods. First, we draft an initial list of human-entities characters mainly building on the large strand of the literature dedicated to the discussion on climate change (Pearce *et al.*, 2014; Hermwille and Sanderink, 2019; Nordensvard and Ketola, 2021; Johnson and Greenwell, 2022). Second, we build on the literature qualitatively analyzing the debate about climate change policies and the relative narratives (Verweij *et al.*, 2006; Fløttum and Gjerstad, 2017; Asayama and Ishii, 2017; Benites-Lazaro *et al.*, 2018) to draft a first list of abstract-entities characters.

We update our initial draft mainly guided by the relevant literature with more data-driven exploratory methods. First, we analyze the most common entities in the set of our tweets. Second, we select a set of words and phrases indicative of a character - e.g. 'nuclear' for the character of nuclear technology - and assess their relative frequency in the corpus of our tweets. Finally, we outline our codebook following the suggestions we provide above. We provide the codebook used by the coders at the end of this appendix (see end of Appendix A).

A.4 Building the Training Dataset

The process of annotation of data can be done internally or externally. The former ensures a cheaper and generally quicker process but could be subject to bias introduced by the authors. The latter is more expensive and time demanding but more transparent. We recruit external coders via the Amazon Mechanical Turk (AMT), a service platform creating a link between a community of workers from all over the world and companies, institutions and private citizens that need employees for short tasks.

Besides the choice about external or internal annotation, researchers should also decide whether the annotation is done by one or multiple coders. The choice on the number of coders annotating data faces a trade-off between bias and performance. A training dataset coming from a single coder can be very consistent and enhance high performance. Nevertheless, the model would learn from a single coder maximizing the risk of bias. Multiple coders might provide a more noisy training set, leading to lower performance. Nevertheless, the model would be less biased towards a unique point of view. In the next paragraphs, we enumerate the main steps we followed to obtain our annotated tweets.

Selecting coders

Once the instructions are ready, the next step consists of selecting the best coders. To do so, we created a customized qualification test for MTurkers (the workers of the AMT platform) to assess their understanding of the instructions and their effort in our task. The selection process started on February 15^{th}, 2022. Each worker willing to participate in the test was asked to answer 15 selected questions. Those workers who obtained more than 85% correct answers (13/15) were then given access to a following task that consisted of annotating 20 tweets. Based on these twenty tweets, we picked the three best coders. We also obtained a small fraction of annotated data from two additional coders, but the number of tweets annotated by them is minimal.

Annotating tweets

There are a few relevant points to consider about the annotation process. First, if the coders are provided with the entirety of tweets they have to annotate all at once, there is a risk they might fall into incorrect patterns. Second, each batch of annotated tweets should be at least partly checked by one of the authors. Third, a subset of the annotated tweets should be exactly the same for all the coders to allow for the computation of one of the different possible coding agreement indexes post-annotation. With these considerations in mind, we provided batches of 100 tweets for each coder. These batches were created by randomly selecting 280 tweets from the data used for model training, and splitting them into three groups of ninety unique and ten common tweets. In section 6.2 of the paper, we explore the determinants of performance, among which the quality of annotation stands out clearly.

Oversampling under-covered character-roles:
Sometimes it might be necessary to over-sample a subset of character-roles. During the analysis of the MTurkers' annotation output, it became apparent that some of the character roles defined ex-ante were not as prominent as expected. This was particularly true for the following roles: BRICS-Hero and Villain, Civil Society Organizations and Movement-Hero and Villain, Scientific Community / Academia-Hero, The Poor / Working Class-Hero and Villain, and Nuclear Technology-Hero and Villain. To address this issue, we took additional sub-samples filtered using character-role-specific keywords and randomly selected sets of these sub-samples for MTurkers to label.

For BRICS, we filtered the full sample of extracted tweets (excluding already labeled tweets) with 'brazil|russia|india|china|chines|south africa'. Similarly, for Nuclear, we filtered the full sample using 'nuclear|fission|fusion|reactor'. For the other character-roles, we used a slightly more complicated approach. Out of all labeled data at that point (7500 tweets), we took the 100 most common words for each character. From these lists, we dropped words that appeared in more than just one list. Then, the ten most common words of the remaining lists of Civil Society Organizations and Movement, Scientific Community / Academia, and The Poor / Working Class were used to filter the tweets. This ten words were, respectively: 'greta OR climatestrike OR thunberg OR strike OR climateaction OR kids OR protest OR extinction OR school OR young', 'scientists OR scientific OR scientist OR facts OR nasa OR experts OR research OR record OR understand OR actually', and 'earth OR poor OR food OR lot OR population OR covid OR feel OR globalwarming OR living OR trying'.

Technical notes on AMT
The AMT platform only allows characters from the Basic Multilingual Plane (BMP).[9] Therefore, going from the tweets extracted via the API to the tweets that we have to systematically exclude any code larger than U+FFFF. Before submitting the tweets to the coders we also rectify the & symbol that the Twitter APIv2 substitutes with ampersands (&). Finally, we replace linebreaks by simple spaces.

A.5 Obtaining and Running the Prediction Model

Before providing information on the prediction model, it is important to notice on which set of data we perform our analysis. We compute predictions only for those tweets that are part of the analysis dataset, and that we could locate in the United States. Below we provide some indications on localization of tweets.

[9] See the second answer in this stackoverflow question.

Geo-localizing tweets

It is important to notice that tweets do not inherently come with localization information and this needs to be retrieved by the researchers, if possible. Many authors using tweets in their analysis developed their own methods to localize tweets (Kirilenko and Stepchenkova, 2014; Baylis, 2020). We build on previous work and structure our own method, that exploits different 'fields' of information provided by the APIv2.

There are two main sources of geographical information available through the Historical API. Among the available user fields, there is one called 'location'. This can be filled in two ways. One method is directly by the user who decides to indicate her location when creating the profile. Another is by the Twitter API algorithm itself, which detects the location if it has been mentioned in the text of the user's self-description. For example, if a user describes herself as 'I am a PhD student based in Zurich', the API would provide Zurich as the user's location. Additionally, among the available tweet fields, there is one called 'geo'. This indicates the location of the tweet if the tweet has been geo-tagged somewhere. In fact, the Twitter application allows users to tag a tweet with a specific location at the time of posting. This simply involves indicating a place to which the user wants to 'tag' the post.

We decide to prioritize the information about the user and hence to assign to each tweet the location of the user that posted it. This is because only a minority of tweets comes with 'geo-tag' information. This might raise the suspect that people tag their tweets only during particular events - such as holidays or work trips - which do not truly represent their environment/location. Only in cases where the user's location information is unavailable do we locate a tweet according to the geo-tag assigned to it, if any. Consequently, our localization pipeline consists of the following steps, which apply to the analysis dataset:

1. We collect all available locations relative to users' profiles.

2. We use the geopy implementation of Nominatim's API which exploits OpenStreetMap data. We query the API inputting the location in string format and obtain its geographical coordinates. We query the API for a total of 340,912 string locations.

3. Once each string location is associated to a set of coordinates we merge the locations back to the main sample of data. The merge is done with the formula 'many to one' so that users with the same location are associated to the same set of coordinates.

4. We repeat step 1, 2 and 3 for those tweets that could not be located by the description location of the users posting them but that present a geo-tag. We query the API for a total of 6,551 geo-tag locations.

5. For all tweets that could be located (either via description or via geo-tag location) we intersect their coordinates with a shapefile of the United States borders and keep only those tweets located within the country. We locate in US a total o 859,738 tweets.

Some important notes on the Nominatim API. First, the API algorithm returns the centroid coordinates of the location, hence when searching for e.g. 'Florida' it would return the coordinates of the centroid of the state of Florida. Second, when the string location is not clear the API returns 'NaN' output. Third, in most of the cases in which the location string is composed of two or more locations - e.g. 'Florida and NY' - the API returns either one of the two or 'NaN' output. This is generally hard to predict, but multiple locations are a minority of the cases in our data. Lastly, the API is not case-sensible hence e.g. 'New York City' and 'new york city' would provide the same coordinates.

Prediction model
Our pipeline harnesses the power of supervised machine learning. We argue this approach makes the most efficient and transparent use of the necessary human coder intervention in the context of narrative analysis. There are however many ways to implement a supervised pipeline. We suggest to combine so called Transformer models and machine learning methods. Moreover, we suggest to use a cluster computer if possible, since these models tend to be computationally demanding, although this is not strictly necessary. In the following paragraphs we outlines some more details about the components of our model.

Transformers are deep learning structures used in various contexts, including Natural Language Processing (NLP), image classification and design. Transformer models revolutionized natural language processing by allowing for the use of deep learning techniques to process language in a more efficient and effective way. Their main advantages include improved accuracy, scalability, and the ability to capture long-term dependencies in language. For those reasons, we strongly propose to include a Transformer model in the process.

We use the Transformer model called RoBERTa proposed by Liu *et al.* (2019), which builds on Google's model BERT. We follow the existing literature and use a learning rate of 5e-4 in pre-training Feng *et al.* (2020) and 1e-5 during fine-tuning Nguyen *et al.* (2020); Joshi *et al.* (2020). Following (Vaswani *et al.*, 2017), we then combine the transformer model predictions with XGBoost, a more traditional ML algorithm using random forest as a basic structure. We implement a Bayesian optimization for XGBoost that takes a number of initial, random combinations of hyper-parameters and then estimates the underlying form of the optimization problem (most commonly through Gaussian process regression). This forms the prior distribution, along which updating happens by continuous machine learning predictions at favorable priors until the algorithm returns the best results after a set number of trials. The hyper parameters of XGBoost that we optimize in this way are the following:

1. The depth of the trees of the algorithm;
2. The relative regularization used in updates (equivalent to a learning rate);
3. The absolute regularization term of weights (L1);
4. The relative regularization term of weights (L2);
5. The absolute regularization used in updates (which can help handling imbalanced data);
6. The size of the sub-sample used for each random tree run;
7. The minimum loss reduction an additional leaf of a tree has to provide to be included.

Prediction steps:

In light of what we said above we apply our prediction model to all tweets that we could locate in the United States with approximation at least at the country level. For any tweet of this sample, the prediction process works as follows:

1. We process the original text of the tweets. We decode emojis into describing words and substitute mentions with the correspondent user name if the mention was among the top 200 most frequent mentions.
2. We input the clean version of the tweet's text into the Transformer model RoBERTa. Within the model the text is vectorized and inputted into the network.
3. We retrieve the output of RoBERTa that consists in a distribution of probabilities indicating the chance that the tweet contains a specific narrative.
4. We augment the information retrieved in the previous step with some metadata concerning the tweet, such as the number of punctuation, the number of emojis, the number of likes etc.
5. The totality of this information (predictions from Transformer added of the metadata) is inputted in XGBoost.
6. Finally, we retrieve the output of XGBoost, consisting in an hot-encoding prediction (either 0 or 1) for each character-role, establishing whether the tweet contains each character-role of interest.

Coder Instructions

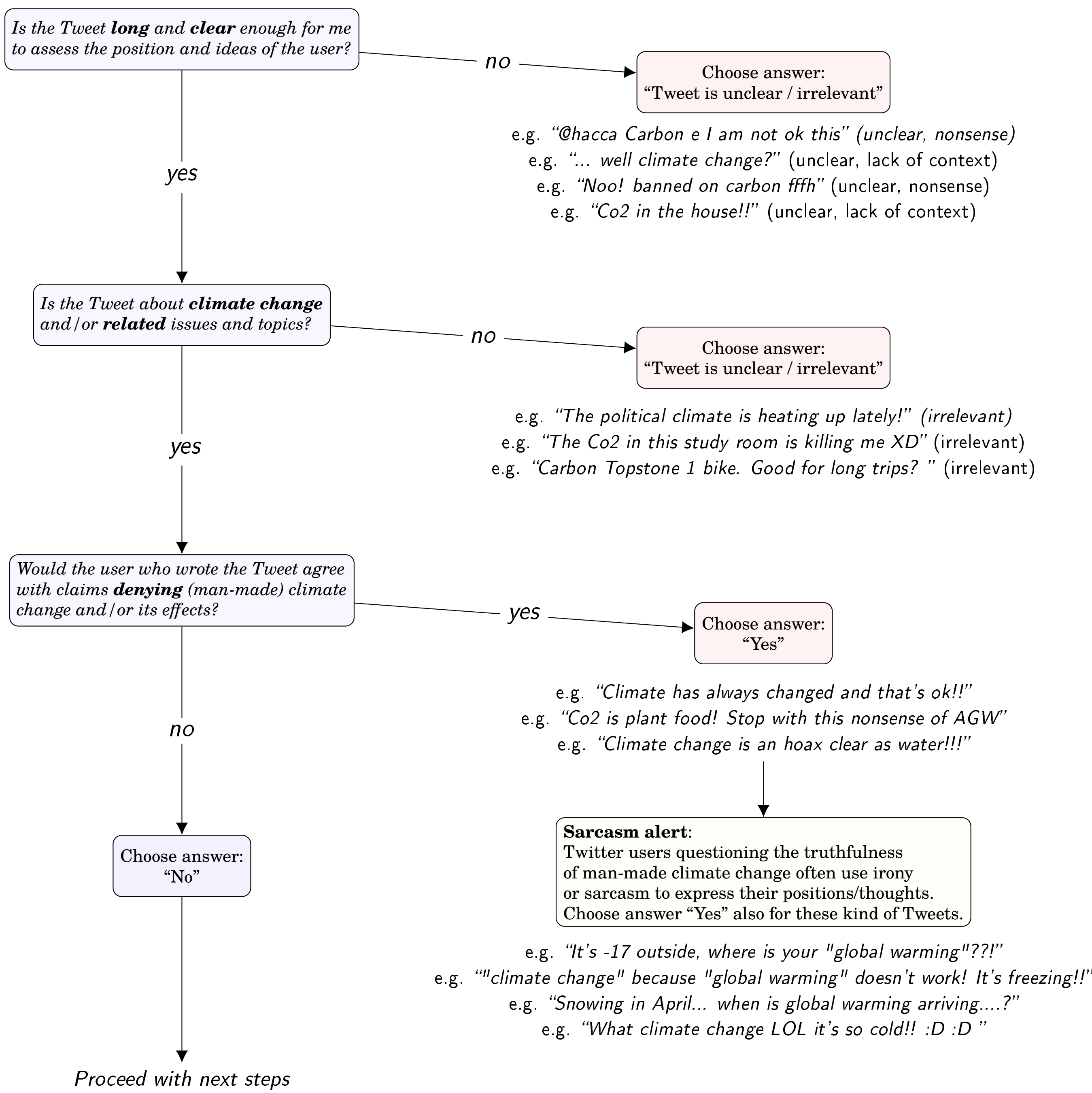
Thank you for reading these instructions carefully. The aim of our project is to extract climate change narratives that spread in the social media platform Twitter. To do so, we plan to compute a Machine Learning model to automatize the process. We will ask you to answer some questions about Tweets. These Tweets will constitute the training set for our model, hence your contribution is highly valuable. At the end of this document we provide some examples of previously coded Tweets (consult at your discretion).

Process and Instructions:

[Step 1] **Is the Tweet conveying that (man-made) Climate Change is not real?**

- ☐ **No**
- ☐ **Yes**
- ☐ **Tweet is unclear / irrelevant**

The aim of this first step is to discern Tweets that are useful to train our model from those that are not. To answer this question we suggest to proceed with the following reasoning:



For the next step, you will be asked the following question:
[Step 2] **Are there actors featured in the Tweet belonging to any of the following categories?**

Think of "actor" as a broad concept indicating a character, entity, institution or individual person that appears and might feature a certain role in the Tweet. We are interested in the following categories of actors:

☐ **BRICS** - Refers solely to Brazil, Russia, India, China and South Africa, the often called "emerging economies". The Tweet might refer to their governments, institutions, citizens or any entity identified with their nationality.

☐ **Western Governments / Political Elite** - Refers to rich, usually Western countries (e.g. members of OECD), the often called "Global North". In addition to geographically north/western countries this includes Japan, South Korea, Australia and New Zealand. The Tweet might refer to their governments, institutions, citizens or any entity identified with their nationality. Tweet might also refer to international institutions (World Bank, International Monetary Fund (IMF), United Nations (UN), etc.) and their members.

☐ **Civil Society Organizations and Movements** - Refers to non-governmental non-business organizations and movements or their members expressing the will, the interests and the position of civil society: e.g. NGOs, grassroots movements, etc.

☐ **Scientific Community / Academia** - Refers to scientists, experts and exponents of Academia. The Tweet might also refer to "the science" as general concept.

☐ **Corporations / Economic Elite** - Refers to corporations, businesses and their managers. Tweet might refer also to exponents of the economic and business elite.

☐ **The Poor / Working Class / Common People** - Refers to the poor and less privileged. Tweet might also refer to the working class and the common people as oppose to the political and economic elite.

To answer the question, we ask you to read carefully the Tweet and identify the actors. Then, assign them to the category that provides the best fit. If actors in the Tweet do not fit any category or simply there are no actors, tick *"None of above"*.
Please keep in mind:

- One phrase/expression can refer to more than one category of actors:
  e.g.: *"Chinese industries are the real problem for climate and nature!"* → from 'Chinese industries', 'Chinese' fits **BRICS**, while 'industries' belongs to **Corporations / Economic Elite**. Tick both boxes in such a case.
- Multiple actors from a single category can be featured. Just tick the box only once in such a case.

Pages 4 and 5 provide for more examples.

For the next step, we are interested in the role that actors play in the context of climate change and/or related issues and topics. For each category of actors you indicated as being present in the Tweet, you will be asked the following question:
[Step 3] **How is** *[category chosen]* **portrayed in the Tweet?**

☐ **Villain** - The actor or its actions/policies are portrayed negatively or as a potential obstacle/problem for fighting climate change.

☐ **Hero** - The actor or its actions/policies are portrayed positively or as a potential solution/help for fighting climate change.

☐ **Victim** - The actor is portrayed as someone/something suffering or bearing negative consequences of climate change or policies related to climate change.

☐ **None of above**

To answer this question, we ask you to think how is the actor presented by the author of the Tweet. Please, do not forget to keep an open mind. There might be actors generally considered as positive (e.g. scientists) portrayed negatively in a specific Tweet or vice-versa.

Please, keep in mind:

- If two or more actors from a single category are present and they are portrayed in opposing ways, choose the answer *"None of above"*:

e.g.: *"Democrats would ruin US with their Green New Deal! Trumps approach created jobs and literally saved lives!"* → actor 'Democrats' is portrayed negatively, actor 'Trump' positively. Tick actor **Western Governments / Political Elite**, portrayed as *"None of above"*.

- If two or more actors from a single category are present and one of them is portrayed positively, negatively or as victim, while the other(s) from that same category in a neutral way, choose the answer accordingly to how first actor is portrayed:

e.g.: *"Corporations have an important role in fighting climate change. But unsure if tech companies will do more good than bad."* → for **Corporations / Economic Elite**, tick *"Positively (Hero)"*.

Pages 4 and 5 provide for more examples.

At last, please focus on any solutions, actions or conclusions suggested in the Tweet. Each solution, action or conclusion can be framed and presented as positive, feasible and useful or on the contrary indicated as something negative, unfeasible or even useless.

[Step 4] **Is the Tweet conveying one of the following solutions / actions / conclusions?**

☐ **Emission pricing: Carbon Tax, Carbon Trade Schemes (CTS), etc.** *Positive / Feasible / Useful*

☐ **Emission pricing: Carbon Tax, Carbon Trade Schemes (CTS), etc.** *Negative / Unfeasible / Useless*

☐ **Change (often radical) in consumption, lifestyle, economic or political system** *Positive / Feasible / Useful*

☐ **Change (often radical) in consumption, lifestyle, economic or political system** *Negative / Unfeasible / Useless*

☐ **Banning or supporting specific products, industries or technologies** *Positive / Feasible / Useful*

☐ **Banning or supporting specific products, industries or technologies** *Negative / Unfeasible / Useless*

☐ **Nature-based solutions: reforesting, ocean farming, planting mangroves, etc.** *Positive / Feasible / Useful*

☐ **Nature-based solutions: reforesting, ocean farming, planting mangroves, etc.** *Negative / Unfeasible / Useless*

☐ **Green technology: solar/wind power, electric cars, carbon storage, etc.** *Positive / Feasible / Useful*

☐ **Green technology: solar/wind power, electric cars, carbon storage, etc.** *Negative / Unfeasible / Useless*

☐ **Nuclear technology: fission or fusion** *Positive / Feasible / Useful*

☐ **Nuclear technology: fission or fusion** *Negative / Unfeasible / Useless*

☐ **None of above**

To answer the question, first identify whether a specific solution / action / conclusion is present in the Tweet. Then, assess whether it is portrayed as positive/feasible/useful or on the contrary as negative/unfeasible/useless. Tick the answer accordingly to your assessment. You can choose more than one solution / action / conclusion, but you need to indicate uniquely whether each one is portrayed as positive or negative. If there is no solution / action / conclusion in the Tweet, tick *"None of above"*.

Pages 4 and 5 provide for more examples.

Note on emojis: Emojis represent a relevant tool to express emotions and thoughts in Twitter. Unfortunately, the mTurk Amazon platform is not able to process their encoding. To not lose this important tool we transformed each one of them when possible, in a set of words describing the emoji itself. Hence, where there was a smiling emoji you will read *[smiling face]*; where there was an angry emoji you will read *[angry face]*. These set of words are always put within [brackets] and in italics.

Examples:

Tweet: *"Climate change is real and manmade but we won't change anything unless we bully China and India into reducing their carbon emissions significantly. People in rich countries will probably be mostly fine but developing world is doomed in a long run."*

- Is the Tweet conveying that (man-made) Climate Change is not real? **No**
- Are there actors featured in the Tweet belonging to any of the following categories?

BRICS ('China' and 'India')
Western Governments / Political Elite ('rich countries')
The Poor / Working Class / Common People ('developing world')

- How is *BRICS* portrayed in the Tweet? **Negatively (Villain)**
- How is *Western Governments / Political Elite* portrayed in the Tweet? **None of above**
- How is *The Poor / Working Class / Common People* portrayed in the Tweet? **Victim**
- Is the Tweet conveying one of the following solutions / actions / conclusions? **None of above**

Tweet: *"@SocialMedia411 @shawncreed Pwned. My respect for Bezos just went up from "I know nothing about this guy" to "damn, he should be grouped with the likes of Musk". It is nice to see the rich contributing to the cause. The poor will be the ones affected by climate change the most."*

- Is the Tweet conveying that (man-made) Climate Change is not real? **No**
- Are there actors featured in the Tweet belonging to any of the following categories?

Corporations / Economic Elite ('Bezos' and 'Musk'),
The Poor / Working Class / Common People ('The poor')

- How is *Corporations / Economic Elite* portrayed in the Tweet? **Positively (Hero)**
- How is *The Poor / Working Class / Common People* portrayed in the Tweet? **Victim**
- Is the Tweet conveying one of the following solutions / actions / conclusions? **None of above**

Tweet: *"@h3h3productions Starting with the elites, billionaires and millionaires have to give up their wealth as the top 1% wealthiest account for 40% the CO2 emissions in the world. Degrowth now or die tomorrow, there is no other choice. Even the elites with their trillions won't survive climate change."*

- Is the Tweet conveying that (man-made) Climate Change is not real? **No**
- Are there actors featured in the Tweet belonging to any of the following categories?

Corporations / Economic Elite ('Elites', 'billionaires', 'millionaires')

- How is *Corporations / Economic Elite* portrayed in the Tweet? **Negatively (Villain)**
- Is the Tweet conveying one of the following solutions / actions / conclusions?

Change (often radical) in consumption, lifestyle, economic or political system *Positive / Feasible / Useful*

Tweet: *"The only effective response to fight climate change is to push hard on fusion research. The problem is of a scale where we need to invest around 8x the energy of our entire energy economy for the last 150 years to get things under control. Renewables won't even make a dent."*

- Is the Tweet conveying that (man-made) Climate Change is not real? **No**
- Are there actors featured in the Tweet belonging to any of the following categories? **None of above**
- Is the Tweet conveying one of the following solutions / actions / conclusions?

Green technology *Negative / Unfeasible / Useless*
Nuclear technology: fission or fusion *Positive / Feasible / Useful*

Tweet: *"You can also just say that climate change isn't a reason not to travel, since the incremental emissions of one flight are trivial and even if everyone on the planet stopped flying it wouldn't solve the problem."*

- Is the Tweet conveying that (man-made) Climate Change is not real? **No** → user does not deny existence, simply argues that stopping traveling wouldn't solve the problem.
- Are there actors featured in the Tweet belonging to any of the following categories? **None of above**
- Is the Tweet conveying one of the following solutions / actions / conclusions?

Change (often radical) in consumption, lifestyle, economic or political system *Negative / Unfeasible / Useless*

Tweet: *"#FridayThoughts: US cities are doing their part to battle #climatechange by creating ambitious climate goals, investing in light rail transportation, and banning single-use plastics. Hopefully our federal government catches up."*

- Is the Tweet conveying that (man-made) Climate Change is not real? **No**
- Are there actors featured in the Tweet belonging to any of the following categories?

Western Governments / Political Elite ('US cities' and 'federal government').

- How is *Western Governments / Political Elite* portrayed in the Tweet? **None of above** → actor 'US cities' is portrayed positively while 'federal government' negatively; being two opposed judgment it is safer to choose 'None of above'.
- Is the Tweet conveying one of the following solutions / actions / conclusions?

Banning or supporting specific products, companies or technologies *Positive / Feasible / Useful*

Tweet: *"I have no objection to a carbon tax. CT was proposed by free market economists bc it's a fair way to implement a financial incentive to reduce emissions."*

-Is the Tweet conveying that (man-made) Climate Change is not real? **No**
- Are there actors featured in the Tweet belonging to any of the following categories?

Scientific Community / Academia ('economists')

- How is *Scientific Community / Academia* portrayed in the Tweet? **Positively (Hero)**
- Is the Tweet conveying one of the following solutions / actions / conclusions?

Emission pricing: Carbon Tax, Carbon Trade Schemes (CTS), etc. *Positive / Feasible / Useful*

Tweet: *"@MikeFriesen10 Canadian government eliminating oil and gas subsidies would save 3.2 billion every year. Stop expanding the tar sands for a start. Invest in renewable energy and insulate existing buildings."*

-Is the Tweet conveying that (man-made) Climate Change is not real? **No**
- Are there actors featured in the Tweet belonging to any of the following categories?

Western Governments / Political Elite ('Canadian government')

- How is *Western Governments / Political Elite* portrayed in the Tweet? **None of above** → the user might think that Canada is not doing well but the judgment is expressed only implicitly.
- Is the Tweet conveying one of the following solutions / actions / conclusions?

Banning or supporting specific products, industries or technologies *Positive / Feasible / Useful*,
Green technology: solar/wind power, electric cars, carbon storage, etc. *Positive / Feasible / Useful*

Tweet: *"The reason I say "forestation" is because we need some way to remove CO2 from the atmosphere over the next few decades to really do anything."*

-Is the Tweet conveying that (man-made) Climate Change is not real? **No**
- Are there actors featured in the Tweet belonging to any of the following categories? **None of above**
- Is the Tweet conveying one of the following solutions / actions / conclusions?

Nature-based solutions: reforesting, ocean farming, planting mangroves, etc. *positive / Feasible / Useful*

B Requirements and Sources

Table B.1 enumerates all necessary packages for the scripts of our analysis used locally (not in the cluster computer), while table B.2 reports the sources of data used in the analysis.

Table B.1: *List of Scripts and Packages*

Script	Aim	Packages	Notes
Python			
01_tweet_extraction	Downloading tweets via API	os, re, requests, json, time, datetime	
02_data_prep	Data cleaning and managing	os, glob, pandas, re, emoji, langdetect, json, ast, (from collections) Counter, datetime, pickle, numpy	A dictionary of emojis is needed for the emoji package to work.
03_mturk	Preparation of input and output for and from Amazon Mechanical Turk platform	os, glob, pandas, numpy, (from sklearn.preprocessing) import LabelEncoder, (from sklearn.preprocessing) import OneHotEncoder, (from sklearn.feature_extraction.text) CountVectorizer, re, emoji, datetime, (from collections) Counter, pickle, matplotlib.pyplot	The script is used iteratively to create all the input batches for the annotation process.
04_model_training	Preparing input and output for model training	os, glob, pandas, numpy, (from sklearn.preprocessing) LabelEncoder, (from sklearn.metrics import f1_score) classification_report, re, emoji, datetime, matplotlib.pyplot, matplotlib.ticker, seaborn, string	
05_pgeocode	Geo-localizing tweets in the United States	re, glob, numpy, (from scipy.spatial.distance) [cdist, euclidean], pandas, (from geopy.geocoders) Nominatim, (from geopy.exc) GeocoderTimedOut, geonamescache, pycountry, geopandas, ast, (from shapely.geometry) Point, (from geopandas) GeoDataFrame, json, pickle, time	The Nominatim API should not be queried more than once per second.
06_predictions	Preparing input and output of prediction process	os, glob, pandas, numpy, (from sklearn.preprocessing) LabelEncoder, (from sklearn.metrics) [f1_score, classification_report], re, emoji, datetime, matplotlib.pyplot, matplotlib.ticker, seaborn, string	
07_maps_creation	Drawing map with geo-localized tweets	os, re, numpy, (from scipy.spatial.distance) [cdist, euclidean], pandas, (from geopy.geocoders) Nominatim, (from geopy.exc) GeocoderTimedOut, pycountry, geopandas, ast, (from shapely.geometry) Point, (from geopandas) GeoDataFrame, matplotlib.pyplot, (from mpl_toolkits.axes_grid1) make_axes_locatable, (from mpl_toolkits.axes_grid1.inset_locator) inset_axes, matplotlib.patches, scipy.stats, (from matplotlib) cm, (from matplotlib.colors) [ListedColormap, LinearSegmentedColormap], matplotlib.colors, matplotlib.lines	
Stata			
data_prep	Preparing raw data for analysis	egenmore, from(http://fmwww.bc.edu/RePEc/bocode/e)	
analysis	Creating output of analysis	egenmore, from(http://fmwww.bc.edu/RePEc/bocode/e) estout, from(http://fmwww.bc.edu/RePEc/bocode/e) coefplot, from(http://fmwww.bc.edu/RePEc/bocode/c) heatplot, from(http://fmwww.bc.edu/RePEc/bocode/h) palettes, from(http://fmwww.bc.edu/RePEc/bocode/p) colrspace, from(http://fmwww.bc.edu/RePEc/bocode/c) regdhfe, from(http://fmwww.bc.edu/RePEc/bocode/r) ppmlhdfe, from(http://fmwww.bc.edu/RePEc/bocode/p)	

Notes: The table reports all necessary packages for the Python scripts and Stata do-files used in the analysis. In section 4 we describe the process of data preparation.

Table B.2: *Data Sources*

Data	Source	Download Date	Availability
Twitter Data			
Model-tweets dataset	Twitter Historical APIv2	7^{th} - 8^{th} Jan. 2022	Cannot be shared
Analysis-tweets dataset	Twitter Historical APIv2	4^{th} - 7^{th} Dec. 2022	Cannot be shared
GIS Data			
USA Shapefile (v4.1)	GADM website	4^{th} Oct. 2022	Can be shared
Election Data			
Presidential election dataset	FiveThirtyEight repository	17^{th} Nov. 2022	Can be shared

Notes: The table reports a description of the sources of the data used in our analysis and their respective availability. In section 4 we describe the process of data preparation.

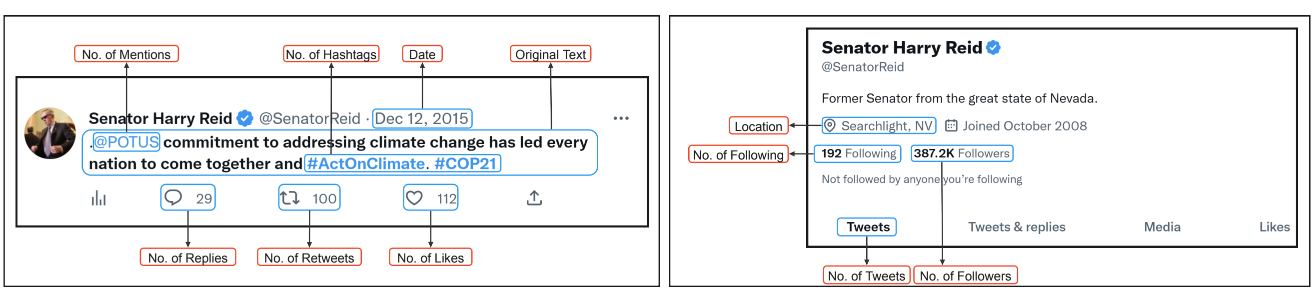
C Descriptives

This appendix collects additional descriptive information on the variables used in our analysis - in section C.1 - on the geo-location of the tweets used in the analysis - in section C.2 - and on the frequency and virality of narratives involved in our analysis, in section C.3.

C.1 Description of Variables

Figure C.1 provides a visual reference to identify the information retrieved via the Twitter APIv2 and used to construct the variables of our analysis. Framed in blue, all the content retrieved. Indicated in red, the variables using that specific content. Table C.1 lists all the variables used in our analysis. For each variable we provide a short description, we provide the values of its categories, scale or interval and we provide the data source from which it is created. Table C.2 reports descriptive statistics for the full sample of tweets extracted to perform the analysis, that is the same sample from which we identify those tweets originating from the United States.

Figure C.1: *Information Scraped via Twitter APIv2*



Notes: The figure shows two screenshots. On the left a tweet posted by the user *@SenatorReid*, on the right the same user's Twitter profile. We frame all the information retrieved via the Twitter APIv2 in blue and indicate the variable for which the information is used in red frames. In section 5 of the paper we describe our data.

Table C.1: *Description of Variables*

Variable	Question/Description	Categories/Scale/Interval	Source
Character-Roles			
Character-Role (predicted)	Variable detecting whether the tweet contains a character-role presenting a character in a specific role	0 = not present, 1 = present	Own computation
Character-Role (annotated)	Variable detecting whether the tweet contains a character-role presenting a character in a specific role	0 = not present, 1 = present	Annotation by human coders
Villain	Variable that indicates at least one villain narrative present in the tweet	0 = villain role not present, 1 = at least one character presented as villain	Own Computation
Hero	Variable that indicates at least one hero narrative present in the tweet	0 = hero role not present, 1 = at least one character presented as hero	Own Computation
Victim	Variable that indicates at least one victim narrative present in the tweet	0 = human role not present, 1 = at least one character presented as human	Own Computation
Human	Variable that indicates at least one human character presented in a role	0 = human character not present, 1 = at least one human character presented in a role	Own Computation
Instrument	Variable that indicates at least one instrument character presented in a role	0 = instrument character not present, 1 = at least one instrument character presented in a role	Own Computation
Simple	Variable that indicates if only one character-role is present	0 = no or more than one character-role present, 1 = one character-role present	Own Computation
Complex	Variable that indicates if more than one character-roles are present	0 = no or one character-role present, 1 = more than one character-roles present	Own Computation
Virality			
No. of Retweets	Number of times the tweet is retweeted	n $\in [0; 194{,}217]$	Twitter APIv2
No. of Replies	Number of times the tweet is replied to	n $\in [0; 30{,}887]$	Twitter APIv2
No. of Likes	Number of times the tweet is liked	n $\in [0; 896{,}759]$	Twitter APIv2
Control Variables			
No. of Words	Number of words in the tweet excluding mentions and hashtags	$\in [0; 117]$	Own computation
No. of Hashtags	Number of hashtags (#) in the tweet	n $\in [0; 71]$	Own computation
No. of Mentions	Number of mentions (@) in the tweet	n $\in [0; 52]$	Own computation
No. of Followers	How many twitter users follow the account of the user who generated the tweet?	n $\in [0; 133{,}245{,}480]$	Twitter APIv2
No. of Following	How many accounts does the user who generated the tweet follow?	n $\in [0; 4{,}066{,}970]$	Twitter APIv2
No. of Tweets	How many tweets the user who generated the tweet produced up until posting the tweet	n $\in [0; 9{,}611{,}963]$	Twitter APIv2
Any Disaster Average	National avg. of natural disasters in the year	n $\in [1.59; 6.06]$	Federal Emergency Management Agency
Other			
Date	Date of creation of the tweet	02.01.2010 - 25.12.2021	Twitter APIv2
Location	Highest level of precision at which the tweet could be located	country = country-level, state = US-state-level, city = city-level	Own computation
Obama's administration	Temporal variable indicating the period of the Obama administration	0 = years from 2017, 1 = years from 2010 to 2016	Own computation
Biden's administration	Temporal variable indicating the period of the Biden administration	0 = years before 2021, 1 = years from 2021	Own computation
Precision	Measurement of predictive performance of a model	n $\in [0.64; 0.92]$	Own computation
Recall	Measurement of predictive performance of a model	n $\in [0.43; 0.92]$	Own computation
F1-Score	Measurement of predictive performance of a model	n $\in [0.51; 0.92]$	Own computation

Notes: The table contains a description of all the variables used to generate the figures and tables in the paper and in the appendix. In section 5 of the paper we describe our data.

Table C.2: *Descriptive Statistics: Analysis Tweets*

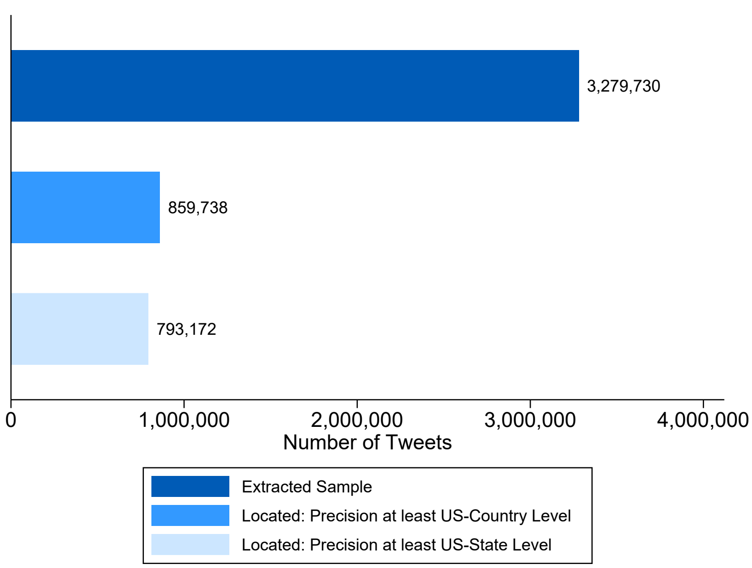
	Mean	Median	St. Dev.	Min.	Max.
Virality					
No. of Retweets (Contagiousness)	1.4	0	174	0	194,217
No. of Replies (Conversation)	.59	0	31	0	30,887
No. of Likes (Popularity)	6.4	0	795	0	896,759
Control Variables					
No. of Words	24	21	13	0	117
No. of Hashtags	.43	0	1.2	0	71
No. of Mentions	1.9	1	5.5	0	52
No. of Followers	9,120	520	374,934	0	133,245,480
No. of Following	1,978	617	8,973	0	4,066,970
No. of Tweets	51,730	15,758	128,700	0	9,611,963
No. of Observations	3,279,730				

Notes: The table reports descriptive statistics for the full sample of tweets extracted via the Twitter APIv2. For all the variables indicated we provide information on average value, median, standard deviation, minimum and maximum values across the sample. The number of words per tweet is calculated disregarding hashtags and mentions. Table 2 reports the same statistics for the subset of these tweets that originate from the United States.

C.2 Geo-localization of Tweets

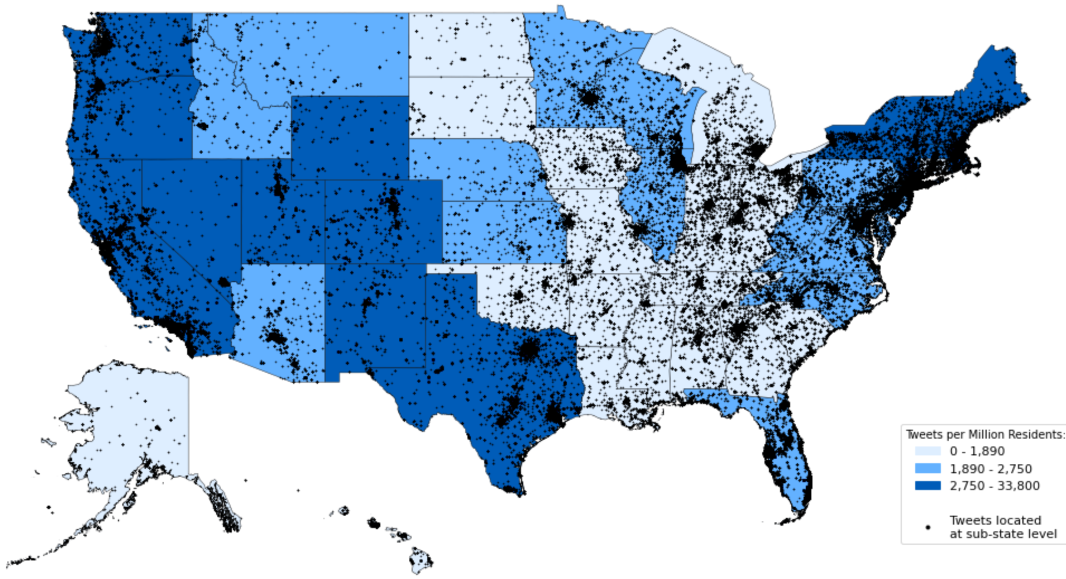
Figure C.2 provides information on the sample of tweets extracted to perform the analysis. In particular, it shows which subset of these tweets originate from the US and which of these tweets could be localize at least at the sub-national / state level. The map in Figure C.3 shows the geographical distribution of tweets in the United States. States are divided in three quantiles: low, medium and high frequency of tweets. Dots show the coordinates of tweets that could be located at the city level.

Figure C.2: *Sample Definition by Precision of Tweet Localization*



Notes: The figure provides information on the samples of tweets extracted via the Twitter APIv2. Of all tweets extracted (n = 3,279,730) we can geo-localize a fraction at least at the US country-level (n = 859,738). Of those tweets, we manage to geo-localize the majority at least at state-level (n = 793,172). Appendix section A.5 provides information on the method used to localize tweets. In section 5 we describe our data.

Figure C.3: *Cross-Sectional Distribution of Tweets from the US*



Notes: The map provides information on the cross-sectional distribution for the tweets that we geo-localize at least at the US state-level (n = 793,172 tweets). The colors of the states show three quantiles of the distribution of tweets. The dots indicate tweets that can be localized at the county or city level. Appendix section A.5 provides information on the method used to localize tweets. In section 5 we describe our data.

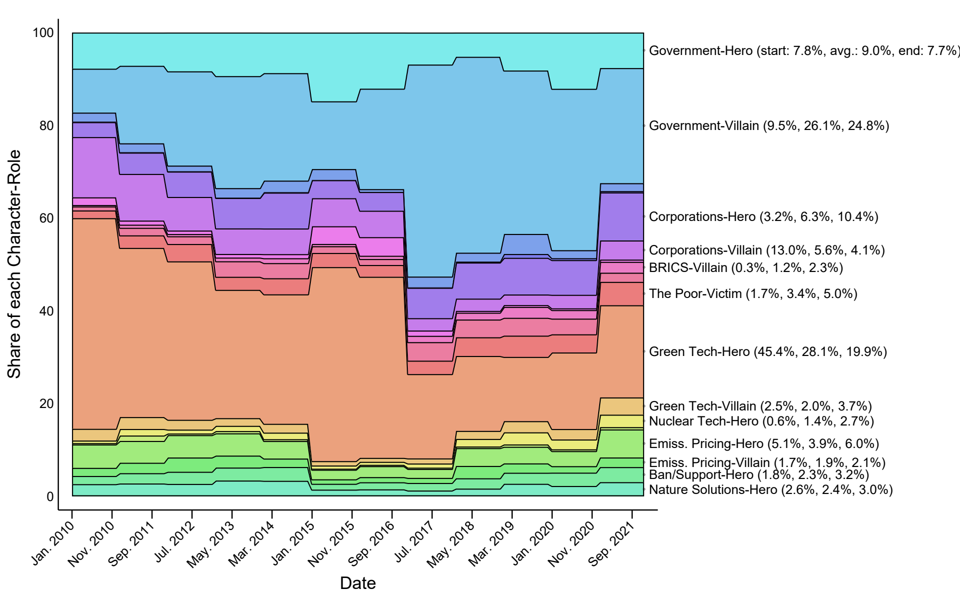
C.3 Narrative Frequency and Virality

Figure C.4 provides insights on the role played by retweets in spreading a particular narrative. Sub-Figure C.4a reports the same image shown in the paper Figure 4, for comparison to Sub-Figure C.4b, that shows narrative shares in time when counting retweets. The measure is obtained by counting the occurrence of each tweet and adding to it the number of retweets it obtained. When doing so, it is even clearer the relevance played by narratives depicting the government as the villain. In particular, in the last period of the analysis, after the year 2016, government-villain becomes consistently the most present character-role.

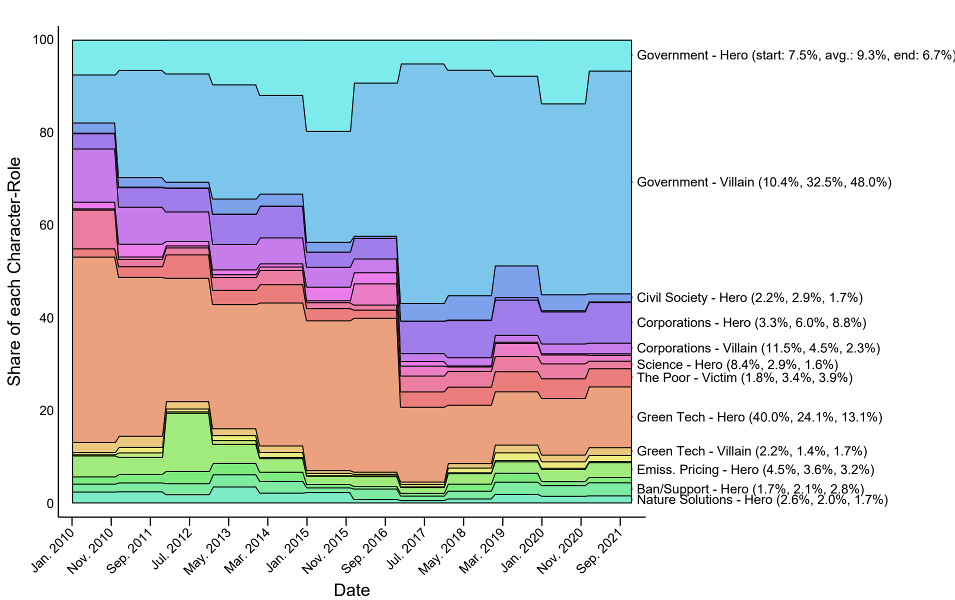
Figure C.5 shows the same information reported in Figure C.4, only using six months intervals instead of yearly intervals. The picture that emerges mirrors what seen above. Government and Green Tech are the most recurrent characters, with the former being consistently more present as villain, in the second part of our period.

Figure C.4: *Frequency of Character-Roles Over Time: Yearly Intervals*

(a) *Frequency Excluding Retweets*



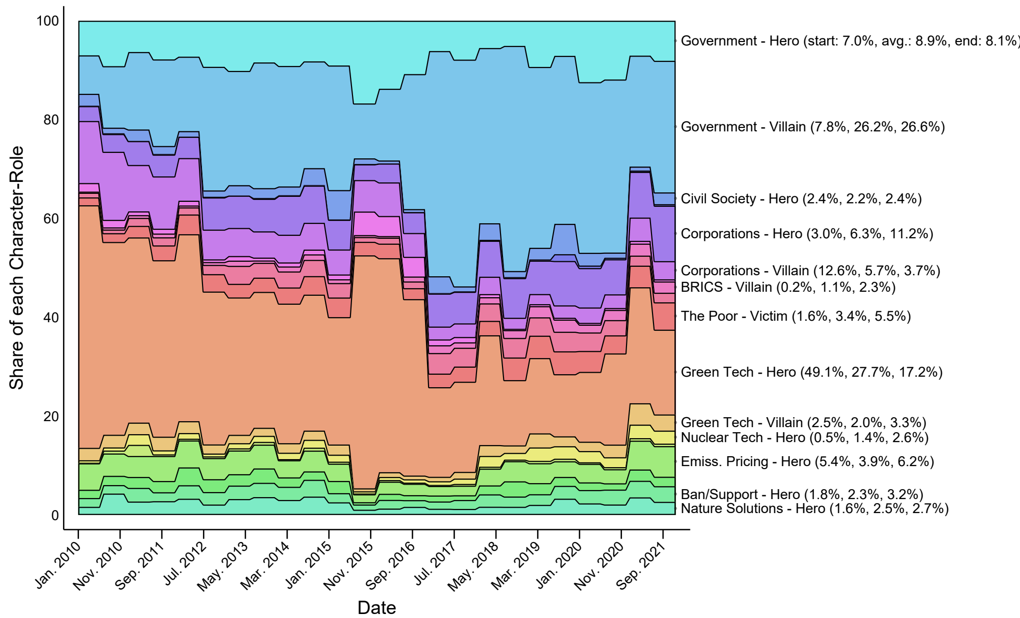
(b) *Frequency Including Retweets*



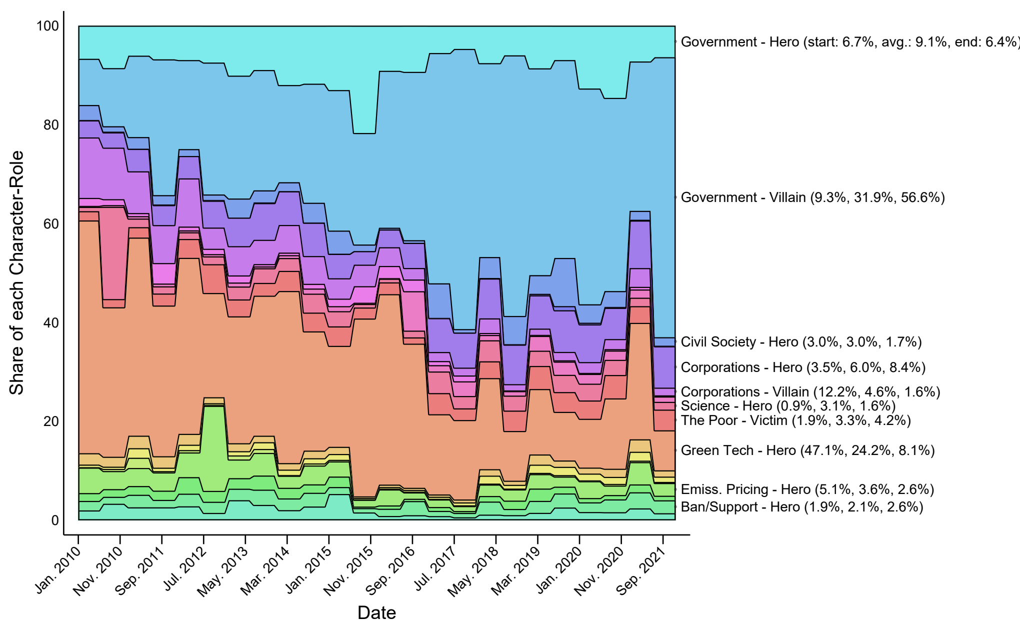
Notes: Panel (a) displays identical results to those presented in Figure 4. We include only character-role combinations for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. The shares of character-roles are cumulative (i.e. sum up to 100) and are computed on a yearly basis. We label the most common character-roles. In parentheses, each label indicates the share of the corresponding character-role in the first period, the average share across all periods, and the share in the last period. Unlabeled character-roles are the following: Civil-Society-Hero, Civil Society-Villain, BRICS-Hero, Science-Hero and Nuclear Tech-Villain. Panel (b) presents comparable information regarding the share of character-roles, while also taking into account the number of retweets each character-role received. Unlabeled character-roles are the following: Civil Society-Villain, BRICS-Hero, BRICS-Villain, Nuclear Tech-Hero, Nuclear Tech-Villain and Emiss. Pricing-Villain.

Figure C.5: *Frequency of Character-Roles Over Time: Six Months Intervals*

(a) *Frequency Excluding Retweets*



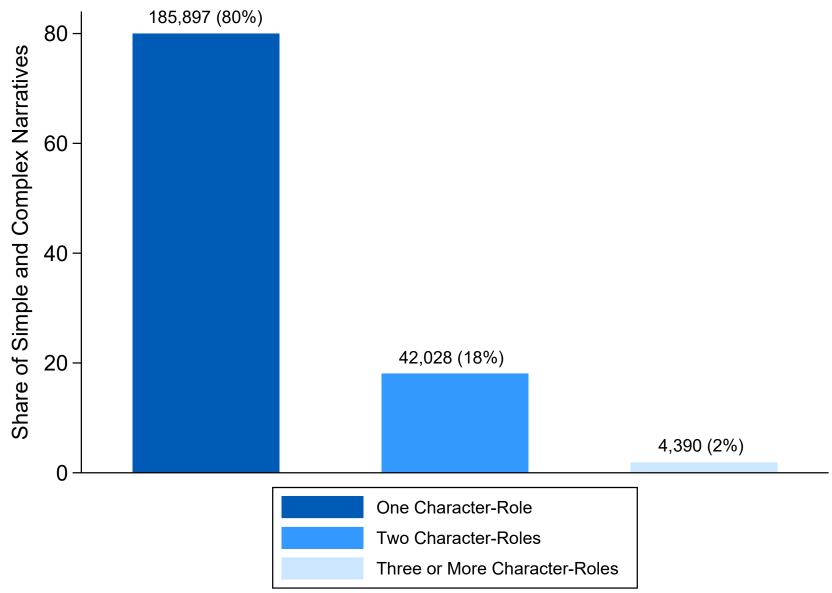
(b) *Frequency Including Retweets*



Notes: Panel (a) shows the six months interval corresponding to Figure 4. We include only character-role combinations for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. The shares of character-roles are cumulative (i.e. sum up to 100) and are computed on a yearly basis. We label the most common character-roles. In parentheses, each label indicates the share of the corresponding character-role in the first period, the average share across all periods, and the share in the last period. Unlabeled character-roles are the following: Civil Society-Villain, BRICS-Hero, Science-Hero, Nuclear Tech-Villain and Emiss. Pricing-Villain. Panel (b) presents comparable information regarding the share of character-roles, while also taking into account the number of retweets each character-role received. Unlabeled character-roles are the following: Civil Society-Villain, BRICS-Hero, BRICS-Villain, Green Tech-Villain, Nuclear Tech-Hero, Nuclear Tech-Villain and Emiss. Pricing-Villain.

Figure C.6 shows the distribution of simple and complex narratives, among the tweets extracted for the analysis that originate from the United States. A simple narrative contains a topic and a single character-role. A narrative is complex when it contains a topic and two or more character-roles. The majority of tweets in our sample contain simple narratives, nevertheless complex narratives containing two character-roles account for roughly 18% of the sample. Narratives containing three or more character-roles rarely occur. Hence, at least within the **Character-Role Narrative Framework**, it is clear that simplicity is more common than complexity in the discussion about climate change policies on Twitter.

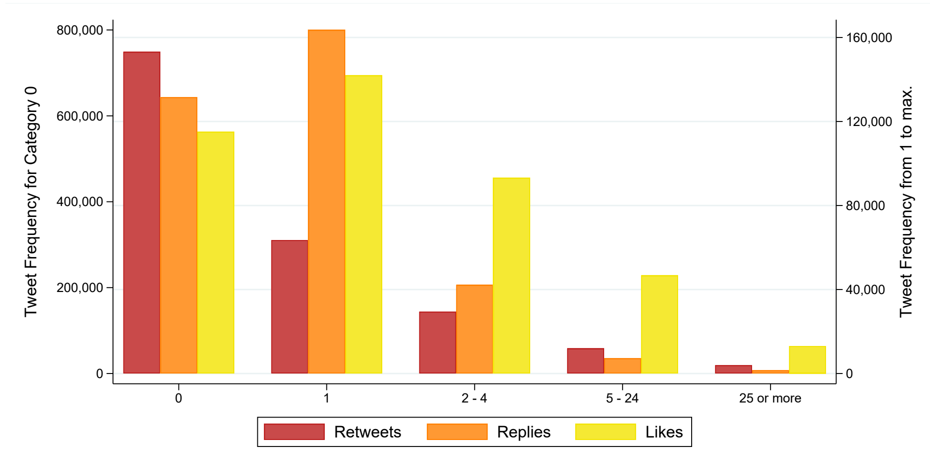
Figure C.6: *Distribution of Simple and Complex Narratives*



Notes: The figure shows the distribution of simple and complex narratives over the total number of narratives detected by the model. We include only character-roles for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. A simple narrative contains a topic and a single character-role. A narrative is complex when it contains a topic and two or more character-roles. Figure 5 shows how frequently each character-role occurs with another one (or by itself).

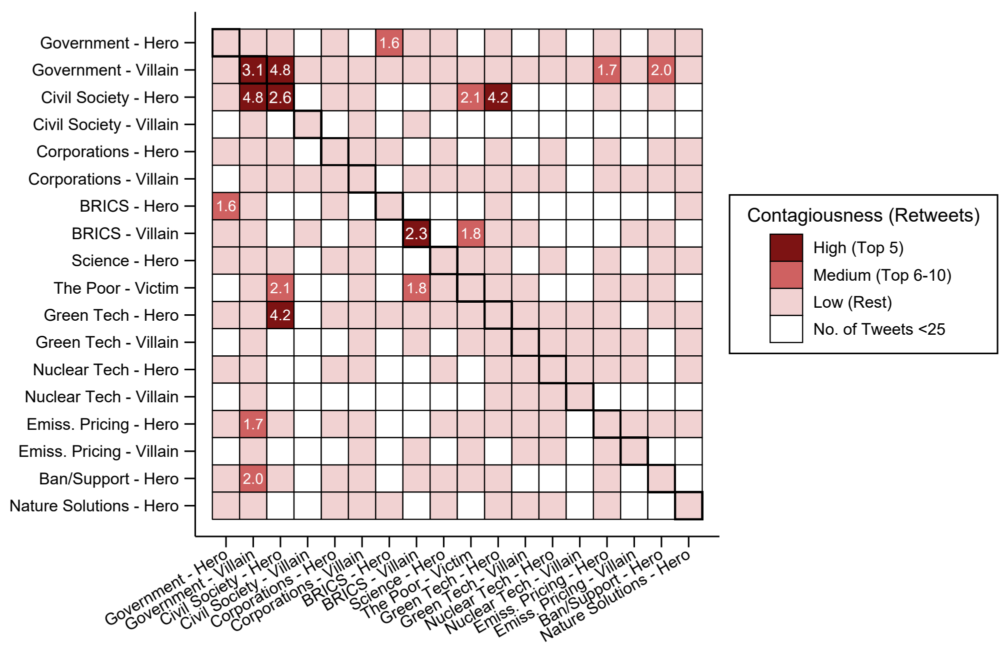
Figure C.7 shows the distribution of retweets, replies and likes across those tweets originating from the United States that contain at least one character-role, thus a simple narrative. Two main insights emerge. First, the majority of tweets do not obtain any retweets, replies and likes. Second, there is heterogeneity across the the three measures of virality, suggesting that distinguishing them is the right decision. Figures C.8, C.9 and C.10 provide information on the virality of character-roles. The diagonal of each matrix shows the average virality of a character-role when occurring by itself in a tweet. All other entries of the matrices show the average virality of the co-occurrences of two character-roles. These figures provide similar information to that shown in the paper - Figures 7, 8 and 9 - but include also those character-role combinations that comprise less than 50 - but more than 25 - tweets in the period of analysis. Changing the threshold from 50 to 25 does not drastically change the results. Once again popularity and contagiousness behave similarly, while conversation is differentiated, with the majority of replies being for tweets that talk about instrumental characters.

Figure C.7: *Distribution of Virality*



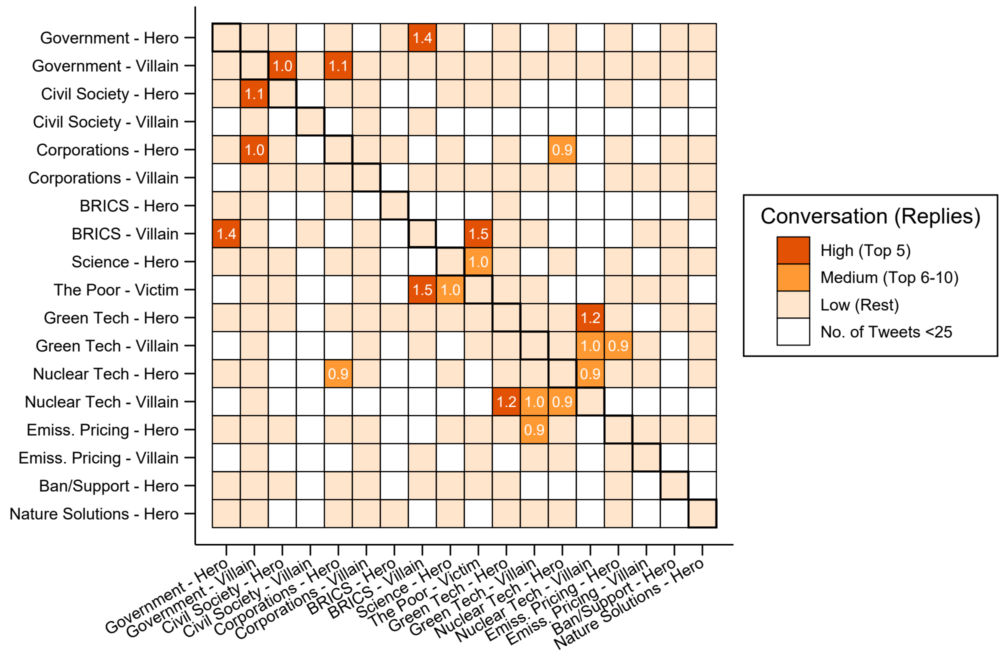
Notes: The figure shows the distribution of tweets containing at least one character-role over the number of retweets/replies/-likes a tweet obtained. We include only character-roles for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. The left y-axis refers to the number of tweets that received 0 retweets/replies/likes. The right y-axis refers to the number of tweets that received 1; 2-4; 5-24; 25 or more retweets/replies/-likes. In Figure 6 we show the virality by character-role.

Figure C.8: *Virality: Contagiousness Measured in Retweets (25)*



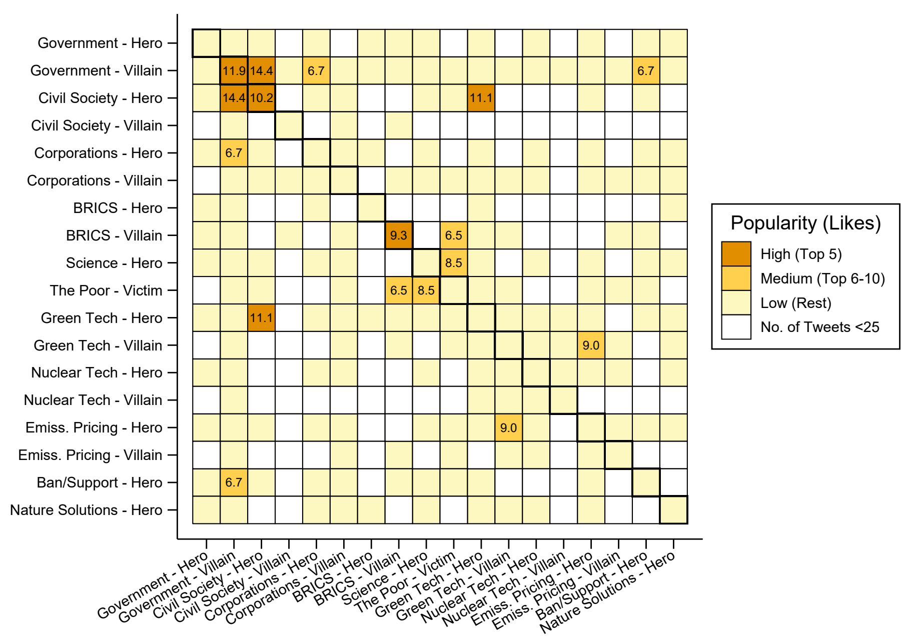
Notes: The figure shows the contagiousness of tweets measured as retweet rate. We include only character-roles for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. We include only character-role combinations that appear in 25 or more tweets in our analysis period. The matrix is symmetric but shown entirely for ease of reading. The diagonale displays the contagiousness of each character-role when appearing in a simple narrative, hence by itself. The intervals of the three levels indicated in the legend are the following: low, below 1.61 retweets; medium, from 1.61 to 2.1; high, above 2.1 retweets. Figure 7 shows the same information considering a threshold of 50 tweets to include the character-role combinations in the matrix.

Figure C.9: *Virality: Conversation Measured in Replies (25)*



Notes: The figure shows the conversation sparked by tweets measured as reply rate. We include only character-roles for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. We include only character-role combinations that appear in 25 or more tweets in our analysis period. The matrix is symmetric but shown entirely for ease of reading. The diagonale displays the conversation sparked by each character-role when appearing in a simple narrative, hence by itself. The intervals of the three levels indicated in the legend are the following: low, below 0.9 replies; medium, from 0.9 to 1.027; high, above 1.027 replies. Figure 8 in the paper shows the same information considering a threshold of 50 tweets to include the character-role combinations in the matrix.

Figure C.10: *Virality: Popularity Measured in Likes (25)*



Notes: The figure shows the popularity of tweets measured as like rate. We include only character-roles for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. We include only character-role combinations that appear in 25 or more tweets in our analysis period. The matrix is symmetric but shown entirely for ease of reading. The diagonale displays the popularity of each character-role when appearing in a simple narrative, hence by itself. The intervals of the three levels indicated in the legend are the following: low, below 6.4 likes; medium, from 6.4 to 9.1; high above 9.1 likes. Figure 9 shows the same information considering a threshold of 50 tweets to include the character-role combinations in the matrix.

D Further Details on the Main Results

Table D.1 shows the correspondent models for the coefficients graphically shown in Figure 10. The columns indicate the specific determinant of virality tested - e.g. villain narratives vs. hero narratives - while the rows report the coefficient of the Poisson pseudo maximum likelihood regression. The dependent variables are respectively the contagiousness exemplified by number of retweets - Panel A - the conversation exemplified by number of replies - Panel B - and popularity exemplified by number of likes, in Panel C. For each coefficient we report also SE and p-value. Table D.2 displays the models for the coefficients in Figure 11. The columns indicate the specific kind of narrative whose frequency we test - e.g. villain narratives - while rows the coefficient of the regression. Panel A shows the results of the regression for the full period while Panel B and C the results for the event study in the 6 months around elections.

Table D.1: *The Determinants of Virality*

Panel A: Contagiousness

	(1) Villain vs Hero	(2) Villain vs Victim	(3) Hero vs Victim	(4) Human vs Instrument	(5) Villain-Human vs Hero-Human	(6) Villain-Instrument vs Hero-Instrument	(7) Simple vs Complex
	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value
Coefficient	0.818	0.852	0.004	0.765	0.744	-0.377	0.554
	(0.162)	(0.110)	(0.145)	(0.137)	(0.148)	(0.198)	(0.113)
	[0.000]	[0.000]	[0.975]	[0.000]	[0.000]	[0.057]	[0.000]
Week-of-the-Year & State FE	yes	yes	yes	yes	yes	yes	yes
Weekly Time Trend	yes	yes	yes	yes	yes	yes	yes
Tweet-Level Controls	yes	yes	yes	yes	yes	yes	yes
Pseudo R-Squared	0.21	0.32	0.13	0.19	0.23	0.16	0.19
N	194,813	110,382	121,561	213,538	134,258	81,426	213,538

Panel B: Conversation

	(1) Villain vs Hero	(2) Villain vs Victim	(3) Hero vs Victim	(4) Human vs Instrument	(5) Villain-Human vs Hero-Human	(6) Villain-Instrument vs Hero-Instrument	(7) Simple vs Complex
	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value
Coefficient	0.151	0.232	0.120	0.125	0.125	-0.005	0.219
	(0.069)	(0.064)	(0.046)	(0.071)	(0.075)	(0.035)	(0.065)
	[0.028]	[0.000]	[0.009]	[0.076]	[0.095]	[0.888]	[0.001]
Week-of-the-Year & State FE	yes	yes	yes	yes	yes	yes	yes
Weekly Time Trend	yes	yes	yes	yes	yes	yes	yes
Tweet-Level Controls	yes	yes	yes	yes	yes	yes	yes
Pseudo R-Squared	0.10	0.14	0.11	0.09	0.10	0.13	0.09
N	194,813	110,382	121,561	213,538	134,258	81,426	213,538

Panel C: Popularity

	(1) Villain vs Hero	(2) Villain vs Victim	(3) Hero vs Victim	(4) Human vs Instrument	(5) Villain-Human vs Hero-Human	(6) Villain-Instrument vs Hero-Instrument	(7) Simple vs Complex
	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value
Coefficient	0.648	0.598	-0.049	0.706	0.550	-0.440	0.590
	(0.144)	(0.113)	(0.137)	(0.112)	(0.150)	(0.103)	(0.103)
	[0.000]	[0.000]	[0.722]	[0.000]	[0.000]	[0.000]	[0.000]
Week-of-the-Year & State FE	yes	yes	yes	yes	yes	yes	yes
Weekly Time Trend	yes	yes	yes	yes	yes	yes	yes
Tweet-Level Controls	yes	yes	yes	yes	yes	yes	yes
Pseudo R-Squared	0.21	0.30	0.16	0.20	0.22	0.17	0.20
N	194,813	110,382	121,561	213,538	134,258	81,426	213,538

Notes: The table shows the coefficients of Poisson pseudo-maximum likelihood regressions testing the determinants of the virality of narratives exemplified by retweets (Panel A), replies (Panel B) and likes (Panel C). We include only character-roles for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. The dependent variable is a tweet-level count of retweets/replies/likes, the independent variable a dummy that equals to one if a particular narrative is present in the tweet. If the model tests e.g., the contagiousness of villain vs. hero, it includes a dummy that equals 1 when the tweet contains a villain narrative, 0 when it contains a hero. The dummy is null when the tweet contains neither one or both simultaneously. All models include state and week-of-the-year FE, a linear weekly time trend and controls for the number of hashtags, mentions and words in the tweet as well as the number of followers, following and tweets ever produced by a user. Standard errors are clustered at the state level. Figure 10 shows the corresponding coefficient plot.

Table D.2: *The Trump Effect*

Panel A: The Trump Effect - Full Time Period

	(1) Villain Coef./SE/p-value	(2) Victim Coef./SE/p-value	(3) Hero Coef./SE/p-value	(4) Human Coef./SE/p-value	(5) Instrument Coef./SE/p-value	(6) Simple Coef./SE/p-value
Trump vs Obama	0.207	-0.001	-0.177	0.120	-0.170	0.032
	(0.008)	(0.000)	(0.007)	(0.004)	(0.010)	(0.001)
	[0.000]	[0.834]	[0.000]	[0.000]	[0.005]	[0.146]
Trump vs Biden	0.084	-0.008	-0.085	0.127	-0.164	0.039
	(0.001)	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
State FE	yes	yes	yes	yes	yes	yes
Week-of-the-Year FE	yes	yes	yes	yes	yes	yes
Weekly Time Trend	yes	yes	yes	yes	yes	yes
Tweet-Level Controls	yes	yes	yes	yes	yes	yes
Adj. R-Squared	0.09	0.01	0.08	0.04	0.07	0.02
N	213,538	213,538	213,538	213,538	213,538	213,538

Panel C: Trump vs. Obama - 6-Month Time Period

	(1) Villain Coef./SE/p-value	(2) Victim Coef./SE/p-value	(3) Hero Coef./SE/p-value	(4) Human Coef./SE/p-value	(5) Instrument Coef./SE/p-value	(6) Simple Coef./SE/p-value
Trump vs Obama	0.086	0.002	-0.083	0.103	-0.140	0.032
	(0.002)	(0.000)	(0.002)	(0.001)	(0.002)	(0.000)
	[0.000]	[0.808]	[0.001]	[0.000]	[0.000]	[0.002]
State FE	yes	yes	yes	yes	yes	yes
Week-of-the-Year FE	no	no	no	no	no	no
Weekly Time Trend	yes	yes	yes	yes	yes	yes
Tweet-Level Controls	yes	yes	yes	yes	yes	yes
Adj. R-Squared	0.19	0.01	0.17	0.08	0.17	0.04
N	22,424	22,424	22,424	22,424	22,424	22,424

Panel D: Trump vs. Biden - 6-Month Time Period

	(1) Villain Coef./SE/p-value	(2) Victim Coef./SE/p-value	(3) Hero Coef./SE/p-value	(4) Human Coef./SE/p-value	(5) Instrument Coef./SE/p-value	(6) Simple Coef./SE/p-value
Trump vs Biden	0.033	0.004	-0.048	0.117	-0.147	0.034
	(0.000)	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)
	[0.020]	[0.442]	[0.002]	[0.000]	[0.000]	[0.006]
State FE	yes	yes	yes	yes	yes	yes
Week-of-the-Year FE	no	no	no	no	no	no
Weekly Time Trend	yes	yes	yes	yes	yes	yes
Tweet-Level Controls	yes	yes	yes	yes	yes	yes
Adj. R-Squared	0.02	0.00	0.02	0.03	0.04	0.01
N	30,281	30,281	30,281	30,281	30,281	30,281

Notes: The table shows the coefficients of regression investigating the impact of being under the Trump administration relative to the Obama and Biden administrations on the frequency of particular narratives, e.g. villain narratives. Panel A shows two coefficients from the same regression covering full period 2010-2021 and reflecting a dummy variable comparing the Trump period with Obama (symbol ●) and with Biden (symbol ▲). The period of reference is the Trump administration. We multiply the coefficients by -1, thus interpreting them as the effect of Trump relative to Obama and to Biden. The regression includes state and week-of-the-year FE, a linear weekly time trend and controls for the number of hashtags, mentions and words in the tweet as well as the number of followers, following and tweets ever produced by a user. Panel B and C show coefficients from separate regressions focusing on an event window 6 months before and after the transition between Trump and Obama (symbol ■) and between Trump and Biden (symbol ⧫). Also in this case, we multiply the coefficients by -1, thus interpreting them as the effect of Trump relative to Obama (Biden). The regression includes state FE and controls for the number of hashtags, mentions and words in the tweet as well as the number of followers, following and tweets ever produced by a user. In all models SE are clustered at the state level. Figure 11 shows the corresponding coefficient plot.

E Additional Results and Robustness Checks

E.1 The Determinants of Virality

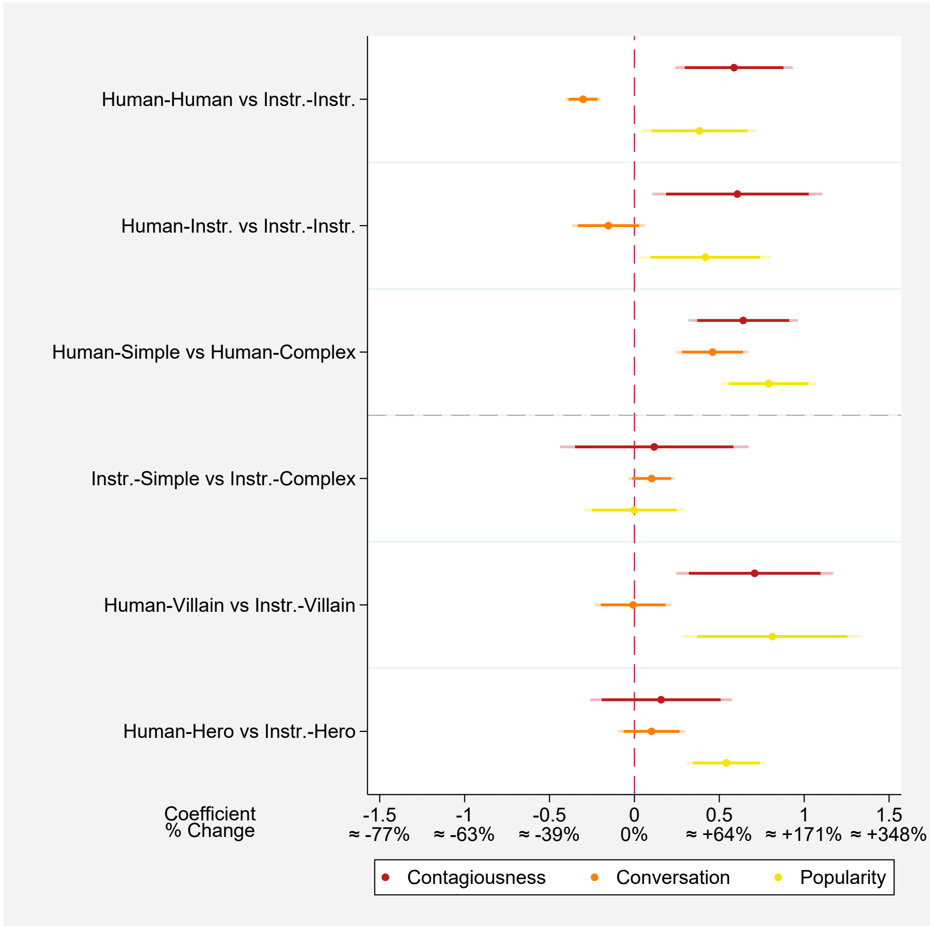
This section reports additional results and robustness checks concerning the regression analysis on the determinants of virality, that our Framework and pipeline allow to explore. In Figure E.1 we show the impact of additional determinants on the three dimensions of virality. The results corroborate what already emerges in Figure 10. Human narratives are generally the most contagious and popular. Nevertheless, tweets containing two instrument characters spark more conversations than those containing two human characters. Additionally, even when associated to human characters narratives, complexity is detrimental for virality. In Figure E.2 we show the same regression models displayed in the paper Figure 10, performed by OLS rather than Poisson pseudo-maximum likelihood regressions. The results do not change considerably from the main specification.

We address the relevant issue of Bots in the social media platform Twitter. In the context of this social media, one could think of a Bot as a user powered by an algorithm that is designed to automate tasks such as liking, retweeting and commenting the content created by other users. Given the automatic nature of their actions, Bots can manage to interact with a large number of other users, potentially becoming quite popular. Hence, we want to test whether the results on the determinant of virality we obtain are somehow driven by the presence and activity of Bots.

There is not a universally accepted way to define a Bot or to spot its activity in social medias. Programs like Botometer provide powerful tools to identify bots on Twitter but, to be very reliable, they necessitate a large amount of tweets for each user, and unfortunately we are limited by our dataset. Hence, we adopt a simpler approach following the related literature. Chu *et al.* (2012) propose to use what they call the *reputation* measure. This consists in an indicator obtained dividing the number of followers of a users, by the sum of the number of followers and number of users followed by the users itself. Generally, Bots will have low reputation, because they follow as many users as possible. We build also on the work by Tabassum *et al.* (2023). The authors indicate the number of tweets as an additional important indicator of the activity of Bots, which generally are extremely prolific.

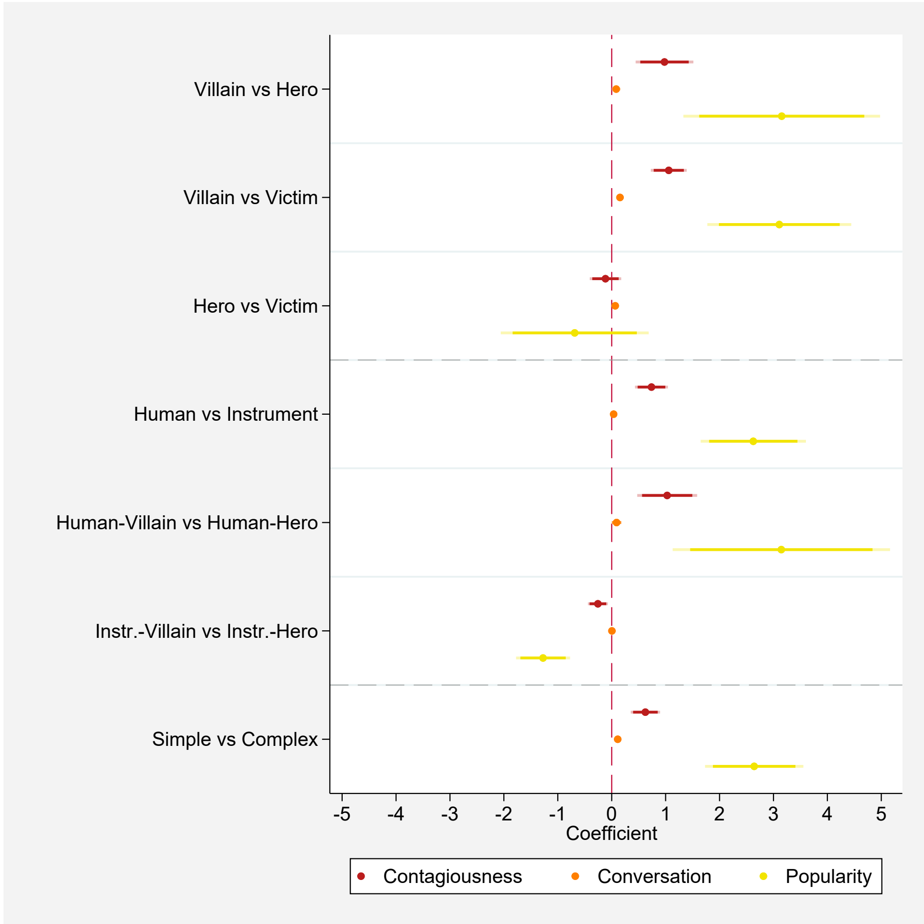
In Figure E.3 we report the same tests shown in the paper Figure 10, but accounting for potential Bots' activity. In particular, we drop those tweets posted by users that are in both the bottom 25% of the reputation measure and in the top 25% of the count of posted tweets. Results do not change considerably from what emerges in the main analysis. We propose two interpretations for this. First, it is possible that Bots do not populate the discussion about climate change policy. Second, it is possible that Bots do take part of the discussion but the popularity of their tweets does not differ from that of human users.

Figure E.1: *Other Determinants of Virality*



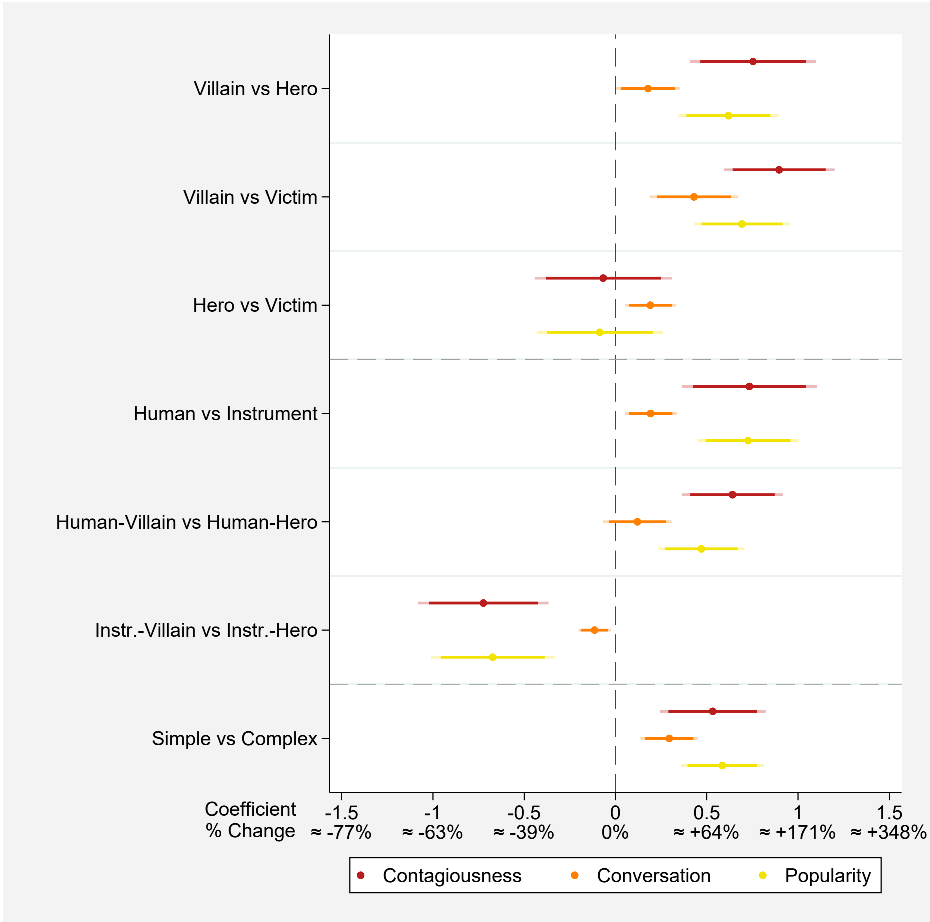
Notes: The figure shows the coefficients of Poisson pseudo-maximum likelihood regressions testing the determinants of the virality of narratives other than the ones reported in Figure 10. We include only character-roles for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. The dependent variable is a tweet-level count of retweets/replies/likes, the independent variable a dummy that equals to 1 if a particular narrative is present in the tweet. If the model tests e.g., the contagiousness of villain vs. hero, it includes a dummy that takes the value 1 when the tweet contains a villain narrative, 0 when it contains a hero. The dummy is null when the tweet contains neither one or both simultaneously. The x-axis reports labels for the coefficient values and the correspondent percentage change computed as follows: $\approx e^{\beta} - 1$. All models include state and week-of-the-year FE, a linear weekly time trend and controls for the number of hashtags, mentions and words in the tweet as well as the number of followers, following and tweets ever produced by a user. Standard errors are clustered at the state level.

Figure E.2: *The Determinants of Virality (OLS)*



Notes: The figure shows the coefficients of OLS regressions testing the determinants of the virality of narratives expressed as the count of retweets, replies and likes. We include only character-roles for which the model performs with an F1-score above .6 and we use tweets that we can geo-localize at least at the US-country level. The dependent variable is a tweet-level count of retweets/replies/likes, the independent variable a dummy that equals to 1 if a particular narrative is present in the tweet. If the model tests e.g., the contagiousness of villain vs. hero, it includes a dummy that takes the value 1 when the tweet contains a villain narrative, 0 when it contains a hero. The dummy is null when the tweet contains neither one or both simultaneously. All models include state and week-of-the-year FE, a linear weekly time trend and controls for the number of hashtags, mentions and words in the tweet as well as the number of followers, following and tweets ever produced by a user. Standard errors are clustered at the state level. Figure 10 in the paper, shows the same tests performed with Poisson pseudo-maximum likelihood regressions.

Figure E.3: *Virality Results Taking Potential Bots into Account*

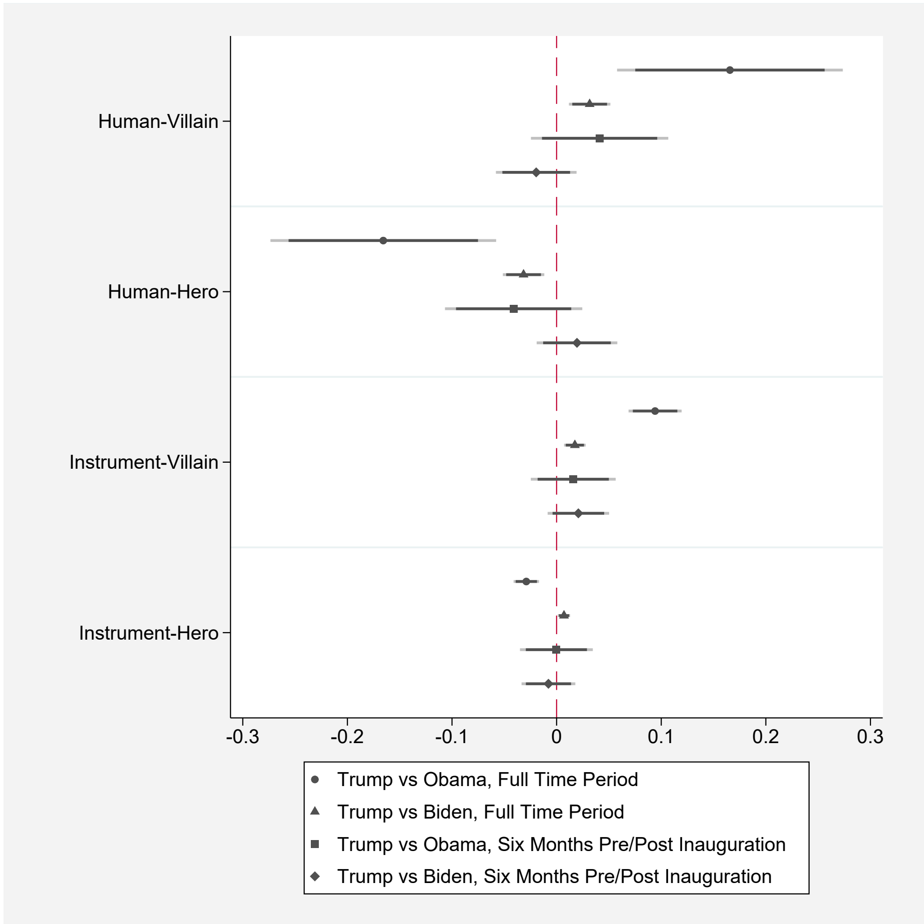


Notes: The figure shows how the main results in Figure 10 change when excluding tweets that are potentially generated by Twitter bots. Chu *et al.* (2012) compute the *reputation* of a Twitter account by dividing the number of followers of that account by its number of followers added to the number of accounts it follows. Additionally, Tabassum *et al.* (2023) identify a high rate of tweet production as a main characteristic of Twitter bots. We incorporate these insights by dropping tweets that are both in the bottom 25% of the *reputation* distribution and the top 25% of the distribution of number of tweets an author generated (n=15,102).

E.2 The Effect of Populism

This section reports additional results and robustness checks concerning the analysis on the effects of populism on the frequency of narratives. In figure 11, we see that there is a clear *Trump Effect* in that the discussion about climate change policy moved during the Trump administration towards simpler narratives about human characters and villains. Figure E.4 shows similar results focusing on additional typologies of narratives. When testing e.g. human-villain narratives, it means we check whether there is a *Trump Effect* on the frequency of narratives that contain a human that is in the role of villain. The statistically significant coefficients show results in line with the *Trump Effect* emerging in Figure 11.

Figure E.4: *The Trump Effect - Further Outcomes*

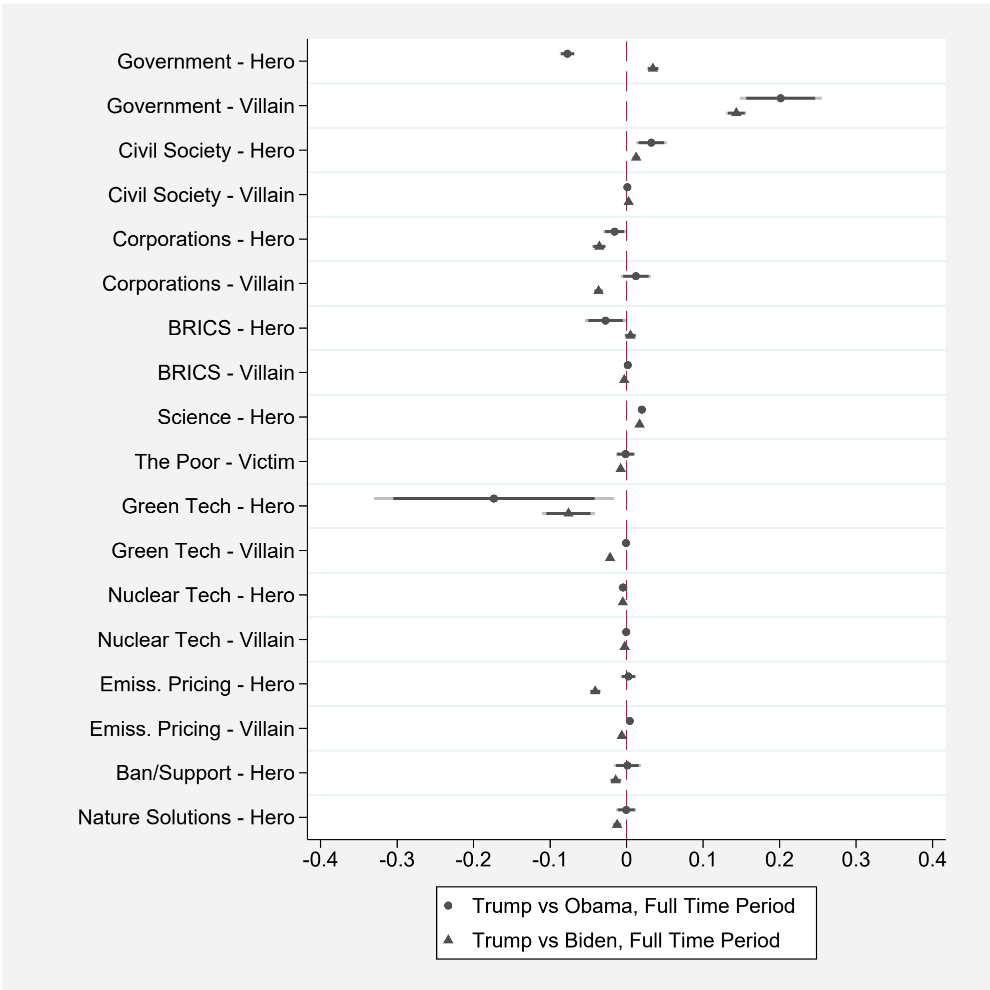


Notes: The figure shows the impact of being under the Trump administration relative to the Obama and Biden administrations on the frequency of particular narratives other than the ones reported in Figure 11. The first two coefficients are from the same regression covering full period 2010-2021 and reflect a dummy variable comparing the Trump period with Obama (symbol ●) and with Biden (symbol ▲). The third and fourth coefficient are from separate regressions focusing on an event window 6 months before and after the transition between Trump and Obama (symbol ■) and between Trump and Biden (symbol ⧫). All regressions include state FE, a linear weekly time trend and controls for the number of hashtags, mentions and words in the tweet, as well as for the number of followers, number of other user followed, and overall number of tweets issued by a user. The full period regressions also controls for seasonality using week-of-the-year FE. In all models SE are clustered at the state level.

Figure E.5 reproduces the same model used for the first two coefficients in Figure 11 (symbols ● and ▲), for each character-role of our analysis. The results corroborate what shown in the paper and some interesting patterns emerge. It seems that the higher frequency of villain narratives during the Trump administration is driven predominantly by those narratives depicting government as the villain. Narratives depicting BRICS as the hero decrease during the Trump administration as well as those depicting Green Tech as the hero. Finally, it is interesting to see how narratives depicting Emission Pricing as the hero, drastically decrease during the Trump administration years relative to the Biden administration years.

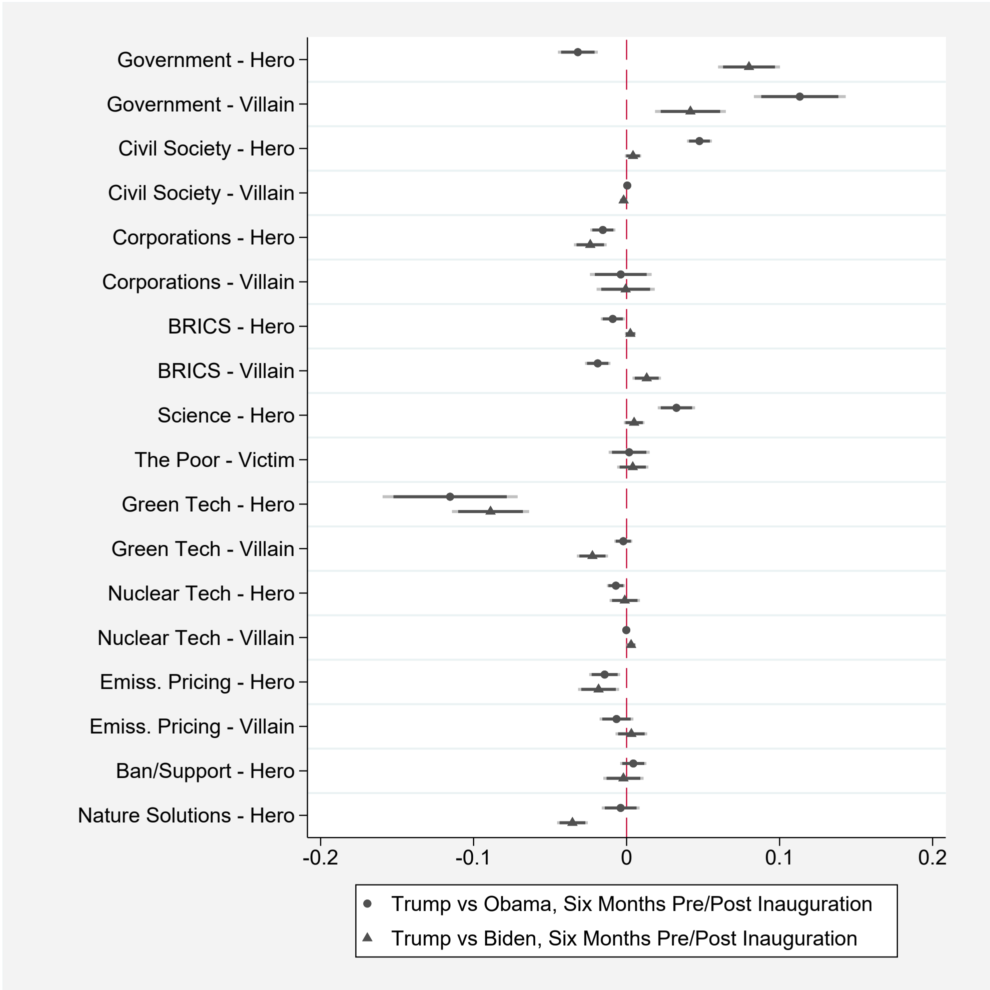
Figure E.6 reproduces the same models used for the last two coefficients in Figure 11 (symbols ■ and ⧫),for each character-role of our analysis. Even when looking at the six months window around the election of Trump after Obama and of Biden after Trump, one can see a strong change in the frequency of narratives. This corroborates our hypothesis of a true *Trump Effect* rather than a simple time trend picked by our full period analysis.

Figure E.5: *The Trump Effect for each Character-Role: Full Period*



Notes: The figure shows the impact of being under the Trump administration relative to the Obama and Biden administrations on the frequency of each character-role of our analysis. The two coefficients are from the same regression covering full period 2010-2021 and reflect a dummy variable comparing the Trump period with Obama (symbol ●) and with Biden (symbol ▲). All regressions include state FE, a linear weekly time trend and controls for the number of hashtags, mentions and words in the tweet, as well as for the number of followers, number of other user followed, and overall number of tweets issued by a user. We also control for seasonality using week-of-the-year FE. In all models SE are clustered at the state level. Figure 11 shows the main results.

Figure E.6: *The Trump Effect for each Character-Role: Six Month Window*



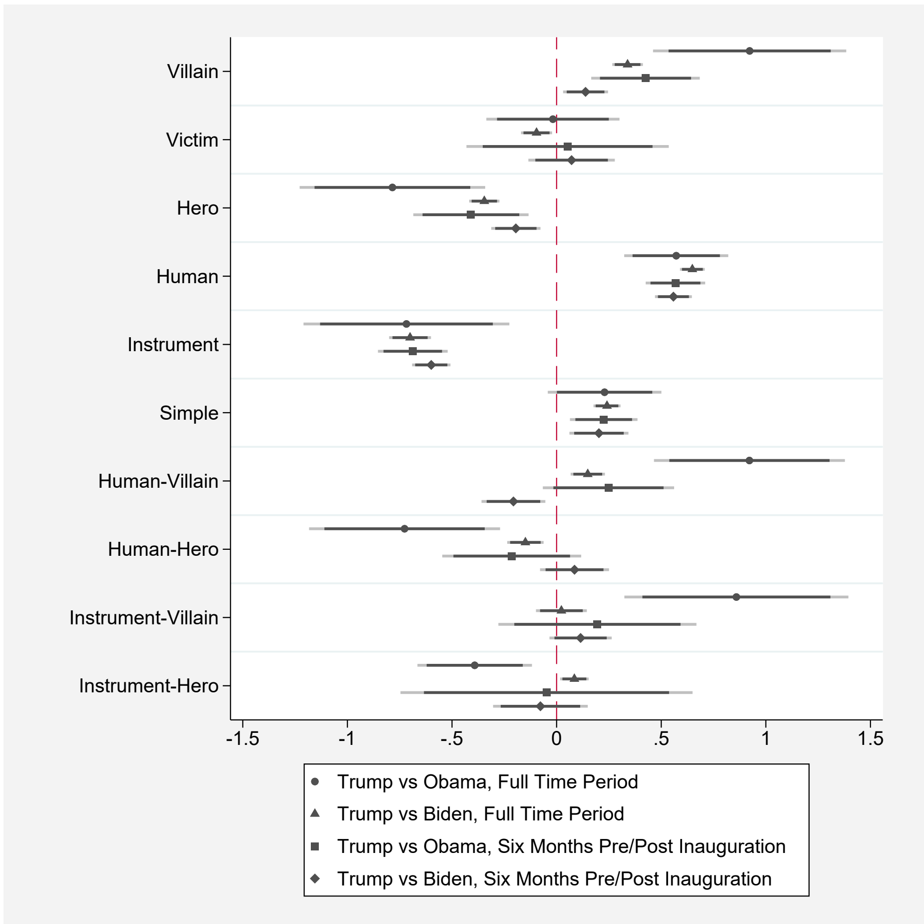
Notes: The figure shows the impact of being under the Trump administration relative to the Obama and Biden administrations on the frequency of each character-role of our analysis. The coefficients come from separate regressions focusing on an event window 6 months before and after the transition between Trump and Obama (symbol ■) and between Trump and Biden (symbol ⧫). All regressions include state FE, a linear weekly time trend and controls for the number of hashtags, mentions and words in the tweet, as well as for the number of followers, number of other user followed, and overall number of tweets issued by a user. Figure 11 shows the main results.

We perform robustness checks to validate our results on the impact of populism. Firstly, we show in Figure E.7 the same tests reported in the paper figure 11, using Logistic Regression instead of OLS. The results do not differ considerably from what is shown in the paper.

One of the concerns we have when exploring the *Trump Effect* on the full period of our analysis - see equation specification 2 and Figure 11 - is that some major events might affect the relationship of interest. The most relevant of those potential events are extreme weather events like hurricanes, floods and fires. In Table E.3, we show in Panel A our main specification to investigate the *Trump Effect* along the full period (see Panel A of Table D.2). We show in Panel B the same specification, adding as a control the national yearly average number of natural disaster. Despite the consistently statistically significant effect, including this control does not affect the main results of the paper.

We also do the same robustness check described above in the section dedicated to the determinants of virality, to rule out the potential impact of Bots on our results. We follow exactly the same approach and, also in this case, eliminating tweets posted by potential Bots does not have a strong impact on the main results.

Figure E.7: *The Trump Effect - Logistic Regression*



Notes: The figure shows the impact of being under the Trump administration relative to the Obama and Biden administrations on the frequency of particular character-role narrative categories, e.g. villain narratives, using Logistic Regression models. The first two coefficients are from the same regression covering the full period 2010-2021 and reflect a dummy variable comparing the Trump period with Obama (symbol ●) and with Biden (symbol ▲). The third and fourth coefficient are from separate regressions focusing on an event window 6 months before and after the transition between Trump and Obama (symbol ■) and between Trump and Biden (symbol ⧫). All regressions include state FE, a linear weekly time trend and controls for the number of hashtags, mentions and words in the tweet, as well as for the number of followers, number of other user followed, and overall number of tweets issued by a user. The full period regressions also controls for seasonality using week-of-the-year FE. In all models SE are clustered at the state level. Table D.2 reports the full specifications. Figure 11 shows the main results.

Table E.3: *The Trump Effect: Controlling for Natural Disasters*

Panel A: Full Time Period without Extreme Weather Event Controls

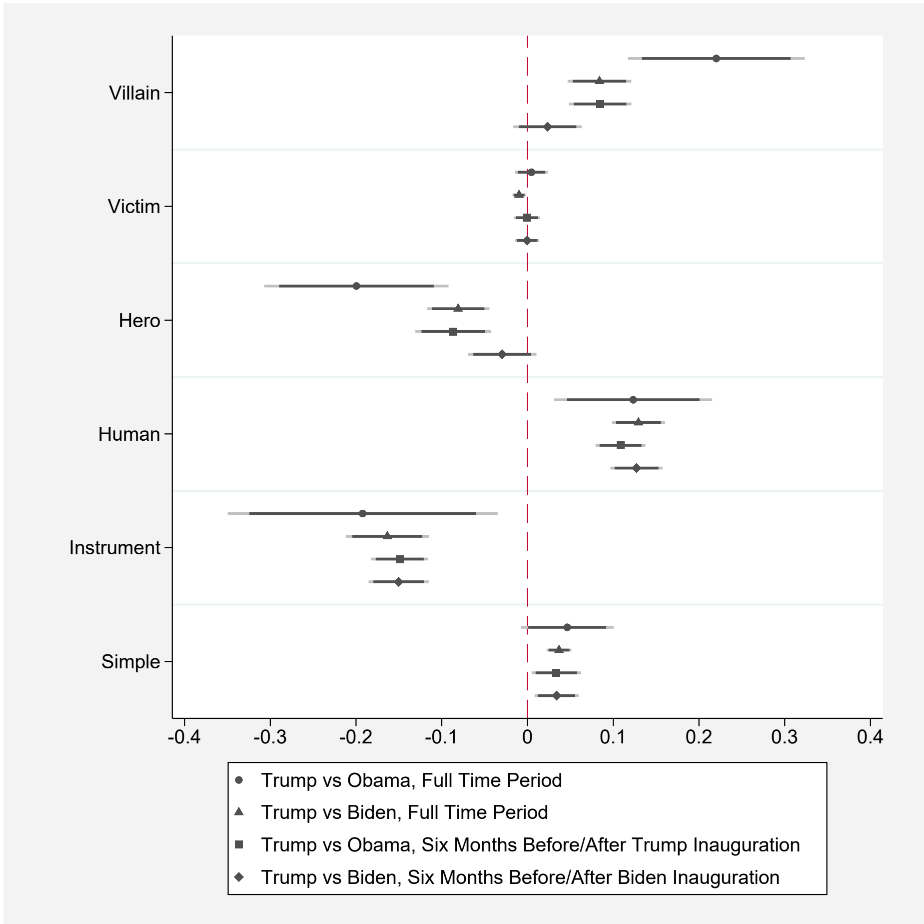
	(1) Villain	(2) Victim	(3) Hero	(4) Human	(5) Instrument	(6) Simple
	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value
Trump vs Obama	0.207	-0.001	-0.177	0.120	-0.170	0.032
	(0.008)	(0.000)	(0.007)	(0.004)	(0.010)	(0.001)
	[0.000]	[0.834]	[0.000]	[0.000]	[0.005]	[0.146]
Trump vs Biden	0.084	-0.008	-0.085	0.127	-0.164	0.039
	(0.001)	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
State FE	yes	yes	yes	yes	yes	yes
Week-of-the-Year FE	yes	yes	yes	yes	yes	yes
Weekly Time Trend	yes	yes	yes	yes	yes	yes
Controls for No. of Mentions and No. of Hashtags	yes	yes	yes	yes	yes	yes
Controls for Extreme Weather Events	no	no	no	no	no	no
Adj. R-Squared	0.09	0.01	0.08	0.04	0.07	0.02
N	213,538	213,538	213,538	213,538	213,538	213,538

Panel B: Full Time Period with Extreme Weather Event Controls

	(1) Villain	(2) Victim	(3) Hero	(4) Human	(5) Instrument	(6) Simple
	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value	Coef./SE/p-value
Trump vs Obama	0.206	-0.002	-0.176	0.119	-0.170	0.032
	(0.009)	(0.000)	(0.008)	(0.004)	(0.010)	(0.001)
	[0.000]	[0.826]	[0.000]	[0.001]	[0.007]	[0.134]
Trump vs Biden	0.102	-0.005	-0.102	0.139	-0.175	0.035
	(0.001)	(0.000)	(0.001)	(0.001)	(0.003)	(0.000)
	[0.000]	[0.013]	[0.000]	[0.000]	[0.000]	[0.000]
National avg. of	0.013	0.002	-0.013	0.009	-0.008	-0.003
natural disasters in	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
the year	[0.000]	[0.000]	[0.000]	[0.000]	[0.011]	[0.034]
State FE	yes	yes	yes	yes	yes	yes
Week-of-the-Year FE	yes	yes	yes	yes	yes	yes
Weekly Time Trend	yes	yes	yes	yes	yes	yes
Controls for No. of Mentions and No. of Hashtags	yes	yes	yes	yes	yes	yes
Controls for Extreme Weather Events	yes	yes	yes	yes	yes	yes
Adj. R-Squared	0.09	0.01	0.08	0.05	0.07	0.02
N	213,538	213,538	213,538	213,538	213,538	213,538

Notes: The table shows how the results in Figure 11 change when natural disasters are taken into account. Panel A is the same shown in Table D.2. Panel B adds a control for the national yearly average of natural disasters in the US.

Figure E.8: *The Trump Effect: Taking Potential Bots into Account*



Notes: The figure shows how the main results in Figure 11 change when excluding tweets that are potentially generated by Twitter bots. Chu *et al.* (2012) compute the *reputation* of a Twitter account by dividing the number of followers of that account by its number of followers added to the number of accounts it follows. Additionally, Tabassum *et al.* (2023) identify a high rate of tweet production as a main characteristic of Twitter bots. We incorporate these insights by dropping tweets that are both in the bottom 25% of the *reputation* distribution and the top 25% of the distribution of number of tweets an author generated (n=15,102).

F Annotation and Model

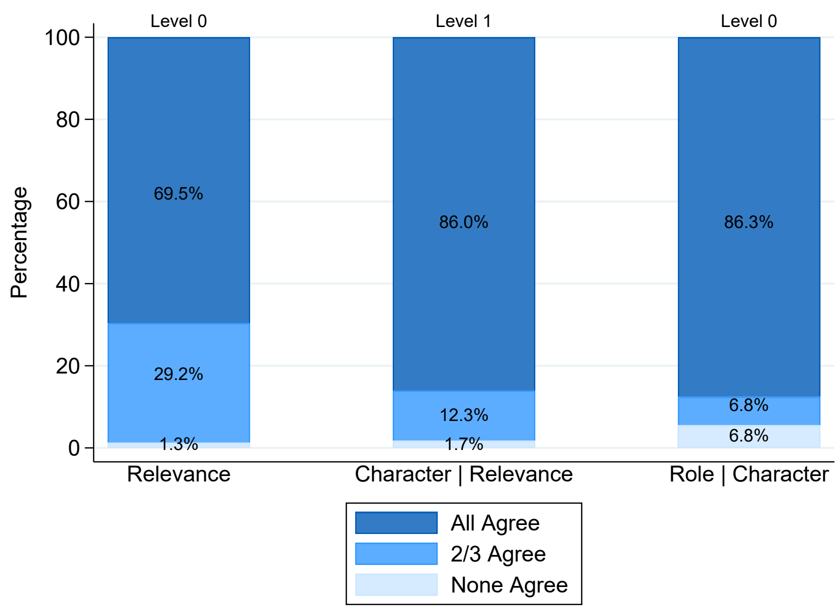
F.1 Intercoder Agreement

In this section we provide insights into the annotation process, in particular we explore coding agreement. The agreement among coders is an important measure to determine the clarity of the annotation codebook and thus, partly, the quality of the annotation itself. In Figure F.1 we show the conditional agreement among the three coders that annotated the majority of tweets in our training sample. As a reminder our annotation process consists of three levels. In Level 0, coders decide whether a tweet is relevant, irrelevant or denying climate change. In Level 1, coders identify the presence of a particular character and if present, in Level 2, they decide which of the three roles in the drama triangle the character plays.

The figure displays three levels of agreement: *All agree*, *2/3 Agree* or *None Agree*. The details on the measurements are explained in the figures' notes.

The descriptive statistics on agreement clearly show that the conditional agreement among the three coders is high. Conditional on agreeing on the relevance of the tweets, coders also agree on the presence or not of a character on average 86% of the times. The same average likelihood applies for the roles. It is important to notice here that in case e.g. the relevance of a tweet was agreed only among two coders, then if the two coders agree on the characters' presence, this observation would fall into the *All agree* count for that specific tweet in the middle bar. This does not distort the results because it applies also for the category *None Agree*. In fact, with the same example, if only two coders agreed on the relevance and then they disagree on character presence, this observation would end up in the count of *None Agree* in the middle bar.

Figure F.1: *Conditional Intercoder Agreement*



Notes: The figure provides insights into the agreement among the three coders that annotated most of the training set. The agreement is computed on a sub-set of tweets that was assigned to all three. The annotation process is divided in three levels, see 1. In level 0, coders classify tweets as relevant, irrelevant or indicating denial of climate change. In level 1, they identify whether one or more of our characters is present in the tweet. In level 2, if a character is present, they decide whether the character assumes one of the three roles of the drama triangle. The figure shows three agreement types: *All agree*, *2/3 Agree* or *None Agree*. The first bar from the left concerns Level 0 of annotation. In this case, *None Agree* means that all three coders selected something different, *2/3 Agree* means two out of three agreed on the relevance, and *All agree* means the three of them agree on the relevance of the tweet. The bar in the middle concerns Level 1 of annotation and the statistics are drawn conditional on the fact that a tweet was found relevant in Level 0. In this case, *None Agree* means that despite coders agree on the relevance of the tweet they do not agree on which characters are present, *2/3 Agree* means that three coders agree on the relevance of a specific tweet and two of them also on which characters are present, and *All agree* means that all three coders agree on the relevance and on which characters are present. The last bar concerns Level 2 of annotation and the statistics are drawn conditionally on the fact that coders agree on the presence of a character. In this case, *None Agree* means that despite coders agree on the presence of the character they do not agree on which role the character plays, *2/3 Agree* means that three coders agree on the presence of a specific character and two of them also on which role the character plays, *All agree* means the three of them agree on the presence of a specific character and also on which role this character has.

F.2 Model Performance

The new generation of language models going under the name of Transformers revolutionized the NLP world reaching ever higher performances in many different tasks. Nevertheless, the complexity of the model and the computational requirements might prevent researchers to adopt them. There are indeed alternative approaches to investigate narratives in text data. These are often faster and less computationally demanding. In this section we argue that the increased performance of Transformer models makes them worth the effort in terms of preparation and computational power.

To investigate the issue above, we compare the performance of our model with a simpler, faster and often used method, the dictionary approach. We report our results in Figure F.2 and Figure F.3. The technical details of the comparison are reported in the notes of the figures. The main idea is to compare the performance of our model against the Hiv4 and NLTK sentiment dictionaries in detecting villain and hero narratives. We use the set of tweets annotated by our three main human coders and restrict it to the ones for which at least two coders agree on the presence of a villain or hero character-role as the *ground truth*. We do this comparison using character-roles that can be easily identified by keywords: BRICS, Green Tech, Nuclear Tech and Emission Pricing, thus making it more favourable for the dictionary approaches.

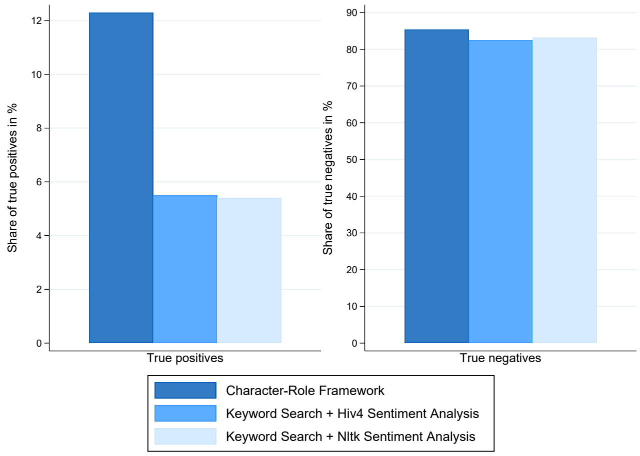
Figure F.2 and Figure F.3 show that our model performs much better than both types of sentiment dictionaries. Not only does it perform better in detecting the presence of a character-role, but it is also more efficient in recognizing false positives and false negatives. It is important to note here that the sentiment analyses can only predict one sentiment per tweet whereas our model can predict multiple characters in different roles. Additionally, one should keep in mind that the comparison is done only using villain and hero narratives. This is because sentiment analysis can hardly be adapted to capture the nuance of the role of a victim and because we do not predict the presence of a character but the absence of a role (which corresponds to a neutral sentiment).

Figure F.4 reports the performance for all character-roles of interest. Alongside the character-roles we report - in red - also the performance of the 'irrelevant' categories. The structure of the model is such that these are are also predicted as any other category at the stage of the machine learning algorithm XGBoost. However, the information obtained in the different steps of annotation is fully exploited for the Transformer model predictions as explained in appendix A.5.

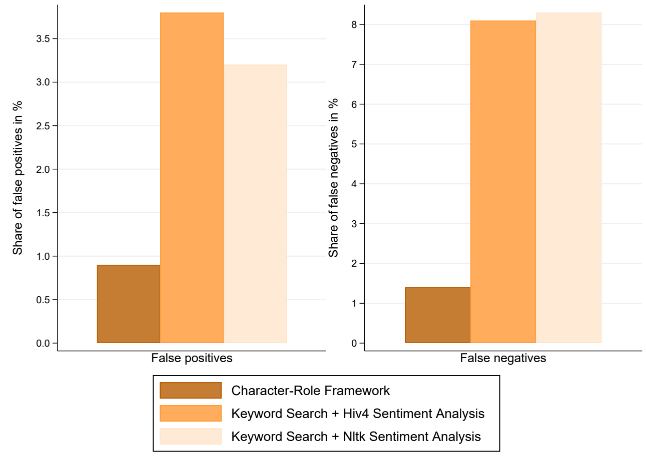
Figure F.5 shows a picture of the factors affecting the performance of the model relative to the annotation process. Figure F.5a shows the relation between the frequency of annotated data per character-role and their respective F1-score. The correlation is slightly negative and not statistically significant. There seems to lack a clear relationship between the number of tweets annotated and the model performance. Figure F.5b shows the relation between the agreement among coders and the performance of the model. The measure of agreement is obtained on a subset of tweets coded by all three coders. The more coders agree the better the performance. The correlation is positive, strong and statistically significant. A final insight concerns the concentration of coders working on a single character-role. The more coders worked on annotating a specific character-role the better the performance. In other words it seems the model works better when it has more diversity from which to learn. These are purely descriptive insights but could represent a good starting point for researchers that want to replicate our pipeline and try to maximize performance.

Figure F.2: *Comparing Performance across Alternatives Conditional on Character-Role Identification*

(a) *Percentage of True Positives and True Negatives*



(b) *Percentage of False Positives and False Negatives*



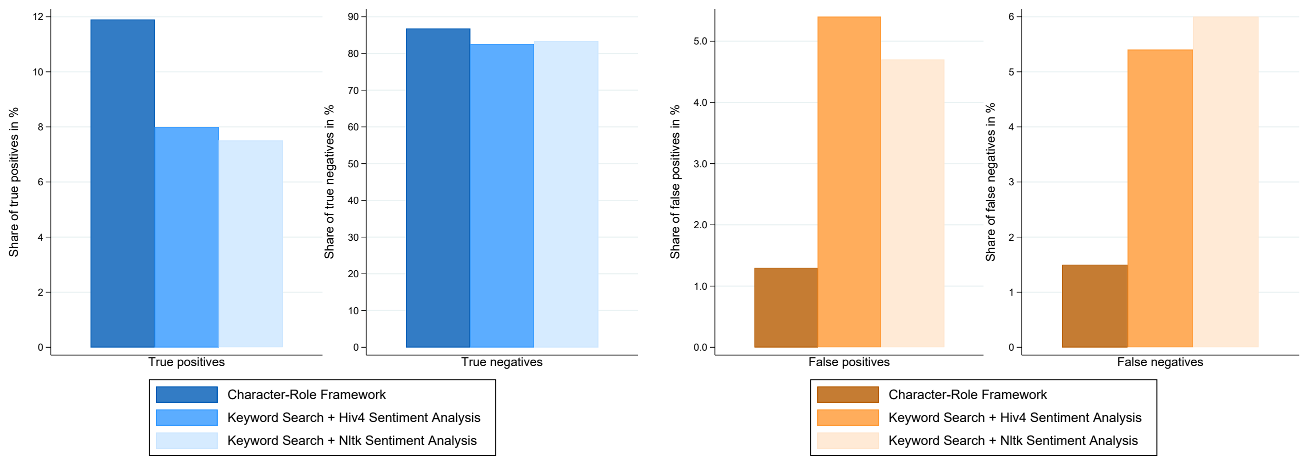
(c) *Table*

	Percentage of true positives	Percentage of true negatives	Percentage of false positives	Percentage of false negatives
Character-Role Framework	12.3	85.4	.9	1.4
Hiv4 Sentiment Analysis	5.5	82.5	3.8	8.1
Nltk Sentiment Analysis	5.4	83.1	3.2	8.3

Notes: The figure provides insights into the performance for detecting narratives of the **Character-Role Narrative Framework** pipeline against simple dictionary approaches. We compare the predictions of those models to what we take as the *ground truth*, namely that subset of tweets annotated by our coders, for which at least two coders agree on which character-roles are present. For the dictionary approaches we do the following: First, we draw a set of keywords to find the character. Second, we look for these keywords in the *ground truth* set of tweets. Third, when a character is found, we predict the sentiment (negative, positive or neutral) using two alternative sentiment dictionaries, the Hiv4 and Nltk dictionaries. When a character is present and the negative sentiment is detected, we treat this prediction as a villain. When a character is present and the positive sentiment is detected, we treat this prediction as a hero. This allows us to find the following statistics: *No. of True Positives* i.e. character-roles that both the predictions and the human coders detect. *No. of True Negatives*, i.e. character-roles that neither the predictions nor the human coders detect, *No. of False Positives*, i.e. character-roles that the predictions detect, but the human coders do not, and *No. of False Negatives*, i.e. character-roles that the predictions do not detect, but the human coders do. We compute the shares of each performance indicator. The share of true negatives is high, since we condition only on the presence of at least one of the eight character-roles listed above in a tweet.

Figure F.3: *Comparing Performance across Alternatives Conditional on Keyword Search*

(a) *Percentage of True Positives and True Negatives* **(b)** *Percentage of False Positives and False Negatives*

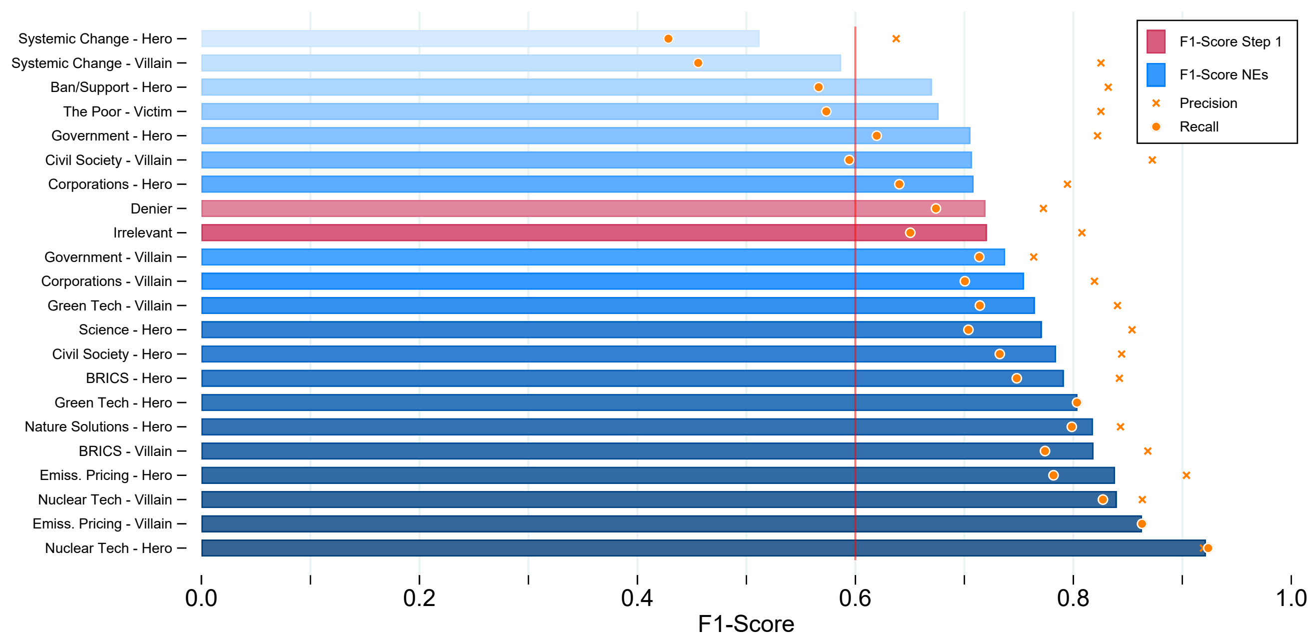


(c) *Table*

	Percentage of true positives	Percentage of true negatives	Percentage of false positives	Percentage of false negatives
Character-Role Framework	11.9	86.8	1.3	1.5
Hiv4 Sentiment Analysis	8	82.6	5.4	5.4
Nltk Sentiment Analysis	7.5	83.4	4.7	6

Notes: The figure shows results identical to what is shown in Figure F.2, but using a different set of tweets. In particular, in Figure F.2 we provide results based on the full set of the *ground truth* tweets. Here, we compute the performance indicators only considering those tweets where the keywords correctly identify the presence of a character. This makes the comparison more favorable for the dictionaries approaches, yet our model still outperforms them by a lot.

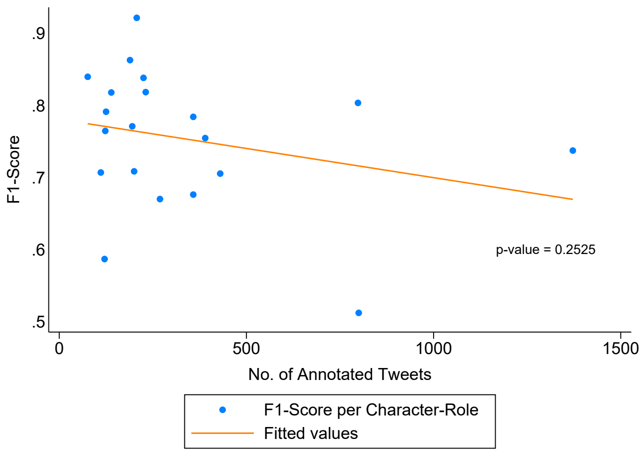
Figure F.4: *Model Performance by Character-Role and Irrelevant Tweets*



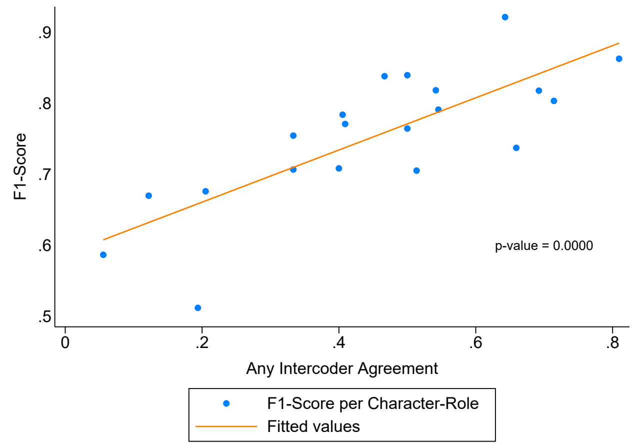
Notes: The figure shows the performance of our prediction model for all character-roles and the categories *Denier* and *Irrelevant* from level 0 of the prediction process. The bars indicate the F1-Score, the harmonic mean of two indicators, precision and recall. Precision is defined as $p = \frac{tp}{tp+fp}$, where tp stands for true positive while fp for false positive. Recall is defined as $r = \frac{tp}{tp+fn}$, where tp stands for true positive while fn stands for false negative. Figure 3 shows the main performance results.

Figure F.5: *Factors of Model Performance*

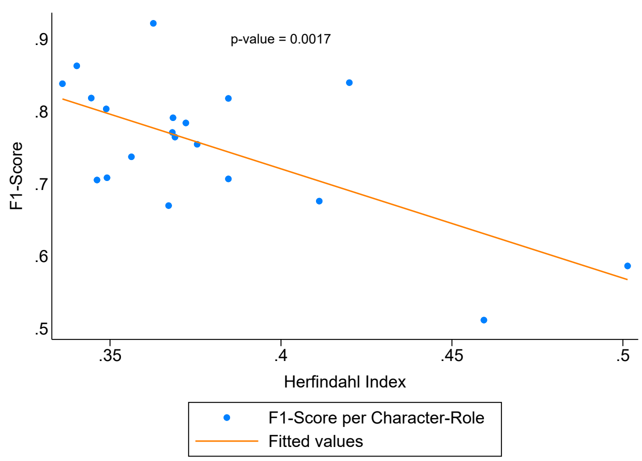
(a) *F1-Score and Frequency of Coding*



(b) *F1-Score and Quality of Coding*



(c) *F1-Score and Concentration of Coders*



Notes: Panel (a) shows a scatterplot of the number of annotated tweets in the training set per character-role against the F1-Score obtained training the model. Panel (b) plots a measure of agreement among the human coders against the F1-Score. The measure of agreement is computed as follows: We use a subset of annotated data that was annotated by all the three main external coders. We define agreement for each character-role as the number of tweets on which at least two coders agree divided by the total number of times the character-role appeared among these shared tweets. We do not include this subset of tweets in the training dataset. Panel (c) shows a scatterplot of the Herfindahl Index per character-role against the F1-Score. The index is a measure of concentration of coders. The higher the Herfindahl Index the fewer coders annotated a specific character-role. In Section 6.2 we analyze the factors affecting the performance of the model.

G The Drama Triangle in Other Disciplines

The analysis of story-telling and in particular the focus on the three main roles of villain, hero and victim has been used for long time in many disciplines. In the following paragraphs we provide an outline of noteworthy examples coming from different research areas.

Public Policy & Political Science:

The political science literature that integrates narratives has grown vastly in recent years. From analyzing stories on human trafficking (O'Brien, 2018), over narratives about gun control (Merry, 2016), to the analysis of hero, villain and victims in the context of US-China trade (Jiangli, 2020), narratives have become a well established tool in the discipline. Narratives play nowadays an important role in the discipline, enough to be integrated as an essential tool in university curricula studying Politics (McBeth and Pearsall, 2021).

The pioneering work by Jones and McBeth (2010) represents a first implementation of a more systematic way to explore narrative. The authors introduce the Narrative Policy Framework that consists in a guideline to identify and analyze narratives. They argue narratives are defined by a context, a plot, their characters and their morale. In particular, the characters involved in the narratives are always alternatively "fixers of the problem (heroes)", "causers of the problem (villain)" or "victims".

The Narrative Policy Framework represents a novelty as it is a falsifiable method and presents a clear set of rules. Nevertheless, the identification of characters in the three main roles of villain, hero and victim was largely in use also before in the discipline. Verweij *et al.* (2006) identify in villains and heroes the core of stories used to frame and explain climate change. These roles are understood as the basic constructs of public stories. Terry (1997) argue that the roles of hero, villain and victim are essential to understand the communication and political strategy of the Reagan administration. The author argues that only through the so called theater metaphor - identifying villain, hero and victim as basic narrative structure - can one understand the anti-government sentiment that built during those years.

Psychology:

In psychology, the first well-known use of the roles triad goes as far back as the 1960's, with the work by Karpman (1968): "Only three roles are necessary in drama analysis to depict the emotional reversals that are drama. These action roles [...] are the Persecutor, Rescuer, and Victim". The author argues that drama starts whenever these roles are present and characters start to switch from one role to another. The victim often reacts with anger and becomes a villain. The villain might be just an outcast that once integrated can become the unexpected hero. The drama triangle as defined by the author has occupied ever since a relevant role in the study of relationships, interactions and in facilitation counseling.

The success of the drama triangle is mainly due to its intuitive design, deeply rooted in social dynamics of almost any context. The drama triangle is "a design program in our heads that enables us to maintain control of relationships, be rational, avoid expressing negative feelings, and be right (win)." (McMahon, 2005, pp. 425). The triangle represents a powerful sense-making tool, that helps analyzing conflict and relations from religion and violence Ganzevoort (2017) to education theories (Kruse and Kruse, 1994). Although many psychologists propose tools to improve interpersonal relationships beyond the - sometimes unhealthy - dynamic of the drama triangle (Shmelev, 2015), all recognize the power of this simple yet comprehensive characterization.

Communication:
Narratives can be intended as both tools to understand / make sense of reality and instruments to build a communication strategy (Jones and McBeth, 2010). The systematic study of storytelling and narratives has not always been present in scholarly work in the communication discipline. Nevertheless, many recognize now the relevance of such approach (Crow and Lawlor, 2016) because of the vast use of narrative structure in many communication tools. In the words of (Gomez-Zara *et al.*, 2018, pp. 311): “Narrative frames are widespread throughout media, be it films, literature, or news.”.

The use of the three archetypal roles of villain, hero and victim is part of the media’s communication strategy in a wide range of different contexts. Brittain (2006) investigate the role of US media in using white femininity stereotypes in justifying the invasion of Iraq. The authors argue the invader is depicted as an heroic liberating force, freeing the innocent victim - indigenous women - by the brutality of the villain men. Anker (2005) investigates the use of melodramatic communication in the elaboration of the collecting trauma of 9/11. In particular, the author claims the three roles of villain, hero and victim represented the ground for almost the entirety of media and political communication about the event. The use of these categorization in the newspapers’ communication is so common that it could be detected in virtually any piece of news about political, economic and social events (Gomez-Zara *et al.*, 2018).

Literature:
One might erroneously think that the study of narratives and literary research are ultimately synonyms of a single discipline or approach. Nevertheless the systematic study of narratives and not of a ’specific’ narrative was introduced in the last century. In the words of (Hyvärinen, 2010, pp. 73): “The abstract, theoretically rich, flexible, and thus quickly moving concept of narrative was a new invention even in literature in the 1960s. [...] The conceptual network was changed, and ‘narrative’ received a higher position in the hierarchy.”

One of the most interesting insights coming from literary studies is the relevance of the characters in the definition of a story. Many authors argue that in a narrative “character is ultimately more important than plot” (Puckett, 2016, pp. 112). In particular a large strand of the literature identifies in the villain, hero, victim archetypal characters, the core of stories. For example, Fog *et al.* (2010) analyze the main factors defining a story, part of which are the characters involved. They find that all stories features a main protagonist that is generally the hero. This character is often antagonized by an adversary, or a villain. The action of the main character are usually helpful for a beneficiary, usually an innocent victim.