

Research Assessment Exercise 2020

Impact Case Study

University: The University of Hong Kong (HKU)

Unit of Assessment (UoA): 29 - Communications and Media Studies

Title of case study: Impact of Social Media Analytics on Journalistic Practice and Think Tank's Research about China

(1) Summary of the impact

While China has rapidly grown in international influence, global society requires reliable information to make informed decisions about China. However, local and international news media and the Internet within the country are subject to vigorous government control, so access to genuine and uncensored information about China is a key challenge. Weiboscope and WeChatscope are aimed at reinterpreting China through the lens of social media. The projects make censored contents publicly accessible, highlight trending matters, and consequently change the processes through which journalists and China-watchers gather information about, and conduct research into, this global power.

(2) Underpinning research

In 2011, Weiboscope was launched by the Journalism and Media Studies Centre (JMSC) at HKU to collect, analyze, and visualize data from Weibo, the Chinese equivalent of Twitter and one of the largest social media platforms in China. An automatic data-collection mechanism was built to routinely track over 100,000 influential social media accounts. A set of social media analytics tools was developed to conduct Chinese text analysis and extract trending patterns, as well as to compare each account's contents every 15-20 minutes to detect 'disappeared' contents. The project has gathered and examined large-scale Chinese social media contents for a variety of applications: emerging event monitoring, public opinion analysis, Internet policy analysis, and censorship detection. As of February 23, 2019, the project has collected over 810 million Weibo posts, within which more than 700,000 posts have been confirmed as censored material.

Using this comprehensive data record, the research team, led by JMSC Associate Professor Dr. King-wa Fu, who joined HKU in 2004, has addressed a set of substantial theoretical and methodological challenges. Here are some examples:

- Develop an innovative random sampling scheme to representatively characterize a “winner-takes-all” distribution of the social media users (3.1);
- Build a quantitative indicator to measure the extent of Internet censorship in China and reveal China's response to public incidents, e.g., the 2014 Umbrella Movement (3.2);
- Demonstrate a theoretical argument that the agenda of grassroots movements is set via social media despite the state's content regulation policy (3.3);
- Establish a useful measure of public awareness and reaction to disease outbreak information released by health authorities (3.4);
- Illustrate a rigorous analysis to examine the role of social media in the exposure of public controversy, e.g., Chinese officials' corruption (3.5);
- Evaluate the Chinese Internet platform's rumor management and censorship policy during a critical period, such as a disaster (3.6).

Building on the Weiboscope, WeChatscope was introduced in early 2018. It is an extension of the previous project's social media content collection function to cover the largest social media provider in China at the time, WeChat public accounts. Using an innovative “App crawling” technique, collecting WeChat public accounts' contents directly from smartphone application's user interface, a panel of several thousand significant WeChat public accounts was selected and their

contents tracked and analyzed to detect whether or not the posts were removed or censored. The findings suggest that, besides posts related to collective action or significant events potentially leading to state instability, China's Internet censorship has recently extended to more subject areas that were not previously deemed as politically sensitive, e.g., the China's economy, the sexual assault of an individual, and a medical product scandal. WeChatscope also provides data visualization and data analytics on its website and the research data are shared with the public via free access to a self-developed Application Program Interface (API).

(3) References to the research

3.1 Fu, K. W., Chan, C.H., & Chau, M. (2013). Assessing Censorship on Microblogs in China: Discriminatory Keyword Analysis and the Real-Name Registration Policy. *IEEE Internet Computing*, 17(3), 42-50.

3.2 Fu, K. W., & Chau, M. (2013). Reality Check for the Chinese Microblog Space: A Random Sampling Approach. *PLoS ONE*, 8(3), e58356.

3.3 Fu, K. W., & Chau, M. (2014). Use of Microblogs in Grassroots Movements in China: Exploring the Role of Online Networking in Agenda Setting. *Journal of Information Technology & Politics*, 11(3), 309-328.

3.4 Fung, I. C.H., Fu, K.W., Ying, Y., Schaible, B., Hao, Y., Chan, C. H., & Tse, Z. T. H. (2013). Chinese social media reaction to the MERS-CoV and avian influenza A(H7N9) outbreaks. *Infectious Diseases of Poverty*, 2(1), 31.

3.5 Nip, J. Y. M., & Fu, K.W. (2016). Challenging Official Propaganda? Public Opinion Leaders on Sina Weibo. *The China Quarterly*, 225, 122-144.

3.6 Zeng J., Chan C.H. and Fu K.W., (2017). How Social Media Construct "Truth" Around Crisis Events: Weibo's Rumor Management Strategies After the 2015 Tianjin Blasts, *Policy & Internet*. Volume 9, Issue 3, Pages 297-320.

(4) Details of the impact

4.1 Global Reach and Open Data

The projects are well recognized internationally, serving as key resources for journalists, researchers, and the general public (including users in Mainland China, even though the project websites are blocked in its territory). Since mid-2014, Weiboscope has generated 85,675 page views from more than 34,502 worldwide users. According to WeChatscope's website survey conducted in 2019 (5.1), among the 557 respondents, Mainland China (43.7%), North America (19.6%), Hong Kong (15.1%), and Europe (12.4%) are the top four main sources of users.

The projects offer data to help global journalists and China-watchers understand online political and social discussions in the country through data analysis, visualization, and open data access. Based on the survey, academic, and students (72.6%), media and journalists (12.5%), civil society organizations and NGOs (3.2%), and governments (1.4%) are the four major user sectors (5.1). The website visitors find the censored data visualization in geographical information (65.7%), wordcloud presentation (36.9%), and trending people (28.9%) the most useful features of the site (5.1).

Updated information is regularly shared on global social media. As of August 2019, 2,544 and 747 Twitter users are following Weiboscope and WeChatscope respectively. The projects have established international data-sharing strategic partnerships with world-renowned human rights-focused think tank The Freedom House, with the in-depth analysis-oriented China Media Project, and with the awarding-winning Internet freedom project GreatFire.org (See below testimonials).

The projects engage actively with international and local media and have helped to make new journalistic practices possible (See below testimonials). For example, in the week when

WeChatscope's 2018 annual report (a yearly summary of key instances of Internet censorship) was published, the report was widely covered by CNN International TV channel, CNBC, BBC, Nature, TIME, and South China Morning Post (5.2). In the Factiva database, 300 entries mentioning "Weiboscope" have been found since 2012 (5.8).

When The Economist cited Weiboscope's findings in October 2014 to illustrate China's online censorship of Hong Kong's "Occupy Movement", the article said, "...the data were compiled by Weiboscope, a censorship-monitoring programme at the University of Hong Kong. Its freedom to pursue such research and publish its findings is one more testament to the rights that the Hong Kong protesters are determined to defend" (5.3). Weiboscope released 1,246 censored Weibo images related to the 1989 Tiananmen Square Crackdown to mark its 30th anniversary (5.9) and the censored pictures were reported in many major international and local media outlets. TIME online reported, "...Beijing's Tiananmen Square remains one of the most censored topics on the Chinese Internet. China's censorship apparatus block all mentions of the event.. But Hong Kong, a semi-autonomous Chinese enclave, lies outside the Firewall, and researchers here were able to archive many posts before they were deleted" (5.4).

The projects make public data sharing via file download and API. Between February and August 2019, over 1.18 million API requests were made to WeChatscope's server. Among the requests, 62% sought a summary of censored posts and 38% accessed the full text of censored posts. China researchers around the world have benefited from the dataset download. A Cornell University doctoral graduate, Dr. Christopher Cairns, who used Weiboscope in his thesis as the main data source, said, "...it is near impossible for individual researchers to scrape the amount of content collected by Weiboscope before it disappears under censors' hands. So for researchers or journalists who need quick information, it is the best, perhaps the only source of its kind to find out what is happening on Weibo" (5.5).

Dr. King-wa Fu is invited to speak on Weiboscope and WeChatscope internationally, including World Economic Forum, the Atlantic Council, and Halifax International Security Forum etc.

4.2 Impacts on Journalistic Practices

Weiboscope and WeChatscope provide one-of-a-kind information to media professionals who are looking for story ideas in China. The data have become an essential tool for journalists and markedly changed the news-gathering practices of the journalists covering China.

The change in journalistic practices is best shown by award-winning journalist Louisa Lim's experience. She says, "Before [Weiboscope and WeChatscope], it was clear that Chinese-language social media was being censored, but no yardsticks were available to quantitatively measure how the censorship process varied from day to day, or geographically in China." She gives newsgathering examples to illustrate the project's uniqueness - "...This type of quantitative measure .. would not have been possible before Weiboscope (2015 Tianjin blast)" - and its time criticality - "the first journalistic piece outlining the sensitivity of Winnie the Pooh/Xi Jinping comparison on the Chinese internet" (5.5).

Patrick Boehler, a staff writer at The New York Times who has used Weiboscope for years, also describes the data source as "an essential useful tool to find trending news stories on Chinese social media and allows me to follow political conversations that would otherwise be concealed by the authorities" (5.5). Mimi Lau, Senior China Reporter at South China Morning Post, also acknowledges Weiboscope as "one of the most important news gathering tools", "one of the essential go to source for journalists covering China seeking to measure the sensitivity of certain subjects and topics" and being "vital for enhancing the quality of our stories as Weiboscope provides an insight and practical metrics of how censors work at what time and intensity" (5.5).

Buzzfeed Media Editor Craig Silverman, who used WeChatscope data in analyzing online discussion about US policy towards China during the trade war, says "... gathering and analyzing

this data in unmatched in the field .. making him and this project uniquely important for research and reporting on WeChat and Chinese censorship” (5.5).

4.3 Impacts on Think Tanks’ Research

Another key user sector is think tanks and NGOs, who require timely and accuracy information about China to inform their activities. Sarah Cook, Senior Research Analyst for the human rights think tank The Freedom House, cites Weiboscope data regularly in the organization’s monthly publication, China Media Bulletin, which typically receive tens of thousands of views. She says, “...one indication of the potential interest by Chinese netizens in deleted images preserved by Weiboscope involved a cartoon of a rabbit with the statement “I’m Taiwanese” under it in English and Chinese, which was featured in the February 2016 issue (5.6). By chance, that issue was posted prominently on the landing page of Freebrowser, a mobile-phone circumvention tool developed by GreatFire.org. As a result, more than 5,000 users clicked on the image and read that issue of the China Media Bulletin...” (5.5).

David Bandurski, Co-director of China Media Project, has used Weiboscope for years and is developing a new, Berlin-based European platform, Echowall, to comprehensively monitor the developing China-Europe relationship. He applauds the contribution of Weiboscope and WeChatscope to his work: “... an entirely original approach to understanding not just the evolving conversation on Chinese social media, but also how it is shaped and constrained by the state.” Another Berlin-based think tank, Mercator Institute for China Studies (MERICS), cites Weiboscope extensively in its report “The Party does not yet rule over everything” (5.7), and says that the project’s “comprehensive and top-quality documentation of the filtering of postings on Sina Weibo is indispensable. We greatly benefit from this project” (5.5). NGOs in China are also beneficiaries of the project. A worker in a China NGO who asked to be anonymous said it was “...helpful to our work by providing back up of our censored contents in accessible and reliable hyperlinks for continuous dissemination [Original in Chinese]” (5.5).

5) Sources to corroborate the impact

5.1 WeChatscope user survey results:

https://github.com/fukingwa/RAE2020_JMSC/blob/master/wechatscope_survey_result.jpg

5.2 Media coverage on WeChatscope: <https://wechatscope.jmcs.hku.hk/media-coverage/>

5.3 The Economist (September 30, 2014), “Where street meets tweet”, url:

<https://www.economist.com/graphic-detail/2014/09/30/where-street-meets-tweet>

5.4 TIME (April 17, 2019), “The Tiananmen Massacre is One of China's Most Censored Topics. Here's a Look at What Gets Banned”, url: <https://time.com/5571372/>

5.5 Testimonials for Weiboscope and WeChatscope:

https://github.com/fukingwa/RAE2020_JMSC/tree/Testimonials

5.6 China Media Bulletin Issue No. 112: February 2016. url: <https://freedomhouse.org/china-media/china-media-bulletin-issue-no-112-february-2016>

5.7 Kristin Shi-Kupfer and Mareike Ohlberg (November 29, 2018), “The Party does not yet rule over everything: Assessing the state of online plurality in Xi Jinping’s “new era””. url:

https://www.merics.org/sites/default/files/2018-11/181029_Merics_Plurality-of-Online-Discussions_english_final_web.pdf

5.8 Factiva (via HKU Library):

https://github.com/fukingwa/RAE2020_JMSC/blob/master/factiva_weiboscope_result.jpg

5.9 Weiboscope’s release of Tiananmen Square crackdown censored pictures:

<https://www.instagram.com/64censoredpics/>