# Pu ZHANG

*Updated November 12, 2024*

**Email**: pzhang012@connect.hkust-gz.edu.cn **HomePage** pzhang.cn **Citizenship**: Chinese

## Research Interests

**Social Media Research, Risk Communications, Computational Social Science, Disaster Risk Perception, LLMs for Social Science**

## Education

**The Hong Kong University of Science and Technology (Guangzhou)** Guangzhou, China
Red Bird Mphil in Innovation, Policy and Entrepreneurship
Supervisors: Corey Kewei XU and Jing TANG Sep 2023 – Jun 2025

**China Agricultural University** Beijing, China
B.Mgt in Regional Development in Rural Areas *GPA: 3.51/4.0*
Supervisor: Associate Professor Feng KONG Sep 2019 – Jun 2023
B.Sc. in Data Science and Big Data Technology *GPA: 3.47/4.0*
Mentor: Director Hui Li Sep 2021 – Jun 2023

## Publications

**Pu ZHANG**, Hao ZHANG, Feng KONG*. Study on the Evolution of Online Public Opinion and Government Response Strategies for the "7-20" extraordinary rainstorm and flooding disaster in Zhengzhou, China. *Natural Hazards*, (2024) doi.org/10.1007/s11069-024-06904-7, **(SCI Q2, IF: 3.7 )**

**Pu ZHANG**, Hao ZHANG, and Feng KONG*. Research on online public opinion in the investigation of the "7–20" extraordinary rainstorm and flooding disaster in Zhengzhou, China. *International Journal of Disaster Risk Reduction*, 105 (2024): 104422. doi.org/10.1016/j.ijdrr.2024.104422 **(SCI Q1, IF: 5.0 )**

**Pu ZHANG**, Hao ZHANG, and Feng KONG*, Yulong KONG. A study on public opinion characteristics of rainstorm flooding disasters based on Sina Weibo data: take the three rainstorm flooding disasters in China in 2021 as an example. *Water Resources And Hydropow Erengineering*, 54.02(2023): 47-59. doi:10.13928/j.cnki.wrahe.2023.02.005. **(In Chinese)**

**Pu ZHANG**, Corey Kewei XU*. How effective is the Shorts Transformation of Traditional Media? An analysis from the perspective of user-generated content. *ChineseCSCW 2024*. (Conference Article, **In Press**)

## Working Papers

**Pu ZHANG**, Zheng WEI, Junxiang LIAO, Changyang HE*. Cultural Narratives and Sentiment Engagement of Game Discourse: Comparing Black Myth: Wukong Discussions on Douyin and TikTok. Submitted to *ACM CSCW 2025*.

This study focuses on user discussion differences of the game "Black Myth: Wukong" on Douyin and TikTok platforms. This study analyzed 200K comments from Douyin and 100K comments from TikTok, utilizing LLM-based Few Shot Learning for classification. The study highlighted the differences in user discussions due to cultural variations, providing insights for future game design and the international expansion of Chinese games.

**Pu ZHANG**, Feng KONG*. Online Public Opinion Analysis of Major Infrastructure Disasters: A Case Study of the 2024 Guangdong Highway Landslide Incident in China. Submitted to *Natural Hazards Review*, **under review**.
This study is focus on user disaster risk perception of the Guangzhou Meida Highway collapse accident on Douyin platform. This study utilized Bertopic for topic modeling and BERT pre-trained model for sentiment analysis, analyzing approximately 50K related comments on Douyin videos. By leveraging ArcGIS for geographical visualization and incorporating Chinese statistical data along with IP geolocation of comments, I elevated the analysis from descriptive to correlational.

**Pu ZHANG**, Feng KONG*. A Study on the Characteristics of Online Public Opinion During the 2023 Jishishan Earthquake. Submitted to *Journal of Natural Disasters* (In Chinese), **under review**.
This study is focus on online public opinion of the Jishishan earthquake disaster. This study utilized BERT pre-trained model for sentiment distribution analysis, TF-IDF for keyword extraction, and Gephi for visualization, combined with manual topic identification. Analyzed approximately 12K user comments.

## Workshop

**Tsinghua Big Data and Causal Inference Seminar** Oct 2023 – Jan 2024
Organized by Tsinghua University, this workshop covers computational social science methods, including text analysis, social network analysis, and double-differencing.

**The International GeoInformatics Summer School** Jun 2024
IGSS 2024 Social Computing Summer School at Wuhan University, focusing on integrating GIS and using social media data for disaster risk perception research.

**ICSC 2024 International Conference on Social Computing** Aug 2024
This workshop provided an opportunity for scholars to exchange advancements in social computing.

## Research Experience

**Research on Internet Public Opinion of Emergency Events Based on Natural Language Processing**
Supervisor: Feng KONG Sep 2022 – Present

I collected social media data from Sina Weibo using a Python-based web crawler, fine-tuning a BERT model to conduct sentiment analysis and generate visualizations. For thematic analysis, I utilized Gephi to perform social network analysis and visually represent the results. Based on these findings, I assessed public opinion trends and provided governance recommendations to inform policy and decision-making.

**Enhancing Urban Resilience through AI: Modeling, Simulating, and Mitigating Catastrophic Risk Scenarios**

Supervisors: Corey Xu and Jing TANG Sep 2023 – Present

In this research project, I utilize social media data to analyze public risk perception in major sudden natural disaster scenarios. Using large language models, I conduct sentiment analysis and thematic analysis to assess the online public opinion characteristics related to these events. I have extensive experience in collecting and processing social media data from multiple platforms, including over one million data points from Douyin, more than two million from TikTok, and additional data from other sources. Using advanced web crawlers, I efficiently gather diverse datasets, and my strong proficiency in data handling allows me to manage and analyze large-scale social media information seamlessly. I leverage large language models (LLMs) with few-shot learning for text classification, which enables me to conduct in-depth, multilingual sentiment and thematic analyses across various platforms. Additionally, I am skilled in using tools like BERTopic for sophisticated topic modeling, which allows me to provide high-accuracy insights into public sentiment and online opinion trends. I am also proficient in visualization software such as ArcGIS and Gephi, enabling me to create detailed spatial and network visualizations. By integrating Chinese statistical data and performing statistical analyses based on the IP geolocation of comments, I advance my research from descriptive to correlational analysis. This approach provides a deeper understanding of regional differences in public opinion, allowing me to identify patterns and relationships between thematic features and demographic factors.

## Teaching Experience

**Teaching Assistant, HKUST (GZ)** Fall 2024

IPEN 5250: Text Analysis and Machine learning

## Selected Honors and Scholarships

| | |
|---|---|
| National Inspiration Scholarship | 2022 |
| China Telecom Scholarship | 2021 |
| Beijing Challenge Cup Second Prize | 2022 |
| Red Bird MPhil Postgraduate Scholarship | 2023-2025 |

## Skills and Software

Python, R, LaTeX, Gephi, Arc GIS, VosViewer

## Languages

Chinese Mandarin (native), English (TOEFL 106, July 2022)