Leveraging Topic Modeling and Toxicity Analysis to understand China-Uyghur Conflicts *

Connice Trimmingham and Nitin Agarwal

University of Arkansas at Little Rock, Little Rock, AR 72204, USA
{ctrimmingham, nxagarwal}@ualr.edu

Abstract. The proliferation of social networking sites, aided by the pervasiveness of mobile technology, has facilitated the propagation of various forms of toxicity. Although social media platforms provide valuable tools for meaningful interactions, political arguments, often fraught with complex mix of emotions, can quickly devolve into flame wars or partisan bickering. This article shifts attention eastward to examine how the media/information environment is being manipulated for advancing political agendas in the Indo-Pacific region. We analyzed 3,239,249 number of tweets discussing issues related to China and Uyghur. We extracted influential topics using a combination of Latent Dirichlet Allocation (LDA) based topic modeling approach. We explained the user relation phenomena by assessing their emerging social structures. Toxicity analysis and bot assessment were performed to examine the nature of discourse about the China and Uyghur issues. Our analysis suggests a high correlation between tweets with high toxicity and bot activities to emerging events such as the existence of internment camps and news about forced Uyghur laborers in China and the Chinese Communist Party in China network.

Keywords: Information operations · Indo-Pacific · Twitter · Social Media · Uyghur · Toxicity Analysis

* This research is funded in part by the U.S. National Science Foundation (OIA-1946391, OIA-1920920, IIS-1636933, ACI-1429160, and IIS-1110868), U.S. Office of Naval Research (N00014-10-1-0091, N00014-14-1-0489, N00014-15-P-1187, N00014-16-1-2016, N00014-16-1-2412, N00014-17-1-2675, N00014-17-1-2605, N68335-19-C-0359, N00014-19-1-2336, N68335-20-C-0540, N00014-21-1-2121, N00014-21-1-2765, N00014-22-1-2318), U.S. Air Force Research Lab, U.S. Army Research Office (W911NF-20-1-0262, W911NF-16-1-0189), U.S. Defense Advanced Research Projects Agency (W31P4Q-17-C-0059), Arkansas Research Alliance, the Jerry L. Maulden/Entergy Endowment at the University of Arkansas at Little Rock, and the Australian Department of Defense Strategic Policy Grants Program (SPGP) (award number: 2020-106-094). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations. The researchers gratefully acknowledge the support.

1 Introduction

China has been in the global spotlight for its economic strategies, investments acquisitions, and policy reinforcement. However, recently China's reputation is globally questioned for its targeted, inhumane, and oppressive policies towards the Uyghur population in Xinjiang [1]. From a geographical standpoint, Xinjiang is an autonomous region that measures one-sixth of China's western border and home to a Chinese Ethnic-Muslim minority. From a political perspective, Xinjiang houses an extensive potential for mineral exploitation in natural resources such as oil, gas, and agricultural production [2].

China's Uyghur conflict has existed for decades, its universal debate however, has recently surfaced with the unprecedented evolution of online social networks. Although religious beliefs, customs, and practices have been tolerated in China to some extent, the degree of tolerance has varied considerably from time to time with the change in the political climate [3]. The use of coercion is not uncommon in Chinese history as far as religious groups are concerned [4]. Extant literature has also shown that policies towards Xinjiang are like the policies that were directed towards Tibet [4]. To cope with these policies, the Uyghur group attempted to separate themselves from the Chinese government and develop their own identity [4]. This independent movement threatens the viability of the unified communist system established by the People's Republic of China [1].

It is pertinent to study the China Uyghur conflict as it highlights the connection between a strong authoritarian state, a terrorist threat, and a minority group [5]. However, the implications of these dynamics are potentially far-reaching, as they promise to complicate China's rise in Central Asia. Many western literatures describe Chinese politics as authoritarian; while this view is not inaccurate, it is incorrect to assume that Chinese citizens have been content to be despotically ruled [6]. As a multi-ethnic state with a vast majority of Han Chinese and various minorities, the Chinese's government considers any nationalist or independence movements as an attack towards China's unified communist system and economic growth [7].

There are significant scholarly works on leveraging the internet to gain more, and better information. Despite these possibilities, extant literature has shown that algorithmic and filtering features of social media platforms have driven users to an "echo chamber" whereby they are exposed to more of what they want and like, as opposed to what they need or should see [8][9]. This can shift their narratives on world issues as users disregard any narrative about topics that are ideologically unpleasant. The pervasiveness of partisan animosity on social media also exacerbates this issue [5]. Researchers have attempted to assess the definition and representation of identities and the leveraging power of minorities versus a superior state in the negotiation process [5][10]. Research by [11] concluded that the approaches utilized by both parties in the 'David and Goliath' duel for a contested region mostly affect the minority group due to low availability of resources and strategies [11].

This article will firstly offer theoretical background about this topic, and then engage in turn with how we leveraged topic modeling, toxicity analysis and bot assessment to understand China-Uyghur issue. The remainder of the article is set out as follows. First, a few extant literature and analytical frameworks relating to China and Uyghur issues are reviewed. Next, the empirical study is described, and the findings are discussed. Lastly, we discuss conclusions, limitations, and directions for future work.

2 Literature Review

This section describes the extant literature on this topic and the theoretical framework we used for this study.

2.1 China and the Uyghurs

It is important for the Uyghur diaspora to establish links with the international community and create awareness in the West, especially amongst non-governmental organizations and human rights activists, so that it can exert some pressure on the Chinese state to correct the plight of the Uyghurs. Researchers have argued that Beijing's strategies in Xinjiang with respect to the Uyghur issues at the domestic, regional, and international levels are characterized with multiple contradictions [5][12]. They further reasoned that China's approach to Xinjiang domestically contributed to the internationalization of the issue [12]. However, others have argued that China faced the prospect of Xinjiang becoming its own West Bank if it fails to re-strategize to a softer approach to integrate the region [12]. They argue that China has explicitly framed episodes in world events such as the 9/11 crisis to shift the narratives towards Uyghur rebellion as "terrorism" and boost their international and regional sympathy [12].

Researchers have also explored how the increasing complexity of the conflicts between Uyghur and China indicates the potential for Uyghur violence to escalate [12][13]. This is specifically in light of the reported inception of a state-initiated mass 'reeducation' campaign for Uyghur and other Muslim minorities across the province [13]. They argued that by reportedly sending Xinjiang's Muslim population to 'vocational education centers,' China's attempts to 'prevent extremism' may lead to a resurgence of ethnic unrest in Xinjiang [13].

2.2 Toxicity analysis on Social Media

Toxicity analysis has been used to understand the pulse of society on hot-button issues [14]. In a study conducted by [14], the researchers evaluated five categories of toxicity on comments posted on pro-and anti-NATO channels on YouTube. They demonstrated anti-NATO channels comments were more toxic when compared to pro-NATO channels comments. Researchers have also aimed to characterize and predict the behavior of toxic users in online discussions [15]. They found topical predictions of toxic response with semantic shifts from parent comments in their study. Another study analyzes online toxicity with a case modeling approach [16]. They developed an epidemiological model to study and evaluate the spread of toxicity on YouTube. They applied the STRS (Susceptible, Toxic, Recovered, Susceptible) model to detect similarities between toxicity propagation on YouTube and the spread of a disease within a population. In another study, the authors evaluated the role of toxicity on tweets about societal issues such as the wearing of face masks during the COVID-19 pandemic [17]. Their results showed that tweets with pro-mask hashtags that supported wearing masks were less toxic compared to tweets who spread news about COVID-19 on YouTube.

2.3 Network Analysis

Tighter government regulations on online activities can make users seek a more democratic channel/outlet. However, Song et al [18] found an increased success of China's Internet repression where the Chinese Twitter proved to be small, lacking an accessible and diverse network due to China's sophisticated Internet content control regime. This coincided with the debate on the Chinese government approach to public diplomacy. Huang et al. [19] demonstrates how the Chinese government utilizes communication channels, specifically a small number of Twitter accounts to amplify its public diplomacy network and promote China's international influence. Huang et al. [19] further explains that China's robust Twitter network function on "timid polyphony" centered around its closest friends with expansion outward to include other alliances. Researchers have also shown how public leaders such as politicians utilize micro-blogging platforms like Twitter to gain rapid attention compared to other traditional ways of communication. Khan et al. demonstrates that understanding the supporters' network of opinion leaders, helps in predicting the type of relationship between supporters of the leaders [20].

2.4 Bot Analysis

Bot and botnet activities have the ability to shift narratives, opinions, and behavior of humans, especially within the political landscape where hot-button issues are debated. Ferrara et al. [21] explain that there are economic and political incentives for injecting social bots into online ecosystems. Some bots act with the objective of forming and growing an audience to exert influence. Further, research in technographic approach argues that the agency of bots should be seen not only as computing units but as interlocutors and informants [22]. Their study of chatbots development in China proved that elevated disruptive technologies such as artificial intelligence and big data are critical factors in state security and narrative control in China [22]. Another study on computational propaganda, domestic automation and opinion manipulation utilizing 1.1 million hashtags on twitter associated with China and Chinese politics showed a large amount of automaton [23]. This automation, however, were more aligned with anti-Chinese state perspectives [23].

3 Methodology

This section focuses on our study design which consists of the data collection and approaches applied for this research.

3.1 Data Collection and Processing

To understand the online universal conversation specific to China and Uyghur, we collected data tailored towards narratives containing a set of preliminary key phrases such as "China" and 'Uyghur'. This allowed us to query and truncate our data to tweets that focus on key issues relating to both China and the Uyghur group. This approach functioned as a filter for refining our data and eliminating any term or outliers irrelevant to our research. We extracted metadata from users and posts on twitter utilizing our in-house twitter API crawler. All tweets collected were posted between 2020–2021. We applied this date range based on peak periods of tweets cross referencing to specific events and news relating to China and Uyghur. Table 1 shows the breakdown of the total tweets extracted for China and Uyghurs, respectively.

Table 1. Frequency of tweets for China and Uyghur

Narrative	Tweets	Users
China	1,508,016	768,855
Uyghur	1,731,233	762,364

3.2 Topic Modeling

To understand the influential topics in our dataset, we applied Latent Dirichlet Allocation (LDA) topic modeling on the extracted tweets. We first tokenized each tweet into sentences, and sentences into words with the removal of punctuation and stop-words. Words were lemmatized and stemmed to their root form. The model was initially trained on a random number of topics and later decreased and ranked to the top 4 topics based on the coherence score of the topic distribution. Topic modeling revealed topic 1 and topic 2 as top topics with distinct overlaps in China narrative. Both topics contained trending words relating to communism, policing and the Chinese Communist Party. Topic 1 with top words such as home, forced had the highest distributions within Uyghur narrative, and relates to the reinforcement of forced Labor on Uyghur Muslims. Table 2 shows top words relating to China and Uyghur along with their respective distributions.

3.3 Toxicity Analysis

Since tweets contain a wealth of information about the thoughts and feelings of people, it is imperative to analyze the toxicity of tweets discussing China-Uyghurs conflicts. Toxic tweets were evaluated using natural language processing techniques specifically, google perspective API which utilizes machine learning and detoxify a pre-trained model used to detect toxic sentences [26][27].

Table 2. Top topics within China and Uyghur narrative.

	China			Uyghur		
Topics	Word1	Word2	%	Word1	Word2	%
1	Communist	CCP	0.683	Home	Forced	0.533
2	Positive	Chinese	0.154	Education	Jalan	0.359
3	AMP	Paupa	0.152	Genocide	Stop	0.077
4	Youth	Muslims	0.009	Uyghur	Chinese	0.032

Both techniques are multilingual and offered a probability score between 0-1. Higher score indicates toxicity. Final toxicity scores were averaged and aggregated monthly within the period of January 2020 to December 2022 then multiply by topic distribution scores to get the toxicity per topics. Fig. 1 shows the volatility of toxic tweets across the top 4 topics relating to China. The most influential topic: Topic 1 had the highest toxicity relative to other topics. This pattern is explainable through the semantics of trending words in topic 1 which revealed top conversations relating to communist, CCP and Chinese government. This signals that events in this period relating to these top words triggered negative interests of twitter users which correlates to the high toxicity of tweets. Distinct events within this period that coincided with various spikes include: “the 50 independent United Nations Human Rights experts highlighting their concern on the situation in China relating but not limited to forced labor [25].”

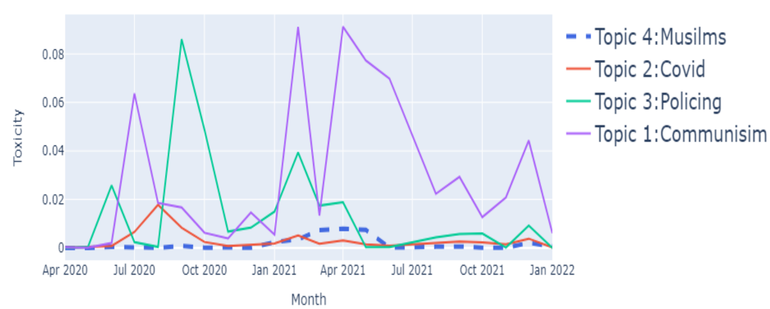

Fig. 1. China Toxicity Trend within the period of 2020 - 2022.

Similarly, a high and volatile toxicity with noticeable spikes across the period was found within Uyghur narrative, demonstrating an ongoing discussion of issues and events on these topics throughout the trend’s lifecycle. Noticeable events in period that coincide with these topics include “Officials denied the existence of internment camps, or alternatively justify them as poverty alleviation and stability maintenance efforts” and “uncovered evidence by the New York Times that reveal that Uyghur laborer’s, many who are interned forcibly, are involved in making personal protective equipment that are shipped all around the world [28][29].”

3.4 Network Analysis and Bot Analysis

Understanding the connective relationships within both narratives helps to discover information flows and any concerted tactics about our topics. We utilized peak points found in our monthly tweets frequency reports to study each narrative social structures. Extreme overlaps were found in tweets posted within various peak points to news events on top topics. We discovered that the behavioral trend of tweets frequency in both narratives increased and/or decreased at the same rate. Due to computational expenses of running network graph on our full data, we applied a random sampling technique to approximate the period each narrative tweets trend began rising.

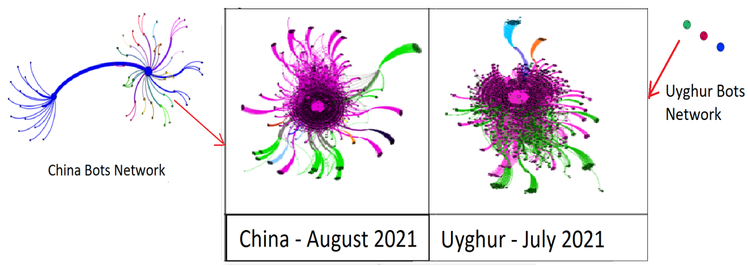

Fig. 2. Network of users with bot CAP scores above 0.90 within period of 2020 - 2021

China network focused on tweets posted in the period of August 2021 to September 2021 and references various events in August 2021 such as "Children of Detained Uyghurs parents held in Welfare schools in China's Xinjiang" [30]. Additionally, Uyghur network looked at July 2021 to August 2021, referencing events such as "president Xi praises Xinjiang armed police for counter terrorism effort in Uyghur territory". The biggest rise for both narratives was seen in July 2021 to November 2021. To detect communities, a modality community detection algorithm was applied, and the top 3 Communities were color coded purple, green, and blue according to ranking (Fig. 2). A total of 16 dense communities were detected within a corpus of 12,292 users in China network from a modularity class of 0.673. Uyghur network was less dense than China with a total of 17 communities and a modularity weight of 0.536 within 5,059 users. Majority of users within both networks had less than 500 connections with a relatively low average degree. However, about 10% of these connections had following count of 1000 or greater. This was seen through China's network top contributor @PaulSmall4eva with 39 followers and connections such as @PinkRangerLB who had 39,100 followers. Uyghur network had top contributor @RAbdiAnalyst maintaining identifiers such as Chief Analyst, geopolitics, and strategy with a following count of 255,000.

The nature and range of bot behaviors makes it universally difficult to define a bot [34]. To balance false positives and negatives, we applied the Complete

Automation Probability (CAP) of 0.90 or higher to raw bot-scores to detect bots. CAP is probability calculation developed by Observatory on Social Media project API Botometer that utilized Bayes' theorem to estimate of the overall prevalence of bots on a score of 0 to 1 [34]. Higher scores equate to higher probability of bot-like activity. Fig. 2 highlights 18 bot communities mirroring china network and 3 bot communities with no relations in Uyghur network while Fig 3 shows bot activities trend co-relating to toxicity on topics within China network. Overall toxicity is directly proportional and highly comparable to bot activity in communism topics while it is relatively low but still comparable to topics on policing. These findings imply that bot activities jolted narratives toxicity and shifted opinions against communism issues in china. further research would need to explore the intention of these accounts' generation.

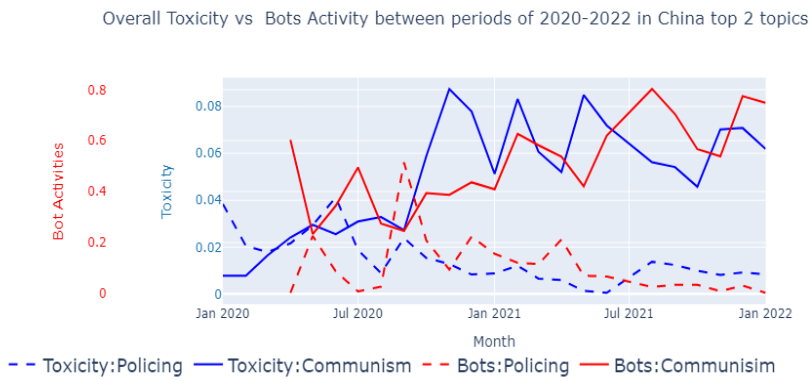

Fig. 3. Overall Toxicity vs Bot Toxicity within topic 1 about Communism and Topic3 about Policing

4 Conclusions and Future Work

In this study, we addressed the prevalence of toxicity in China – Uyghurs dilemma. Understanding the network of users is useful in showing the viewpoint of actors as a collective on issues relating to our key topics. Hence, we focused on top topics on twitter that contain key phrases relating to the two focal narratives. We also utilized network analysis tools such as network-X and Gephi to explore the network of users to detect communities and top users in the network. Our network analysis is grounded on the theoretical framework that several extant literatures have used. For instance, we applied modularity to detect communities.

This paper also contributes to an ongoing body of research aimed to fully understand the online universal dialogue of diplomacy, identity and policies within an authoritative state and the extent to which these policies can affect the rights of a minority group. It defines an interoperable methodology to understand the topics of high relevance, identifying toxicity within these topics and detecting the top contributors within the network. There are many actors pushing the Uyghurs' identity conflict as discussion internationally past China's borders. Some actors are influential based on followers count and only need one tweet to create a ripple effect. There are neutral actors such as news agencies that maintain unbiased tweet profiles. Some actors are initiators, i.e., by taking more responsibility in bringing awareness to Uyghur issues by posting and retweeting. We found that topics within these networks of users range from forced labor, genocide, education, communism, and politics to policing. Their differences provide a unique combination that illustrates the overlap in the China-Uyghur network. Our findings showed probable anti-China communities with a top contributor and smaller connections that discussed the topics highlighted. Future study is required to see how these networks transform over time in relation to the topics discussed. This work paves the way for development of more in-depth research to understand the dimensions of topics, their toxicity and bot activities, and relationships within China-Uyghur on twitter. The dataset for this study was collected retrospectively after the event has unfolded. Future study can investigate the application of various social network analysis techniques to garner real-time inferences about emerging socio-political issues.

References

1. Bovingdon, G. (2004). Autonomy in Xinjiang: Han nationalist imperatives and Uyghur discontent.
2. Glen, S. (2015). Sources of Conflict in the Xinjiang Uyghur Autonomous Region: History, Power, and Uyghur Identity Flux?. Griffith Asia Quarterly, 3(1).
3. Marc Julienne, Moritz Rudolf and Johannes Buckow, “Beyond Doubt: The Changing Face of Terrorism in China,” The Diplomat, 28 May 2015; URL: https://thediplomat.com/2015/05/beyond-doubt-the-changing-face-of-terrorism-in-china/.
4. Raza, Z. (2019). China's 'political re-education'camps of Xinjiang's Uyghur Muslims. Asian Affairs, 50(4), 488-501.
5. Lecours, André, and Geneviève Nootens, eds. Dominant nationalism, dominant ethnicity: Identity, federalism, and democracy. No. 15. Peter Lang, 2009.
6. Mukherjee, K. (2010). The Uyghur question in contemporary China. Strategic Analysis, 34(3), 420-435.
7. Hyer, E. (2006). China's policy towards Uighur nationalism. Journal of Muslim Minority Affairs, 26(1), 75-86.
8. Kent, M. (2001). Managerial rhetoric as the metaphor for the World Wide Web. Critical Studies in Media Communication, 18(3), 359-375.
9. Kent, M. L. (2013). Using social media dialogically: Public relations role in reviving democracy. Public relations review, 39(4), 337-345.
10. Davis, E. V. W. (2008). Uyghur muslim ethnic separatism in Xinjiang, China. Asian Affairs: An American Review, 35(1), 15-30.
11. Glen, S. (2015). Sources of Conflict in the Xinjiang Uyghur Autonomous Region: History, Power, and Uyghur Identity Flux?. Griffith Asia Quarterly, 3(1).
12. Clarke, M. (2015). China and the Uyghurs: the “Palestinization” of Xinjiang. Middle East Policy, 22(3), 127-146.

13. Soliev, N. (2019). UYGHUR VIOLENCE AND JIHADISM IN CHINA AND BEYOND. Counter Terrorist Trends and Analyses, 11(1), 71–75.
14. Obadimu, A., Khaund, T., Mead, E., Marcoux, T., and Agarwal, N., 2021. "Developing a socio-computational approach to examine toxicity propagation and regulation in COVID-19 discourse on YouTube," Information Processing and Management, pp. 58, https://doi.org/10.1016/j.ipm. 2021.102660.
15. Almerekhi, H., Kwak, H., Jansen, B. J., Salminen, J. (2019, September). Detecting toxicity triggers in online discussions. In Proceedings of the 30th ACM conference on hypertext and social media (pp. 291-292).
16. Obadimu, A., Mead, E., Maleki, M., and Agarwal, N., 2020b. "Developing an Epidemiological Model to Study Spread of Toxicity on YouTube," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 266–276,
17. Pascual-Ferrá, P., Alperstein, N., Barnett, D., and Rimal, R., 2021. "Toxicity and verbal aggression on social media: Polarized discourse on wearing face masks during the COVID-19 pandemic," Big Data and Society, pp.8, https://doi.org/10.1177/20539517211023533
18. Song, S. Y., Faris, R., Kelly, J. (2015). Beyond the Wall: Mapping Twitter in China. Berkman Center Research Publication, (2015-14).
19. Huang, Z. A., Wang, R. (2019). Building a network to "tell China stories well": Chinese diplomatic communication strategies on Twitter. International Journal of Communication, 13, 24.
20. Khan, A., Zhang, H., Shang, J., Boudjellal, N., Ahmad, A., Ali, A., Dai, L. (2020). Predicting politician's supporters' network on twitter using social network analysis and semantic analysis. Scientific Programming, 2020.
21. Ferrara, E., Varol, O., Davis, C., Menczer, F., Flammini, A. (2016). The rise of social bots. Communications of the ACM, 59(7), 96-104.
22. Xu, Y. (Joe). (2018). Programmatic Dreams: Technographic Inquiry into Censorship of Chinese Chatbots. Social Media + Society. https://doi.org/10.1177/2056305118808780
23. Bolsover, G., Howard, P. (2019). Chinese computational propaganda: Automation, algorithms and the manipulation of information about Chinese politics on Twitter and Weibo. Information, communication society, 22(14), 2063-2080.
24. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
25. Ferrara, E., Varol, O., Davis, C., Menczer, F., Flammini, A. (2016). The rise of social bots. Communications of the ACM, 59(7), 96-104.
26. Xu, Y. (Joe). (2018). Programmatic Dreams: Technographic Inquiry into Censorship of Chinese Chatbots. Social Media + Society. https://doi.org/10.1177/2056305118808780
27. Bolsover, G., Howard, P. (2019). Chinese computational propaganda: Automation, algorithms and the manipulation of information about Chinese politics on Twitter and Weibo. Information, communication society, 22(14), 2063-2080.
28. https://news.un.org/en/story/2020/06/1067312
29. https://perspectiveapi.com
30. https://pypi.org/project/detoxify/
31. https://xinjiang.sppga.ubc.ca/timelines/forced-labour/
32. https://www.nytimes.com/2020/07/19/world/asia/china-mask-forced-labor.html
33. https://www.rfa.org/english/news/uyghur/welfare-schools-08162021174124.html: :text=More
34. https://botometer.osome.iu.edu/faq