Emerald Insight

Internet Research

Googling environmental issues: Web search queries as a measurement of public attention on environmental issues

Jie Qin Tai-Quan Peng

Article information:

To cite this document:
Jie Qin Tai-Quan Peng , (2016),"Googling environmental issues", Internet Research, Vol. 26 Iss 1 pp. 57 - 73
Permanent link to this document:
http://dx.doi.org/10.1108/IntR-04-2014-0104

Downloaded on: 09 November 2016, At: 20:29 (PT)
References: this document contains references to 65 other documents.
To copy this document: permissions@emeraldinsight.com
The fulltext of this document has been downloaded 365 times since 2016*

Users who downloaded this article also downloaded:

(2016),"Analyzing user perspective on the factors affecting use intention of mobile based transfer payment", Internet Research, Vol. 26 Iss 1 pp. 38-56 http://dx.doi.org/10.1108/IntR-05-2014-0143

(2016),"The effect of intrinsic and extrinsic motivations on mobile coupon sharing in social network sites: The role of coupon proneness", Internet Research, Vol. 26 Iss 1 pp. 101-119 http://dx.doi.org/10.1108/IntR-05-2014-0136

Access to this document was granted through an Emerald subscription provided by emerald-srm:563821 []

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

The current issue and full text archive of this journal is available on Emerald Insight at:
www.emeraldinsight.com/1066-2243.htm

Googling environmental issues

Web search queries as a measurement of public attention on environmental issues

Jie Qin
School of Liberal Arts, SCAD Hong Kong, Hong Kong S.A.R., China, and
Tai-Quan Peng
Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore, Singapore

Received 18 April 2014
Revised 26 August 2014
21 October 2014
3 December 2014
Accepted 9 December 2014

Abstract

Purpose – Queries as a pioneering measure of public attention on various social issues have elicited considerable scholarly attention. The purpose of this paper is to address two fundamental questions, as follows: first, how do we identify niche queries that internet users search for on specific social issues?; and second, what are the measurement properties of queries data in gauging public attention on social issues?

Design/methodology/approach – The present study uses public attention on environmental issues in the USA as the empirical setting of research. An iterative framework is developed to identify niche queries to measure public attention on environmental issues. The measurement properties of queries data are assessed by comparing the dynamics of public attention on environmental issues captured by queries data with that measured by the "most important problem" (MIP) question in Gallup opinion polls.

Findings – A list of 39 niche queries that internet users search for on environmental issues is identified. The dynamics of public attention on environmental issues determined by the search trends of these 39 queries is found to positively correlate with that measured by Gallup MIP polls, whereas both dynamics can forecast each other well in a 12-month time frame.

Originality/value – The findings of the study possess methodological and practical implications. The study shows that queries data are complementary to, rather than substitutes of, public opinion polls in measuring public attention on environmental issues. The iterative framework developed in the study can be applied in future studies to help researchers identify valid queries to measure public attention on other social issues, as it can minimize researchers' subjective biases in selecting search queries. Policymakers and environmentalists can utilize our approach to monitor the status of public attention on environmental issues and implement campaigns to mobilize favorable public opinion when the decline of public attention is predicted by the trends of web search queries.

Keywords Google trends, Environmental issues, Public attention, Web search queries

Paper type Research paper

Introduction

Web search queries, or web queries or search queries or online search queries (hereinafter "queries"), refer to the words or the combination of multiple words and/or symbols that individuals enter into an online search engine for expected search results. Despite different goals people tend to accomplish in using web search engines (Jansen *et al.*, 2008), self-initiated queries represent what people are truly curious or worried about at the moment, which have been used to understand the state of the art (i.e. nowcasting) and the future (i.e. forecasting) of public attention on various social objects (Zhu *et al.*, 2012; Schoen *et al.*, 2013).

The authors would like to thank the editor and anonymous reviewers for their valuable comments. The study was supported by a Start-Up Grant (M4081281.060) from Nanyang Technological University.

Internet Research
Vol. 26 No. 1, 2016
pp. 57-73
© Emerald Group Publishing Limited
1066-2243
DOI 10.1108/IntR-04-2014-0104

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

Research interests in using queries as an alternative measurement device in social research have boomed since the launch of Google Flu Trends (GFT) in 2008. GFT tracks the search volumes of 45 queries people tend to search for during the flu season, and is currently the most visible yet most controversial application of queries as a real-world indicator of public attention (Lazer *et al.*, 2014). Ginsberg *et al.* (2009) reported that GFT could alert the general public to the flu pandemics faster than the official statistics from the Centers for Disease Control and Prevention in the USA. This study has prompted scholars to apply queries data to predict epidemic trends of various diseases, such as dengue fever (Althouse *et al.*, 2011), stroke (Walcott *et al.*, 2011), and tuberculosis (Zhou *et al.*, 2011). Moreover, scholars have extended this line of research and attempted to adopt queries to measure public attention on commercial products, such as automobiles (Carrière-Swallow and Labbé, 2013) and stocks (Curme *et al.*, 2014; Preis *et al.*, 2013), and public attention on social issues, such as youth unemployment (Fondeur and Karamé, 2013), campaign performance (Huang *et al.*, 2013), and public participation in political issues (Reilly *et al.*, 2012).

Despite upsurging research interests in using queries as an alternative measurement device in social research, scholars have recently been reminded to handle queries data with great caution because of the algorithmic black box in generating queries data (Rutherford, 2014; Lazer *et al.*, 2014) as well as different usage patterns of web search engines across societies (Carrière-Swallow and Labbé, 2013; Cook *et al.*, 2011; Mellon, 2013). We agree that user-generated data from search engines has enormous potential, as well as practical limitations, in measuring public attention on various social issues (Castillo *et al.*, 2013; Jungherr and Jürgens, 2013; Kalampokis *et al.*, 2013; Schoen *et al.*, 2013). To ensure that queries can be a sound measurement of public attention, two fundamental questions need to be addressed: first, what queries internet users search for on specific social issues; and second, how public attention measured with queries data are related to that measured with other established devices.

The present study will use public attention on environmental issues in the USA as a research context to address the two aforementioned questions. Specifically, we aim to develop an approach to identify queries internet users will search for about environmental issues and examine the interrelationship between the dynamics of public attention on environmental issues captured by queries data and that measured by the "most important problem" (MIP) question in Gallup opinion polls, which is argued to be the most established measure of public attention (Smith, 1985).

Measuring public attention on environmental issues in empirical research

The notion of environmental issues, which the general public has focussed attention on, is a changing construct with many facets. In the 1970s in the USA, along with Carson's (1962) warning on the pesticide DDT and the emerging antinuclear movements, public attention on environmental issues was significantly limited to pollution and nuclear-related topics (Downs, 1972). In the 1980s, air pollution and "smog" became major environmental concerns among the general public (Baker and Bagozzi, 1982). In the 1990s, public attention on the environmental issues incorporates concerns over an extensive range of topics, such as water, wildlife, energy, and health (Zimmer *et al.*, 1994). At present, along with the process of globalization, concerns over transnational environmental issues, such as global warming, climate change, and greenhouse gas, are among the most salient environmental concerns (Nisbet and Myers, 2007).

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

Verifying the downward public attention on environmental issues is a useful strategy to keep the public aware of current unfavorable environmental situations and motivate supportive public opinion on environmental issues (Agnone, 2007; Daniels *et al.*, 2013). In the past, opinion polls have been an important source for researchers or policymakers to track the dynamics of public attention on environmental issues (e.g. Downs, 1972; Dunlap, 1991, 1995; Nisbet and Myers, 2007). By describing the rise and fall of public attention on environmental issues from the 1960s to the 1990s with opinion poll data, Dunlap (1991) concluded that the 1990s was a decade in which "public concern for environmental quality reached unprecedented levels" (p. 285). However, environmental issues are now at the bottom of the "issue-attention cycle" (Downs, 1972). Environmental issues have been losing its priority to economic crisis (Project for Improved Environmental Coverage, 2013; Yeager *et al.*, 2011), given the fact that "people's concern for the environment rises when they feel economically more secure" (Worcester, 1998, p. 164).

Although opinion polls have been a popular tool among social researchers to observe public attention on social issues, the criticisms associated with opinion polls have been well documented in the literature, ranging from the quality of responses (Bishop *et al.*, 1980; Price and Roberts, 1987; Price and Neijens, 1997), to the limited number of predetermined issues (Newig, 2004), and to the static nature of responses and lack of regularity (Ripberger, 2011). Recently, certain studies have adopted queries to determine the dynamics of public attention on environmental issues (e.g. Mccallum and Bury, 2013; Ripberger, 2011).

Ripberger's (2011) study is among the first attempts to assess the validity of queries as a measurement of public attention on environemental issues. Considering that environmental issues are a "sensational issue" that is significantly driven by media coverage (Soroka, 2002b), public attention on environmental issues measured by queries data were hypothesized to be correlated with the number of media coverage on the issue. The hypothesis was supported in Ripberger's study; a positive correlation is observed between search volume of one single query "global warming" and *New York Times*' coverage on this issue. Riperger then concluded that queries data possess convergent validity in measuring public attention. However, one fatal flaw in Ripberger's study is that only one single query was used to measure public attention on complicated social issues. Mccallum and Bury (2013) recognized the limitation in Ripberger's study and attempted to use a series of relevant queries to detect how public attention on environmental issues, such as conservation, pollution, and climate change, will rise or fall with the search volumes of different queries. They found that casual selection of queries will lead to contradictory conclusions concerning the general trend of public attention on environmental issues. They identified certain queries with an upward search trend such as "sustainability" and "climate change," others with a relative stable trajectory such as "global warming," and some with a downward trend such as "habitat fragmentation" and "pollution."

These two studies have shown the promising as well as challenging aspects in using queries as a new measure of public attention on environmental issues. Considering that queries are "sensitive to fads, whims or other stochastic changes in public perception and language" (Mccallum and Bury, 2013, p. 1361), a major challenge lies in the identification of niche queries to measure public attention on environmental issues. In the above two studies, niche queries were actually identified in a top-down approach (i.e. assigned by researchers in a priori way), which substantially devalue the self-initiated behavioral nature of queries data. Moreover, exclusive reliance on single

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

query will lead to unreliable conclusions. The crux of the problems lies in our limited understanding about "How do they search the Web? What do they search for on the Web?" (Jansen *et al.*, 2000, p. 208).

Although the search log scandal of America Online in 2006 refrained search engine companies from sharing original search logs with the public, aggregated data sets provided by search engine companies offer an alternative way for researchers to extract abundant social and behavioral information embedded in queries. Therefore, developing a novel approach to identify multiple user-generated queries to measure public attention on complex social issues in a valid manner is of methodological necessity and practical value (Mccallum and Bury, 2013; Mellon, 2013). Moreover, it is methodologically necessary and significant to empirically assess the measurement properties of queries data in gauging public attention on environmental issues by comparing it with that measured in traditional public opinion polls. This kind of comparison can advance our understanding of measurement properties of queries data as an operationalization of public attention in future empirical research, as "sound measurement properties can produce valuable insight into the structure and interrelationships among complex variables while poor ones can result in erroneous conclusions regarding the existence, magnitude, and direction of association between constructs" (Segars, 1997, p. 108).

Research method

Data collection

We limit our research focus to the USA where 78 percent of the population is internet users (Internet World Stats, 2013). Moreover, Google is the most popular search engine among internet users in the USA (comScore, 2013), which justifies our choice of Google Trends as the source to retrieve queries data to address our research concerns. Moreover, the USA has a long history of public opinion polls and has a good archive for opinion polls data, which provides a solid and reliable benchmark in our assessment of measurement properties of queries data.

To assess the measurement properties of queries data, the search trends of web queries are compared with the trends gauged in Gallup MIP polls (hereinafter "Gallup MIP"). The Gallup MIP question (i.e. "What do you think is the most important problem facing this country today?") has been extensively acknowledged and used as a valid measure of public attention on various social issues. The monthly Gallup MIP data are retrieved from Roper Center iPOLL Databank at the University of Connecticut, resulting in a data set composed of 66 monthly survey data from January 2008 to June 2013. Gallup MIP measures public attention on environmental issues by the percentages of respondents who mention "Environment-Pollution" as the MIP facing the country in each of 66 surveys.

An iterative framework, which is a mixture of top-down and bottom-up approaches, is developed in the study to identify niche queries internet users search for online about environmental issues. Four seed queries (i.e. environment protection, pollution, ecological problems, and global warming), which are directly relevant to public attention on environmental issues, are first derived from the original codes of the MIP questions in Gallup opinion polls (McCombs and Zhu, 1995). Then, these four seed queries are submitted to Google Correlate, another Google product, to obtain a longer list of queries, which yields 693 correlated queries for our four seed queries[1].

However, as Google (2013) explains in its documentation, "Google Correlate is like Google Trends in reverse. With Google Trends, you type in a query and get back a data

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

series of activity (over time or in each US state). With Google Correlate, you enter a data series (the target) and get back a list of queries whose data series follows a similar pattern," the returned results from Google Correlate are quite messy and may include some queries totally irrelevant to environmental issues. Therefore, the authors manually code the 693 queries retrieved from Google Correlate in order to identify valid queries in our study. In the first round of manual coding, the authors code the queries independently with inter-coder reliability equal to 0.79. The authors then have a discussion about those inconsistent queries and finally reach consensus on 39 niche queries.

These 39 queries are then submitted to Google Trends to retrieve their search trends. The time frame is specified from January 2008 to June 2013, and the location is constrained to the USA. Due to the fact that Google Trends imposes two transformations on the raw search volume data prior to public release, the magnitudes in the retrieved search trends are unit-free and are not directly comparable across queries (Google, 2014). To resolve this issue, a benchmark approach is adopted in the study. Briefly speaking, we use one irrelevant query (i.e. "IT Job" in the study) as a benchmark query and submit the benchmark query as well as 39 queries to Google Trends in pairs. Several benchmark queries have been tested in the study. Our criterion in benchmark query selection is that the benchmark query should have a stable and predictable search volume over time (Carneiro and Mylonakis, 2009; Reilly *et al.*, 2012). "IT Job" outperforms other candidate queries in this regard.

The search magnitudes of 39 queries are returned from Google Trends based on the same benchmark (i.e. the search magnitude of "IT Job"). In other words, the search magnitudes of 39 queries obtained via the benchmark approach share the same scale and can be compared with each other. Furthermore, to "control for artificial trends in substantive queries under study" (Zhu *et al.*, 2012, p. 3), a weekly ratio is calculated for each niche query by dividing the search magnitude of the query by that of the benchmark query. To generate consistent time units between Google Trends and Gallup MIP, the weekly ratios of all 39 niche queries are aggregated to the monthly level by taking the arithmetic mean of the weekly ratios in a month.

Analytical design

In order to assess the measurement properties of queries data in gauging public attention on environmental issues, the dynamics of public attention on environmental issues measured by Google Trends (hereinafter "GT dynamics") is compared with that measured by Gallup MIP (hereinafter "Gallup dynamics"). Vector autoregressive (VAR) modeling approach, which is a macroeconometric framework introduced by Christopher Sims (1980), is adopted in the study. VAR has been used in empirical studies to analyze the dynamics of public attention on different social issues (Green-Pedersen and Stubager, 2010; Ripberger, 2011; Soroka, 2002a, b; Wood and Peake, 1998). Specifically, one reduced-form VAR model will be fitted to GT dynamics and Gallup dynamics, which can help us examine whether multiple time series are correlated with each other and if they can forecast each other (Stock and Watson, 2001) when their own past values are controlled. We will rely on Granger causality analysis and impulse response function to assess if GT dynamics and Gallup dynamics are correlated with each other. As a new measure of public attention on social issues, GT dynamics is expected to positively correlate with Gallup dynamics on the same issue.

Beyond investigating the correlation between GT dynamics and Gallup dynamics, we attempt to examine if the two dynamics can forecast each other. Forecast error variance decomposition analysis in VAR is used to examine the percentage of the

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

variance of the error generated in forecasting a variable (e.g. Gallup dynamics) because of a specific shock (e.g. the error term in GT dynamics) at a given time frame (e.g. 12 months). Moreover, forecasting accuracy is assessed by dividing the entirety of the 66-month data into two subsamples: a test sample including the first 54 months (i.e. January 2008-June 2012) and a holdout sample including the last 12 months (i.e. July 2012-June 2013). The holdout sample will be used to cross-validate the VAR model we estimate in the test sample (Yaffee, 2000). Specifically, we make 12-month forecasts (i.e. from July 2012 to June 2013) for both time series (i.e. GT dynamics and Gallup dynamics) based on the test sample with the aforementioned VAR models and compared the forecasted trend with the observed trend in the holdout sample.

Analytical findings

Queries that internet users search for about environmental issues

Table I summarizes the 39 queries that people tend to search for about environmental issues. Two notable phenomena deserve our attention. First, some of the queries are essentially the same with each other, such as "global warming," "global warming is," and "the global warming." However, instead of combining them into one, we retain them as separate queries due to both conceptual and empirical considerations. Conceptually, these queries are the raw ones submitted by internet users and retrieved from Google. In other words, internet users not only use "global warming" but also use queries like "global warming is" and "the global warming" while they search information about global warming on the internet. We want to keep these raw queries as they are and have no intention to over-manipulate the data. Empirically, the search trends of these similar queries are a bit different from each other. We have tried different ways to contact Google

		Breath of public attention on environmental issues		
		Causes	Consequences	Solutions
Depth of public attention on environmental issues	Generic	About global warming; causes global warming; causes of global warming; for global warming; global warming; global warming is; the global warming; what is global warming; global warming?	Effects of global warming; global warming effect; global warming effects; global warming news	Global warming science
	Specific	Atom bomb; nuclear waste; toxic waste	Acid rain; air pollution; biomes; deforestation; desert plants; endangered animal; endangered animals; endangered species; environmental problems; greenhouse effect; ocean pollution; ozone depletion; ozone layer; pollution; tropical rain; vegetation; warming; water pollution	Clean air act; environmental law; environmental protection; environmental education

Table I.
The breadth and depth of public attention on environmental issues captured by 39 search queries

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

(e.g. submitting request on their website, posting questions on their blogs, and reading publicly available white papers) to ask for clarifications about the different search trends of similar queries. However, we have not got any responses from Google, just like Lazer *et al.* (2014) encountered in their study. To be conservative, we retain those similar queries as separate ones in our study. Second, although our focus in the study is public attention on environmental issues in USA, most of the queries turn out to be related with global warming. This finding echoes Nisbet and Myers' (2007) argument that transnational environmental issues, such as global warming, climate change, and greenhouse gas, are among the most salient environmental issues nowadays.

Following Stamm *et al.*'s (2000) conceptualization of public understanding in global warming, these 39 queries are classified into two dimensions which can represent the breadth and depth of public attention on environmental issues, respectively. The breadth of public attention on environmental issues refers that people will search for queries which are related to different domains of environmental issues, namely, the causes, consequences, and solutions of various environmental problems (Stamm *et al.*, 2000). For example, the query "causes global warming" concerns the causes of global warming, "effects of global warming" concerns the consequences, whereas "environmental protection" represents public concerns over the solutions to global warming. The depth of public attention on environmental issues refers that internet users will search queries that are relevant to generic or specific aspects of environmental issues. Queries, such as "acid rain," "deforestation," "greenhouse effect," and "ozone depletion," represent internet users' concern over specific aspects of environmental issues. By contrast, queries like "effects of global warming" merely show internet users' concern over general problems related to the environment.

The search trends of these 39 queries are highly correlated with each other, as shown in Figure 1, which describes the distribution of 741 bivariate Pearson correlation coefficients between search trends of 39 queries. The bivariate correlation coefficients range from 0.69 to 0.99, with a mean equal to 0.90 and a standard deviation equal to 0.06. The Cronbach's α of the search trends of 39 queries is equal to 0.88, suggesting that the dynamics of these multiple queries can be combined as a one-dimensional factor to gauge public attention on environmental issues (Cronbach, 1951; Babbie, 2013). Althought it is conceptually desirable and theoretically interesting to measure public attention on various facets of environment issues, it is statistically inappropriate to retain multiple facets of public attention on environmental issue in our study. Therefore, a composite measurement of public attention on environmental issues is developed by averaging the monthly ratio of 39 queries. The composite GT dynamics can help researchers fully observe public attention on environmental issues which cannot be fulfilled by any single query (e.g. Curme *et al.*, 2014). The dynamics of the composite measurement will be compared with that measured by Gallup MIP to assess the measurement properties of Google Trends.

Correlation between GT dynamics and Gallup dynamics

A reduced-form VAR model is fitted to GT dynamics and Gallup dynamics. The time lags for the VAR model that are four months which are determined based on likelihood ratio tests between models with different time lags. The VAR model results are summarized in Table II. The Durbin-Watson *d*-statistics of the VAR model is close to 2, suggesting that no autocorrelations are observed among residuals of two time series (Durbin and Watson, 1950). Unlike traditional multiple regression analysis, hypotheses testing in VAR is not based on traditional statistics, such as the statistical significance

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

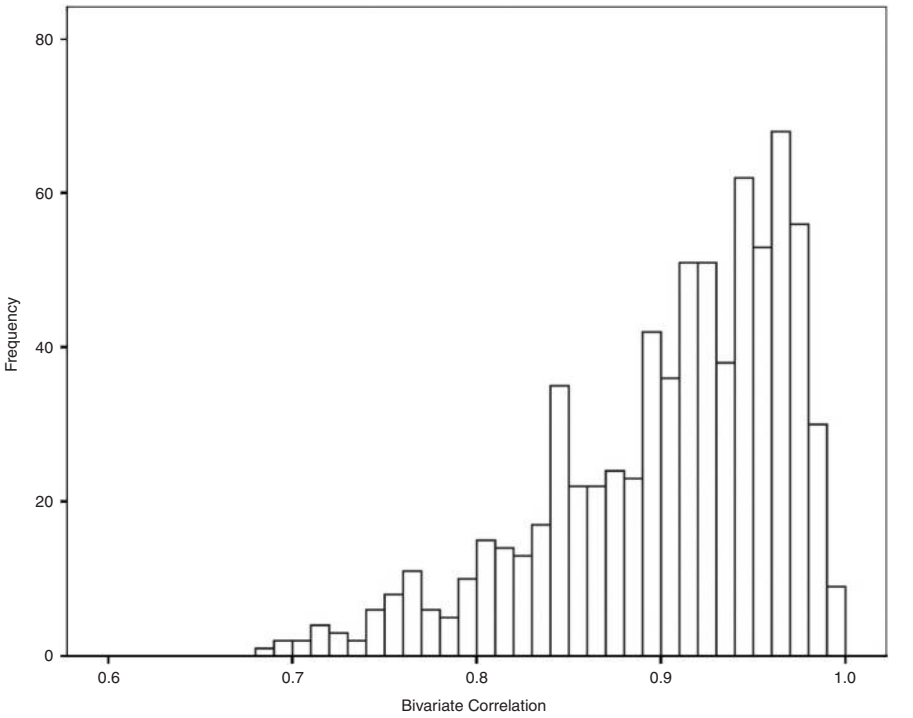

Figure 1.
Distribution of bivariate correlation coefficients between search trends of 39 queries

and magnitude of individual coefficients and the overall fit of the model (e.g. Freeman *et al.*, 1989; Stock and Watson, 2001). VAR tests hypotheses by "assessing the joint statistical significance of the coefficients on single variables or on blocks of endogenous variables" (Hannan and Freeman, 1984, p. 845), which is fulfilled by Granger causality analysis and impulse response function analysis.

Granger causality analysis examines whether lagged values of one variable help to predict another (Granger, 1969). We found that GT dynamics Granger-causes Gallup dynamics on environmental issues and Gallup dynamics also Granger-causes GT dynamics on environmental issues. Although the Granger causality analysis provides evidence concerning the joint effects of lagged Gallup dynamics on GT dynamics and vice versa, it does not provide an indication of the polarity of relations (Wood and Peake, 1998). Impulse response function is used in this study to identify the direction of relations between GT dynamics and Gallup dynamics, by examining how the current and future values of one variable (e.g. Gallup dynamics) will respond to a one-unit increase in the current values of another variable (e.g. GT dynamics) in the model. The results of the impulse response function analysis for VAR models are shown in Figure 2.

As shown by the solid line in Figure 2, in response to a one-unit increase in GT dynamics on environmental issues, Gallup dynamics will consistently shift upward by approximately 0.01-0.13 standard deviations in the following 12 months. In response to a one-unit increase in Gallup dynamics on environmental issues, GT dynamics will first shift downward by approximately 0.02 standard deviations on the first month and then shift upward by approximately 0.01-0.04 standard deviations in the following 11 months, as shown by the dashed line in Figure 2. Therefore, we conclude that a positive

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

VAR estimates

	Dependent variable = Gallup dynamics$_t$		Dependent variable = GT dynamics$_t$	
	Coefficients	SE	Coefficients	SE
Gallup dynamics$_{t-1}$	0.08	0.16	−0.01	0.04
Gallup dynamics$_{t-2}$	0.12	0.16	0.08	0.04
Gallup dynamics$_{t-3}$	−0.02	0.16	−0.05	0.04
Gallup dynamics$_{t-4}$	−0.01	0.15	0.07	0.03
GT dynamics$_{t-1}$	0.37	0.52	1.15	0.12
GT dynamics$_{t-2}$	0.24	0.81	−0.54	0.19
GT dynamics$_{t-3}$	0.56	0.79	−0.09	0.18
GT dynamics$_{t-4}$	−0.15	0.52	0.19	0.12
Constant	0.19	0.19	0.03	0.04

Granger causality test

Granger-causal relation	GT dynamics→Gallup dynamics	Gallup dynamics→GT dynamics
F-statistics	$F = 4.0$, $p < 0.01$	$F = 3.8$, $p < 0.01$

Forecast error variance decomposition

	Dependent variable = Gallup dynamics		Dependent variable = GT dynamics	
Forecast period	Gallup dynamics (%)	GT dynamics (%)	Gallup dynamics (%)	GT dynamics (%)
1st month	100	0	2	98
2nd month	99	1	3	97
3rd month	97	3	3	97
4th month	90	10	4	96
5th month	87	13	9	91
6th month	86	14	11	89
7th month	86	14	12	88
8th month	86	14	11	89
9th month	85	15	11	89
10th month	85	15	11	89
11th month	84	16	11	89
12th month	84	16	12	88

Notes: The VAR model contains four lags (i.e. months) and 54 monthly observations in the test sample. Unlike traditional multiple regression analysis, the individual estimated coefficients and their standard errors convey limited information in VAR (Freeman *et al.*, 1989; Stock and Watson, 2001). Therefore, these statistics are reported here for the readers' reference

Table II.
VAR model results

correlation exists between Gallup dynamics and GT dynamics on environmental issues. In short, Granger causality analysis and item response function analysis provide consistent evidence that a positive correlation exists between GT dynamics and Gallup dynamics on environmental issues. In other words, when an increase (or a decline) in public attention on environmental issues is observed in public opinion polls, the same upward (or downward) pattern can be detected by examining the composite search trends of these 39 queries.

Forecasting performance between GT dynamics and Gallup dynamics

After testing correlations between GT dynamics and Gallup dynamics on environmental issues, further considering how well GT dynamics and Gallup dynamics can forecast each other is necessary. First, forecast error variance decomposition analysis is used to

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

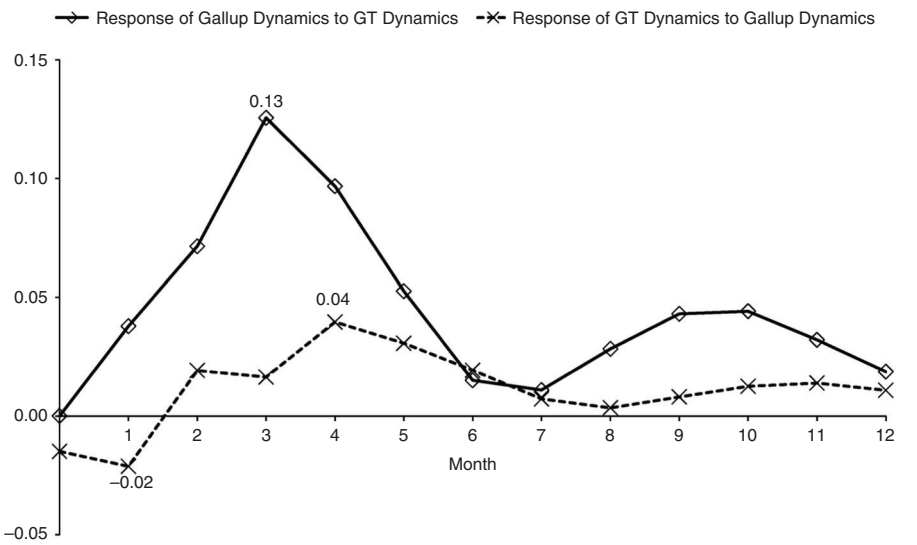

Figure 2.
Impulse response function analysis results

examine the forecasting performance of one variable (e.g. GT dynamics) on future values of the other variable (e.g. Gallup dynamics) in VAR models (Stock and Watson, 2001). The forecast error variance decomposition, like a partial R^2 for the forecast error, refers to the percentage of the variance of the error generated in forecasting a variable (e.g. Gallup dynamics) because of a specific shock (e.g. the error term in GT dynamics) at a given time frame (e.g. 12 months) (Stock and Watson, 2001).

As shown in the bottom part of Table II, GT dynamics accounts for an increasing amount of variance of the error in the forecast of Gallup dynamics on environmental issues, from 0 percent at the first month of the forecast period to 10 percent at the fourth month, to 13 percent at the fifth month, and maintained at approximately 15 percent throughout the remaining seven months. Gallup dynamics has a role in forecasting GT dynamics on environmental issues. At the first month of the forecast period, 2 percent of the variance of the error in the forecast of GT dynamics on environmental issues can be attributed to Gallup dynamics, which increases to 4 percent at the fourth month, to 9 percent at the fifth month, and maintained at approximately 12 percent throughout the remaining seven months.

Finally, the forecasting performance of one measure (e.g. Gallup dynamics) on another measure (e.g. GT dynamics) is examined in a holdout sample. The aforementioned VAR models are used to conduct a 12-month forecast for Gallup dynamics and GT dynamics. Figure 3 shows the fitness between the forecasted trends and observed trends of GT dynamics and Gallup dynamics.

As regards Gallup dynamics on environmental issues, the forecasted trend seems deviate from the observed trend, as shown in the left panel of Figure 3. Nevertheless, all the observed values fall within the forecasted interval at 95 percent confidence level. More importantly, the overall pattern of forecasted and observed trends in the whole 12-month forecasting period is upward. Considering the forecast error variance decomposition results, we conclude that GT dynamics can forecast Gallup dynamics on environmental issues. A perfect fit between forecasted trend and observed trend is observed for GT dynamics on environmental issues. As shown in the right panel of Figure 3, the forecasted

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

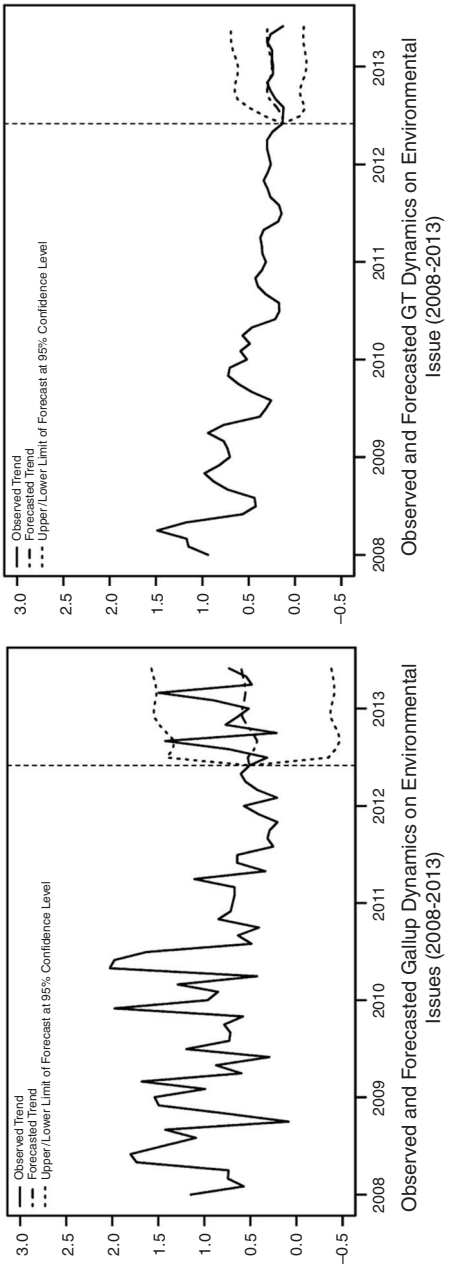

Figure 3.
Forecasting of GT dynamics and Gallup dynamics on environmental issues in holdout samples

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

trend of GT dynamics is almost duplicated by the observed trend in the forecasting time period (i.e. July 2012-June 2013). Moreover, as suggested in earlier forecast error variance decomposition, Gallup dynamics has an important role in forecasting GT dynamics on environmental issues. Therefore, Gallup dynamics can well forecast GT dynamics on environmental issues.

Conclusions and discussion

Social scholars have recently turned to web queries as an alternative measure of public attention on various social issues, as web queries is a continuously updating source of information about what issues and concerns the general public are thinking about. Despite its empirical applications in various domains, two fundamental questions remain unresolved, as follows: first, how to identify niche queries that can be used to measure public attention on various social issues; and second, how to assess the measurement properties of queries as a new device of an established concept (i.e. public attention). Focussing on public attention on environmental issues in the USA, the current study aims to fill in the gap by designing an approach to generate a list of queries relevant to public attention on environmental issues and comparing the attention dynamics measured by queries data with that measured by public opinion polls. The findings of the study advance our understanding of the prospects and challenges in employing web queries as a measurement tool to observe public attention on social issues.

First, our findings suggest that queries data are complementary to, rather than substitutes of, opinion polls in measuring public attention. We make the first attempt, to our knowledge, to empirically assess the measurement properties of web queries in observing public attention on social issues. As a new device to gauge public attention on environmental issues, the measurement properties of queries are well established in the study when public attention on the same issue measured by Gallup MIP is used as a benchmark. Public attention on environmental issues measured by Google Trends is observed to positively correlate with that measured by Gallup MIP. Public attention on environmental issues measured by Google Trends and that measured by Gallup MIP can forecast each other well in a 12-month time frame. Moreover, Google Trends can improve the prediction of Gallup MIP on environmental issues, in addition to the past values of Gallup MIP.

In a practical sense, our study suggests that queries can be used not only as a surveillance of public attention on environmental issues among the general public in the USA *talis qualis* but also as a tool used to forecast public attention on environmental issues in the USA. Considering that Google Trends is a publicly available and regularly updated data source, policymakers or environmentalists can utilize it to monitor the status of public attention on environmental issues and implement some campaigns to mobilize favorable public opinion when the decline of public attention can be predicted by Google Trends (Mccallum and Bury, 2013).

Second, an iterative framework is developed to generate valid queries to measure public attention on certain social issues. The iterative framework developed in the current study combines top-down and bottom-up approaches, which can help researchers identify a glossary of valid queries that internet users will search for on specific issues (i.e. environmental issue in the study). The iterative framework is implemented with three steps. First, a top-down approach is used to generate seed queries directly relevant to environmental issues from an extensively accepted measure of public attention on environmental issues (i.e. Gallup MIP). Then, a bottom-up

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

approach is adopted by submitting these seed queries to Google Correlate to produce a longer list of queries relevant to environmental issues. Finally, the retrieved queries are manually screened to ensure the face validity of retained queries.

This iterative framework is of methodological implication which can minimize researchers' subjective biases in selecting search queries and restore the self-initiated nature of individuals' web search behavior. In previous studies, researchers adopted an a priori determined list of queries based on their subjective judgment (i.e. top-down approach) to measure public attention on various social issues, as the original search log files are not accessible to them. However, this top-down approach will devalue queries data in measuring public attention, since "the frequency of searches for a given keyword can grow and decline for various reasons, some of which may or may not be related to a real-world event of interest" (Curme *et al.*, 2014, p. 11604). More importantly, as users' online information-seeking behavior is self-initiated by nature. The queries generated from this framework can approximate the real life and real needs of real users (Jansen *et al.*, 2000, 2011), which can help improve the validity of web queries in measuring public attention. Moreover, multiple queries identified through this iterative framework, just like composite measures in traditional survey research (Babbie, 2013), can help fully observe public attention on diverse dimensions of environmental issues. In addition, as Curme *et al.* (2014) have noticed, "valuable information may be contained in search engine data for keywords with less-obvious semantic connections to the event in question (p. 11604)." In our classification, specific queries have less-obvious semantic connections to global warming. Scholars in the future are expected to compare the performance of different categories of queries and try to find the most reliable and the most cost-effective combination of queries to measure public attention on environmental and various issues.

Third, the study develops a benchmark approach to normalize the search magnitudes of different queries retrieved from Google Trends to a same scale, which facilitate the processing and analysis of the big data retrieved from Google Trends by a bigger audience. Utilizing web queries data released by Google and other search engine companies is just like a cat-and-mouse game (Huberman, 2012; Lazer *et al.*, 2014). On the one hand, search engine companies try to build an algorithm black box, in which they process the raw search volumes before public release, in order to protect their commercial interests. On the other hand, researchers try to figure out those algorithms to verify the results and make better use of the data. As we argued earlier in the article, Google Trends normalizes the search trend of each query and calibrates the trend on a 100 scale before public release (Google, 2014). After the normalization and scaling, researchers can only make within-query comparison (i.e. observing the rise and fall of search trends of specific queries over time), while it is technically impossible for researchers to make between-query comparisons (i.e. comparing the search trends between different queries), which seriously limits the use of search queries data in social research. Our benchmark approach assigns a shared scale, which is the search trend of the benchmark query (i.e. "IT Job" in our study), to the trends of all search queries by submitting the search queries as well as the benchmark query to Google Trends in a pairwise way. This benchmark approach allows researchers to make between-query comparisons between the search trends of different queries, which empower researchers to process the data in a more creative way (e.g. averaging the search trends of different queries). Therefore, we believe that the benchmark approach offers an alternative and feasible way in utilizing the Google Trends data without knowing what are present in the algorithm black box, which can help improve the applicability of Google Trends data in various research contexts.

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

Last but not least, we cannot exaggerate our findings to claim that researchers can unconditionally accept queries as a new device to gauge public attention on various issue domains. Instead, researchers are advised to use web queries with caution by taking issue peculiarity and characteristics of web queries into account (Mellon, 2011). As query trends are constrained by the wording of queries (e.g. Mccallum and Bury, 2013; Ripberger, 2011; Zhu *et al.*, 2012), future empirical studies are expected to examine how different wordings, different combinations of queries, and queries in different languages will affect their measurement validity on various social issues. Future research is encouraged to utilize and verify the list offered in this study and, furthermore, to enrich the glossary of search queries concerning environmental issues in different regions, as individuals in different societies may have different information-seeking patterns on the same issue.

Note

1. The original list of 693 queries will be provided on request.

References

Agnone, J. (2007), "Amplifying public opinion: the policy impact of the US environmental movement", *Social Forces*, Vol. 85 No. 4, pp. 1593-1620.

Althouse, B.M., Ng, Y.Y. and Cummings, D.A.T. (2011), "Prediction of dengue incidence using search query surveillance", *PLoS Neglected Tropical Diseases*, Vol. 5 No. 8, pp. 1-7.

Babbie, R. (2013), *The Practice of Social Research*, 13th ed., Wadsworth Cengage Learning, Boston, MA.

Baker, D.A. and Bagozzi, R.P. (1982), "Attitudes toward public policy alternatives to reduce air pollution", *Journal of Public Policy & Marketing*, Vol. 1 No. 1, pp. 85-94.

Bishop, G.F., Oldendick, R.W., Tuchfarber, A.J. and Bennett, S.E. (1980), "Pseudo-opinions on public affairs", *Public Opinion Quarterly*, Vol. 44 No. 2, pp. 198-209.

Carneiro, H.A. and Mylonakis, E. (2009), "Google Trends: a web-based tool for real-time surveillance of disease outbreaks", *Clinical Infectious Diseases*, Vol. 49 No. 10, pp. 1557-1564.

Carrière-Swallow, Y. and Labbé, F. (2013), "Nowcasting with Google Trends in an emerging market", *Journal of Forecasting*, Vol. 32 No. 4, pp. 289-298.

Carson, R. (1962), *Silent Spring*, Houghton Mifflin Company, Boston, MA.

Castillo, C., Mendoza, M. and Poblete, B. (2013), "Predicting information credibility in time-sensitive social media", *Internet Research*, Vol. 23 No. 5, pp. 560-588.

comScore (2013), "comScore releases July 2013 US search engine rankings", available at: www.comscore.com/Insights/Press-Releases/2013/8/comScore-Releases-July-2013-US-Search-Engine-Rankings (accessed February 12, 2014).

Cook, S., Conrad, C., Fowlkes, A.L. and Mohebbi, M.H. (2011), "Assessing Google Flu Trends performance in the United States during the 2009 influenza virus A (H1N1) pandemic", *PLoS ONE*, Vol. 6 No. 8, pp. 1-8.

Cronbach, L. (1951), "Coefficient alpha and the internal structure of tests", *Psychometrika*, Vol. 16 No. 3, pp. 297-334.

Curme, C., Preis, T., Stanley, H.E. and Moat, H.S. (2014), "Quantifying the semantics of search behavior before stock market moves", *Proceedings of National Academy of Sciences of the United States of America*, Vol. 111 No. 32, pp. 11600-11605.

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

Daniels, D.P., Krosnick, J.A., Tichy, M.P. and Tompson, T. (2013), "Public opinion on environmental policy in the United States", in Kamieniecki, S. and Kraft, M. (Eds), *The Oxford Handbook of US Environmental Policy*, Oxford University Press, Oxford, pp. 1-19.

Downs, A. (1972), "Up and down with ecology: the issue attention cycle", *Public Interest*, Vol. 28 No. 1, pp. 38-50.

Dunlap, R.E. (1991), "Trends in public opinion toward environmental issues: 1965-1990", *Society & Natural Resources*, Vol. 4 No. 3, pp. 285-312.

Dunlap, R.E. (1995), "Public opinion and environmental policy", in Lester, J.P. (Ed.), *Environmental Politics and Policy: Theories and Evidence*, Duke University Press, Durham, NC, pp. 63-114.

Durbin, J. and Watson, G.S. (1950), "Testing for serial correlation in least squares regression: I", *Biometrika*, Vol. 37 Nos 3/4, pp. 409-428.

Fondeur, Y. and Karamé, F. (2013), "Can Google data help predict French youth unemployment?", *Economic Modelling*, Vol. 30 No. 1, pp. 117-125.

Freeman, J.R., Williams, J.T. and Lin, T. (1989), "Vector autoregression and the study of politics", *American Journal of Political Science*, Vol. 33 No. 4, pp. 842-877.

Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S. and Brilliant, L. (2009), "Detecting influenza epidemics using search engine query data", *Nature*, Vol. 457 No. 7232, pp. 1012-1014.

Google (2013), "FAQ of Google Correlate", available at: www.google.com/trends/correlate/faq (accessed October 10, 2014).

Google (2014), "How Trends data is normalized", available at: https://support.google.com/trends/answer/4365533?hl = en&ref_topic = 13975&rd = 1 (accessed October 15, 2013).

Granger, C.W.J. (1969), "Investigating causal relations by econometric models and cross-spectral methods", *Econometrica*, Vol. 37 No. 3, pp. 424-438.

Green-Pedersen, C. and Stubager, R. (2010), "The political conditionality of mass media influence: when do parties follow mass media attention?", *British Journal of Political Science*, Vol. 40 No. 3, pp. 663-677.

Hannan, M.T. and Freeman, J. (1984), "Structural inertia and organizational change", *American Sociological Review*, Vol. 49 No. 2, pp. 149-164.

Huang, J., Zheng, R. and Emery, S. (2013), "Assessing the impact of the national smoking ban in indoor public places in China: evidence from quit smoking related online searches", *PLoS ONE*, Vol. 8 No. 6, pp. 1-10.

Huberman, B.A. (2012), "Sociology of science: big data deserve a bigger audience", *Nature*, Vol. 482 No. 7385, 308pp.

Internet World Stats (2013), "Internet usage statistics", available at: www.internetworldstats.com/stats.htm (accessed January 15, 2014).

Jansen, B.J., Booth, D.L. and Spink, A. (2008), "Determining the informational, navigational, and transactional intent of web queries", *Information Processing & Management*, Vol. 44 No. 3, pp. 1251-1266.

Jansen, B.J., Spink, A. and Saracevic, T. (2000), "Real life, real users, and real needs: a study and analysis of user queries on the web", *Information Processing & Management*, Vol. 36 No. 2, pp. 207-227.

Jansen, B.J., Liu, Z., Weaver, C., Campbell, G. and Gregg, M. (2011), "Real time search on the web: queries, topics, and economic value", *Information Processing & Management*, Vol. 47 No. 4, pp. 491-506.

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)

Jungherr, A. and Jürgens, P. (2013), "Forecasting the pulse: how deviations from regular patterns in online data can identify offline phenomena", *Internet Research*, Vol. 23 No. 5, pp. 589-607.

Kalampokis, E., Tambouris, E. and Tarabanis, K. (2013), "Understanding the predictive power of social media", *Internet Research*, Vol. 23 No. 5, pp. 544-559.

Lazer, D., Kennedy, R., King, G. and Vespignani, A. (2014), "The parable of Google Flu: traps in big data analysis", *Science*, Vol. 343 No. 6176, pp. 1203-1205.

Mccallum, M. and Bury, G. (2013), "Google search patterns suggest declining interest in the environment", *Biodiversity and Conservation*, Vol. 22 Nos 6-7, pp. 1355-1367.

McCombs, M.E. and Zhu, J.-H. (1995), "Capacity, diversity, and volatility of the public agenda: trends from 1954 to 1994", *Public Opinion Quarterly*, Vol. 59 No. 4, pp. 495-525.

Mellon, J. (2011), "Search indices and issue salience: the properties of Google Trends as a measure of issue salience", sociology working papers, University of Oxford, Oxford, available at: www.sociology.ox.ac.uk/materials/papers/2011-01-1.pdf (accessed December 3, 2013).

Mellon, J. (2013), "Where and when can we use Google Trends to measure issue salience?", *Political Science & Politics*, Vol. 46 No. 2, pp. 280-290.

Newig, J. (2004), "Public attention, political action: the example of environmental regulation", *Rationality and Society*, Vol. 16 No. 2, pp. 149-190.

Nisbet, M.C. and Myers, T. (2007), "The polls – trends: twenty years of public opinion about global warming", *Public Opinion Quarterly*, Vol. 71 No. 3, pp. 444-470.

Preis, T., Moat, H.S. and Stanley, H.E. (2013), "Quantifying trading behavior in financial markets using Google Trends", *Scientific Reports*, Vol. 3, pp. 1-6, available at: www.nature.com/articles/srep0168

Price, V. and Neijens, P. (1997), "Opinion quality in public opinion research", *International Journal of Public Opinion Research*, Vol. 9 No. 4, pp. 336-360.

Price, V. and Roberts, D.F. (1987), "Public opinion process", in Chaffee, S.H. and Berger, C.R. (Eds), *Handbook of Communication Science*, Sage, Beverly Hills, CA, pp. 781-816.

Project for Improved Environmental Coverage (2013), "Environmental coverage in the mainstream news: we need more", available at: http://greeningthemedia.org/wp-content/uploads/Environmental-Coverage-in-the-Mainstream-News.pdf (accessed December 3, 2013).

Reilly, S., Richey, S. and Taylor, J.B. (2012), "Using Google search data for state politics research: an empirical validity test using roll-off data", *State Politics & Policy Quarterly*, Vol. 12 No. 2, pp. 146-159.

Ripberger, J.T. (2011), "Capturing curiosity: using internet search trends to measure public attentiveness", *Policy Studies Journal*, Vol. 39 No. 2, pp. 239-259.

Rutherford, A. (2014), "Thoughts on the Google Flu Trends", United Nations Global Pulse, available at: http://unglobalpulse.org/google-flu-trends (accessed May 13, 2014).

Schoen, H., Gayo-Avello, D., Metaxas, P.T., Mustafaraj, E., Strohmaier, M. and Gloor, P. (2013), "The power of prediction with social media", *Internet Research*, Vol. 23 No. 5, pp. 528-543.

Segars, A.H. (1997), "Assessing the unidimensionality of measurement: a paradigm and illustration within the context of information systems research", *Omega*, Vol. 25 No. 1, pp. 107-121.

Sims, C.A. (1980), "Macroeconomics and reality", *Econometrica*, Vol. 48 No. 1, pp. 1-48.

Smith, T.W. (1985), "The polls: America's most important problems part I: national and international", *Public Opinion Quarterly*, Vol. 49 No. 2, pp. 264-274.

Soroka, S.N. (2002a), *Agenda-Setting Dynamics in Canada*, UBC Press, Vancouver.

Soroka, S.N. (2002b), "Issue attributes and agenda-setting by media, the public, and policymakers in Canada", *International Journal of Public Opinion Research*, Vol. 14 No. 3, pp. 264-285.

Stamm, K., Clark, F. and Eblacas, P. (2000), "Mass communication and public understanding of environmental problems: the case of global warming", *Public Understanding of Science*, Vol. 9 No. 3, pp. 219-237.

Stock, J.H. and Watson, M.W. (2001), "Vector autoregressions", *Journal of Economic Perspectives*, Vol. 15 No. 4, pp. 101-115.

Walcott, B.P., Nahed, B.V., Kahle, K.T., Redjal, N. and Coumans, J.-V. (2011), "Determination of geographic variance in stroke prevalence using internet search engine analytics", *Neurosurgical Focus*, Vol. 30 No. 6, pp. 1-4.

Wood, B.D. and Peake, J.S. (1998), "The dynamics of foreign policy agenda setting", *The American Political Science Review*, Vol. 92 No. 1, pp. 173-184.

Worcester, R. (1998), "Public opinion and the environment", in Jacobs, M. (Ed.), *Greening the Millennium: The New Politics of the Environment*, Political Quarterly Monograph Series, Wiley, Oxford, pp. 160-173.

Yaffee, R. (2000), *Introduction to Time Series Analysis and Forecasting*, Academic Press, New York, NY.

Yeager, D.S., Larson, S.B., Krosnick, J.A. and Tompson, T. (2011), "Measuring Americans' issue priorities: a new version of the Most Important Problem question reveals more concern about global warming and the environment", *Public Opinion Quarterly*, Vol. 75 No. 1, pp. 125-138.

Zhou, X., Ye, J. and Feng, Y. (2011), "Tuberculosis surveillance by analyzing Google Trends.", *IEEE Transactions on Bio-medical Engineering*, Vol. 58 No. 8, pp. 2247-2254.

Zhu, J.J.H., Wang, X., Qin, J. and Wu, L. (2012), "Assessing public opinion trends based on user search queries: validity, reliability, and practicality", paper presented at the Annual Conference of the World Associate for Public Opinion Research, Hong Kong, June 12-14.

Zimmer, M.R., Stafford, T.F. and Stafford, M.R. (1994), "Green issues: dimensions of environmental concern", *Journal of Business Research*, Vol. 30 No. 1, pp. 63-74.

Further reading

Cronbach, L.J. and Meehl, P.E. (1955), "Construct validity in psychological tests", *Psychological Bulletin*, Vol. 52 No. 4, pp. 281-302.

About the authors

Dr Jie Qin is a Professor of Communications at the School of Liberal Arts, SCAD Hong Kong, Hong Kong S.A.R., China.

Dr Tai-Quan Peng currently is an Assistant Professor at the Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore. His research outputs have appeared in SSCI/SCI-listed journals, including *Communication Research*, *New Media & Society*, *Computers in Human Behavior*, *IEEE Transactions on Visualization and Computer Graphics*, *Journal of Informetrics*, *Journal of the American Society for Information Science and Technology*, *Scientometrics*, and elsewhere. Dr Tai-Quan Peng is the corresponding author and can be contacted at: winsonpeng@gmail.com

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**

Downloaded by TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES At 20:29 09 November 2016 (PT)