## Terms and Application Form of using Weibo-COV

Beijing Institute of Technology & University of Science and Technology of China

Contacts: Yong Hu, Anfan Chen (email: <u>huyong@bit.edu.cn</u>; <u>caf16@ustc.edu.cn</u>)

**Weibo-COV**: Weibo-COV is a large-scale COVID-19 social media dataset, covering 40,893,953 tweets from 1 December 2019 to 30 April 2020. We build this dataset based on 20 million Weibo active user pool, and first collect a total of 692,792,816 original tweets posted by users in the specified period, and then filter these tweets by selected keywords. More detail can be accessed at: <a href="https://arxiv.org/abs/2005.09174">https://arxiv.org/abs/2005.09174</a>. In addition, to promote extensive and in-depth researches, we also release all original tweets with GEO tag from original 692,792,816 tweets without COVID-19 related keywords filtering, named Weibo-COVALL-GEO-45M, covering 45,902,071 tweets.

## Terms of use:

**Application Form:** 

- 1. Respect the privacy of personal information of the original source.
- 2. The original copyright of all the data of the Weibo-COV belongs to users on Weibo and Weibo platform. Yong Hu et al. collect, organize, filter them. Weibo-COV is free to the public for only academic research.
- 3. The dataset is only for the specified applicant or study groups for research purposes. Without permission, it should not be used for any commercial purposes.
- 4. If you want to use the dataset for further research, data providers, Yong Hu et al. should be identified in your results.
- 5. If the terms changed, the latest online version shall prevail.
- 6. If you want to use the corpus, please download the application form, fill sign it respectively, then e-mail the scan version to Yong Hu and Anfan Chen (email: <a href="https://huyong@bit.edu.cn">huyong@bit.edu.cn</a> and <a href="mailto:caf16@ustc.edu.cn">caf16@ustc.edu.cn</a>)

## 

Signature : Date: