doi:xx.xxxx/xxxxx

Sentiment in Statecraft: A Natural Language Processing Analysis of China's Digital Diplomacy

HANNAH BAILEY *University of Oxford*

The Chinese Communist Party (CCP) boasts a vast online presence, with multiple CCP-controlled media outlets among the top fifty most followed Facebook pages. This article offers the first large-scale analysis of CCP narrative strategies and audience responses, revealing that the CCP's approach mirrors electoral campaigns by emphasizing China's positive aspects and critiquing global rivals. Employing machine learning, I analyze nearly 800,000 Facebook posts from CCP-controlled outlets between 2011-2022, introducing a method for quantifying relative topic-level sentiment in text data using topic modeling with BERT word embeddings. Contrary to the conventional wisdom that negative campaign content attracts more attention, this research finds that posts praising CCP's domestic affairs see higher engagement, while more negative posts on international politics also garner significant interaction. This domestic-international distinction underscores effectiveness of the CCP's tailored use of sentiment in online campaigns, and advancing our understanding of online campaigning as a statecraft tool.

Word Count: 10,398

Hannah Bailey is PhD Candidate, Oxford Internet Institute, University of Oxford (hannah.bailey@oii.ox.ac.uk)

For helpful comments and suggestions I thank Philip Howard, Tom Nicholls, Michael Jensen, and Marc Rigter. I would also like to thank audiences at APSA, at Oxford University, and at the ICA conference.

This is a manuscript submitted for review.

INTRODUCTION

Controlling the narrative is a commonplace mantra. But its significance is magnified with the use of online discourse, particularly when authoritarian states are the agents attempting to influence international narratives. How can we pinpoint the underlying agenda of these states, identify their methods of persuasion, and evaluate their success in engaging audiences? This article examines the Chinese Communist Party's (CCP) efforts to shape international online discourse and evaluates audience responses. CCP-controlled media outlets operate social media accounts on international platforms to increase the People's Republic of China's (PRC) discourse power to effectively "tell China's stories well" (*jiang hao zhongguo gushi*) (Xinhua 2013; Huang and Wang 2019; Tambe and Friedman 2022; Cook 2020). With five CCP-controlled media accounts ranked among the top fifty most followed on Facebook (Social Blade 2023), their influence is considerable. Since Facebook and other international social media platforms like Twitter are blocked in the PRC, and given that the content produced by these CCP-controlled media accounts is predominantly in English, it is reasonable to assume that these accounts primarily target international users. These Party-controlled outlets disseminate various controversial and misleading messages toward international audiences, such as conspiracy theories about COVID-19's origin (Bright et al. 2020), narratives discrediting international reporting on Xinjiang's genocide (U.S. Department of State 2022), and claims that the US is deliberately prolonging the Russia-Ukraine conflict (Bailey 2023).

Building on the understanding that the CCP seeks to influence international narratives, this article reveals a dual-tracked agenda: these outlets promote positive narratives regarding the PRC's domestic affairs while simultaneously pushing negative narratives about international politics. Within both narratives, sentiment is employed as a persuasive tool. Sentiment, in this context, is defined as the emotional tone accompanying a campaign's messages, subtly shaping public opinion and enhancing the narrative's impact. The use of sentiment as a tool is a consistent strategy observed throughout the twelve years of CCP-controlled media activity on Facebook covered in this study, spanning a wide range of topics. Previous research on the CCP's attempts to shape international discourse has often been limited to specific topics, events, audiences, or moments in time (Min and Luqiu 2021; Lams 2018), typically involving small-scale data analysis or qualitative assessments (Hagström and Nordin 2020;

Nordin 2016; Callahan 2015; Edney 2014). This article's comprehensive, large-scale analysis builds on this existing work, exemplifying the CCP's exercise of statecraft through campaigning and enabling a deeper understanding of its broader diplomatic objectives. Insights gained from identifying the CCP's narratives and sentiment strategies at scale provide a framework for understanding nation-state objectives and strategies when influencing international audiences.

This article contends that digital public diplomacy campaigns, such as those conducted by the CCP, can be viewed as a form of statecraft through permanent campaigning, resembling political campaigns in democratic settings (Blumenthal 1982; Larsson 2016; Ornstein 2000). Both digital public diplomacy and democratic political campaigns aim to sway public opinion and capture attention, using comparable communication resources to achieve their respective political objectives. In this context, this article investigates whether digital public diplomacy campaigns use narrative and sentiment as persuasive tools in line with campaign literature predictions, or whether they display distinct characteristics. The campaigning literature notes that narrative and sentiment usage depends on the political actor's position. Incumbents typically adopt positive messages, highlighting policy improvements and benefits under their governance, while challengers often use negative messages to criticise the incumbent's actions and the overall state of affairs (Crabtree et al. 2020; Rheault et al. 2016). This article investigates the extent to which the CCP's approach aligns with this logic.

The key finding is that CCP-controlled media outlets employ both incumbent and challenger strategies by positively addressing domestic affairs under its governance and negatively critiquing international politics, challenging the global political incumbents. Drawing on insights from campaign literature, this article examines how current political communications theories apply or diverge in the context of social media, which has broadened the range and variety of political actors conducting campaigns (Gilardi et al. 2022).

In this study I also explore online audience responses to the CCP's narrative and sentiment strategies. Although research on election campaigns indicates that negative messages often capture more attention and online engagement (Marcus et al. 2000; Lau 1982; Mutz and Reeves 2005), this study finds a different pattern. Audiences engage more with CCP domestic affairs posts which feature positive sentiment and with international politics posts which contain negative sentiment. This finding emphasises the apparent effectiveness of the dual-tracked approach. I posit that these findings reflect differences in messaging content, actors, and targeted audiences in digital public diplomacy, which may shape audience responses.

The methodology employed in this research presents a novel minimally-supervised approach to quantifying topic-level sentiment in large-scale text data, building on existing computational sentiment detection techniques (Rice and Zorn 2021; Osnabrügge et al. 2021; Weiss et al. 2022; Jerzak et al. 2023). I use machine learning to fine-tune pre-trained Bidirectional Encoder Representations from Transformers (BERT) embeddings (Devlin et al. 2019) on a corpus of 797,793 Facebook posts from CCP-controlled media outlets between 2011 and 2022. By combining these embeddings with a topic model, I examine the relationship between various topics and sentiment. This minimally supervised method can be applied to any large body of text and, unlike traditional dictionary-based sentiment analysis, incorporates the entire text when calculating document-level sentiment. This method captures *relative sentiment*, which enables a comparison of sentiment usage across topics and time periods. I use this approach to identify the narratives and sentiment strategies employed by CCP-controlled media in their international digital public diplomacy, as well as the corresponding reactions of international audiences to each of these narratives and sentiments.

DIGITAL PUBLIC DIPLOMACY CAMPAIGNS

The proliferation of social media has allowed political actors to bypass traditional media gatekeepers, communicate directly with their audiences, and strategically exercise greater influence over online discourse to achieve political goals (Jungherr et al. 2020; Persily and Tucker 2020; Lu and Pan 2021). This shift has triggered a surge in the diversity and number of actors aiming to sway public opinion (Bradshaw et al. 2021).

Nation-state governments have emerged as key actors, adeptly using social media to shape international public opinion and advance their diplomatic objectives (Kampf et al. 2015; Cowan and Arsenault 2008). This strategy, referred to more broadly as public diplomacy, involves shaping foreign citizens' attitudes to indirectly sway the actions of their governments (Malone 1985, 1988; Gilboa 1998, 2001; Tuch 1990). Public diplomacy can be used to achieve a variety of political goals, including advocating for particular policies or interests, facilitating information exchange, reducing prejudices, promoting soft power resources and constructing alliances (Cull 2008). Within this broad range of objectives, the rapid and volatile nature of online discourse necessitates an approach akin to permanent campaigning. Just as political campaigns in a democratic system often operate continuously to influence public sentiment, digital public diplomacy campaigns also frequently maintain an ongoing presence to effectively manage and direct international narratives (Bradshaw and Howard 2018).

Within public diplomacy, the emergence of social media has led to the rise of digital public diplomacy as a prominent new form of engagement. This approach involves nation-states strategically using digital platforms to shape international public attitudes and thereby promote their diplomatic goals (Bjola et al. 2019, 2020; Cull 2011). Such digital platforms allow governments to reach foreign audiences directly and more effectively (Entman 2008; Golan and Viatchaninova 2014). Digital public diplomacy involves the use of a variety of tools, such as state-backed or government-aligned media outlets (Bright et al. 2020; Rebello et al. 2020; Bradshaw et al. 2022), social media accounts operated by diplomatic spokespeople (Huang 2022), and paid online advertisements (Tambe and Friedman 2022). More subversively, some operations may employ inauthentic social media accounts (Bailey and Howard 2022; Schliebs et al. 2021a,b). Together, these approaches allow governments to directly influence international audiences.

I contend that digital public diplomacy is best understood as a form of political campaigning, grounded in the shared aim of influencing public attitudes to achieve political objectives. Much like political campaigns in democratic settings, digital public diplomacy campaigns can either direct their efforts toward immediate issues (Salvo and Soula 2017), such as COVID-19 (Bright et al. 2020; Rebello et al. 2020), or aim at broader, long-term objectives (Kragh and Åsberg 2017). Similarly, democratic political campaigns may focus on an impending election, or function as an ongoing effort, a strategy often referred to as permanent campaigning (Blumenthal 1982; Larsson 2016; Ornstein 2000).

Operating within the same media-saturated environment, both digital public diplomacy and political campaigns in a democratic context increasingly rely on online platforms to reach audiences (Jungherr 2023; Bradshaw et al. 2021). The political status of the actor conducting the campaign, either as a political insider or outsider, is also important in both contexts. In a democratic campaign, this likely corresponds with whether the actor is an incumbent or challenger. Meanwhile, in digital public diplomacy, status may be determined by the actor's position as an ally or adversary. Another shared characteristic lies in the substantial budgets and communication resources used by both digital public diplomacy campaigns and political campaigns within democratic settings (Chester and Montgomery 2017; Bradshaw and Howard 2019). In light of these shared goals, environments, targeted audiences, and tools, this article argues that digital public diplomacy can be effectively studied as a form of political campaigning.

To understand digital public diplomacy as a form of political campaigning one must explore the strategies that underpin both narratives and sentiments. This article adopts this approach by investigating: (1) the narratives employed in digital public diplomacy campaigns, and (2) the sentiment used to convey these narratives. The goal is to determine the extent to which the narratives and the use of sentiment within digital public diplomacy align with the current literature on campaigning, which studies campaigns predominantly in democratic contexts.

Narrative and Sentiment in Political Campaigns

Political campaigns are designed to influence and persuade audiences. Central to achieving this goal are two elements of a political message. The first is the narrative content, which contains the key themes and messages the campaign wants to convey. The second is the sentiment that accompanies and enhances the narrative (Osnabrügge et al. 2021; Jung 2020). These two campaign elements are often strategically tailored to appeal to audiences, and contingent on the political position and objectives of the campaign actor. This section explores each of these components in detail, and assesses how the insights drawn from the campaigning literature might apply to the context of digital public diplomacy campaigns.

The choice of narrative content in a political campaign is pivotal as it steers the audience's attention toward the preferred issues of the campaign and away from other issues. This narrative selection typically depends on the political standing of the campaign actor. The political campaigning literature, which has mainly focused on democratic contexts, often defines the political standing of a campaigning political actor as either an incumbent or a challenger. Incumbent political actors tend to emphasize their own track record, underlining the incumbent's accomplishments and competence in governance (Müller 2022; Petrocik 1996; Geer and Vavreck 2014, p. 219). This approach aims to reinforce their image of reliability and success, bolstering their claim for continued tenure. Conversely, challengers often orient their narrative toward the incumbent, aiming to critique, undercut, or expose flaws in the incumbent's track record (Dolezal et al. 2018; Petrocik 1996; Mayer 1996, p. 451). Absent a governance track record, challengers tend to focus on the incumbent's performance, laying out reasons for a change in leadership. This strategy is particularly prevalent among politicians or candidates trailing in polls, who are more inclined to position themselves as the underdog relative to the incumbent party (Auter and Fine 2016; Gross and Johnson 2016; Nai 2020).

Alongside the narrative content, sentiment, which is defined as the emotional tone accompanying a campaign's messages, plays a critical role in political campaigns. Sentiment is often used to elicit specific emotional responses from audiences and shape public opinion, thereby enhancing the impact of the narrative (Osnabrügge et al. 2021; Kosmidis et al. 2019; Rheault et al. 2016; Ridout and Searles 2011). Negative messaging, especially messages that evoke emotions such as anger, fear, and anxiety, is particularly widespread in modern political campaigns (Lau et al. 1999; Fridkin and Kenney 2011; Nai 2021; Ansolabehere and Iyengar 1995). In the context of a political campaign, the interplay between emotions and cognitive processing can enhance the persuasiveness of a particular message (Petty and Briñol 2015; Chaiken 1987; Sinclair et al. 1994).

As with narrative choices in political campaigns, the strategic use of sentiment often aligns with the political standing of the campaign actor. Incumbents, with a governance record to showcase, generally employ positive sentiment to highlight their achievements while in office (Nai 2020; Crabtree et al. 2020; Müller 2022). Conversely, challengers frequently use negative sentiment to criticise the status quo and place blame on the incumbent party (Müller 2022; Auter and Fine 2016; Gross and Johnson 2016; Nai 2020). This approach is also often adopted by smaller, extremist, or populist parties, who align themselves as political outsiders and echo public grievances (Widmann 2021; Crabtree et al. 2020).

In short, the existing campaigning literature indicates that in democratic contexts, the choice of narrative and sentiment in political campaigns aligns with the campaign actor's political position.

That is, incumbents often use positive sentiment to highlight achievements, while challengers employ negative sentiment to critique incumbents.

This article adapts this understanding of narrative and sentiment in political campaigns to digital public diplomacy campaigns. However, the roles of incumbent and challenger do not fit neatly into the context of digital public diplomacy conducted by nation-state actors. This prompts the question: how do these roles manifest in the context of digital public diplomacy?

I theorise that nation-state governments, when conducting digital public diplomacy campaigns, can adopt the roles of incumbent or challenger, or both. An incumbent strategy, similar to that of an incumbent in a democratic election, would involve the nation-state government focusing its narrative on domestic affairs under its governance. This might include highlighting its domestic tourism, culture or economic accomplishments, accompanied by positive sentiment to underline the positive aspects of the country under its governance. The aim of this approach could be to promote the nation-state's soft power and establish alliances.

In contrast, a challenger strategy in digital public diplomacy may involve the nation-state government focusing its narrative on the state or political system it seeks to challenge. This could be a rival nation-state, a group of states, or the wider international political system, depending on the nation-state's objectives. In this role, the nation-state would aim to upset or challenge the status quo, much like a challenger in a democratic campaign. This narrative would likely be accompanied by negative sentiment, as the challenger nation-state critiques the status quo and places blame on the political entity it seeks to challenge.

I argue that the incumbent-challenger roles within a democratic campaign can be adopted by nation-states within digital public diplomacy campaigns. Moreover, by examining the narratives and sentiment used by nation-states in digital public diplomacy campaigns, we can identify which roles they are adopting and the political objectives they are aiming to achieve. Are they seeking alliances, or are they seeking to challenge the status quo? Analysing the narratives and sentiment in a digital public diplomacy campaign provides a framework for understanding the broader statecraft objectives and strategies of the nation-state conducting the campaign.

Audience Response

In addition to examining the narratives and sentiment strategies employed in digital public diplomacy campaigns, it is equally important to understand how online audiences respond to these messages. Audience's responses to a political campaign can serve as an indicator of the campaign's ability to engage its target audience.

Negative messaging often dominates political campaigns, particularly online (Klinger et al. 2022; Bobba 2019). This negativity captures audience attention and encourages engagement, leading to more shares and likes (Fridkin and Kenney 2011; Lau 1982; Mutz and Reeves 2005). Affective intelligence theory provides a possible explanation for this phenomenon, suggesting that fear-inducing messaging can motivate individuals to seek new information and challenge existing beliefs, whereas positive and enthusiastic messaging merely reinforces an individual's commitment to their prior beliefs (Marcus and MacKuen 1993; Brader 2005, 2020).

I extend this premise to digital public diplomacy campaigns, investigating whether audiences respond similarly to negative messages in these types of campaigns. The key question is, do audiences engage more with negative digital public diplomacy messages than positive ones? Or does the unique context of digital public diplomacy elicit different audience responses? Understanding how audiences engage with these campaigns provides valuable insights into the effectiveness of different narrative and sentiment strategies.

CHINA'S DIGITAL PUBLIC DIPLOMACY

To investigate the dynamics of digital public diplomacy campaigns and audience responses, I examine the CCP's digital public diplomacy efforts. The period for this study spans from 2011 to 2022, a time during which the CCP significantly expanded its online public diplomacy capacities and amassed a large international audience for its online messaging (Walker et al. 2021; Martin 2021; Chang 2021).

CCP-controlled media outlets have substantial followings on international social media platforms such as Facebook and Twitter (Olesen 2015; Tambe and Friedman 2022; The Economist 2019). Since these platforms are blocked within the PRC, they serve primarily as a channel for the CCP to reach and influence international audiences. The scale, funding, and reach of these campaigns signal their potential for wide impact, however, the motivations and objectives underpinning these efforts remain unclear.

Under Xi Jinping's leadership, the CCP's public diplomacy efforts have gained momentum. Xi's 2016 call for a "flagship media with strong international influence" (Central People's Government of the People's Republic of China 2016), coupled with the guiding principle "tell China stories well" (*jiang hao zhongguo gushi*) further reflects this focus (Xinhua 2013; Huang and Wang 2019; Tambe and Friedman 2022; Cook 2020). This commitment is also evident in the CCP's substantial public diplomacy investment. In 2020, the CCP's estimated annual expenditure on public diplomacy was around $8 billion, or four times that of the US (Walker et al. 2021; Martin 2021).

Part of this hefty investment has been directed toward expanding the CCP's global media reach through Party-controlled outlets. These include *China Daily*, *Xinhua News Agency*, *China Global Television Network (CGTN)*, *China Daily Group*, and *Global Times* (Brady 2009). In a notable example, *China Central Television*'s international arm underwent a rebranding to *CGTN* in 2016. Now broadcasting in five languages from three countries, this move has significantly extended its global reach (Reporters Without Borders 2019; Hamilton and Ohlberg 2020, p. 168). Adding to this media landscape, the *Global Times*, established in 2009, plays a distinct role by amplifying the Party's more assertive messages (Hamilton and Ohlberg 2020 p. 168). All these outlets are supervised by the State Council Information Office (Shambaugh 2007; Creemers 2015), and have successfully garnered substantial followings on international social media platforms. For instance, on Facebook, five CCP-controlled outlets are among the top fifty most 'liked' accounts (Social Blade 2023).

Here, I explore the narratives and sentiment strategies employed by CCP-controlled media outlets on international social media platforms. The key question is whether these outlets, in conducting digital public diplomacy campaigns, use narratives and sentiment in a manner similar to political campaigns in democratic settings. Specifically, do they adopt the roles of incumbent or challenger? Given what we know about the CCP's diplomatic objectives, we might expect these outlets to employ both strategies. As an incumbent party within its nation-state, the CCP might aim to positively promote its national achievements toward international audiences and showcase its governance capabilities in a bid to curry favour and seek alliances. As the ruling party of a nation-state that has only recently risen to world-power status, the CCP may also take on the role of a challenger within the international political system. In doing so, the CCP might negatively critique and seek to undermine the status-quo international political system within its public diplomacy campaigns.

The literature thus far indicates that the CCP is likely engaging in both incumbent and challenger strategies in its public diplomacy campaigns. One perspective argues that the CCP uses positive messaging to promote harmony, enhance the PRC's soft power, and improve relations with other nation-states (Hagström and Nordin 2020; Mingjiang 2008; Nordin 2016; Heng 2010). This was evident during instances such as Russia's initial invasion of Ukraine and the COVID-19 outbreak, where CCP-controlled outlets projected peace-promoting CCP spokespeople and discussed aid efforts from the PRC, respectively (Bailey 2023; Bright et al. 2020). From this perspective, the CCP's digital public diplomacy seeks to construct a positive image of the nation and government, thereby exercising soft power and conforming with its tradition of harmony (Brady 2015). This would align with an incumbent strategy.

The second perspective is that the CCP adopts negative, nationalist sentiment to challenge the Western-led world order. This strategy involves asserting the CCP's goals within this system, often through criticism of Western nation-states, while simultaneously reinforcing domestic nationalist narratives. Domestic nationalism is, in part, rooted in resentment toward the Western-led international political system, a sentiment stemming from the so-called "Century of Humiliation" (Gries 2004, p. 47, Weiss 2014; Wang 2004; Zhao 2004). Anti-US rhetoric is a common theme in these narratives (Brady 2006; Min and Luqiu 2021). The CCP perceives itself as facing a hostile foreign enemy with greater influence within the international political system (Lams 2018; Brady, 2009, p. 151). Its foreign propaganda activities are designed to combat this perceived enemy and garner international support, while undermining the domestic foundations of support within Western nation-states themselves (Callahan 2015, p. 219-220). Certain instances of this challenger strategy are evident in the CCP's digital public diplomacy. For example, CCP-controlled outlets blamed Western nations for perpetuating the Ukraine conflict (Bailey 2023) and criticised democratic states' COVID-19 responses, questioning their leadership abilities (Bright et al. 2020).

Empirical evidence for these two perspectives is thin. It remains unclear which of these two strategies is prioritized, whether either or both of these strategies have changed over time and under Xi Jinping's leadership, and importantly, whether the these two strategies are in conflict or whether they might comprise a single, coordinated public diplomacy strategy. I investigate these uncertainties and pose three key research questions: (1) How does the CCP employ narratives and sentiment in its digital public diplomacy campaigns? (2) Does the CCP's strategy align with those used in political campaigns conducted in democratic settings? (3) How do audiences react to the CCP's strategic use of narratives and sentiment?

By conducting a comprehensive quantitative analysis of the narratives and sentiment used by CCP-controlled media outlets, this article makes two valuable contributions. First, it offers an empirically based analysis of the underlying statecraft objectives pursued by the CCP. Second, it assesses the impact of this public diplomacy campaign by examining the responses of global audiences to these messages.

DATA

To test the strategic use of narratives and sentiment by CCP-controlled media, and the responses of audiences to these narratives, I build a comprehensive dataset of Facebook posts produced by internationally-focused CCP-controlled media outlets between 2011 and 2022. It's important to note that while inauthentic or paid accounts are a known issue on social media, Facebook asserts such accounts constitute a small fraction of total followers of these CCP-controlled media pages (The Economist 2019). Given the opaqueness of Facebook's account authenticity data, this analysis is unable to verify the authenticity of the engaging accounts.

The data begins in 2011, since data on CCP-controlled media prior to this are limited in terms the number of Facebook posts and audience engagements, thus constraining any meaningful natural language processing analysis. This dataset includes Facebook posts from prominent CCP-controlled media outlets that: (a) have an active Facebook presence; (b) target global audiences; (c) do not have a specific subject area focus; and (d) post primarily in English. The media outlets that fulfil these criteria at the time of data gathering in January 2023 are: *China Xinhua News*; *CGTN*; *China Daily*; *Global Times*; *People's Daily*; *CCTV*; *Beijing Review*; and *China.org.cn*. The post text and metadata from the Facebook pages operated by these outlets are extracted using the Facebook CrowdTangle API (CrowdTangle n.d.; Schliebs 2020).

Here I focus on Facebook, as it a platform where CCP-controlled media have amassed particularly large followings. Facebook is also one of the most popular social media platforms for internet users in predominantly English-speaking countries (YouGov 2022; Auxier and Anderson 2021), and is therefore a valuable source of data on the responses of these audiences to CCP narratives.

Table 1 summarises the data on CCP-controlled media Facebook posts. In total, these outlets created 797,793 posts, receiving over 1.9 billion likes, 111 million shares, and 30 million comments. The volume of posts steadily increased over this period, with a notable surge between 2014 and 2016.

However, the pattern of audience engagement, as measured through likes, shares, and comments, did not consistently follow the same trend. It rose between 2014 and 2016, dipped in 2017 and 2018, then peaked in 2020 with over 377 million likes, 17 million shares, and 7 million comments. Despite a record 121,972 posts in 2021, engagement fell sharply and continued to decline in 2022. This suggests that factors other than post frequency, such as perhaps the post content, or Facebook's internal algorithms might influence engagement. Further descriptive information is available in Appendix A.

TABLE 1. Data on Posts by CCP-Controlled Media on Facebook Between 2011-2022

Year	Number of Posts	Number of Likes	Number of Shares	Number of Comments
2011	4,031	12,016	1,236	2,961
2012	2,918	4,625	1,167	560
2013	8,627	427,634	7,635	4,207
2014	14,699	7,921,443	650,901	117,952
2015	59,055	171,767,442	12,348,730	1,578,927
2016	85,568	299,271,313	20,979,221	2,984,914
2017	83,403	278,069,150	17,630,105	3,296,549
2018	91,818	199,156,298	14,501,538	2,707,625
2019	100,929	295,004,389	17,175,072	4,735,162
2020	116,974	377,032,583	17,988,976	7,011,143
2021	121,972	200,842,536	5,614,839	3,800,113
2022	107,799	144,128,405	4,665,474	4,058,806
Total	797,793	1,973,637,834	111,564,894	30,298,919

Note: The values for likes, shares and comments refer to the sum of the likes, shares and comments received by all Facebook posts issued in the specified year.

MEASURING SENTIMENT

My approach to assessing the CCP's use of narrative and sentiment in its digital public diplomacy campaigns requires the measurement of data on these two components in the text produced by CCP-controlled media. I introduce a novel method for quantifying sentiment within topic narratives in large-scale textual data, enabling the evaluation of sentiment usage within topics. The proposed novel approach consists of two steps: (1) building a topic model to identify distinct narratives; and (2) training word embeddings and merging these embeddings with the topic model to measure sentiment within distinct topic narratives. This section provides the rationale for adopting the proposed methodology, examines its relationship to existing literature, and highlights its contribution to narrative and sentiment analysis in large-scale text data.

Two approaches are commonly used to measure sentiment in a corpus of text: machine learning and dictionary methods. Supervised machine learning approaches use text annotated with sentiment information to train a classifier, which then predicts the sentiment of unlabeled text. (e.g., Pang and Lee 2004, Maas et al. 2011, Tang et al. 2014). Dictionary methods use pre-defined lists of positive and negative sentiment words and count their frequency in the corpus (McCarthy and Boonthum-Denecke 2012). Both approaches have limitations. Supervised machine learning is costly and dependent on human annotation. Pre-defined dictionaries may falter outside their development domain due to contextual inaccuracies (Grimmer and Stewart 2013, p. 274). Both methods lack the ability to adapt to specific text contexts, which can then affect the accuracy of the sentiment analysis.

This paper contributes to a growing field of research aimed at reconciling these two approaches by developing minimally supervised methods for detecting sentiment (Widmann and Wich 2022; Gennaro and Ash 2022; Esberg and Siegel 2022; Jerzak et al. 2023). Specifically, I build a sentiment vector tailored to the corpus by using word embeddings trained on CCP-controlled media Facebook posts. This customised vector represents both positive and negative sentiment, and is used to measure changes in sentiment expression in CCP-controlled media over time.

Uniquely, this evolving, customized sentiment vector accurately tracks sentiment over time within the corpus of CCP-controlled media Facebook posts. This novel method not only captures sentiment changes in individual documents but also across broader topics, marking a significant innovation in sentiment analysis.

Sentiment in Word Embeddings

Language representation models convert text meanings into numerical representations for analysis, serving as a powerful tool for various natural language processing tasks (Pennington et al. 2014; Bengio et al. 2003). In these models, words are represented as dense real-value vectors, or word embeddings, produced by a neural network model (Devlin et al. 2019). These embeddings capture semantic and syntactic relationships among words in the corpus by predicting a word's likelihood based on its context. The distances between vectors indicate word or phrase similarity in the corpus context (Collobert and Weston 2008; Mikolov et al. 2013), aligning with the distributional hypothesis that a word's meaning can be inferred from its context (Firth 1957, p. 11). This context-based approach makes embedding models well-suited to sentiment analysis, as the sentiment attached to a word can typically be determined from the context in which it appears.

Political scientists have used word embeddings to capture and quantify latent concepts, such as ideology in parliamentary corpora (Rheault and Cochrane 2020), track concepts such as human rights and equality over time in newspaper corpora (Rodman 2020), and examine how lexical markers of social class shift over time (Kozlowski et al. 2019). In several cases, social scientists use word embeddings to develop minimally supervised methods for the detection of specific concepts, emotions, or sentiment (Rice and Zorn 2021; Rodriguez et al. 2023; Osnabrügge et al. 2021; Choi et al. 2022). The supervisory element of this method involves identifying *seed* words to target a particular point in the vector space of the corpus that represents a desired concept. After identifying the vector for the concept, the researcher can use it to measure the proximity of the concept to any other word or concept in the corpus vector space. This article develops a novel method that builds on the above cited earlier work by combining minimally supervised concept detection with topic modeling to identify sentiment expression at the topic level, and track changes in topic-level sentiment over time.

Generating temporal word embeddings presents a further challenge, as conventional models are trained on the entire corpus and are not inherently sensitive to time. To overcome this, researchers use diachronic word embedding models, which generate separate embeddings for each word in each time period (Kutuzov et al. 2018; Tahmasebi et al. 2021). Typically, this method involves generating independent vector space models for each period and aligning them onto a common vector space (Szymanski 2017). However, this approach can be unreliable if the corpus for each period is small and is dependent on the researcher to determine the ideal positioning of the corpus time periods.

To address the difficulties posed by the conventional method of constructing diachronic embeddings through alignment, I expand on the continuously evolving embedding technique introduced by Horn (2021). The method proposed by Horn (2021) involves training the entire corpus while calculating a weighted running average of contextualized embeddings produced by the transformer model BERT (Devlin et al. 2019). This method enables researchers to examine semantic shifts in the embeddings at preselected snapshot intervals. For the present analysis, this is particularly useful as the size of the corpus of CCP-controlled media posts varies over time, with earlier periods (i.e., between 2011 and 2014) having considerably fewer posts relative to later periods. The continuously evolving embedding approach enables embedding snapshots within these earlier time periods without compromising the reliability of the embeddings.

To implement this approach, I train continuously evolving embeddings computed using a pre-trained BERT model. Within this model, I take embedding snapshots every three months between January 2011 and December 2022. For details on the cleaning, fine-tuning and snapshot processes, see Appendix B.

Corpus Specific Sentiment Vectors

To quantify sentiment expression within the corpus of CCP-controlled media posts, I use these word embeddings to create a corpus-specific dictionary of positive and negative words. This leverages the ability of word embedding models to identify words within the corpus of study that are similar to any chosen word or concept.

Constructing this dictionary requires several steps. First, I use two sets of four unambiguously positive and negative seed words to form the basis of the sentiment dictionary. The positive seed words are $\widehat{\mathcal{S}}_{\text{pos}} = \{\text{“}great\text{”}, \text{“}amazing\text{”}, \text{“}good\text{”}, \text{“}excellent\text{”}\}$, and the negative words are $\widehat{\mathcal{S}}_{\text{neg}} = \{\text{“}bad\text{”}, \text{“}negative\text{”}, \text{“}wrong\text{”}, \text{“}horrible\text{”}\}$. These seed words are selected based on their similarity to those used in sentiment dictionaries for word embedding analysis by Rice and Zorn (2021), and Bellodi (2022). Owing to the corpus's small size in earlier periods, some seed words do not appear until July 2011. As a result, the sentiment analysis begins in July 2011.

Second, I calculate a centroid positive and negative vector at each snapshot period. Let v_w^t indicate the word embedding for word w at snapshot t, and $|\mathcal{A}|$ denote the number of elements in set $\mathcal{A}$. The positive centroid vector, represented as $\hat{v}_{\text{pos}}^t$, is generated by subtracting the average of the embeddings of the negative seed words from the average of the embeddings of the positive seed words:

$$\hat{v}_{\text{pos}}^t = \frac{\sum_{w \in \widehat{\mathcal{S}}_{\text{pos}}} v_w^t}{|\widehat{\mathcal{S}}_{\text{pos}}|} - \frac{\sum_{w \in \widehat{\mathcal{S}}_{\text{neg}}} v_w^t}{|\widehat{\mathcal{S}}_{\text{neg}}|}. \tag{1}$$

The negative centroid vector, represented as $\hat{v}_{\text{neg}}^t$, is the negative of the positive centroid vector:

$$\hat{v}_{\text{neg}}^t = -\hat{v}_{\text{pos}}^t. \tag{2}$$

I denote the cosine similarity between embeddings a and b by $cos(a, b)$. For all words w at each snapshot, I compute the cosine similarity with respect to the seed centroid embedding. For the positive seed centroid, I compute the cosine similarity by:

$$cos(\hat{v}_{\text{pos}}^t, v_w^t) = \frac{\hat{v}_{\text{pos}}^t \cdot v_w^t}{||\hat{v}_{\text{pos}}^t|| \times ||v_w^t||}, \tag{3}$$

where "$a \cdot b$" denotes the dot product between a and b, and $||a||$ denotes the norm of a.

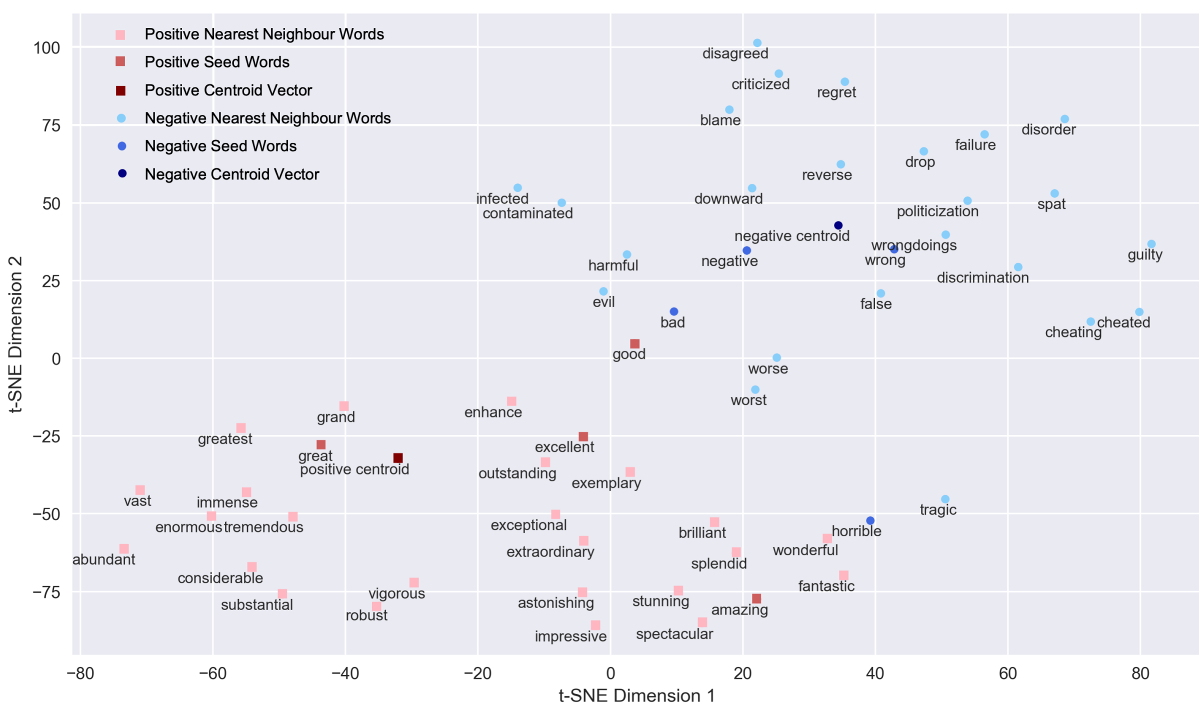
From here, I identify the 28 words with the highest cosine similarity to the positive centroid vector and another 28 to the negative centroid vector, across all snapshot intervals.[1] For additional methodological details and replications using different seed words and number of neighbors, refer to Appendices C and D.

The two final sets of corpus-specific positive and negative sentiment words, each containing 28

[1]Adhering to the convention outlined by Pierrejean and Tanguy (2019), which involves selecting the 25 nearest words to each centroid vector, I apply the same count per snapshot interval. As the nearest 25 neighbour words vary across snapshot intervals, this gives a larger list of words that appear as a nearest neighbour within any snapshot interval. From these, words appearing consistently in the data across all intervals were retained, resulting in two final sets of 28 words.

words, are denoted as $\mathcal{S}_{\text{pos}}$, and $\mathcal{S}_{\text{neg}}$, respectively. Figure 1 employs a t-distributed stochastic neighbour embedding (t-SNE) plot to visualize these words in the vector space at the data's temporal midpoint (2017-01-31). This plot compresses the 768-dimensional vector space into a two-dimensional graph, providing a visual approximation of the word-embeddings' proximity to each other. With negative words on the top-right and positive words on the bottom-left, it reveals distinct sentiment clusters. This clustering is consistent across snapshot periods (see Appendix E).

FIGURE 1. Illustrative t-SNE Plot of Corpus-Specific Sentiment Word Embeddings in the Midpoint Temporal Vector Space (2017-01-31)



Building on the distinct sentiment clusters observed, the final sentiment vector, v_{sent}^t, is computed using $\mathcal{S}_{\text{pos}}$ and $\mathcal{S}_{\text{neg}}$ as follows:

$$v_{\text{sent}}^t = \frac{\sum_{w \in \mathcal{S}_{\text{pos}}} v_w^t}{|\mathcal{S}_{\text{pos}}|} - \frac{\sum_{w \in \mathcal{S}_{\text{neg}}} v_w^t}{|\mathcal{S}_{\text{neg}}|}. \tag{4}$$

This single sentiment vector uses the *corpus-specific* nearest neighbour sentiment words to generate a *single* measure for sentiment, such that word vectors that have a higher cosine similarity relative to the sentiment vector are more positive, and words that have a lower cosine similarity are more negative.

Measuring Sentiment in Documents

Having identified a single corpus-specific sentiment vector, in the following sections I demonstrate how this vector can be used to quantify sentiment: (1) at the document level (i.e., in an individual Facebook post); and (2) at the topic level.

To measure topic-level sentiment, each Facebook post first must be assigned a sentiment score. This score is obtained by computing the average embedding for all the words in each document at the nearest snapshot interval to the publication date of the post. This simple calculation is a common method for calculating document embeddings as it outperforms more complicated methods and accurately summarises the overall document sentiment (Arora et al. 2017; Rossiello et al. 2017).

Here, I let $\mathcal{W}_d$ denote the set of words in document d. The corresponding word embedding model snapshot at the document's Facebook posting time is denoted by t_d. The centroid vector of the words in the document, v_d, is computed as follows:

$$v_d = \frac{\sum_{w \in \mathcal{W}_d} v_w^{t_d}}{|\mathcal{W}_d|}. \tag{5}$$

This calculation assigns each Facebook post a single embedding in the corpus embedding space that summarises the words contained in post.

In order to better understand the sentiment present within each post, I compare the calculated document embeddings with the sentiment vector $v_{\text{sent}}^{t_d}$ from the post's publication snapshot. Let S_d represent the sentiment score for document d. S_d is equal to the cosine similarity between the document centroid and the corresponding snapshot sentiment vector $v_{\text{sent}}^{t_d}$:

$$S_d = cos(v_d, v_{\text{sent}}^{t_d}). \tag{6}$$

Document sentiment scores are normalised between zero and one, where values closer to zero or one indicate more negative or positive sentiment, respectively.[2] Though effective, this approach doesn't

[2]While the theoretical range of sentiment scores is between zero and one, the real-world range within this dataset is constrained, owing to documents being a mixture of positive, negative, and neutral words, rather than being entirely positive or negative. As such, the cosine similarity scores vary over a more condensed range, with a

account for intricate word interplay, like double negatives. Despite limitations, this method, tailored to the specific text, offers higher precision than standard approaches like the dictionary method. The impact of these limitations on results is minimal.

Table 2 lists a selection of Facebook posts produced by CCP-controlled media that have the highest and lowest sentiment scores at various snapshot time periods. The highly negative and positive posts in Table 2 demonstrate the ability of word embeddings to quantify sentiment in text. The most negative posts in Table 2 discuss intuitively negative topics, such as "bloody crimes" and an "imported coronavirus variant case reported in Shanghai". The most positive posts similarly discuss topics associated with positive sentiment, such as a "outstanding natural beauty" and "splendid aerial shots", using positive descriptors. Appendix F provides further examples of negative and positive posts within the corpus at a variety of snapshot time periods, and further context for the posts listed in Table 2.

TABLE 2. Examples of Negative and Positive Posts

Date	Most Negative Posts	Document Sentiment Score
2011-07-31	Bloody crimes prompt debate over savage justice	0.494
2013-01-31	Shenzhen Universiade causes $1.9b loss	0.470
2015-01-31	Court rules against, fines 'gay conversion' clinic	0.477
2017-01-31	Koreans protest against #THAAD deployment	0.491
2019-01-31	Goodbye, Maruko.	0.475
2021-01-31	Imported coronavirus variant case reported in Shanghai	0.488
Date	**Most Positive Posts**	**Document Sentiment Score**
2011-07-31	Xinjiang's outstanding natural beauty gives it enormous tourism potential	0.666
2013-01-31	Beautiful China	0.646
2015-01-31	10 impressive panda moments in 2014	0.652
2017-01-31	Amazing active #volcano	0.681
2019-01-31	Sydney welcomed 2019 in a spectacular style with huge fireworks display	0.664
2021-01-31	Speldid aerial shots in 2020	0.679

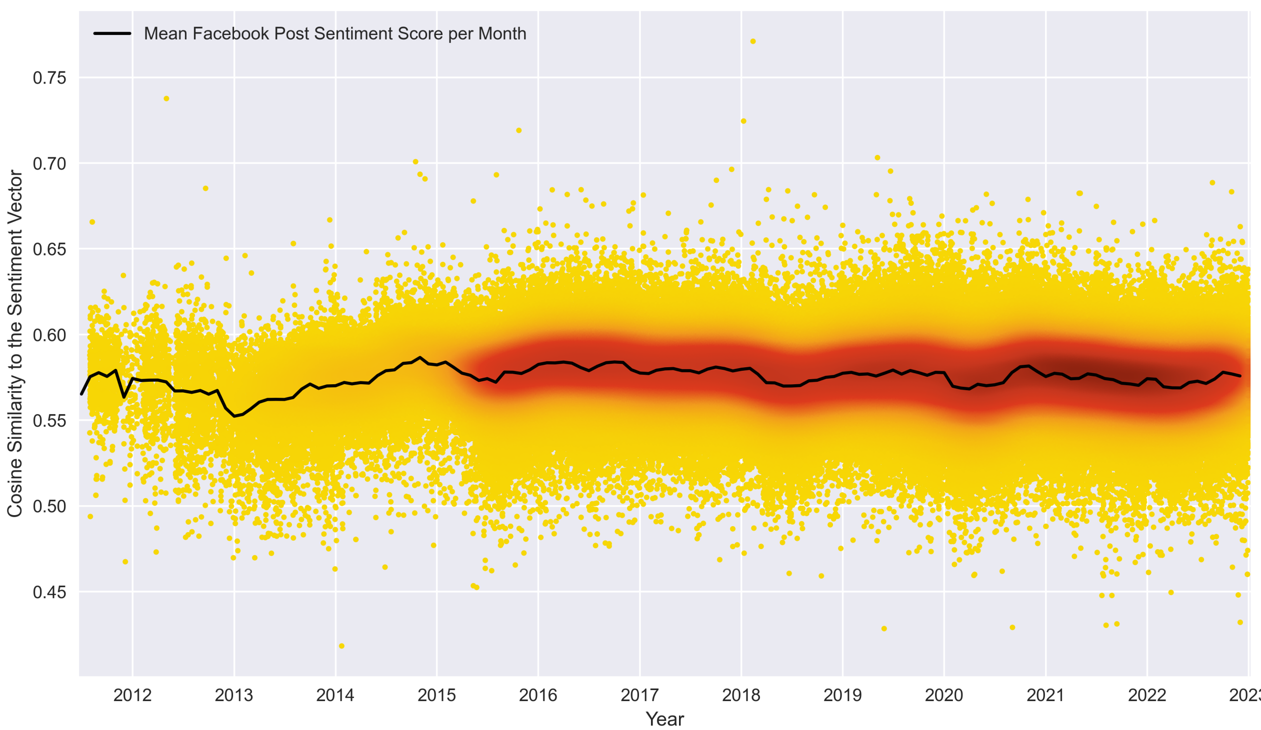
Having established that this method for measuring sentiment corresponds to our intuitive conception of sentiment, Figure 2 shows how sentiment use in Facebook posts published by CCP-controlled outlets varies over time. Each point signifies a post, its colour intensity reflecting scatter point density. Higher values denote a more positive sentiment, while lower ones suggest negativity. A surge in post

minimum value of 0.418 and maximum value of 0.771.

concentration towards later time periods illustrating the increasing volume of posts from CCP-controlled outlets between 2015 and 2022. Although sentiment score fluctuations exist, average sentiment remains fairly steady without a distinct upward or downward trend. A plot of the overall distribution of Facebook post cosine similarity to the sentiment vector can be found in Appendix G.

Figure 2 provides a comprehensive temporal sentiment analysis, but does not explore the topics experiencing these sentiment shifts. The following sections examine the topics discussed within these Facebook posts, their alignment with challenger or incumbent campaign strategies, and pinpoint trends in within-topic sentiment variations.

FIGURE 2. CCP-Controlled Media Facebook Post Sentiment, 2011-2022



Note: Colour intensity signifies scatter point density with red indicating higher and yellow lower. The y-axis measures each document's proximity to the sentiment vector: higher values for positive sentiment, lower for negative. The black line denotes average post sentiment.

Measuring Sentiment in Topics

Sentiment has, thus far, been measured at the document level. This section extends this investigation to explore sentiment within specific topics, linking these to the document-level sentiment scores. This novel approach extends sentiment detection into specific topics, offering insights into how sentiment is used within specific topic narratives in a large body of text. In the context of this analysis, this method offers insights into the use of sentiment within specific narratives in the CCP's digital public diplomacy.

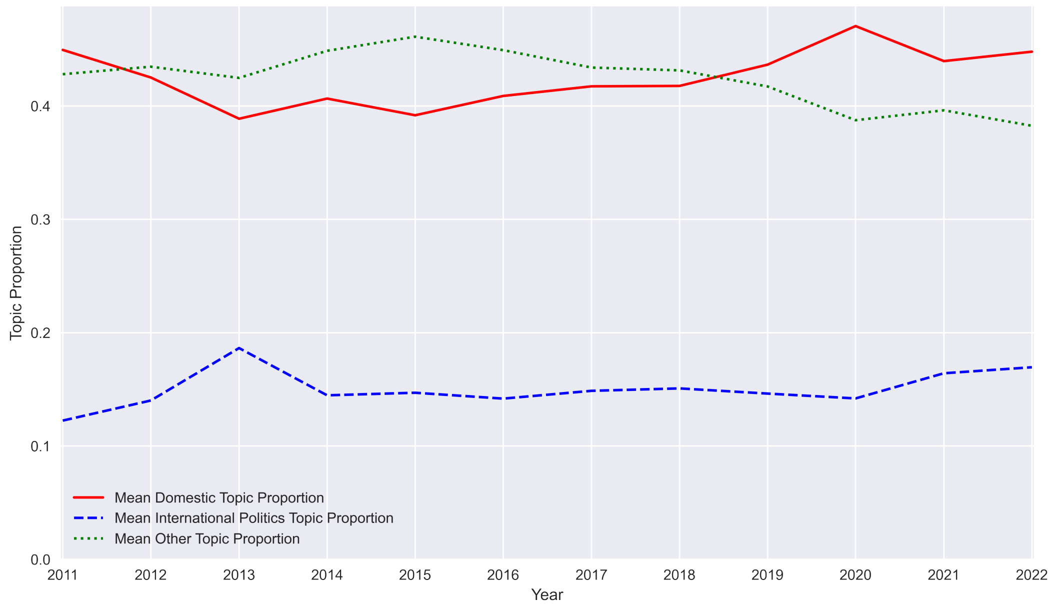
To identify sentiment topic-level sentiment, I combine the document sentiment scores with a Structural Topic Model (STM) with 100 topics of CCP-controlled media Facebook posts (Roberts et al. 2014). STMs are a category of topic modeling that applies unsupervised machine learning to organize textual data into topics, generating semantically coherent themes within the documents (Blei 2012). Additional information on topic model training and topic number selection can be found in Appendix H.

To examine the narrative approach taken within CCP digital public diplomacy messaging, I divide the topics into three categories: domestic, international politics, and miscellaneous. As noted earlier, an incumbent strategy in public diplomacy would be expected to emphasise domestic affairs under the political actor's governance, while a challenger strategy would be expected to target the state or political system it seeks to challenge. For the CCP this is likely the Western-led international political system.

Here, these two strategies are operationalised by categorising topics into domestic and international politics groups. These labels are assigned by manually examining the fifty posts that most closely align with each topic. Topics predominantly discussing PRC events, people or locales are assigned to the domestic group, while those that overwhelmingly focus on international politics or the PRC's global relations are tagged international politics. All remaining topics fall under the miscellaneous "other" category. Complete topic lists, prevalence, and group assignments can be found in Appendix I.

Figure 3 shows the yearly average topic proportions per group. Topics discussing domestic affairs and miscellaneous other topics account for a much larger share of CCP-controlled media posts than international politics topics. Since mid-2018, CCP-controlled outlets have also heightened their focus on domestic topics, posting a higher volume of content on domestic affairs than either other miscellaneous topics or international politics.

FIGURE 3. Mean Yearly Topic Proportions For International Politics Topics, Domestic Topics, and Other Topics



To measure sentiment within topic groups, I compute an annual sentiment score for each topic via a weighted average of documents from the respective year. Let $\mathcal{D}_y$ be the document set for year y, and τ be a topic. The sentiment score for topic τ at year y is S_τ^y. The is computed as follows:

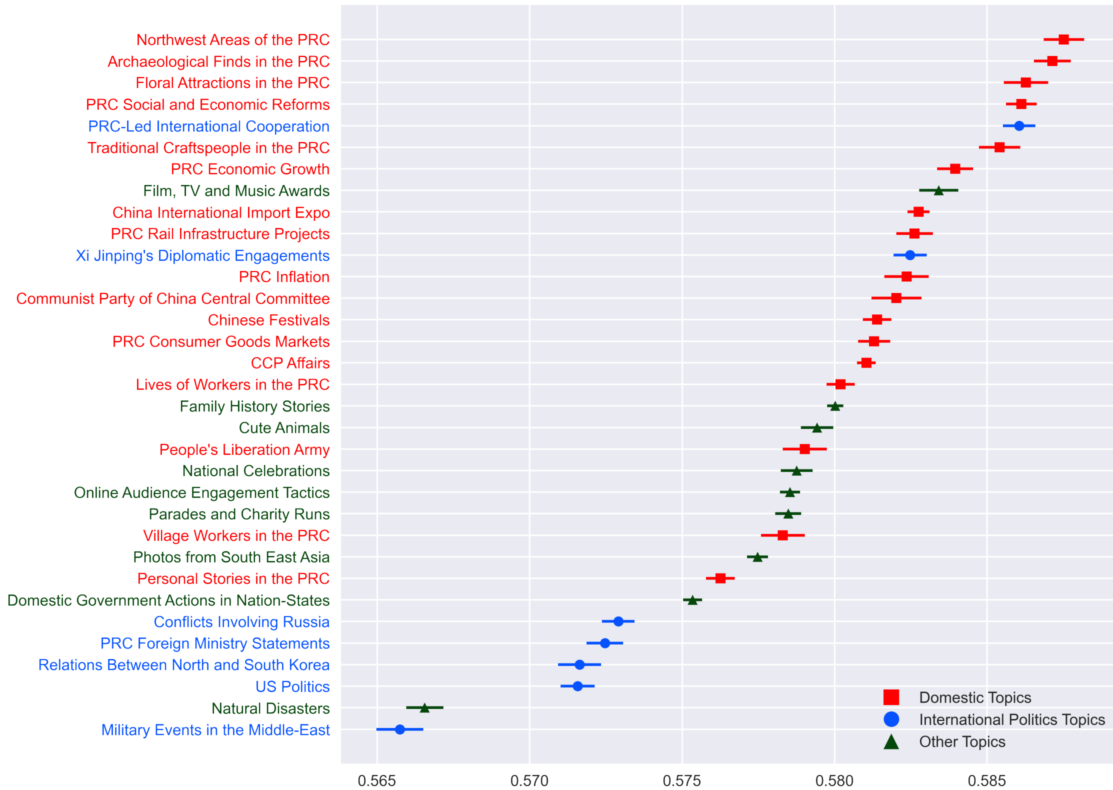
$$S_\tau^y = \frac{\sum_{d \in \mathcal{D}_y} S_d \cdot P(\tau \mid d)}{\sum_{d \in \mathcal{D}_y} P(\tau \mid d)}, \tag{7}$$

where each weighting, $P(\tau \mid d)$, is the probability that document d is assigned to topic τ according to the topic model. To generate 95 percent confidence intervals for each topic-level sentiment score, I apply a bootstrap approach, resampling yearly document populations with replacement, and calculate topic-level sentiment scores. I repeat this process 10,000 times per year.

Figure 4 shows the 2017 sentiment scores for the most prevalent 50 percent of topics, arranged by decreasing sentiment, where again, the larger scores denote relatively more positive sentiment. This year is chosen as it is the temporal midpoint in the data; graphs for earlier and later years reveal similar patterns and can be found in Appendix J.

Figure 4 also illustrates the distinction between the sentiment used to discuss domestic topics versus international politics topics. Five of the six most negatively portrayed topics discuss international politics: "Military Events in the Middle East", "US Politics", "Relations Between North and South

FIGURE 4. Topic-Level Sentiment Scores for the Most Prevalent Topics in 2017



Note: This plot displays the sentiment scores for the top 50 percent of most prevalent topics in 2017, arranged in descending order of sentiment value. Error bars illustrate 95% confidence intervals.

Korea", "PRC Foreign Ministry Statements", and "Conflicts Involving Russia". For context, the topic "PRC Foreign Ministry Statements" often refers to inflammatory or accusatory comments made by the Foreign Ministry, or other diplomatic representatives on international affairs. In contrast, the other prevalent topics mainly address domestic issues or other topics, with only two exceptions: "Xi Jinping's Diplomatic Engagements" and "PRC-Led International Cooperation". Unlike the previously mentioned international topics, these more positive topics describe the PRC's actions in the international sphere.

The remaining topics relate to domestic or miscellaneous issues. Domestic topics cluster at the top of the figure, many displaying highly positive sentiment scores. The most positive 2017 topics include "Northwest Areas of the PRC", "Archaeological Finds in the PRC", "Floral Attractions in the PRC", and "PRC Social and Economic Reforms". Overall, Figure 4 shows that CCP-controlled outlets demonstrate a tendency towards positive language for domestic topics and negative language for topics discussing the actions of other nation-states within the context of international politics.

Figure 5 extends this analysis, using a violin plot to compare sentiment in CCP-controlled outlets across the three topic categories: international politics, domestic topics, and miscellaneous other topics. In this figure, I conduct an independent t-test to compare sentiment scores for domestic topics and international politics for each year from 2011 to 2022.[3] The figure shows the yearly difference in mean sentiment scores ($\Delta\mu$) between these groups, along with corresponding p-values. There is a statistically significant difference in sentiment, with domestic topics consistently conveying more positive sentiment than international politics topics.

FIGURE 5. Topic Sentiment Shifts Between 2011 and 2022

Note: This plot illustrates the difference in sentiment scores for domestic topics, international topics, and other topics. Here, $\Delta\mu$ represents the mean of domestic topics minus international politics topics. The p-values refer to the t-test on the mean sentiment scores between domestic topics and international politics for each year.

The large negative tails in domestic topics from 2020 to 2022 relate to domestic COVID-19 outbreak topics, detailed in Appendix J. In short, Figure 5 indicates a consistent trend: CCP-controlled media maintain significantly more positive discourse on domestic topics than on international politics, a pattern that holds throughout the presence of these outlets on Facebook.

Supplementing Figure 5, Table 3 provides some practical context to the topic sentiment scores by

[3] The miscellaneous other topic group is visually included in the plot, but not in the t-test.

featuring examples of highly negative international politics posts and highly positive domestic posts.

TABLE 3. Examples of Highly Negative International Politics Posts and Highly Positive Domestic Posts

International Politics Posts		
Date	**Post Text**	**Document Sentiment Score**
2020-07-28	Mike #Pompeo: Lies & slanders #WeLiedWeCheatedWeStole	0.462
2016-09-12	#Editorial: Carter wrong to blame China for #NorthKorea nuclear issue	0.478
2013-04-23	Op-ed: Unyielding stance deadlocks Iran-US talks	0.482
2021-10-22	EU, US send wrong signals on Taiwan, risking miscalculation	0.483
2021-11-09	#ChinaDailyCartoon Imaginary enemy	0.486
Domestic Posts		
Date	**Post Text**	**Document Sentiment Score**
2016-08-17	Magnificent Shenxianju Mountain in eastern China's Zhejiang province Shenxianju is a typical rhyolite landform	0.648
2018-09-04	#ChinaTravelGuide Spectacular scenery of Mount Huangshan	0.646
2019-12-19	Eggshell carving is a delicate and exquisite art! #ChinaStory	0.646
2015-02-28	The vast, majestic scenery of #Turpan, #Xinjiang #Uygur Autonomous Region. #UygurCulture #AmazingChina	0.645
2016-09-08	Magnificent view of Kuitun Grand Canyon Photos show the magnificent landscape of Kuitun Grand Canyon at the northern foot of Tianshan Mountain in Kuitun, Yili Prefecture, Xinjiang Uyghur Autonomous Region. #AmazingPlacesinChina	0.643

These findings offer support for theorising that the CCP is employing a dual-track strategy in its digital public diplomacy discourse, depicting the PRC as both an incumbent and a challenger. CCP-controlled outlets typically use more positive language to highlight CCP's governance in domestic affairs, while also using more negative language in international politics discourse to challenge and critique the existing international order. Notwithstanding, the CCP places greater emphasis on domestic affairs, as reflected by the larger volume of content on this topic produced by these outlets.

AUDIENCE ENGAGEMENT WITH SENTIMENT

The focus thus far has been on messages published by CCP-controlled media on Facebook. Yet, what is left to be determined is the extent to which audiences engage with these narratives and their associated sentiments. To investigate this, I use regression analysis to examine the relationship between Facebook post sentiment and online audience engagement. Table 4 summarises the results from four regression models, each with different specifications. Appendix K provides further model details and robustness checks.

Each model incorporates separate dependent variables to quantify the effect of sentiment on: (1) user likes; (2) user comments; (3) post shares; and (4) overall engagement.[4] Exploring varied engagement measures as dependent variables ensures robustness of the results. Owing to overdispersion, I apply a log transformation to each of these dependent variables. The four models assess the relationship between a distinct measure for audience reaction and domestic and international topic proportions, as well as the interaction between these proportions and the document sentiment scores. Model controls include the number of media attachments (e.g., images or videos) in the post, the number of URLs contained in the post, post length (in words), the Facebook account's subscriber count for the CCP-controlled media outlet (scaled in units of 100,000 for interpretability), and a linear time trend. Full model output is provided in Appendix K. Standardised coefficients are presented in the model results, enabling a clear comparison of the effect magnitudes across different explanatory variables. Figure 6 visualises these standardised coefficients.

Within each of the models, I focus on the interaction terms between domestic and international politics topic proportions and sentiment, as these variables best capture the combined impact of narrative and sentiment on audience engagement. The variables "Domestic Topic×Sentiment" and "Intl. Politics Topic×Sentiment" respectively represent these interactions.

These interaction terms highlight a key nuance, that is, that the influence of topic on audience engagement is modulated by the sentiment within the post. In all four models, positive sentiment regarding domestic topics leads to higher audience engagement, while positive sentiment on international politics tends to decrease engagement.

Simply put, CCP-controlled media outlets exhibit greater online engagement when they discuss

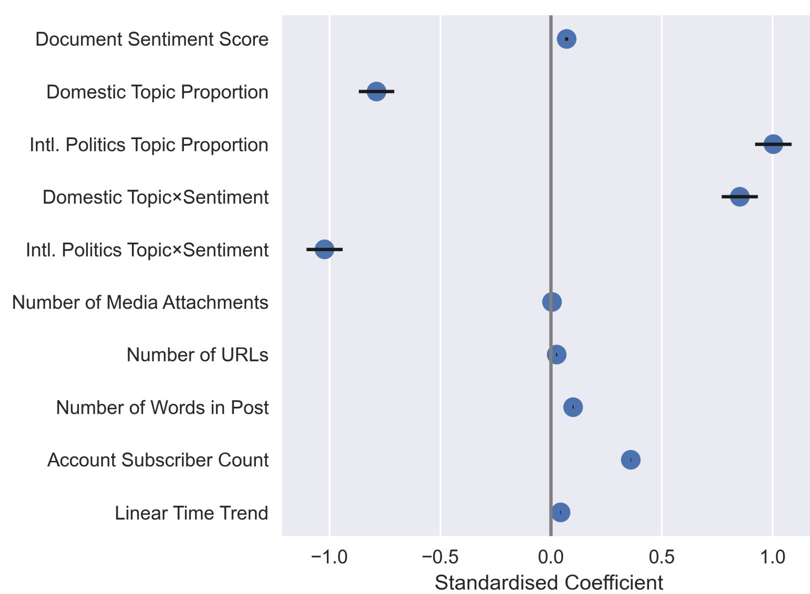
[4]In this study, "overall engagement" refers to all interactions received by each Facebook post, such as likes, comments, shares, and all types of emoji reactions. However, emoji reactions, introduced on February 24, 2016 (Facebook 2016), were not available for the entirety of the study period. The model using overall engagement primarily serves as a robustness check. The results from this model align with other models that consider engagement types available throughout the entire study (likes, shares, comments).

TABLE 4. Regression Analysis of User Engagement with Topic and Sentiment

	Dependent Variables			
	log(Like Count)	log(Comment Count)	log(Share Count)	log(Total Engagement)
Independent Variables	**(1)**	**(2)**	**(3)**	**(4)**
Document Sentiment Score	8.418*** (0.404) [0.070]	1.958*** (0.306) [0.022]	10.450*** (0.342) [0.104]	7.983*** (0.402) [0.067]
Domestic Topic Proportion	-7.789*** (0.404) [-0.788]	-6.259*** (0.306) [-0.836]	-6.577*** (0.343) [-0.794]	-8.221*** (0.403) [-0.834]
Intl. Politics Topic Proportion	11.346*** (0.473) [1.004]	17.726*** (0.358) [2.072]	20.182*** (0.401) [2.131]	12.333*** (0.471) [1.094]
Domestic Topic×Sentiment	14.396*** (0.702) [0.852]	11.021*** (0.532) [0.862]	12.021*** (0.595) [0.849]	15.057*** (0.700) [0.894]
Intl. Politics Topic×Sentiment	-20.252*** (0.825) [-1.022]	-30.081*** (0.624) [-2.007]	-36.333*** (0.699) [-2.190]	-21.974*** (0.819) [-1.168]
Controls	X	X	X	X
Linear Time Trend	X	X	X	X
Constant	-2.063*** (0.232)	-1.408*** (0.176)	-5.152*** (0.197)	-1.761*** (0.231)
N	781,260	781,260	781,260	781,260
R^2	0.160	0.159	0.140	0.160

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors in parentheses and standardised coefficients in square brackets. 'Controls' include media attachments, URLs, post length, and subscriber count; see Appendix K, Table K.2 for detailed results.

FIGURE 6. Standardized Coefficients of Predictors for Like Counts



Note: The standardised coefficients for all variables and their confidence intervals are taken from Model 1.

domestic affairs more positively and international politics more negatively. Other factors, such as the number of media attachments, URL links, post text length, account subscriber count, and the linear time trend are statistically significant in all models.

Collectively, all four models suggest that likes, comments, shares, and other reactions to CCP-controlled Facebook posts hinge on both the discussed topic and its associated sentiment. In previous sections, I discussed the dual strategy messaging of these outlets: positive sentiment for domestic affairs and negative for international politics. The results from Table 4 and Figure 6 indicate that online audiences reward *both* of these approaches with higher levels of online engagement. Notably, the interaction terms between topics and sentiment have larger standardised coefficients than the control variables, underscoring their relative importance. This suggests that the persuasive strategies employed by the CCP in its digital public diplomacy are effective in garnering attention and engagement from its target audience.

These findings offer an intriguing departure from expectations from by campaigning literature, which proposes that audiences engage more with negative campaign content that criticises rivals. In the context of the CCP's digital diplomacy campaigns, however, audience engagement levels are elevated not only in response to negative messaging, but also positive sentiment about domestic conditions under CCP governance. Three plausible distinctions between digital public diplomacy and typical democratic campaigns may explain the elevated engagement with discussions of domestic affairs painted in a positive light: (1) the audience for this campaign, (2) the nature of the nation-state conducting the campaign, and (3) the nature of the message content. I briefly discuss each of these here.

First, while measuring audience engagement with Facebook posts lends insight into audience responses to different messaging, it is important to note that these online audiences are not a representative sample of the international population. Many individuals who are exposed to CCP-controlled Facebook posts likely self-select into this exposure by liking or following these media accounts. Moreoever, there are also concerns that the engagement statistics may not reflect genuine user engagement due to potential inauthentic or paid accounts. Facebook have responded to these concerns by reporting that inauthentic accounts represent less than 0.001 percent of the total followers of CCP-controlled media pages, suggesting that the majority of engagement likely reflects genuine user interaction (The Economist 2019). While it would be beneficial to independently verify this claim, unfortunately the opaqueness of Facebook data access presents challenges for outside researchers in determining user authenticity. However, it is important to note that any engagement, authentic or not, amplifies the content via social media algorithms and enhances the visibility of these messages to wider audiences.

Second, the nature of the political actor conducting the campaign may influence audience engagement. Here, an authoritarian nation-state's ruling party, rather than a democratic political party, leads the campaign, potentially affecting how audiences interact with positive content on international platforms.

Third, the positive content produced by this campaign is subtly different from that typically seen within democratic contexts. It emphasises the nation-state's domestic conditions, focusing on soft power resources like language and culture. In democratic contexts, incumbents may highlight improvements under their governance, such as economic advancements. This subtle difference in focus may account for varied audience behaviour.

While beyond this article's scope, future research could explore these campaign features to understand why audience reactions to sentiment within digital public diplomacy campaigns are different from the reactions observed from campaigns in a democratic context.

CONCLUDING REMARKS

The novel data and methods presented in this article offer important insights into the strategies and goals of the CCP's large-scale international influence campaign, advancing our understanding beyond previous, topic-limited studies (Nordin 2016; Callahan 2015; Min and Luqiu 2021; Edney 2014). These results uncover both the persuasive strategies employed by CCP-controlled media outlets, and international audience responses to this campaign.

Analogous to political campaigns, digital public diplomacy shares the ambition of swaying public opinion for political gain. I have argued that digital public diplomacy campaigns should be understood as a form of political campaigning, and have tested whether these campaigns use narratives and sentiment similarly to campaigns studied within democratic contexts (Nai 2020; Crabtree et al. 2020; Müller 2022; Auter and Fine 2016). In digital public diplomacy campaigns, sentiment choice mirrors the campaign actor's political position: incumbents spotlight achievements via positive sentiment while challengers criticise incumbents negatively.

The roles of incumbents and challengers are, however, less defined in public diplomacy, thereby allowing nation-states flexibility in role adoption. By analysing public diplomacy strategies, this article pinpoints the adopted roles and associated objectives in the context of the CCP's ongoing campaign. Here, the CCP emerges as both incumbent and challenger, using positive sentiment to extol domestic conditions under its governance and negative sentiment to criticise international politics. Despite this duality, CCP-controlled media place greater focus on domestic affairs, signalling the CCP's intent to portray itself predominantly in a positive manner as a competent incumbent.

Intriguingly, international audiences appear to reward both strategies with higher engagement, attesting to the efficacy of the CCP's approach. Unlike conventional wisdom from campaign literature, positive sentiment about domestic conditions under CCP governance also raises engagement levels. Future research might investigate this discrepancy by examining features that distinguish digital public diplomacy campaigns from campaigns in a democratic context, such as the nature of the audience, the political actor conducting the campaign, and the message content.

Methodologically, this article has introduced a novel minimally-supervised method to quantify topic-level sentiment in text data. This technique combines word embeddings with a topic model to capture relative sentiment, thus enabling cross-topic and temporal comparisons. Uniquely, this method accommodates entire text for document-level sentiment calculation, thereby providing more precision and versatility than dictionary-based approaches.

As online campaigning becomes a key statecraft tool, other nations may well also deploy sustained online public diplomacy campaigns. The campaign literature proves insightful but insufficient, as nations may adopt campaign strategies that differ from those used in democratic contexts. The evidence from the CCP suggests that audiences react differently to digital public diplomacy campaigns, relative to campaigns studied in a democratic context. By identifying authoritarian objectives in digital public diplomacy narratives and demonstrating their ability to engage audiences, this study offers a robust framework for future research.

REFERENCES

Ansolabehere, Stephen and Shanto Iyengar (1995). *Going negative: How campaign advertising shrinks and polarizes the electorate*. New York: Free Press.

Arora, Sanjeev , Yingyu Liang, and Tengyu Ma (2017). A Simple But Tough-To-Beat Baseline for Sentence Embeddings. *International Conference on Learning Representations*.

Auter, Zachary J. and Jeffrey A. Fine (2016). Negative Campaigning in the Social Media Age: Attack Advertising on Facebook. *Political Behavior 38*(4), 999–1020.

Auxier, Brooke and Monica Anderson (2021), *Social Media Use in 2021*. https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/.

Bailey, Hannah (2023). Peaceful World Leadership or Nationalist Advocate? How China Uses Social Media to Project its Stance on Russia's Invasion of Ukraine. *Case Western Reserve Journal of International Law 55*(1), 321.

Bailey, Hannah and Philip N Howard (2022). The Instigators and Targets of Organised Social Media Manipulation: Global Index 2022. *The Programme on Democracy and Technology Working Paper 2022.1*.

Bellodi, Luca (2022). A Dynamic Measure of Bureaucratic Reputation: New Data for New Theory. *American Journal of Political Science n/a*(n/a), 1–18.

Bengio, Yoshua , Réjean Ducharme, Pascal Vincent, and Christian Jauvin (2003). A Neural Probabilistic Language Model. *Journal of Machine Learning Research 3*, 1137–1155.

Bjola, Corneliu , Jennifer Cassidy, and Ilan Manor (2019). Public Diplomacy in the Digital Age. *The Hague Journal of Diplomacy 14*(1-2), 83–101.

Bjola, Corneliu , Jennifer A. Cassidy, and Ilan Manor (2020). Digital Public Diplomacy: Business as Usual or a Paradigm Shift? In *Routledge Handbook of Public Diplomacy* (2 ed.). Routledge.

Blei, David M. (2012). Probabilistic topic models. *Communications of the ACM 55*(4), 77–84.

Blumenthal, Sidney (1982). *The permanent campaign*. Simon & Schuster.

Bobba, Giuliano (2019). Social media populism: features and 'likeability' of Lega Nord communication on Facebook. *European Political Science 18*(1), 11–23.

Brader, Ted (2005). Striking a responsive chord: How political ads motivate and persuade voters by appealing to emotions. *American Journal of Political Science 49*(2), 388–405.

Brader, Ted (2020). *Campaigning for Hearts and Minds: How Emotional Appeals in Political Ads Work*. University of Chicago Press.

Bradshaw, Samantha , Hannah Bailey, and Philip N. Howard (2021), *Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation*. https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/127/2021/01/CyberTroop-Report20-FINALv.3.pdf.

Bradshaw, Samantha , Renée DiResta, and Carly Miller (2022). Playing Both Sides: Russian State-Backed Media Coverage of the# BlackLivesMatter Movement. *The International Journal of Press/Politics*.

Bradshaw, Samantha and Philip N. Howard (2018). Challenging truth and trust: A global inventory of organized social media manipulation. *The Computational Propaganda Project 1*, 1–26.

Bradshaw, Samantha and Philip N. Howard (2019). The global disinformation order: 2019 global inventory of organised social media manipulation. *The Computational Propaganda Project Working Papers*.

Brady, Anne-Marie (2006). Guiding hand: The role of the CCP Central Propaganda Department in the current era. *Westminster Papers in Communication and Culture 3*(1), 58–77.

Brady, Anne-Marie (2009). *Marketing Dictatorship: Propaganda and Thought Work in Contemporary China*. Rowman & Littlefield Publishers.

Brady, Anne-Marie (2015), *China's Foreign Propaganda Machine*. https://www.wilsoncenter.org/article/chinas-foreign-propaganda-machine.

Bright, Jonathan , Hubert Au, Hannah Bailey, Mona Elswah, Marcel Schliebs, Nahema Marchal, Christian Schwieter, Katarina Rebello, and Philip N Howard (2020). Coronavirus Coverage by State-Backed English-Language News Sources. Technical report, The Computational Propaganda Project, Oxford, UK. https://demtech.oii.ox.ac.uk/research/posts/coronavirus-coverage-by-state-backed-english-language-news-sources/.

Callahan, William A. (2015). Identity and Security in China: The Negative Soft Power of the China Dream. *Politics 35*(3-4), 216–229.

Central People's Government of the People's Republic of China (2016), *Xi Jinping: jianchi zhengque fangxiang chuangxin fangfa shouduan tigao xinwen yulun chuanbo li yindao li*. http://www.gov.cn/xinwen/2016-02/19/content_5043970.htm.

Chaiken, S. (1987). *The heuristic model of persuasion*. Hilsdale, NJ: Erlbaum.

Chang, Yung-Yung (2021). The Post-Pandemic World: between Constitutionalized and Authoritarian Orders – China's Narrative-Power Play in the Pandemic Era. *Journal of Chinese Political Science 26*(1), 27–65.

Chester, Jeff and Kathryn C. Montgomery (2017). The role of digital marketing in political campaigns. *Internet Policy Review 6*(4), 1–20.

Choi, Donghyun Danny , J. Andrew Harris, and Fiona Shen-Bayh (2022). Ethnic Bias in Judicial Decision Making: Evidence from Criminal Appeals in Kenya. *American Political Science Review 116*(3), 1067–1080.

Collobert, Ronan and Jason Weston (2008). A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning. In *Proceedings of the 25th International Conference on Machine Learning*, Helsinki, Finland.

Cook, Sarah (2020). Beijing's Global Megaphone. Technical report, Freedom House. https://freedomhouse.org/report/special-report/2020/beijings-global-megaphone.

Cowan, Geoffrey and Amelia Arsenault (2008). Moving from monologue to dialogue to collaboration: The three layers of public diplomacy. *The annals of the American academy of political and social science 616*(1), 10–30.

Crabtree, C. , M. Golder, T. Gschwend, and I. H. Indriđason (2020). It Is Not Only What You Say, It Is Also How You Say It: The Strategic Use of Campaign Sentiment. *The Journal of Politics 82*(3), 1044–1060.

Creemers, Rogier (2015, July). Never the twain shall meet? Rethinking China's public diplomacy policy. *Chinese Journal of Communication 8*(3), 306–322.

CrowdTangle, *CrowdTangle - Content Discovery and Social Monitoring Made Easy*. https://www.crowdtangle.com//.

Cull, Nicholas J. (2008). Public Diplomacy: Taxonomies and Histories. *The ANNALS of the American Academy of Political and Social Science 616*(1), 31–54.

Cull, Nicholas J (2011). WikiLeaks, public diplomacy 2.0 and the state of digital public diplomacy. *Place Branding and Public Diplomacy 7*(1), 1–8.

Devlin, Jacob , Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT 1*, 4171–4186.

Dolezal, Martin , Laurenz Ennser-Jedenastik, Wolfgang C Müller, Katrin Praprotnik, and Anna Katharina Winkler (2018). Beyond salience and position taking: How political parties communicate through their manifestos. *Party Politics 24*(3), 240–252.

Edney, Kingsley (2014). *The globalization of Chinese propaganda: International power and domestic political cohesion*. Springer.

Entman, Robert M. (2008). Theorizing Mediated Public Diplomacy: The U.S. Case. *The International Journal of Press/Politics 13*(2), 87–102.

Esberg, Jane and Alexandra A. Siegel (2022). How Exile Shapes Online Opposition: Evidence from Venezuela. *American Political Science Review*, 1–18.

Facebook (2016, February), *Reactions Now Available Globally*. https://about.fb.com/news/2016/02/reactions-now-available-globally/.

Firth, John Rupert (1957). *Studies in linguistic analysis*. Wiley-Blackwell.

Fridkin, Kim L. and Patrick Kenney (2011). Variability in Citizens' Reactions to Different Types of Negative Campaigns. *American Journal of Political Science 55*(2), 307–325.

Geer, John G. and Lynn Vavreck (2014, April). Negativity, Information, and Candidate Position-Taking. *Political Communication 31*(2), 218–236.

Gennaro, Gloria and Elliott Ash (2022, April). Emotion and Reason in Political Language. *The Economic Journal 132*(643), 1037–1059.

Gilardi, Fabrizio , Theresa Gessler, Maël Kubli, and Stefan Müller (2022, January). Social Media and Political Agenda Setting. *Political Communication 39*(1), 39–60.

Gilboa, Eytan (1998). Media diplomacy: Conceptual divergence and applications. *Harvard International Journal of Press/Politics 3*(3), 56–75.

Gilboa, Eytan (2001). Diplomacy in the media age: Three models of uses and effects. *Diplomacy and statecraft 12*(2), 1–28.

Golan, Guy J. and Evhenia "Zhenia" Viatchaninova (2014). The Advertorial as a Tool of Mediated Public Diplomacy. *International Journal of Communication 8*(0), 21.

Gries, Peter Hays (2004). *China's new nationalism: Pride, politics and diplomacy*. Univ of California Press.

Grimmer, Justin and Brandon M. Stewart (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis 21*(3), 267–297.

Gross, Justin H. and Kaylee T. Johnson (2016). Twitter Taunts and Tirades: Negative Campaigning in the Age of Trump. *PS: Political Science & Politics 49*(4), 748–754.

Hagström, Linus and Astrid H M Nordin (2020). China's "Politics of Harmony" and the Quest for Soft Power in International Politics. *International Studies Review 22*(3), 507–525.

Hamilton, Clive and Mareike Ohlberg (2020). *Hidden Hand: Exposing How the Chinese Communist Party is Reshaping the World*. Simon and Schuster.

Heng, Yee-Kuang (2010). Mirror, mirror on the wall, who is the softest of them all? Evaluating Japanese and Chinese strategies in the 'soft' power competition era. *International Relations of the Asia-Pacific 10*(2), 275–304.

Horn, Franziska (2021). Exploring Word Usage Change with Continuously Evolving Embeddings. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: System Demonstrations*, Online, pp. 290–297.

Huang, Zhao Alexandre (2022, March). "Wolf Warrior" and China's digital public diplomacy during the COVID-19 crisis. *Place Branding and Public Diplomacy 18*(1), 37–40.

Huang, Zhao Alexandre and Rui Wang (2019). Building a Network to "Tell China Stories Well": Chinese Diplomatic Communication Strategies on Twitter. *International Journal of Communication 13*, 2984–3007.

Jerzak, Connor T. , Gary King, and Anton Strezhnev (2023). An Improved Method of Automated Nonparametric Content Analysis for Social Science. *Political Analysis 31*(1), 42–58.

Jung, Jae-Hee (2020). The Mobilizing Effect of Parties' Moral Rhetoric. *American Journal of Political Science 64*(2), 341–355.

Jungherr, Andreas (2023). Digital campaigning: How digital media change the work of parties and campaign organizations and impact elections. In *Research Handbook on Digital Sociology*, pp. 446–462. Edward Elgar Publishing.

Jungherr, Andreas , Gonzalo Rivero, and Daniel Gayo-Avello (2020). *Retooling politics: How digital media are shaping democracy*. Cambridge University Press.

Kampf, Ronit , Ilan Manor, and Elad Segev (2015). Digital Diplomacy 2.0? A Cross-national Comparison of Public Engagement in Facebook and Twitter. *The Hague Journal of Diplomacy 10*(4), 331–362.

Klinger, Ulrike , Karolina Koc-Michalska, and Uta Russmann (2022). Are Campaigns Getting Uglier, and Who Is to Blame? Negativity, Dramatization and Populism on Facebook in the 2014 and 2019 EP Election Campaigns. *Political Communication*, 1–20.

Kosmidis, Spyros , Sara B. Hobolt, Eamonn Molloy, and Stephen Whitefield (2019). Party Competition and Emotive Rhetoric. *Comparative Political Studies 52*(6), 811–837.

Kozlowski, Austin C. , Matt Taddy, and James A. Evans (2019). The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings. *American Sociological Review 84*(5), 905–949.

Kragh, Martin and Sebastian Åsberg (2017, September). Russia's strategy for influence through public diplomacy and active measures: the Swedish case. *Journal of Strategic Studies 40*(6), 773–816.

Kutuzov, Andrey , Lilja Øvrelid, Terrence Szymanski, and Erik Velldal (2018). Diachronic word embeddings and semantic shifts: a survey. *Proceedings of COLING 2018 abs/1806.03537*.

Lams, Lutgard (2018). Examining Strategic Narratives in Chinese Official Discourse under Xi Jinping. *Journal of Chinese Political Science 23*(3), 387–411.

Larsson, Anders Olof (2016). Online, all the time? A quantitative assessment of the permanent campaign on Facebook. *New Media & Society 18*(2), 274–292.

Lau, Richard (1982). Negativity in political perception. *Political Behavior 4*(4), 343–77.

Lau, Richard R. , Lee Sigelman, Caroline Heldman, and Paul Babbitt (1999). The Effects of Negative Political Advertisements: A Meta-Analytic Assessment. *American Political Science Review 93*(4), 851–875.

Lu, Yingdan and Jennifer Pan (2021, March). Capturing Clicks: How the Chinese Government Uses Clickbait to Compete for Visibility. *Political Communication 38*(1-2), 23–54.

Maas, Andrew L , Raymond E Daly, Peter T Pham, Dan Huang, Andrew Y Ng, and Christopher Potts (2011). Learning Word Vectors for Sentiment Analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oregon, pp. 142–150.

Malone, Gifford D. (1985). Managing Public Diplomacy. *The Washington Quarterly 8*(3), 199–213.

Malone, Gifford D. (1988). *Political advocacy and cultural communication: Organizing the nation's public diplomacy*, Volume 11. Lanham, MD: University Press of America.

Marcus, George E. and Michael B. MacKuen (1993). Anxiety, Enthusiasm, and the Vote: The Emotional Underpinnings of Learning and Involvement During Presidential Campaigns. *American Political Science Review 87*(3), 672–685.

Marcus, George E , W. R. Neuman, and M. Mackuen (2000). *Affective intelligence and political judgment*. Chicago, IL: University of Chicago Press.

Martin, Peter (2021). *China's civilian army: The making of wolf warrior diplomacy*. Oxford University Press.

Mayer, William G. (1996, September). In Defense of Negative Campaigning. *Political Science Quarterly 111*(3), 437–455.

McCarthy, Philip M. and Chutima Boonthum-Denecke (Eds.) (2012). *Applied Natural Language Processing: Identification, Investigation and Resolution*.

Mikolov, Tomas , Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeff Dean (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems 26*.

Min, Bumgi and Luwei Rose Luqiu (2021). How Propaganda Techniques Leverage Their Advantages: A Cross-national Study of the Effects of Chinese International Propaganda on the U.S. and South Korean Audiences. *Political Communication 38*(3), 305–325.

Mingjiang, Li (2008). China Debates Soft Power. *The Chinese Journal of International Politics 2*(2), 287–308.

Mutz, Diana C. and Byron Reeves (2005). The New Videomalaise: Effects of Televised Incivility on Political Trust. *American Political Science Review 99*(1), 1–15.

Müller, Stefan (2022). The Temporal Focus of Campaign Communication. *The Journal of Politics 84*(1), 585–590.

Nai, Alessandro (2020). Going Negative, Worldwide: Towards a General Understanding of Determinants and Targets of Negative Campaigning. *Government and Opposition 55*(3), 430–455.

Nai, Alessandro (2021, April). Fear and Loathing in Populist Campaigns? Comparing the Communication Style of Populists and Non-populists in Elections Worldwide. *Journal of Political Marketing 20*(2), 219–250.

Nordin, Astrid H M (2016). *China's international relations and harmonious world: time, space and multiplicity in world politics*. Routledge.

Olesen, Alexa (2015), *Where Did Chinese State Media Get All Those Facebook Followers?* https://foreignpolicy.com/2015/07/07/china-facebook-peoples-daily-media-soft-power/.

Ornstein, Norman J. (2000). *The permanent campaign and its future*. American Enterprise Institute.

Osnabrügge, Moritz , Sara B. Hobolt, and Toni Rodon (2021). Playing to the Gallery: Emotive Rhetoric in Parliaments. *American Political Science Review 115*(3), 885–899.

Pang, B. and L. Lee (2004). A sentimental education: Sentiment analysis using subjectivity analysis using subjectivity summarization based on minimum cuts. In *Proceedings of ACL*, Volume 10.

Pennington, Jeffrey , Richard Socher, and Christopher Manning (2014). Glove: Global Vectors for Word Representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, pp. 1532–1543.

Persily, Nathaniel and Joshua A. Tucker (2020). *Social media and democracy: The state of the field, prospects for reform*. New York: Cambridge University Press.

Petrocik, John R. (1996, August). Issue Ownership in Presidential Elections, with a 1980 Case Study. *American Journal of Political Science 40*(3), 825.

Petty, Richard E. and Pablo Briñol (2015). Emotion and persuasion: Cognitive and meta-cognitive processes impact attitudes. *Cognition and Emotion 29*(1), 1–26.

Pierrejean, Bénédicte and Ludovic Tanguy (2019). Investigating the Stability of Concrete Nouns in Word Embeddings. In *Proceedings of the 13th International Conference on Computational Semantics - Short Papers*, Gothenburg, Sweden, pp. 65–70. Association for Computational Linguistics.

Rebello, Katarina , Christian Schwieter, Marcel Schliebs, Kate Joynes-Burgess, Mona Elswah, Jonathan Bright, and Philip N Howard (2020). Covid-19 News and Information from State-Backed Outlets Targeting French, German and Spanish-Speaking Social Media Users. *Project on Computational Propaganda*.

Reporters Without Borders (2019, March), *China's Pursuit of a New World Media Order*. https://rsf.org/en/rsf-report-chinas-pursuit-new-world-media-order.

Rheault, Ludovic , Kaspar Beelen, Christopher Cochrane, and Graeme Hirst (2016). Measuring Emotion in Parliamentary Debates with Automated Textual Analysis. *PLOS ONE 11*(12), e0168843.

Rheault, Ludovic and Christopher Cochrane (2020). Word Embeddings for the Analysis of Ideological Placement in Parliamentary Corpora. *Political Analysis 28*(1), 112–133.

Rice, Douglas R. and Christopher Zorn (2021). Corpus-based dictionaries for sentiment analysis of specialized vocabularies. *Political Science Research and Methods 9*(1), 20–35.

Ridout, Travis N. and Kathleen Searles (2011). It's My Campaign I'll Cry if I Want to: How and When Campaigns Use Emotional Appeals. *Political Psychology 32*(3), 439–458.

Roberts, Margaret E. , Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand (2014). Structural Topic Models for Open-Ended Survey Responses. *American Journal of Political Science 58*(4), 1064–1082.

Rodman, Emma (2020). A Timely Intervention: Tracking the Changing Meanings of Political Concepts with Word Vectors. *Political Analysis 28*(1), 87–111.

Rodriguez, Pedro L. , Arthur Spirling, and Brandon M. Stewart (2023). Embedding Regression: Models for Context-Specific Description and Inference. *American Political Science Review*, 1–20.

Rossiello, Gaetano , Pierpaolo Basile, and Giovanni Semeraro (2017). Centroid-based Text Summarization through Compositionality of Word Embeddings. In *Proceedings of the MultiLing 2017 Workshop on Summarization and Summary Evaluation Across Source Types and Genres*, Valencia, Spain, pp. 12–21.

Salvo, David and Etienne Soula (2017, October), *Russian Government's Fission Know-How Hard at Work in Europe*. https://securingdemocracy.gmfus.org/russian-governments-fission-know-how-hard-at-work-in-europe/.

Schliebs, Marcel (2020), *rtangle: R Interface fro Crowdtangle Facebook API*. https://schliebs.github.io/rtangle/.

Schliebs, Marcel , Hannah Bailey, Jonathan Bright, and Philip N. Howard (2021a). China's public diplomacy operations: understanding engagement and inauthentic amplifications of PRC diplomats on Facebook and Twitter. *Programme on Democracy and Technology, Oxford University*.

Schliebs, Marcel , H. Bailey, J. Bright, and P. N. Howard (2021b). China's inauthentic UK twitter diplomacy: A coordinated network amplifying PRC diplomats. *Programme on Democracy and Technology, Oxford University*.

Shambaugh, David (2007). China's propaganda system: Institutions, processes and efficacy. *The China Journal 57*, 25–58.

Sinclair, Robert C. , Melvin M. Mark, and Gerald L. Clore (1994). Mood-Related Persuasion Depends on (Mis)Attributions. *Social Cognition 12*(4), 309–326.

Social Blade (2023, February), *Top 100 Facebook pages sorted by Likes - Socialblade Facebook Statistics*. https://socialblade.com/facebook/top/100/likes.

Szymanski, Terrence (2017). Temporal Word Analogies: Identifying Lexical Replacement with Diachronic Word Embeddings. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, Vancouver, Canada, pp. 448–453.

Tahmasebi, Nina , Lars Borin, and Adam Jatowt (2021). Survey of computational approaches to lexical semantic change detection. *Computational approaches to semantic change 6*, 1. Publisher: Language Science Press Berlin.

Tambe, Arjun M. and Toni Friedman (2022). Chinese state media Facebook ads are linked to changes in news coverage of China worldwide. *Harvard Kennedy School (HKS) Misinformation Review*.

Tang, Duyu , Furu Wei, Nan Yang, Ming Zhou, Ting Liu, and Bing Qin (2014). Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Baltimore, Maryland, pp. 1555–1565.

The Economist (2019, April), *China is using Facebook to build a huge audience around the world*. https://www.economist.com/graphic-detail/2019/04/20/china-is-using-facebook-to-build-a-huge-audience-around-the-worl

Tuch, Hans N. (1990). *Communicating with the world: US public diplomacy overseas*. Springer.

U.S. Department of State (2022). PRC Efforts To Manipulate Global Public Opinion on Xinjiang. Technical report, Global Engagement Center. https://www.state.gov/prc-efforts-to-manipulate-global-public-opinion-on-xinjiang/.

Walker, Vivian S , Shawn Baxter, and Kristina Zamary (2021). 2020 Comprehensive Annual Report on Public Diplomacy & International Broadcasting: Focus on FY 2019 Budget Data. Technical report, United States Advisory Commission on Public Diplomacy. https://www.state.gov/2020-comprehensive-annual-report-on-public-diplomacy-and-international-broadcasting/.

Wang, Zheng (2004). *Never forget national humiliation: Historical memory in Chinese politics and foreign relations*. Columbia University Press.

Weiss, Chagai M. , Alexandra A. Siegel, and David Romney (2022). How Threats of Exclusion Mobilize Palestinian Political Participation. *American Journal of Political Science*, 1–16.

Weiss, Jessica Chen (2014). *Powerful patriots: nationalist protest in China's foreign relations*. Oxford University Press.

Widmann, Tobias (2021). How Emotional Are Populists Really? Factors Explaining Emotional Appeals in the Communication of Political Parties. *Political Psychology 42*(1), 163–181.

Widmann, Tobias and Maximilian Wich (2022). Creating and Comparing Dictionary, Word Embedding, and Transformer-Based Models to Measure Discrete Emotions in German Political Text. *Political Analysis*, 1–16.

Xinhua (2013), *Xi jinping: Jiang hao zhongguo gushi chuanbo hao zhongguo shengyin*. http://www.xinhuanet.com/zgjx/2013-08/21/c_132648439.htm.

YouGov (2022), *The most popular social networks in the UK*. https://yougov.co.uk/ratings/technology/popularity/social-networks/all.

Zhao, Suisheng (2004). *A nation-state by construction: Dynamics of modern Chinese nationalism.* Stanford University Press.