INTELLIGENCE · OPEN-SOURCE INTELLIGENCE

Social Media Intelligence and Social Listening

Social media carry a constant stream of opinion, complaint and eyewitness reporting. This article explains how organisations use social listening and social media intelligence, what the research shows about its predictive power, and the biases, fakes and ethical limits that must be managed.

ModQu research article · 1,650 words · 17 references (APA 7th edition)

Every day, millions of people post opinions about products, complain about service, share news and discuss what matters to them. For organisations, this public conversation is a source of intelligence that did not exist a generation ago. **Social listening** is the commercial practice of monitoring and analysing social media conversation about brands, products, competitors and issues. **Social media intelligence**, a term introduced by Omand et al. (2012) and often shortened to SOCMINT, describes the broader use of social media as an intelligence source, with methods and safeguards suited to its particular nature. This article explains what social media can reveal, what research shows about its value, and the pitfalls and ethical limits that apply.

What social media offer

Kaplan and Haenlein (2010) define social media as internet applications that allow the creation and exchange of user-generated content, ranging from collaborative projects and blogs to content communities, social networking sites and virtual worlds. Each type produces different material: short public posts, long reviews, videos, forum discussions and comments.

For intelligence purposes, social media offer several things that other sources do not:

- **Speed.** Problems, events and opinions often appear on social media before they reach the news or formal channels.
- **Unprompted opinion.** People express views in their own words, without the framing of a survey question.
- **Eyewitness material.** Photos and videos from the scene of events.
- **Scale.** Large volumes of conversation allow patterns to be detected across many people and places.
- **Competitor visibility.** Conversation about competitors' products and service is as visible as conversation about one's own.

Uses in business

Brand and reputation monitoring

The most common use is tracking what is said about an organisation and its products: volume of mentions, tone and the topics raised. A sudden rise in negative comment can signal a product fault, a service failure or a developing reputational crisis.

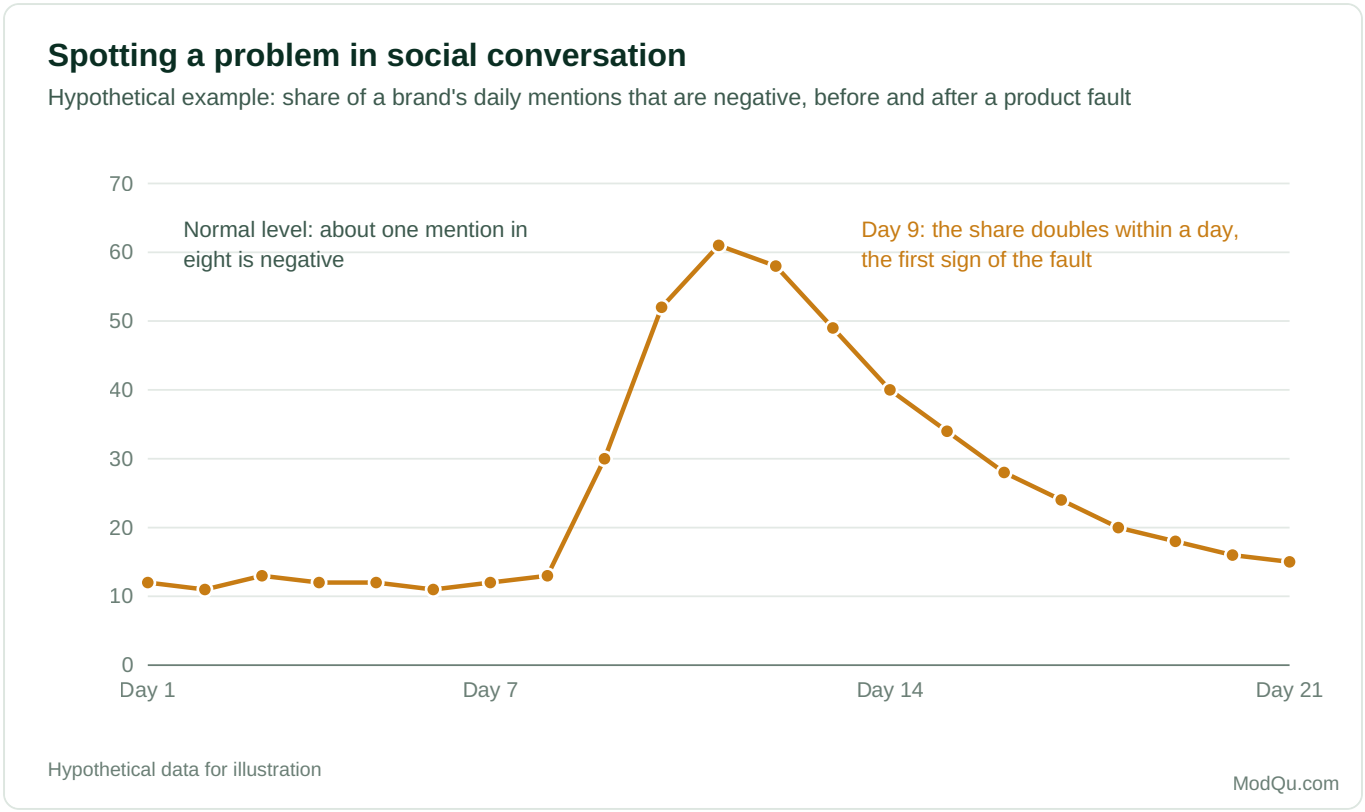


Figure 1. Spotting a problem in social conversation. A sudden change in the share of negative mentions can reveal a fault before formal complaints accumulate. Hypothetical data for illustration; chart by ModQu.

Understanding customers and markets

Text in reviews and posts reveals what customers value and dislike. Tirunillai and Tellis (2014) analysed large numbers of online product reviews with a statistical method that discovers topics automatically, identifying the dimensions of quality that consumers discussed for different brands and how these changed over time. Netzer et al. (2012) mined discussions in online forums to map which car brands consumers mentioned together and in what terms, producing a picture of market structure that closely matched one derived from traditional surveys.

Reviews and sales

Online reviews influence what people buy. Chevalier and Mayzlin (2006) studied book reviews on two large online retailers and found that improvements in a book's reviews on one site led to relative increases in its sales there, and that negative reviews had a larger effect than positive ones. Reviews are therefore both a source of intelligence about customer satisfaction and a factor shaping sales.

Forecasting

Researchers have tested whether social media can predict outcomes. Asur and Huberman (2010) found that the rate at which films were mentioned on a microblogging platform before release predicted box-office revenues well. Bollen et al. (2011) reported that measures of public mood derived from posts improved predictions of daily movements in a stock market index. Such findings attracted wide attention, but they have not always held up in other settings. Gayo-Avello (2012) reviewed attempts to predict elections from social media and concluded that many had been no better than chance or had been identified only after the results were known.

Methods: from posts to insight

Social listening typically involves defining search terms for brands, products and topics; collecting matching posts; filtering out irrelevant material; classifying posts by topic and sentiment; and reporting trends. **Sentiment analysis**, the automated classification of text as positive, negative or neutral, is central. Liu (2012) explains that sentiment analysis is harder than it looks: sarcasm, comparisons, negation and domain-specific language all cause errors, and accuracy varies widely by subject.

Schweidel and Moe (2014) showed that sentiment depends on where people post. Comments in forums, blogs and microblogs about the same brand differ systematically in tone, partly because different people use different platforms and partly because each format encourages different kinds of expression. A shift in the mix of platforms can therefore look like a change in sentiment when nothing has changed. They recommend adjusting for the venue before drawing conclusions.

Biases and pitfalls

Who is speaking

Social media users are not the population. Mellon and Prosser (2017), comparing survey data on British users of two major platforms with the general population, found that users were younger and more educated, and differed in their political attitudes. Only a small share of users post regularly, and they may be unrepresentative even of users.

Ruths and Pfeffer (2014) and Tufekci (2014) set out further methodological pitfalls: each platform's design shapes what people post; data obtained through platforms' interfaces may be incomplete samples; studies often rely on hashtags or keywords that capture only part of the relevant conversation; and results from one platform may not generalise to others.

Bots and fakes

Automated accounts can produce large volumes of posts. Ferrara et al. (2016) describe how such bots can imitate human behaviour, inflate the apparent popularity of views and manipulate trends. Reviews can also be fake. Luca and Zervas (2016) found that a substantial share of reviews on a popular review platform were filtered as suspicious, and that businesses were more likely to post fake positive reviews when their reputation was weak or competition was intense.

False information

Vosoughi et al. (2018), studying rumours on a major platform over many years, found that false stories spread further, faster and more widely than true ones, and that humans, not bots, were mainly responsible. Analysts must therefore verify claims before acting on them, particularly in a crisis.

Ethics and law

Many posts are public, but that does not settle how they may be used. Zimmer (2010) showed that collecting and combining publicly visible social media information can violate users' reasonable expectations, even when no privacy setting is breached. Omand et al. (2012) argue that social media intelligence should meet tests of necessity, proportionality and proper authority. For businesses, privacy law, platforms' terms of service and good practice point in the same direction: analyse conversation in aggregate, avoid profiling private individuals, do not collect more personal information than needed and be transparent about what is done.

Setting up a listening programme

A useful social listening programme starts with questions rather than tools. Typical questions are: Are customers reporting problems with our products? How is a new launch being received? What do people say about competitors' service? Which issues in our industry are gaining attention? Each question determines which terms, platforms and languages to monitor.

Practical steps include:

1. **Build careful search terms.** Include brand and product names, common misspellings, nicknames and competitor names, and exclude terms that bring in irrelevant material, such as other meanings of a brand name.
2. **Establish a baseline.** Record normal levels of volume and tone for several weeks, so that changes can be recognised.
3. **Check the classifier.** Have people read a random sample of posts each month and compare their judgements with the automated sentiment labels. Where the tool performs poorly, adjust it or rely on human reading.
4. **Set alert thresholds.** Decide in advance what change in volume or tone triggers a review, and who receives the alert.
5. **Read the posts.** Numbers show that something has changed; reading the posts shows what and why.
6. **Connect to action.** Route product issues to quality teams, service complaints to customer service and reputational issues to communications, with agreed response times.

Measuring what matters

Counting mentions is easy but rarely useful on its own. More informative measures include the share of mentions that are negative, the topics driving negative comment, the share of conversation an organisation holds relative to competitors, and the speed with which issues are resolved once raised. Comparisons over time and against competitors matter more than absolute figures, because platform changes and seasonal patterns affect everyone. Linking social measures to business outcomes, such as returns, complaints, sales or survey satisfaction scores, shows which social signals are worth watching and which are noise.

An illustration

A maker of small kitchen appliances monitors mentions of its brands and those of three competitors. One weekend, the share of negative mentions of its newest kettle doubles overnight, while mentions of other products are unchanged. Reading a sample of posts reveals several customers describing the same fault: the lid opens unexpectedly when the kettle boils. Some posts include photos. The company's customer service team has received only two complaints so far.

The firm verifies the reports by testing units from the same production batch, confirms a defect in a hinge component and stops shipments of that batch within days. It responds publicly to the posts, offering replacements. Without social listening, the problem would have taken weeks to surface through formal complaints. The firm also reviews what went right and wrong: the alert worked, but it reached the quality team only on Monday morning because no one was assigned to weekend monitoring. It changes the rota so that serious alerts are reviewed within hours on any day of the week, and adds the kettle's component supplier to its list of suppliers to watch.

Conclusion

Social media provide a fast, large and unprompted record of what people think and experience. Used well, they warn of problems, reveal customer priorities and show how competitors are perceived. But the people who post are not the population, sentiment tools make errors, platforms shape what is said, and bots, fake reviews and false stories distort the picture. Social media intelligence works best as one source among several, with careful attention to who is speaking, verification before action and respect for the people whose words are being analysed.

References

Asur, S., & Huberman, B. A. (2010). Predicting the future with social media. In *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology* (Vol. 1, pp. 492–499). IEEE. https://doi.org/10.1109/WI-IAT.2010.63

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1–8. https://doi.org/10.1016/j.jocs.2010.12.007

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, *43*(3), 345–354. https://doi.org/10.1509/jmkr.43.3.345

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM, 59*(7), 96–104. https://doi.org/10.1145/2818717

Gayo-Avello, D. (2012). No, you cannot predict elections with Twitter. *IEEE Internet Computing, 16*(6), 91–94. https://doi.org/10.1109/MIC.2012.137

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons, 53*(1), 59–68. https://doi.org/10.1016/j.bushor.2009.09.003

Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool.

Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science, 62*(12), 3412–3427. https://doi.org/10.1287/mnsc.2015.2304

Mellon, J., & Prosser, C. (2017). Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. *Research & Politics, 4*(3). https://doi.org/10.1177/2053168017720008

Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine your own business: Market-structure surveillance through text mining. *Marketing Science, 31*(3), 521–543. https://doi.org/10.1287/mksc.1120.0713

Omand, D., Bartlett, J., & Miller, C. (2012). Introducing social media intelligence (SOCMINT). *Intelligence and National Security, 27*(6), 801–823. https://doi.org/10.1080/02684527.2012.716965

Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior. *Science, 346*(6213), 1063–1064. https://doi.org/10.1126/science.346.6213.1063

Schweidel, D. A., & Moe, W. W. (2014). Listening in on social media: A joint model of sentiment and venue format choice. *Journal of Marketing Research, 51*(4), 387–402. https://doi.org/10.1509/jmr.12.0424

Tirunillai, S., & Tellis, G. J. (2014). Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent Dirichlet allocation. *Journal of Marketing Research, 51*(4), 463–479. https://doi.org/10.1509/jmr.12.0106

Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. In *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media* (pp. 505–514). AAAI Press.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Zimmer, M. (2010). "But the data is already public": On the ethics of research in Facebook. *Ethics and Information Technology, 12*(4), 313–325. https://doi.org/10.1007/s10676-010-9227-5