Dissertation Summary

Allison Koh[1]

[1]**Centre for International Security, Hertie School, Berlin**

My dissertation, titled "Grey Areas of Digital Transnational Repression on Western Media Platforms", leverages information from platforms with a global readership to understand the contours of *grey propaganda*—or that which cannot be definitively attributed to governments—in media attacks targeting diaspora communities from repressive contexts. The use of social media content by state-aligned entities and their supporters comprise a broader strategy of engaging in *digital transnational repression*—or the use of digital tools for implementing repressive policies with the aim of silencing or coercing the diaspora located outside of territorial borders. With an empirical focus on pro-CCP Twitter's attacks on prominent regime critics in the Uyghur diaspora, I demonstrate the importance of understanding how digital transnational repression materializes in the public sphere.

Online media have transformed how repressive governments and their supporters target exile and diaspora critics in the online opposition abroad. To investigate how platform affordances and identities can be weaponized by state-aligned users, I am working with Professors Joshua Tucker and Jonathan Nagler at NYU's Center for Social Media and Politics on a case study focusing on pro-CCP tweets targeting regime critics in the Uyghur diaspora. In this project, we investigate sources of variation in the *toxicity*—or the probability that a tweet will lead to a target limiting their activity on the platform[1]—of pro-CCP tweets in their targeted media attacks. Our primary expectation is that, among all pro-CCP tweets, posts authored by overtly state-affiliated accounts[2] are—on average—less likely to be toxic compared to content produced by other pro-CCP accounts without clear ties to the state. To assess possible sources of variation in toxicity due to cultural aspects, we also hypothesize that women among the diaspora regime critics are more likely to be targeted with toxic content, and that English-language tweets are less likely to be toxic compared to tweets published in Chinese. Our findings have important implications for identifying how state-aligned actors can manipulate new media to engage in transnational smear campaigns.

To further delve into the implications of pro-government social media for transnational repression, another paper in my dissertation investigates the relationship between pro-CCP tweets targeting diaspora regime critics and domestic policies that target ethnic minorities in China. Relevant scholarship outlines the key role of prominent dissidents in diasporas, but little is known about how pro-government actors use targeted media attacks to limit global communication about state repression. With a focus on China's repression of the Uyghur diaspora, I argue that pro-government actors add noise to relevant conversations involving prominent dissidents, and that this manipulation of the online environment is collaborative. I expect that the overall volume of relevant content increases, topical focus decreases, and tweets are more coordinated. I test my hypotheses with: (1) all tweets mentioning keywords relevant to "Xinjiang" and the Uyghur region, and (2) pro-CCP tweets identified by a fine-tuned BERT + Active Learning model. I identify potential shifts in the online activity of pro-CCP tweeters after April 2017, when the Chinese government started arbitrarily detaining Uyghurs, ethnic Kazakhs, and minority Muslim groups in internment camps. The findings of this study illustrate how pro-government actors manipulate the transnational information environment around diaspora discussions of domestic repression.

Finally, to synthesize the contributions of this research, I outline a novel conceptual and empirical basis for studying digital transnational repression with social media data. I argue that research on digital

---

[1]https://perspectiveapi.com/

[2]e.g. accounts with "China state-affiliated media" labels

transnational repression has predominantly investigated covert strategies, with an emphasis on states' use of domestic technologies to subjugate diaspora regime critics who play a crucial role in the politics of their home countries. With this focus on domestic channels of communication, relevant research has overlooked the strategic use of global media platforms for openly attacking prominent regime critics. Accordingly, I develop an analytical framework for studying digital transnational repression in the public sphere. Combining theoretical perspectives from the relevant scholarship on digital repression and foreign influence operations, I outline (1) how state-aligned media attacks can vary based on who is targeted, (2) the extent of plausible deniability in state attribution, and (3) whether media attacks are temporally aligned with other instances of state repression. I highlight how these sources of variation manifest in the transnational information environment and outline how researchers can leverage social media data to understand these dynamics. Overall, this paper contributes a novel conceptual and empirical basis for studying the strategic repertoires of digital transnational repression on global media platforms.