Appendix 1: Validation of automatic location classification
We manually checked a random sample of 100 locations to validate our location identification approach with OpenStreetMaps (https://www.openstreetmap.org/), which was accessed directly from R with the package *RgoogleMaps* (Loecher & Ropkins, 2015). The country was then identified based on the location's coordinates with the package *maps* in R (Becker et al., 2018). Our analysis showed that 81% of the locations were correctly classified. The 19% wrongly classified locations did not show a systematic bias as the misclassification is quite random. The misclassification is due to the use of a fictional location name (e.g., Terminus, King's Landing) or just a vague description (e.g., Indigenous Land, in the wild).

We also checked a random sample of 100 locations that OpenStreetMap could not identify. 68% of these locations were indeed not "real" locations (e.g., party in my head, Currently between dimensions). However, 32% mentioned a potentially real location that OpenStreetMap could not identify. Most of these unidentified cases mention more than one location abbreviation in the description (e.g., "tpe, nyc, dc, hkg") or refer to the location it in a way that is not recognized by OpenStreetMap (e.g., Occupied Muwekma Ohlone Lands, Left coast Appalachian).

Appendix 2: Keyword-based issue identification
We used regex patterns to identify tweets that cover an issue for each broader topic. Tweets can cover different issues. For us, the true positive rate is more important as we always compare the tweets that include an issue with all other tweets in the model. Due to the large sample size, we always have a very robust baseline value based on all other tweets, that do not cover an issue. This means that the tweets that cover an issue but were not identified (false negatives), do not really have an impact on the baseline value.
We always took a random sample of 20 tweets that cover the issue and 20 tweets that do not cover the issue for the validation.

Table: Validation of regex patterns to identify topics

Topic	Pattern used in R with grepl()	True positives	True negatives
China	"China|Chinese|Diaspora|中国|Jinping|Huawei", ignore.case = T	20 (100%)	19 (95%)
US	"Americ|United States|\\bUS\\b|stopasianhate|美国|Trump|Biden", ignore.case = F	20 (100%)	18 (90%)
COVID-19	COVID	20 (100%)	20 (100%)
Hong Kong	"HK|Hong Kong", ignore.case = F	20 (100%)	20 (100%)
Black Lives Matter	"\\bBLM\\b|\\bBlack", ignore.case = T	18 (90%)	20 (100%)
Uyghurs and Xinjiang	"xinjiang|Uyg", ignore.case = T	20 (100%)	20 (100%)
Tiananmen Protest	"Tiananmen", ignore.case = T	20 (100%)	20 (100%)

Communism and Socialism	"Communis|socialis|marx", ignore.case = T	17 (85%)	17 (85%)

References

Loecher, M., & Ropkins, K. (2015). RgoogleMaps and loa: Unleashing R Graphics Power on Map Tiles. Journal of Statistical Software, 63(4), 1–18.

Becker, O. S. code by R. A., Minka, A. R. W. R. version by R. B. E. by T. P., & Deckmyn, A. (2018). maps: Draw Geographical Maps. https://CRAN.R-project.org/package=maps