

## Introduction to Big Data-Driven Social Media Management Minitrack

Xiangbin Yan

University of Science and Technology Beijing, China  
[xbyan@ustb.edu.cn](mailto:xbyan@ustb.edu.cn)

Mingxin Gan

University of Science and Technology Beijing, China  
[ganmx@ustb.edu.cn](mailto:ganmx@ustb.edu.cn)

### 1. Introduction

Big data-driven social media has always been an important medium for linking individual behaviors and community activities. Especially since the beginning of 2020, the global pandemic of COVID-19 has caused tremendous changes in communication, working, and entertainment behaviors. Many real-world activities turned to online. This has led to the rapid penetration of big data-driven social media into various industries, as well as some challenges. On the one hand, how to discover the potential knowledge from the accumulated massive data has become an important topic for promoting the sustainable development of social media management. Big data-driven social media should not only focus on individual activities and business behaviors, but also have knowledge discovery methods based on behavior analysis and data analysis. On the other hand, it is necessary to discover and explore practical problems through big data-driven social media, and give reasonable suggestions and measures, forming positive feedback to real-world industry to promote sustainable economic development.

This year, three accepted papers explored how to discover potential and valuable knowledge in social media from a big data-driven perspective. They cover the fields of economics management, entertainment and scientific research, involving the techniques in text analysis, sentiment analysis and network analysis. Topics are focused on Big data-driven impact analysis of remote work on economic, Big data-driven factor analysis and industry advice for film quality and Big data-driven knowledge discovery on scientific research network. The introductions of the papers are as below:

### 2. Big data-driven impact analysis of remote work on economic

**Title:** Data-driven Analysis of Remote Work in China during the COVID-19 Pandemic

Telecommuting has become popular since February 2020 primarily due to the pandemic, and people have been slowly returning to their office from May 2020. A paper in the minitrack leverages online content to investigate the biggest impact of COVID-19 – remote work, by using China as a primary case study. The authors focus on two time windows in the year 2020 to calculate the growth of different job sectors. The results indicate the negative impact of teleworking in manufacturing industry, but shows that information technology-related industries are less affected by working from home. The authors also investigate the impact of COVID-19 on the stock market and discussed what plan of action the policy makers should take to provide a good economic environment for the country. In addition to the overall economic situation, this paper revealed how the psychological situation of employees could affect their job performance, indirectly affecting the development of certain industry sectors. Therefore, misinformation in certain Chinese social media channels was also studied in this paper specifically examining the rumors and their latent topics. This work may initiate a dialogue between scientists, policy makers and government officials to consider the observations highlighted in this paper.

### 3. Big data-driven factor analysis and industry advice for film quality

**Title:** Exploring the Influencing Factors of IP Film Rating by Sentiment Analysis and GMM

Recently, intellectual property (IP) film has become an important accessory for entertainment, and its rating has become the focus of quality evaluation. However, existing research seldom conducts study on influencing factors of rating. To explore the factors that affect IP film rating, a paper in the minitrack uses sentiment analysis and generalized method of moments (GMM), taking advantage of production, broadcast, genre and

audience feedback to construct six explanatory variables, including actor influence, screenwriter participation, broadcast time, broadcast platform, genre, and adaptation satisfaction. The authors use LLC, IPS and Sargan tests to conduct variable stability test and model setting test. From the regression results of 134 IP films that obtained by sample filtering, the impact of each influencing factor on the rating is obtained. Empirical study found that short-term historical rating, actor influence, adaptation satisfaction and screenwriter participation positively affect current rating. While, long-term historical rating has a negative impact on current rating. In addition, broadcast time and broadcast platform have imposed positive impact on IP film rating, and genre has only a weak impact on rating. This work may provide advice for IP film producers, prompting them to improve quality by emphasizing celebrity effects and author participation.

#### **4. Big data-driven knowledge discovery on scientific research network**

**Title:** IRIS: Learning the underlying information of scientific research interests using heterogeneous network representation

Understanding scientific research fields and finding potential relations between seemingly distinct fields can help researchers rapidly grasp their most interested topics with expertise. To represent both author and keyword nodes in heterogeneous research network, a paper in the minitrack built an “Integrated Research Interest Space (IRIS)” to construct a heterogeneous network which contains authors, keywords, papers and institutions. Some interesting observations were discovered: similar keywords in the sense of research interest and research manner can obvious aggregate together. Authors that are interested in different keywords distributed in different IRIS areas, with strongly associated with research objectives and methodologies of the keywords. The average similarities between authors and their real used keywords are significantly higher than that of randomly chosen author-keyword pairs. On this basis, the authors propose a simple algorithm which attempts to recommend potential interested keywords for researchers, and got meaningful results. The study on a dataset downloaded from PubMed may also give useful hints for understanding research interests and discovering potential cross disciplines.