Depicting US-China Disputes on Tech Giants through Social Media:

Evidence from Large-scale Twitter Data

下一步修改：突出只是介绍一个promising的领域，然后提供了一个demo（可以不出现在标题中）



**Existing Works**

This section summarizes existing works in understanding the dynamics of public opinion, pointing out its shift from traditional media to social media and the impetus behind it. We show that large-scale social media data ensures a promising future for as well as social simulating and modelling.

**1 Traditional media as the beginning: Efforts by social scientists**

@铭琪

The Verb In Context System (VICS) is also an early attempt to [1]. Based on the 1969 study of the ‘operational code’ [2],

**2 The advent of social media: Previous works in depicting online public opinion**

Opposite to social scientists’ focus on xxx, the advent of social media and the rapid development of computer science attract researchers to discover more latent and complex information from social media data.

Xu et al. [3] and Chen et al. [4] are both event-based country image study with Twitter data, observing online public opinion during the 70th anniversary of the People’s Republic of China and the COVID-19 pandemic, respectively. Their data was retrieved through Twitter Streaming API, and sentiments towards China (positive, negative, neutral) were analysed with machine learning algorithms trained on manually labelled data. Their features include: 1) Xu et al. collected and compared English and Chinese data, while Chen et al. focused on English discourse. It was found that a significant opposition existed between the online public opinions of the two languages. 2) Chen et al. provided fine-grained sentiment analysis by dividing online public opinion towards China into seven categories: Politics, Economy, Foreign affairs, Culture, Epidemic situation, Anti-epidemic measures, and Racism. They revealed that the gradual increase in negative politics-, foreign affairs-, and racism-related tweets and the decrease in non-negative epidemic situation-, anti-epidemic measures-related tweets resulted in the overall sentiments’ transition from non-negative to negative towards China. 3) Chen et al. displayed the different patterns in the attitudes of Congress members, media, and social bots, showing that social bots were more likely to spread negative sentiments towards China, while media were usually non-negative. For US congress members, the Republicans were more negative than the Democrats. 4) Xu et al. explored how positive and negative tweets were distributed among different countries and found that states enjoying better diplomatic relations with China generally had a positive view towards China. 5) Xu el al. obtained word vectors for the top 100 frequently and uniquely used words for both English and Chinese, positive and negative tweets through word2vec. Preferred topics of distinct languages and sentiments were analysed, e.g., positive Chinese tweets mostly focused on celebration activities while negative Chinese tweets tended to talk about broader topics like Hong Kong.

Other papers addressing online public opinion towards political events include

**3 The era of large-scale Internet data: Promises for data scientists**

Despite their considerable contributions to the emerging field of large-scale social media data analysis, the shortcomings of these works are also prominent. Their contributions to social challenges are vague:

Lack real-time or ‘nowcast’ analysis, which has the potential to detect major events at an early stage and provide governments and the society with necessary notifications.

[1] M. Schafer and S. G. Walker, "Operational code analysis at a distance: The verbs in context system of content analysis," in *Beliefs and leadership in world politics*: Springer, 2006, pp. 25-51.

[2] A. L. George, "The 'operational code': A neglected approach to the study of political leaders and decision-making," *International studies quarterly,* vol. 13, no. 2, pp. 190-222, 1969.

[3] Y. Xu, Q. He, and S. Ni, "Understanding Online Public Sentiments: A Machine Learning-Based Analysis of English and Chinese Twitter Discourse during the 2019 Chinese National Day," presented at the The 2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC 2020), Kimberley, South Africa, 2020.

[4] H. Chen *et al.*, "Country Image in COVID-19 Pandemic: A Case Study of China," *IEEE Transactions on Big Data,* pp. 1-1, 2020, doi: 10.1109/TBDATA.2020.3023459.