Proceedings of CONF-SPML 2026 Symposium: The 2nd Neural Computing and Applications Workshop 2025
DOI: 10.54254/2755-2721/2026.TJ29488

Application of Data Visualization in Social Media Analysis

Yueyao Wang

School of Economics, Qingdao University, Qingdao, China
13655360037@163.com

Abstract. With the rapid development of social media, massive user-generated data continues to emerge, making it increasingly challenging for enterprises and researchers to extract effective insights from unstructured information. Data visualization, as a key technology that transforms complex data into intuitive graphics, plays a vital role in addressing this issue. This study adopts a combined approach of theoretical analysis and case research to systematically explore the application of data visualization in social media analytics. It outlines basic types and commonly used tools of data visualization suitable for social media scenarios. Subsequently, through specific cases involving international sports brands, chain restaurants, and cosmetics companies, it conducts an in-depth analysis of three core application areas: user behavior analysis, public opinion monitoring and crisis management, and marketing effectiveness evaluation. The findings indicate that the proper use of visualizations such as network diagrams, time series charts, and heat maps can accurately identify opinion leaders, predict user activity patterns, and quantify marketing return on investment (ROI). This research provides practical references for enterprises to enhance social media analysis efficiency and formulate data-driven decisions.

Keywords: data visualization, social media analysis, user behavior, public opinion monitoring, marketing effect

1. Introduction

In the digital age, social media platforms like TikTok, YouTube, and Facebook have become crucial channels for information dissemination and user interaction, generating over 500 million new data entries daily. However, these datasets are characterized by massive volume, high redundancy, and rapid updates, rendering traditional text-based analytical methods inefficient and prone to errors, leading to a paradox of "data deluge with scarce insights." Current research in social media analytics primarily focuses on optimizing data mining algorithms, while lacking in-depth exploration of how to visually present analytical results to support decision-making.

This study investigates the most suitable data visualization types for various social media analysis scenarios and explores how visualization technology can enhance the accuracy and efficiency of user insights, public opinion response, and marketing evaluation. The research methodology combines literature review with case studies. Using literature review and case studies, it identifies mainstream visualization tools and validates their practical effectiveness. Theoretically, it enriches

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.TJ29488

scenario-based visualization research; practically, it provides enterprises with operable visualization solutions to transform abstract social media data into implementable strategies.

2. Theoretical overview

Data visualization refers to the presentation of data through charts, graphs, and other visual formats. Visualization technology provides users with interactive exploration tools and data analysis methods that are more intuitive [1]. It helps people better grasp information from data and offers meaningful support for making informed decisions.

2.1. Basic types of data visualization

Data visualization is a crucial tool in social media analytics, encompassing various chart types such as bar charts, line graphs, pie charts, scatter plots, and radar charts. These visualizations are designed for specific application scenarios and help convey relationships between data formats [2].

Trend analysis primarily visualizes data patterns over time, with line charts and area charts being the most commonly used formats. These visualizations effectively illustrate temporal changes in metrics like topic popularity, follower growth, and post publication frequency. For example, a line chart can clearly show fluctuations in brand mentions before and after marketing campaigns, helping analysts pinpoint critical events such as "peak mentions at 8 PM on campaign day."

Comparison and distribution charts are used to analyze numerical differences between categories or observe data distribution patterns. Bar charts are ideal for comparing interactions (such as shares, comments, likes) across different posts or user groups. Pie charts or doughnut charts effectively illustrate proportional relationships, like geographic distribution of social media users. Box-and-whisk charts visually demonstrate data dispersion, helping identify outliers such as "abnormally high interaction counts for individual posts."

The relational network type is the most distinctive visualization category in social media analytics, representing users as nodes and their interactions as edges. This force-directed algorithmic layout helps reveal group structures and core opinion leaders. Point distribution maps or heat maps visualize the geographic density of user engagement and topic discussions, revealing the concentration of discussions in the city center area.

Text content analysis tools process massive text data (e.g., posts, comments) by using word clouds to visually display keyword frequency through font sizes. High-frequency terms like "fitness" and "running" instantly reveal user interests. Advanced semantic network diagrams further reveal word co-occurrence relationships, such as the correlation between "new lipstick" and "long-lasting" or "moisturizing," helping uncover hidden thematic connections.

2.2. Data visualization tools

Data visualization tools are divided into chart library, business intelligence analysis, visualization big screen, and professional categories, catering to various industries' functional needs [2]. Entry-level and business intelligence tools provide powerful visualization capabilities through low-code or no-code interfaces. Tableau eliminates learning barriers, enabling users to create data visualizations entirely without coding. With just a mouse click, users can craft sophisticated dashboards within its interface, making it ideal for business analysts and decision-makers to conduct exploratory data analysis [3]. News organizations can also invest in developing low-code/no-code data visualization

DOI: 10.54254/2755-2721/2026.TJ29488

platforms, allowing editors to transform complex data into dynamic charts, interactive maps, or infographics without programming expertise [4].

When developers require higher customization and automated analysis processes, they typically utilize programming libraries. In the Python ecosystem, tools like Matplotlib and Seaborn provide fundamental static chart plotting capabilities, while Plotly and Bokeh support interactive chart creation. On the web front, D3.js stands out as the most powerful and flexible JavaScript library, capable of generating nearly any type of visualization—though its steep learning curve demands serious commitment. China's open-source ECharts also delivers diverse chart types and intuitive interaction features, making it a go-to choice for social media analytics projects across the country.

Specialized social media analytics platforms such as Brandwatch, Sprout Social, Hootsuite Insights, and domestic solutions like Sina Weibo Hot Topics and Qingbo Big Data integrate data collection, cleaning, analysis, and visualization capabilities. These platforms provide pre-configured models and dashboards tailored for social media analysis. For instance, Qingbo Big Data automatically generates user profile word clouds and sentiment trend charts, significantly lowering the technical threshold for users.

3. Application of data visualization in social media analysis

3.1. User behavior analysis: insight into user portrait and interaction mode

The core objective of application is to gain deep insights into target audiences by addressing key questions such as "Who are the users?", "What do they like?", and "How do they interact with the brand and other users", thereby establishing a foundation for content strategy development. Data visualization tools leverage visual analytics to dynamically display user behavior patterns in real-time, providing decision-makers with clear and intuitive evidence to optimize business strategies and enhance user experience [5]. Taking the international sports brand "X" as an example, the company sought to analyze its user demographics on an emerging short video platform's official account. The process began with collecting over 50,000 data points through the platform's API within a 30-day period, including videos, comments, and user profiles related to brand topics. Subsequently, three visualization methods were employed for comprehensive analysis.

By constructing an interactive network graph with comments and @relationships as edges, we created a user network comprising over 10,000 nodes. Through visualization using Gephi, three distinct communities emerged: a large community centered around five prominent fitness bloggers (node degree>500), a group of casual users sharing daily exercise routines (node degree <100), and a specialized community of sports equipment reviewers (node degree 100-500). This network diagram directly reveals the "opinion leaders" and interest clusters within brand fan communities.

The line chart visualizes user engagement patterns, illustrating the temporal distribution of comments and video uploads. The peak interaction periods occur after 8 PM on weekdays (averaging over 2,500 comments per hour) and between 2-4 PM on weekends (averaging over 3,000 comments per hour), while the off-peak period is between 2-4 AM (averaging fewer than 100 comments per hour). This data provides insights for brands to optimize content release schedules.

To create a user tag cloud, we extracted labels from profiles of highly active users (over 50 comments) and generated a visual representation. The most prominent tags ("fitness", "running", "yoga", "basketball") with over 1,000 occurrences were displayed in larger fonts, while secondary tags ("outdoor", "fat loss") appearing 500-1,000 times were shown in smaller fonts. This approach clearly outlines the core user group's interests through this visual analysis.

DOI: 10.54254/2755-2721/2026.TJ29488

Through the combination of the above three types of visualization, brand X not only identifies key communication nodes (fitness bloggers and equipment reviewers), but also grasps the active rules and interest preferences of users.

3.2. Public opinion monitoring and crisis management: real-time perception and rapid response

The application objective is to conduct real-time monitoring of online reputation for brands, products, or public events, enabling timely detection of negative trends, tracking dissemination paths, and providing decision support for crisis management. For instance, machine learning algorithms can predict future sales trends, data visualization tools can visually present financial ratio analysis results, and natural language processing technology can analyze public opinion trends on social media regarding the company. This approach yields richer information input, more accurate predictions, and faster decision-making support [6]. Take the well-known chain restaurant 'Y' as an example. When a branch was exposed for food safety issues, the related information began to spread rapidly on social media. The company used Tableau to build a real-time public opinion monitoring dashboard and trace the information dissemination path, effectively controlling the crisis spread. The public sentiment monitoring dashboard features three core visualizations. First is the sentiment trend line chart, which displays real-time changes in the proportion of positive, negative, and neutral emotions across online discussions. Within two hours of a crisis outbreak, the proportion of negative emotions surged from 5% to 75%, triggering a red alert. Six hours after the company issued an apology statement, the negative sentiment ratio dropped to 40%.

Secondly, the real-time topic cloud dynamically displays the most frequently discussed terms. During the initial phase (0-4 hours), "Brand Y", "XX Branch", and "food poisoning" showed the highest frequency. In the intermediate phase (4-12 hours), new terms like "management failure" and "apology" emerged, reflecting shifts in public discourse focus. By the final phase (12-24 hours), "rectification plan" and "refund" became the core topics.

The final step involves creating a geographic heat map to pinpoint locations with negative discussions. The incident-affected city shows the highest heat intensity (60% coverage), while surrounding areas show 20% coverage and other regions remain relatively cool (coverage <5%). This analysis guides companies in strategically deploying regional PR resources: focusing on the affected city and its vicinity, with dedicated teams assigned to address customer inquiries.

To trace the information dissemination path, we selected a key post with 100,000+ retweets as the initial exposure source and visualized its spread through a tree diagram. The analysis revealed that the post was first reposted by a local food blogger (1 million+ followers), then amplified by three entertainment KOLs (each with over 5 million followers) before going viral. The company contacted both the blogger and KOLs to implement corrective measures. After the blogger removed the original retweets, secondary dissemination of negative information was reduced by 30%.

With the help of this visual "command room", the corporate PR team can quickly capture key signals from massive information, assess the severity of the situation and formulate targeted countermeasures, control the crisis within 24 hours, reduce the proportion of negative emotions to 20%, and avoid further damage to brand reputation.

3.3. Marketing effect evaluation: quantifying ROI and optimizing strategies

The application goal is to scientifically measure the input-output ratio of marketing activities (such as new product release, topic challenge), analyze the influence of activities, user engagement and

DOI: 10.54254/2755-2721/2026.TJ29488

conversion effect, and provide a basis for future strategy optimization. Take beauty brand "Z" as an example. In order to promote its new lipstick, the brand launched a hashtag challenge titled "#MyExclusiveColor" on Weibo and invited five beauty bloggers to collaborate. After the campaign lasted for seven days, the brand used a multi-indicator marketing dashboard to evaluate the effect.

The core KPI card highlights the key indicators such as the total number of participants, topic reading volume, the total video playback volume, and the brand mention volume. The activity effectiveness comparison bar chart uses grouped column charts to analyze the performance of this campaign versus last month's "Lipstick Makeup Tutorial" event, along with collaborative outcomes among different bloggers. Compared to last month's campaign, this activity saw a 60% increase in total participants and a 40% rise in topic engagement. Among the five bloggers, Blogger A and Blogger B generated the highest interaction volume, while Blogger E recorded the lowest engagement level. The Sanki Flow Chart illustrates the complete user conversion path from watching the blogger's video to becoming brand followers. Blogger A generated the highest initial traffic, with the highest conversion rate from "clicking the topic page to publishing original content." This indicates that Blogger A demonstrates the highest traffic conversion efficiency, with the "fan conversion" stage experiencing the most significant user attrition.

The visualized user-generated content showcased 20 high-quality original images and video thumbnails from the event, which served as promotional material that garnered over 10,000 reposts on the brand's official account. Through this visual analytics solution, Brand Z transcended vanity metrics like "1 billion reads" by gaining deep insights into the true impact of marketing campaigns and user engagement levels.

4. Current challenges and future development

Current challenges in social media data visualization include data privacy and ethical concerns, which pose significant challenges in social media data security. For instance, evolving usage patterns may expose new privacy risks in models built on social media data [7]. These platforms contain sensitive personal information such as user locations and consumption preferences. The critical legal and ethical challenges now include implementing effective anonymization (e.g., user ID obfuscation) during data collection, analysis, and visualization processes, along with obtaining informed consent from users.

Social media data is plagued by massive amounts of spam, bot accounts, and paid content manipulation. These noise elements severely contaminate data sources, distorting visualization results. To address this, a governance approach guided by archival science principles should be implemented. This includes monitoring data archiving processes, overseeing quality control measures, and coordinating existing social media data management systems. Such efforts enhance data quality while enabling effective sharing and value realization of social media content [8]. For instance, in a brand's user profile word cloud, the frequent appearance of "discounts" through bot-generated posts led the brand to misjudge users' primary concern as price optimization, thereby neglecting their genuine demand for product quality.

Information overload and visualization-induced misinformation have made data visualization increasingly challenging to implement [9]. In China, the development of data visualization misinformation primarily occurs in news media, driven by data journalism and its visualization techniques, as well as the viral spread of such content on social media platforms [10]. Poorly designed visualizations may become new forms of "noise". Overly complex charts increase interpretation difficulty, while intentional distortion of axes or inappropriate chart types can unintentionally or deliberately mislead audiences, conveying misleading information.

DOI: 10.54254/2755-2721/2026.TJ29488

The technical and cognitive barriers of advanced customization visualization often require programming skills (e.g., Python) and data analysis expertise, which excludes many business decision-makers. Moreover, interpreting complex visualizations demands specific training and experience. A survey indicates that 60% of corporate managers find 'difficulty in understanding visualization results' as a major obstacle to decision-making.

Future development trends in data visualization will be significantly influenced by AI, which will be deeply integrated into visualization workflows to automatically identify analytical intent, recommend, and generate suitable chart types. This integration will leverage natural language processing to extract insights from charts, enabling conversational data analysis. For example, 20th Century Fox uses AI to analyze social media data for personalized marketing strategies [11]. Additionally, advancements in computing power will support real-time processing and visualization of social media data streams, allowing for minute-level responsiveness in monitoring public sentiment. VR and AR technologies will also transform data visualization by providing immersive experiences, such as overlaying data charts onto physical environments. Predictive and prescriptive analytics integrated with visual interfaces will go beyond traditional visualization by forecasting trends and simulating decision-making outcomes.

5. Conclusion

This study systematically explores the application of data visualization in social media analytics through theoretical frameworks and case analyses. Key findings include that different visualization types have distinct application boundaries in social media scenarios. Network graphs are effective for identifying opinion leaders, time series charts for analyzing activity patterns, and heat maps for visualizing geographic distribution. Second, combining multiple visualization methods (e.g., "network graph + word cloud + time series chart") can form a comprehensive insight system that outperforms single visualization approaches. Third, tools like Tableau, Gephi, and ECharts each have unique advantages, and enterprises should select tools based on their technical capabilities and analytical needs.

Future research could focus on two key directions: First, exploring the application of emerging technologies like VR/AR in social media visualization to enhance analytical immersion; Second, investigating solutions for data privacy issues in visualization, such as developing privacy-preserving visualization algorithms that maintain insight accuracy while ensuring data security. Overall, with technological advancements, data visualization will play an increasingly vital role in social media analysis, helping enterprises transform their data advantages into competitive edges.

References

[1] Zhao, Y. (2024). Research on the application practice of mobile internet data visualization. Software, 45(11), 127-129. https: //doi.org/CNKI: SUN: RJZZ.0.2024-11-027

[2] Chen, S. (2022). Research on multi-source heterogeneous data management methodology for discrete manufacturing processes based on data visualization. Dalian University of Technology. https: //doi.org/10.26991/d.cnki.gdllu.2022.1098

[3] Qian, H. (2023). Design of a LIBS-based monitoring and early warning platform with visualized information for heavy metal in cultivated soil. Northwest Normal University. https: //doi.org/10.27410/d.cnki.gxbfu.2023.000903

[4] Lu, W. (2025). Research on the transformation and development of news editors in the era of convergent media. Journalism and Communication, (17), 11-13. https: //doi.org/CNKI: SUN: YWCB.0.2025-17-003

[5] Zhang, H. (2025). Design of a data mining-based user behavior analysis system. Software, 46(01), 37-39. https: //doi.org/CNKI: SUN: RJZZ.0.2025-01-009

DOI: 10.54254/2755-2721/2026.TJ29488

[6] Zhang, Q. (2024). Building a corporate financial analysis system from a value perspective. Business Observation, 10(03), 33-36. https: //doi.org/CNKI: SUN: SYCG.0.2024-03-004
[7] Zhou, H. (2024). Platform governance of security risks in AI foundational models. Journal of Financial Law, (05), 3-22. https: //doi.org/10.16823/j.cnki.10-1281/d.2024.05.010
[8] Liu, L., & Jin, B. (2024). Archival governance of social media data: A case study of the Library of Congress's Twitter project. Zhejiang Archives, (07), 42-46+63. https: //doi.org/10.16033/j.cnki.33-1055/g2.2024.07.006
[9] Zuo, Q. (2024). Digital justice in the context of algorithmic uncertainty. Guilin University of Electronic Technology. https: //doi.org/10.27049/d.cnki.ggldc.2024.000493
[10] Hao, S. (2024). Research on the misleading nature and development trends of data visualization information. Journal of Information Science, 43(02), 170-176.
[11] Chen, Z. (2024). The correlation between artificial intelligence and film from the perspective of media technology philosophy. Nanjing University of the Arts. https: //doi.org/10.27250/d.cnki.gnjyc.2024.000237