A Preliminary Study on the Applicability Assessment of Network Behavior Data (Postprint)

Authors: Nie Lei, Wang Yanfei

Date: 2023-07-26T00:00:00+00:00

Abstract

[Objective/Significance] This study investigates the connotation, influencing factors, and evaluation methods of online behavioral data applicability, providing references for related research to facilitate the scientific use of such data. [Method/Process] Through literature analysis, the core influencing factors of online behavioral data applicability are identified. Building upon this foundation and integrating existing achievements in the fields of intelligence material evaluation and social survey data assessment, this study explores approaches to evaluating the applicability of online behavioral data. [Results/Conclusion] An applicability assessment framework and methods tailored to the characteristics of online behavioral data are ultimately proposed, and the usability of the proposed methods is preliminarily validated through case studies.

Full Text

Preamble

A Preliminary Study on the Adequacy Evaluation of Internet Behavior Data

Nie Lei, Wang Yanfei
Department of Information Management, Peking University, Beijing 100871

Abstract

[**Purpose/Significance**] This paper explores the meaning, influencing factors, and assessment methods of internet behavior data adequacy, providing references for related research to promote the scientific use of such data. [**Method/Process**] Using the literature method, this study first identifies the core influencing factors of internet behavior data adequacy. Based on these factors, combined with existing achievements in information material evaluation and social survey data evaluation, it explores how to assess the adequacy of internet behavior data. [**Result/Conclusion**] Finally, an adequacy evaluation framework and method consistent with the characteristics of internet behavior data are proposed, and the usability of the proposed method is preliminarily verified through cases.

Keywords: Internet behavior data; Data adequacy; Representativeness; Validity; Information awareness
Classification Number: G250
DOI: 10.13266/j.issn.0252-3116.2019.06.004

Obtaining sufficient samples is a fundamental guarantee for quantitative research in social sciences, yet this is often difficult to achieve in many studies due to practical constraints [?]. The rise of big data provides new opportunities to solve this problem, enabling researchers to observe social phenomena at an unprecedented level and conduct more diverse and in-depth studies [?]. As a typical representative of big data, internet behavior data related to search behavior, social media discourse, and e-commerce consumption is being increasingly used in economic and social research. However, it remains unclear whether these data are applicable in specific research contexts or whether they may bias research results.

Intelligence awareness is the core business of intelligence work, requiring intelligence professionals to clearly recognize, interpret, and express intelligence user needs, intelligence object content, and intelligence task organization [?]. In recent years, internet behavior data, as one of the important sources of open-source intelligence, is playing an increasingly important role in intelligence work under the new era context [?]. The introduction of these complex and diverse internet behavior data presents new challenges for data perception in intelligence awareness, such as whether these data can be used to meet the specific needs of intelligence users.

2. The Meaning and Research of Internet Behavior Data Adequacy

This paper uses "data adequacy" to characterize the degree of fit between data and research questions—that is, the extent to which specific data are suitable for answering specific research questions. There is currently no unified definition of the specific content of this concept. Referencing existing research on social survey data evaluation, this paper discusses the adequacy of internet behavior data around the following two aspects: Sample representativeness, or the degree to which samples in the data can be used to effectively estimate the population; and Measurement validity, or the extent to which measurement based on specific data can reflect the true meaning of research concepts [?].

Rigorous research cannot be separated from effective, high-quality measurement, yet this is often difficult to achieve in the application of internet behavior data and the big data they represent [?]. Most commonly used big data were not generated for scientific research [?], and their production and collection processes often lack advance design for data validity [?]. Once removed from applicable contexts, big data lose meaning and value [?]. J. Merino et al. point out that in application processes, adequacy is the main criterion for evaluating big data quality [?]. The European Commission's Big Data Quality Task Team notes that the representativeness of big data for the target population is an important quality issue [?]. J. Liu et al. point out that big data, especially internet big data, bring not only possibilities but also five "big errors" to research, including the non-authoritativeness of data sources and data representativeness issues [?]. Z. Tufekci focuses on the applicability issues of social media data, noting that the representativeness and validity of social media data will pose significant challenges to research effectiveness [?].

Faced with imperfect data, researchers need to be aware of existing problems and use scientific methods to reduce their impact on research. Regarding how to address the adequacy issues of internet behavior data or big data, there have been some discussions. J. Merino et al. propose a measurement framework and process for big data adequacy but do not specify how to conduct each measurement [?]. D. Lazer et al. suggest that researchers should understand the algorithms affecting big data generation and combine "big data" with "small data" to reduce bias in research [?]. Z. Tufekci points out that methods such as locating non-social dependent variables, resampling, conducting baseline measurements, and collaborating with industry can address the inadequacy of representativeness and validity in social media data [?]. J. Liu notes that researchers need to better understand and evaluate the "big errors" brought by big data, and that methods such as collecting data rigorously in cooperation with data providers, supplementing big data with traditional data, and expanding multi-source data can help mitigate the side effects of these errors [?]. Huang Hengjun et al. propose evaluating online merchant data through methods such as single-source channel authority evaluation, data generation mechanism analysis, technical inspection, verification by alternative data sources, verification by complementary data sources, and information availability screening [?].

Thus, the adequacy issues of internet behavior data and the big data they represent may cause serious consequences in research and application, which has attracted high attention from academia. Regarding how to address this issue, existing discussions are mostly directional, and systematic research at the operational level remains insufficient. At the operational level, both information material evaluation in information science [?] and survey data evaluation in sociology [?, ?] can provide references for evaluating internet behavior data adequacy. However, due to characteristics such as large volume and high update frequency of internet behavior data, existing evaluation methods are not directly applicable. This paper will focus on the characteristics of internet behavior data to consider the core influencing factors of data adequacy issues and explore how to evaluate the adequacy of internet behavior data at the operational level based on existing evaluation methods, so that internet behavior data can better serve social science research and intelligence work.

3. Core Influencing Factors of Internet Behavior Data Adequacy

The adequacy issues of internet behavior data stem from their production or generation mechanisms. Internet behavior data were not created for statistics but were recorded by organic systems according to business needs [?]. Those who determine data content and form are often no longer researchers but system platforms that are completely independent of researchers in most cases. Moreover, the term "internet" naturally limits behavior initiators to internet users. Platforms and users are the main factors affecting the adequacy of internet behavior data.

3.1 Platform Factors

Currently, the most commonly used internet behavior data, such as social media data and search data, are mostly collected, stored, and displayed by specific platforms, most of which have commercial attributes. This affects data adequacy in the following three aspects:

3.1.1 Platform-User Bidirectional Selection Affects Data Representativeness

Internet behavior data are not full-sample data. Taking Sina Weibo as an example, its monthly active users reached 392 million in December 2017 [?], a huge volume but only accounting for 28.2% of China's total population. Increased data volume cannot guarantee data representativeness [?], especially for data from commercial platforms [?]. The bidirectional selection between users and platforms, such as differentiated marketing strategies among platforms and user preferences, may result in different platforms representing different user groups. Therefore, once the research object population exceeds the platform scope, the adequacy of specific platform data becomes problematic. Notably, this is not just a sampling issue but a mechanisms issue, and sampling on a single platform cannot solve this problem [?].

3.1.2 Platforms Influence User Behavior, Thereby Affecting Data Measurement Validity

Platforms often use various methods to attract and retain users, such as precision marketing on various platforms, keyword completion and recommendations by search engines, and trending event pushes by social media. These all influence user behavior, causing the meaning behind behavior data to change, which in turn causes seemingly effective or even once-effective measurements to become invalid. Google Flu Prediction is a typical case. D. Lazer et al. point out that the process of engineers optimizing services and users accepting services creates algorithm dynamics problems—that is, user behavior changes with algorithms—which is one of the main reasons why Google Flu Prediction could not sustain its success [?].

3.1.3 Platform Data Management Behavior Affects Researchers' Data Access

As a typical representative of big data, internet behavior data are characterized by large volume and high update speed [?]. To address these characteristics, platforms have corresponding management models for data recording, storage, and retrieval, including the standardization and completeness of metadata records, data warehousing cycles, data ranges available for retrieval by different users, and data interface call ranges. These models are determined by platforms, usually not publicly disclosed, and may change at any time, with their formation and changes often determined by commercial purposes. These models inevitably affect the quantity, form, and even content of data that researchers can obtain, thereby affecting data adequacy. For example, non-standardized records of data attributes in metadata may cause researchers to misunderstand data and extract data unsuitable for specific research.

3.2 User Factors

Internet behavior data are generated by internet users, but the objects of social science research are not limited to internet users. Therefore, the impact of user factors on data adequacy is often unavoidable, mainly from the following three aspects:

3.2.1 Demographic and Social Attributes of Internet Users

Even without considering platform factors, the overall characteristics of internet users affect data representativeness—that is, the demographic and social attributes of internet users may be inconsistent with research objects. For example, as of December 2017, urban internet users accounted for 73% and rural internet users for 27% of China's internet users [?], while during the same period, urban population accounted for 58.52% and rural population for 41.48% of China's total population [?]. The two distributions differ significantly. If network behavior data are directly used to reflect urban-rural differences on a certain issue without processing, the results may be biased.

3.2.2 Same Behavior Patterns May Represent Different Meanings

On the one hand, different or even the same types of users may implement the same behavior according to completely different logics. Mixing various types of behavior data together may invalidate measurement results, thereby greatly reducing their adequacy. On the other hand, if the way researchers obtain data is inconsistent with user behavior patterns, it may also lead to inapplicable results. For example, obtaining social media data through hashtags excludes users who do not like to use hashtags [?].

3.2.3 Users' Reactions to Being Measured

The measurability of internet behavior data is no longer a secret. If users do not want to be measured, they may adopt corresponding strategies to make their behavior "invisible." For example, G. Lotan's research shows that some Twitter users use various strategies to counter Twitter and, in this process, well understand and utilize Twitter's trending topic algorithm [?]. Similar practices are also common in China, such as expressing opinions through images, emojis, and coded language, all of which increase measurement difficulty, making it hard for researchers to obtain data suitable for specific research, and may even lead to obtaining unsuitable data without realizing it.

4. Evaluation Framework and Methods for Internet Behavior Data Adequacy

Through the above analysis, it is not difficult to find that platforms and users have important impacts on the adequacy of internet behavior data. Therefore, evaluation can be based on platforms and users. At the same time, we can also work backward from results—that is, evaluate the measurement results obtained using internet behavior data to judge data adequacy in reverse. The overall evaluation approach is shown in Figure 1 [Figure 1: see original paper].

The evaluation framework under this approach is shown in Table 1 , where evaluation based on influencing factors corresponds to the content derived from platform and user impacts on internet behavior data adequacy, and evaluation based on measurement results references social survey data validity evaluation [?].

4.1 Evaluation Based on Influencing Factors

4.1.1 Evaluating Platform and User Characteristics

For information materials, evaluating the characteristics of their sources helps judge their degree of adequacy [?]. For internet behavior data, the source consists of platforms and users. As discussed above, both internet platforms and users have representativeness issues, which can easily cause errors when using internet behavior data to study populations. However, for researchers, the important thing is not the absence of errors but the ability to know and control the magnitude of errors [?], especially when errors may occur in core variables of the research. For evaluating data representativeness, the following methods can be referenced:

(1) **Evaluation based on platform official data.** Large platforms often analyze their user characteristics and publicly release them, including relatively accurate data such as user numbers in company financial reports, as well as estimated data such as user profiles. These data can provide directional judgments about data representativeness.

(2) **Evaluation based on existing survey data or research findings.** On the one hand, data representativeness can be evaluated using data released by institutions such as the China Internet Network Information Center (CNNIC). On the other hand, judgment basis can be obtained through

existing empirical research, such as M. Duggan and J. Brenner's analysis of Twitter, which helps researchers evaluate Twitter user characteristics [?].

Taking G. Doyle's research on English dialect evolution using Twitter data as an example, the author first pointed out based on existing research that Twitter data skew toward younger groups and slightly toward urban areas. However, linguistic research shows that urban youth are the main drivers of language change, and the informality of language on self-media also meets the requirements of language change research. Therefore, the author believes that user behavior data on Twitter is applicable to this study [?].

4.1.2 Evaluating Platform Impact on Users

Research on platform impact on user behavior is currently relatively limited. Moreover, this impact may be dynamic and irregular, as platforms irregularly launch new functions. Therefore, in addition to referencing existing research, researchers more need to evaluate platform impact on users in specific research cases through experiments.

(1) **Time-series comparison on a single platform.** The basic idea is that if the measurement object has certain time-series characteristics, "abnormal" changes in its time-series data are mostly the result of external influences. If platform behavior is known, we can evaluate whether it has affected time-series data representing user behavior; if platform behavior is unknown, we can use "abnormal" time-series data identification for auxiliary judgment. For example, we can judge the degree of variation through variance of difference values, random factor values after time series decomposition, and other indicators within specific time periods, or identify data anomalies through outlier detection algorithms. K. H. Brodersen et al.'s research results show that time series models can verify whether an advertisement has affected Google users' search and click behavior [?], a method that can be extended to adequacy evaluation of internet behavior data.

(2) **Multi-platform comparison.** The basic idea of this method is that for measurement of the same issue, the higher the consistency of data obtained from multiple platforms, the lower the possibility that a single platform has a specific impact on users. It should be noted that when measurement results from different platforms are highly consistent, it may be that each platform has the same impact on users. Therefore, the more platforms used and the greater the differences between platforms, the more valuable the comparison results. When Huang Hengjun et al. studied using internet data to build a business directory, they verified the authenticity and comprehensiveness of merchant information by comparing data from Dianping.com and Nuomi.com [?]. The essence is comparing merchant behavior on different websites, so their approach is also applicable when evaluating platform impact on users.

4.1.3 Evaluating Platform and User Behavior Patterns

The problems researchers face when evaluating platform data management mechanisms and user behavior pattern impacts are essentially the same: establishing the connection between platform and user behavior and research. For this issue, the following evaluation methods can be referenced:

(1) **Evaluating the impact of platform data management models on adequacy through official information.** Some platforms publish their technical information on developer platforms, industry forums, and other channels, and researchers can also consult platform customer service for relevant technical information. Based on the author's research, although some useful information can be found, such as data interface sampling ratios and update cycles, the information obtained through these channels is often insufficient and not timely, which may be why relevant scholars often recommend that researchers cooperate with data providers [**?**].

(2) **Evaluating data adequacy based on research related to internet behavior patterns.** Internet behavior patterns are a hot research topic in recent years, and disciplines such as information science, computer science, and psychology have made many achievements in this field. Reviewing existing research will help researchers evaluate data adequacy. For example, when Sun Yi, Lv Benfu, et al. studied consumer confidence using search engine data, they verified the association between internet search behavior and consumer confidence through existing research, and then built a consumer confidence index based on search data [**?**].

(3) **Evaluating platform and user behavior patterns through experiments.** If there is a lack of existing research and official information as evaluation basis, researchers can only study user behavior patterns and platform data management models in reverse through experiments. Although different research purposes will lead to different experimental directions, the core idea is to first mine data that has already been affected by platform and user behavior, discover patterns within it, and then understand its generation process in reverse. For example, clustering users based on user behavior and studying differences in behavior patterns between categories to explore the meanings represented by different behavior patterns, or studying platform data release patterns through dynamic tracking and analysis of data. To illustrate this approach, this paper conducted a simple example: the author used a social media platform's internal search function with "research" as the keyword and region limited to "Beijing," conducted multiple searches at different time points with the same search conditions, and obtained the number of data entries shown in Table 2 . The search time range was from 8:00 to 11:59 on May 16, 2018. It is not difficult to find from the table that the number of data entries shows a decreasing trend, with a 2.1% reduction within five days. Although a single experiment itself cannot prove the existence of a pattern, it provides a direction. If this conclusion is verified through a

large number of experiments, whether due to some users tending to delete behavior data or due to restrictions on the internal search engine, it means that historical data obtained this way by researchers may be incomplete. If the research question is sensitive to this point, especially when studying events that occurred long ago, the adequacy of this data will be greatly reduced.

4.2 Evaluation Based on Measurement Results

If researchers cannot evaluate platforms and users, they can also deduce the adequacy of internet behavior data by evaluating measurement results. This aspect can draw on predictive validity, concurrent validity, and construct validity assessment methods from social survey data validity evaluation.

4.2.1 Evaluating Predictive Validity

Predictive validity is "comparing obtained measurement results with what actually happens in the future to check the consistency between the two" [**?**]. When measurement has time-series attributes, predictive validity can be used for evaluation. Specifically, there are two different methods:

(1) **Measure first, wait for results to appear, and then evaluate.** Taking statistical indicator proxies as an example, the release of most statistical indicators is lagged—for instance, releasing May measurement results in mid-June. If researchers complete measurement through internet behavior data before the release of statistical indicators, they can verify it after the statistical data are released. A single measurement can calculate the difference between predicted and actual values, and multiple measurements can calculate mean square error and other indicators to judge data adequacy. This is an ideal situation but is subject to certain limitations in actual use. For example, if the timing of result occurrence is uncertain, the research cycle may be infinitely extended. Therefore, the second method is sometimes used in practice.

(2) **Evaluation using historical data.** The basic idea is that if a historical result can be accurately predicted using data from before its occurrence, it indicates that the data were once adequate, and following the idea of time series extrapolation, the data are still considered to have some adequacy now. If the historical data sequence is long enough, indicators such as mean square error and cumulative mean square error can be used for evaluation. In practice, researchers often use major historical events as prediction objects. This method is easy to implement but belongs to post-hoc explanation. Coupled with the inherent defects of extrapolation methods, its scientific nature is easily questioned and is therefore often used as an auxiliary evaluation method.

Economic prediction research based on internet behavior data often uses predictive validity to evaluate data adequacy. For example, when Google scientist

S. L. Scott and economist H. R. Varian used Google Trends data for economic prediction research, they found through comparison between pure time series models and time series models incorporating Google Trends that the latter had smaller prediction errors and could better predict the 2008-2009 economic crisis, thus indicating that Google Trends data have important value in such research [?].

4.2.2 Evaluating Concurrent Validity and Construct Validity

In addition to predictive effectiveness, validity can also be evaluated through correlation. The basic idea is that if the measured object is the same as or highly correlated with a known object conceptually, then when measurement results are highly correlated with data representing the known object, we can be more confident that the measurement is valid. According to different types of correlation, evaluation basis can be divided into concurrent validity and construct validity.

(1) **Concurrent validity.** Concurrent validity is used to judge whether a new measurement can replace an existing measurement [?]—that is, using internet behavior data to measure a known variable. If the new measurement results are highly correlated with known results (e.g., large correlation coefficient, significant regression coefficient), it can be considered valid. Concurrent validity is often used when existing measurements are highly recognized but difficult to implement, such as large-scale social surveys. If researchers want to use internet behavior data to measure concepts that are the same as or similar to existing indicators, concurrent validity can be used.

Concurrent validity has been used in empirical research. For example, G. Doyle verified the adequacy of SeeTweet in his dialect research by comparing SeeTweet calculation results with high-quality but extremely time-consuming Atlas of North American English and Harvard Dialect Survey results [?]. Sun Yi, Lv Benfu, et al. verified the correlation between network inflation expectations based on search data and the Consumer Price Index through regression analysis [?].

(2) **Construct validity.** Construct validity evaluation aims to understand "whether the measurement tool reflects the internal structure of concepts and propositions" [?]. The basic idea is that when the measurement object is theoretically highly correlated with another object, if correlation can be proven at the empirical level, we can be more confident that the measurement is valid. For example, if a researcher uses internet behavior data to measure concept A for the purpose of studying the relationship between concept A and concept B but is uncertain about measurement validity, and if concept A is theoretically highly correlated with concept C and concept C has been quantified, then the significance test of the regression coefficient of A on C can be used to judge whether the measurement of A is valid, thereby determining whether the data are adequate.

Internet behavior data brings new opportunities to social science research and intelligence work while also bringing huge risks. If researchers use inadequate internet behavior data without realizing it, it may not only lead to biased or invalid results but also produce negative economic and social impacts. Therefore, research on the adequacy of internet behavior data is of great significance.

This study explores the core influencing factors of internet behavior data adequacy issues and, based on this foundation and combined with existing research results in information science and sociology, proposes a scientific, systematic, and operable evaluation framework for internet behavior data adequacy, fills in specific evaluation methods, and preliminarily verifies its usability through cases, hoping to provide reference basis for related research.

This study is still in the initial exploration stage, and subsequent continuous verification and expansion of current conclusions through extensive literature research and experiments are still needed. This process will inevitably involve many theoretical and methodological difficulties, such as whether it is scientific to use internet behavior data to verify its own adequacy. However, as these problems are solved, internet behavior data and the internet big data they represent will contribute more to the development of social sciences.

References

[?] King, Keohane, Verba. *Research Design in Social Sciences* [M]. Translated by Chen Shuo. Shanghai: Gezhi Press, 2014.

[?] Golder SA, Macy MW. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures [J]. *Science*, 2011, 333(6051): 1878-1881.

[?] Wang Yanfei, Zhao Keran, Chen Meihua, et al. Research analysis of intelligence awareness [J]. *Information Studies: Theory & Application*, 2018, 41(8): 1-4.

[?] Williams HJ, Blum I. Defining second generation open source intelligence (OSINT) for the defense enterprise [EB/OL]. [2018-06-20]. https://www.rand.org/pubs/research_{reports}/RR1964.html.

[?] Babbie. *The Practice of Social Research* [M]. Translated by Qiu Zeqi. 11th ed. Beijing: Huaxia Publishing House, 2009.

[?] Power DJ. Using 'big data' for analytics and decision support [J]. *Journal of Decision Systems*, 2014, 23(2): 222-228.

[?] Lazer D, Kennedy R, King G, et al. The parable of Google flu: traps in big data analysis [J]. *Science*, 2014, 343(6176): 1203-1205.

[?] Kaisler S, Armour F, Espinosa JA, et al. Big data: issues and challenges moving forward [C]//Proceedings of 2013 46th Hawaii International Conference on System Sciences. HI: IEEE, 2013: 995-1004.

[?] Boyd D, Crawford K. Critical questions for big data [J]. *Information, Communication & Society*, 2012, 15(5): 662-679.

[?] Merino J, Caballero I, Rivas B, et al. A data quality in use model for big data [J]. *Future Generation Computer Systems*, 2016, 63(C): 123-130.

[?] United Nations Economic Commission for Europe Big Data Quality Task

Team. A suggested framework for the quality of big data [EB/OL]. [2018-05-21]. http://www1.unece.org/stat/platform/download/attachments/108102944/Big%20Data%20Quality%20Framew
%20final%20Jan08%202015.pdf?version=1&modificationDate=1420725063663&api=v2.

[?] Liu J, Li J, Li W, et al. Rethinking big data: a review on the data quality and usage issues [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 115(C): 134-142.

[?] Tufekci Z. Big questions for social media big data: representativeness, validity and other methodological pitfalls [C]//Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media. MI: AAAI, 2014.

[?] Huang Hengjun, Tao Ran, Fu Deyin. Business directory updating: internet big data sources and their data quality assessment [J]. *Statistical Research*, 2017, 34(1): 12-22.

[?] Wang Yanfei, Qin Tiehui. *Information Analysis and Decision Making* [M]. 2nd ed. Beijing: Peking University Press, 2010.

[?] Zhou Jun. On the identification of social science intelligence materials [J]. *Information and Documentation Services*, 2005(3): 46-48.

[?] Yuan Fang. *Social Research Methods* [M]. Beijing: Peking University Press, 2004.

[?] Groves RM. Three eras of survey research [J]. *Public Opinion Quarterly*, 2011, 75(5): 861-871.

[?] Sina Technology. Sina releases unaudited financial results for Q4 and full year 2017 [EB/OL]. [2018-05-21]. http://tech.sina.com.cn/i/2018-02-13/doc-iyrpeie3108640.shtml.

[?] Gartner Inc. What is big data? [EB/OL]. [2018-05-21]. https://www.gartner.com/it-glossary/big-data.

[?] China Internet Network Information Center. The 41st Statistical Report on China's Internet Development [EB/OL]. [2018-05-21]. http://www.cnnic.net.cn/hlwfzyj/hlwxzbg/hlwtjbg/201803/P020180305409870339136.pdf.

[?] National Bureau of Statistics of the People's Republic of China. Statistical communique of the People's Republic of China on the 2017 national economic and social development [EB/OL]. [2018-05-28]. http://www.stats.gov.cn/tjsj/zxfb/201802/t20180228_{158563

[?] Lotan G. Data reveals that "occupying" twitter trending topics is harder than it looks [EB/OL]. [2018-05-21]. http://blog.socialflow.com/post/712024374/data-reveals-that-occupying-twitter-trending-topics-is-harder-than-it-looks.

[?] Duggan M, Brenner J. The demographics of social media users-2012 [EB/OL]. [2018-05-21]. http://www.pewinternet.org/2013/02/14/the-demographics-of-social-media-users.

[?] Doyle G. Mapping dialectal variation by querying social media [C]//Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics. Gothenburg: Association for Computational Linguistics, 2014: 98-106.

[?] Brodersen KH, Gallusser F, Koehler J, et al. Inferring causal impact using Bayesian structural time-series models [J]. *Annals of Applied Statistics*, 2015, 9(1): 247-274.

[?] Sun Yi, Lv Benfu, Chen Hang, et al. Research on constructing and applying

consumer confidence index based on internet search behavior [J]. *Management Review*, 2014, 26(10): 117-125.
[?] Scott SL, Varian HR. Predicting the present with Bayesian structural time series [J]. *Social Science Electronic Publishing*, 2012, 5(1): 4-23.
[?] Sun Yi, Lv Benfu, Chen Hang, et al. Research on measuring and applying inflation expectations from a big data perspective [J]. *Management World*, 2014(4): 171-172.

Author Contributions:
Nie Lei: Paper writing;
Wang Yanfei: Paper revision.

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