Developing a Quantitative Instrument for Positive Publicity Effects: Guangdong Foshan's Innovative Application of Media Big Data to Launch an Online Positive Energy Index –Postprint

Authors: Fan Chunbai

Date: 2023-10-08T00:00:00+00:00

Abstract

Is there a way to conduct a relatively comprehensive statistical analysis of positive publicity reports about a certain entity on the internet? Can the effectiveness of positive cyberspace energy dissemination be quantified? To address the above needs, the Foshan Cyberspace Security and Informatization Association has innovatively developed a network positive energy index system utilizing media big data, artificial intelligence, and machine learning, conducting index evaluations for over 270 Party and government organs throughout the city. This provides scientific data support and robust information backing for further grasping the patterns of online communication, strengthening the timing, measure, and effectiveness of positive cyberspace energy dissemination, and enabling municipal, district, and township levels to carry out positive energy dissemination work and jointly foster a clear and bright cyberspace.

Full Text

Building a Quantitative Instrument for Positive Publicity Effects: Foshan, Guangdong Innovatively Applies Media Big Data to Launch a Network Positive Energy Index

Fan Chunbai
Foshan Cybersecurity and Informatization Association, Foshan, Guangdong 528000

Abstract: Can we develop a method to comprehensively 统计 positive publicity reports about an organization across the internet? Can we quantify the effectiveness of online positive energy dissemination? Addressing these needs, the

Foshan Cybersecurity and Informatization Association has innovatively developed a network positive energy index system using media big data, artificial intelligence, and machine learning to evaluate more than 270 Party and government agencies citywide. This system provides scientific data support and robust information assistance for municipal, district, and township-level units to conduct online positive energy dissemination and jointly cultivate a clean cyberspace, while furthering our understanding of network communication patterns and enhancing the timeliness, scope, and effectiveness of positive energy propagation.

Keywords: network positive energy index; network communication patterns; media big data; machine learning; weight

Chinese Library Classification: G641 **Document Code:** A **Article ID:** 1671-0134(2021)08-029-03 **DOI:** 10.19483/j.cnki.11-4653/n.2021.08.006

Citation Format: Fan, C. Building a Quantitative Instrument for Positive Publicity Effects: Foshan, Guangdong Innovatively Applies Media Big Data to Launch a Network Positive Energy Index [J]. China Media Technology, 2021(08): 29-31.

For many Party and government agencies, publicity work is critically important —not only because an organization's achievements require public recognition, but also because it reflects how effectively they implement the mass line online. After conducting positive energy activities and initiatives, many organizations remain unclear about which media outlets published what content, often resulting in incomplete statistics. This is partly because some news articles are difficult to find through Baidu, and partly because dissemination channels have multiplied: PC terminals represent just one outlet, alongside mobile apps, WeChat, Weibo, and platform-hosted accounts. Is there a way to comprehensively 统计 positive publicity reports about a given organization across all these channels?

Furthermore, thoroughly implementing Xi Jinping Thought on Socialism with Chinese Characteristics for a New Era, particularly General Secretary Xi Jinping' s important thoughts on building China into a cyber powerhouse [1], and advancing the "Network Positive Energy Dissemination Enhancement Project" under the theme of "Gathering Positive Energy Online" constitutes vital work for cyberspace administration offices at all levels. To assess the effectiveness of online positive energy dissemination and identify which organizations perform well requires a quantifiable metric.

Defining Network Positive Energy Content

First, we must clearly define what constitutes online positive energy content. This includes relevant media reports that play a guiding role in disseminating positive energy during major policy implementations, significant themes, important activities, major events, critical time nodes, hot-button issues, and emer-

gencies—reports that advocate for netizens to consciously spread and promote positive energy to foster a clean cyberspace, as well as content that helps promote socialist core values. In simple terms, any "widely socially relevant event reported through media and online channels that exerts a positive promoting effect" qualifies as positive energy content. However, routine functional work of the organization or department—including study tours, meetings, seminars, and policy releases—does not count.

Using internet data collection technology, we gather the latest relevant online information into our database, then employ semantic analysis combined with keyword libraries and other rules to aggregate all Foshan-related positive energy content for subsequent data calculation. The process involves five steps: (1) through backend configuration, establish targeted collection tasks; (2) for different target data sources, configure collection according to content tags, using either API data interfaces or regular expression-based page content decomposition; (3) schedule collection tasks to synchronize with target sites for incremental updates; (4) after data collection, complete heterogeneous-to-homogeneous conversion; and (5) feed the collected content into the next analysis stage.

Under the guidance of the Foshan Municipal Cyberspace Administration Office and in response to these dual needs combined with media big data applications, the Foshan Cybersecurity and Informatization Association innovatively developed a network positive energy index system [2] to evaluate Party and government agencies at the municipal, district, and township levels. The index evaluation system received its software copyright certificate in November 2019. After nearly two years of operation, the system has fully leveraged its indicator and guidance functions, providing scientific data support and robust information assistance for municipal, district, and township units to conduct online positive energy dissemination, jointly cultivate a clean cyberspace, and further grasp network communication patterns while enhancing the timeliness, scope, and effectiveness of positive energy propagation. This innovative approach has also received recognition from both the Guangdong Provincial Cyberspace Administration Office and the central Cyberspace Administration Office.

2. How is the Index Value Calculated?

After acquiring relevant content, we must analyze and process it to identify Foshan-related positive energy material from the crawled information. The analysis layer comprises two main components: page element analysis and relevance analysis.

2.1 Page Element Analysis

Once pages are collected, we extract links, metadata, body text, titles, and summaries for subsequent filtering and processing. Title extraction, for example, is illustrated in Figure 1 [Figure 1: see original paper]. To explain the network positive energy index algorithm, we must first define a key concept: "coverage

count" is calculated based on article "node appearances." If one article is published on Media A, B, and C, with Media A recommending it three times across different columns and B and C each recommending it once, the coverage count equals five.

2.2 Relevance Analysis

First, we perform word segmentation on collected content, extract high-frequency keywords, and preprocess them for database entry (building a vocabulary reserve for subsequent "keyword learning"). After segmentation, content keywords are compared against preset keywords; if they match the rules, the content is identified as target material—Foshan-related positive energy content—and entered into the database. The preset keyword rule model supports three combination states: "AND," "NOT," and "OR," requiring simultaneous appearance of two or more keywords (e.g., landmark [Foshan] + department [Cyberspace Administration Office or Civilization Office or···]), or excluding certain keywords.

The network positive energy index system consists of three components: crawling, analysis, and presentation. The public sees the presentation layer—rankings and specific values of each organization' s network positive energy index. Top-ranked organizations also receive concise data analysis to guide their future positive energy dissemination efforts (Figure 2 [Figure 2: see original paper]).

2.2.1 What is Weight? Weight is a crucial metric in index calculation. We establish corresponding weights based on different departmental levels to achieve fairer and more impartial scoring. According to the administrative levels of the city' s five districts and various township units, we assign different weights to corresponding media tiers. Data sources are divided into six levels: core central media, key central media, provincial media, municipal media, district media, and township media, encompassing websites, apps, WeChat public accounts, and WeChat channels.

The data source classification is as follows: - **Core central media:** People' s Daily Online, Xinhua Net, CCTV.com, China National Radio Online, etc. - **Key central media:** Guangming Online, China Youth Online, Global Times Online, China.org.cn, China News Service, etc. - **Provincial media:** Nanfang Net, Southern Metropolis Daily (Oeeee), Nanfang+ (homepage), The Paper, Jiemian, Beijing Time, Cover News, etc. - **Municipal media:** Foshan Daily (official website), TV station (official website), radio station (official website), Toutiao provincial channel, Nanfang+ (channel, local recommendations), etc. - **District media:** District TV stations, Toutiao (city, hosted accounts), portal site (Sina, Tencent, NetEase, Sohu) city channel reprints, hosted accounts, etc. - **Township media:** Community newspapers, township (street) official websites and public accounts.

The weight hierarchy from highest to lowest is: core central media > key central

media > provincial media > municipal media > district media > township media, as shown in Figure 4 [Figure 4: see original paper].

For example (Figure 3 [Figure 3: see original paper]), when the data pool contains substantial information, we can use this data for standard quantification —this forms our index algorithm model. Corresponding weights differ for media at various levels and units at various levels. Therefore, under equivalent conditions, it is relatively more difficult for township and district-level units to obtain coverage from central media compared to municipal-level units, necessitating correspondingly higher weights.

After several rounds of correction and machine learning, the system' s content hit rate reached over 90% by the end of the trial run, meaning more than 90% of reports 统计 from media at all levels qualified as positive energy coverage. The "Foshan Network Positive Energy Index" has gradually attracted broader attention, and in late October 2019, the "Network Positive Energy Index Evaluation System" received a software copyright certificate from the National Copyright Administration (Figure 6 [Figure 6: see original paper]). Currently, the Foshan Network Positive Energy Index conducts targeted monitoring of 71 municipal-level units, 202 district-level units, and 32 township (street) units.

2.2.3 Media Weight Also Varies by Column According to calculation rules, media weight is based on homepage, rolling, and push placements. Content appearing in secondary or tertiary columns has its weight correspondingly downgraded to provincial or municipal level. In the 统计 rules, "Study Xi" platform is classified as key central media, but if published content appears on hosted accounts or local columns, its weight is calculated at the municipal level.

3. Manual Correction and Machine Learning

The "Foshan Positive Energy Index" has been released for nearly two years. The transition from "probationary period" (trial operation) to "formal implementation" took over six months, during which we continuously solicited feedback from various departments and units, constantly revising system parameters and filling gaps to ensure fairness and impartiality.

Keyword rule design constitutes the most critical system component, affecting content acquisition accuracy, completeness, and weighting. Therefore, after keyword rule configuration, we optimize through manual correction and machine learning [1]. Phase one involves a fixed model: manually selecting corresponding keywords for the system to automatically acquire content according to fixed algorithms while building a vocabulary library. This phase prioritizes broad acquisition but with relatively low accuracy. Phase two involves the system automatically identifying "confirmed" and "candidate" content, with manual screening of candidate material that is then fed back to the system, which automatically optimizes output based on evaluation. After extensive training over long periods with large datasets, output precision progressively increases.

The index' s reflection of societal hotspots and the city' s major events over the past year shows high correspondence (Figure 5 [Figure 5: see original paper]).

References: [1] People' s Daily. "Party Building Network Micro-platform" WeChat Public Account. Xi Jinping on Building China into a Cyber Power [EB/OL]. (2019-03-24) http://politics.people.com.cn/GB/n1/2019/0324/c1001-30991900.html.
[2] Foshan Daily Official Account. Release of 2020 Foshan Network Positive Energy Index [EB/OL]. (2021-01-20) http://epaper.fsonline.com.cn/fsrb/html/2021-01/20/content_{{{35955}}}_{{185146}}}.htm.
[3] Li, N. On the Impact of Big Data on News Communication [J]. China Media Technology, 2020(12): 20-22.

Author Bio: Fan Chunbai (1974-), male, from Foshan, Guangdong, is Vice President and Secretary-General of the Foshan Cybersecurity and Informatization Association, member of the Foshan Municipal Political Consultative Conference, Vice President of the Foshan New Social Stratum Federation, and a journalist. His research focuses on network information security.
(Editor: Zhang Xiaojing)

Note: Figure translations are in progress. See original paper for figures.