



---

## Google chemtrails: A methodology to analyze topic representation in search engine results

by Andrea Ballatore

---

### Abstract

Search engine results influence the visibility of different viewpoints in political, cultural, and scientific debates. Treating search engines as editorial products with intrinsic biases can help understand the structure of information flows in new media. This paper outlines an empirical methodology to analyze the representation of topics in search engines, reducing the spatial and temporal biases in the results. As a case study, the methodology is applied to 15 popular conspiracy theories, examining type of content and ideological bias, demonstrating how this approach can inform debates in this field, specifically in relation to the representation of non-mainstream positions, the suppression of controversies and relativism.

### Contents

- [1. Introduction](#)
  - [2. Search engines and their effects](#)
  - [3. A case study: Conspiracy theories](#)
  - [4. Methodology](#)
  - [5. Analysis of search engine representation](#)
  - [6. Findings from the case study](#)
  - [7. Conclusion](#)
- 

### 1. Introduction

Far from being neutral aggregators of information, search engines are replacing the manual models of content filtering and gatekeeping of previous media with complex automated tools. In the process of crawling, indexing, and structuring Web content, search engines create an informational infrastructure with specific characteristics and biases. In parallel to the emergence of these new global information gatekeepers, the production and spread of media content has also changed dramatically. The reduced barriers to online publication and the explosion of blogging, forums, and social media have opened uncharted territory to content producers whose narratives have found new audiences.

Once confined to the technical spheres of computer science and information retrieval, search engines are now notable objects of study for several, complementary disciplines. Social scientists analyze mainstream search engines for their cultural, cognitive, and political implications (Spink and Zimmer, 2008; Halavais, 2009; Vaidhyanathan, 2011; Graham, *et al.*, 2014; König and Rasch, 2014). Large-scale, general-purpose search engines have become powerful gatekeepers of information, having enormous impact on flows of information, beliefs, and ideas. As Hillis, *et al.* (2013) pointed out, search technologies exert powerful socio-economic and political influence on society. Considering their ubiquity, Grimmelmann (2010) states that "search engines are the new mass media ... capable of shaping public discourse itself." [1] As the boom of the search engine optimization (SEO) field shows, search engines are already a central part of the media landscape, and the nature and effects of their biases should be taken seriously. As media and communication scholars have long investigated the biases of newspapers, radio and TV channels (Entman, 2007), it is worth conducting the same enterprise on search engines.

This paper contributes to this area of research by designing a methodology to collect and analyze such search engine results, reducing the geographic and temporal biases in the data to extract stable,

representative Web content. Such stable content can be then treated as editorial content and analyzed along multiple dimensions. As a case study, a selection of 15 conspiracy theories is used to illustrate the methodology on controversial topics, focusing on Google Search and Microsoft Bing as popular search engines. The methodology is deployed to study the representation of these 15 topics, answering the following questions:

- (i) what type of content is returned by search engines when searching for conspiracy theories?
- (ii) what is the bias of the search results towards the conspiratorial, neutral, or debunking Web sites?
- (iii) are there differences between search engines?
- (iv) what differences exist between conspiracy theories?
- (v) how polarized are the results between conspiratorial and debunking results?



---

## 2. Search engines and their effects

The ubiquity and dominance of search engines has attracted attention about their broad effects on society. Grimmelmann (2013) frames the role of search engines, and Google in particular, according to three complementary perspectives. Engines are seen as conduits in traditional communication networks, delivering content to consumers, and as advisors, suggesting content to users based on their specific informational needs. Alternatively, and more importantly for this paper, search engines can be compared to newspaper editors, selecting content to be shown to readers by calibrating their algorithms. In this sense, when Google's search quality team meets to optimize its algorithms, it resembles "a newspaper staff debating which stories to put on the front page of the metro section" [2], embedding specific biases that in turn shape the reality of their users.

A number of claims have been made about search engines' effects. Mainstream search engines might either empower marginal groups, or reinforce the dominant position of the powerful (Introna and Nissenbaum, 2000); according to Vaidhyanathan (2011), they arbitrarily organize and serve up content, while claiming that they adopt "objective" criteria of relevance, showing worrying degrees of capital concentration, and lack of transparency. Moreover, search engines are accused of favoring popularity over the content's trustworthiness and credibility (Lewandowski, 2012). Others have claimed that search results present geographic coverage biases (Vaughan and Thelwall, 2004), while suppressing controversial topics (Gerhart, 2004). In principle, they provide new platforms to marginal groups, but Reilly (2008) claims that old media remain dominant. From a more optimistic viewpoint, Goldman (2008) observes that search engine bias is intrinsic to the optimization of results, and that search personalization will mitigate the negative aspects of the engines' hegemony. While these studies demonstrate the breadth of interest in the topic, the empirical evidence presented in these arguments tends to be limited and often inconclusive, primarily because of the difficulties of extracting reproducible results. Wouters and Gerbec (2003) observed that search engines "do not present the results in a way that is suitable for the creation of data sets", a statement that the present paper aims at countering.

While information and library scientists have investigated search behavior for a long time (Hargittai, 2007), the impact of search engines *qua* media on opinion formation has been only marginally studied, and deserves much more interdisciplinary empirical investigation (Brossard and Scheufele, 2013). The credibility (and perception of credibility) of search results bears a crucial role in opinion formation and yet remains largely unexamined. When using search engines, users are prone to many subconscious biases, which heavily influence information processing and perception. When examining Web pages, users consider the quality of their design and ease of use, while the author's profile, credentials, and affiliation tend to be ignored (Eysenbach and Köhler, 2002; Fogg, *et al.*, 2003). Experimental investigations by Pan, *et al.* (2007) and Keane, *et al.* (2008) suggest that users tend to trust search engine's ranking in a strikingly uncritical way, providing indirect arguments in favor of the study of the representation of topics in search engine results.



---

## 3. A case study: Conspiracy theories

In order to illustrate the methodology to extract a stable representation of a topic in a search engine, conspiracy theories were selected as a case study providing highly controversial and polarized media content. Because of their ubiquity and enduring popularity, conspiracy theories attract psychologists and political scientists, interested in the variety of social and psychological conditions that might favor

“conspiratorial thinking,” such as dispossession, powerlessness, political alienation, social exclusion, and low levels of education (Knight, 2000; Clarke, 2002). While conspiratorial thinking has been strongly present in the public sphere since the nineteenth century (Hofstadter, 1964), conspiracy theories have found in the Web their ideal medium for the twenty-first century. Web 2.0 platforms have generally increased the visibility of fringe beliefs, which used to be confined to local media ecosystems.

In the context of the explosion of unfiltered blogs, news Web sites, amateur videos, and mash-ups, Wood (2013) has recently surveyed the scarce research on this topic. In his view, on the one hand, Web-based communication might dissolve conspiracy theories into highly polarized and radical (but overall irrelevant) narratives. On the other hand, conspiracy theories might arise to new levels of mainstream visibility and legitimacy, as observed in the case of the 9/11 truth movement (Wood, *et al.*, 2012). While the diffusion of false news stories on social media has received attention (*e.g.*, Mocanu, *et al.*, 2014), search engines are understudied in their potential to spread and reinforce beliefs. Conspiracy theorists often invite readers to avoid traditional, mainstream newspapers and TV, and to do their research on search engines (see, for example, [Figure 1](#)).



**Figure 1:** “Google chemtrails” banner (Source: <http://hartkeisonline.com>). Chemtrails are believed to be secret (and poisonous) geo-engineering practices.

As Clarke (2002) has pointed out, conspiracy theories escape a clear-cut definition. In a Wittgensteinian family resemblance, the label “conspiracy theory” is used, usually in a derogatory way, to refer to narratives with a number of recurring features, which include: (i) overly complex and implausible explanations for phenomena that can be otherwise explained satisfactorily; (ii) arbitrary causal relations between unrelated individuals and events; (iii) reliance on poor quality evidence, epistemological fallacies, self-referential sources; (iv) focus on spectacular and popular events (*e.g.*, large disasters, deaths of celebrities); (v) under-estimation of coordination costs for the perpetrators; (vi) apophenia, pareidolia, and exaggerated perception of agency behind events, coupled with dismissal of non-deterministic and unintentional effects; and, (vii) a single-cause explanation is preferred to complex, multi-causal explanations. Usually, conspiracy theories are considered false by the scientific consensus, although exceptions exist in the forms of real conspiracies, which indeed occur in many arenas, usually at a small scale (Sunstein and Vermeule, 2009).

A wave of psychological, sociological, and media studies have tackled specific conspiracy theories in recent decades, particularly conspiracy theories deemed to pose tangible societal threats. Stempel, *et al.* (2007) investigated beliefs in 9/11 conspiracies, finding positive correlation between belief in conspiracy theories and consumption of blogs and tabloids, and membership of disadvantaged groups. More recently, Mocanu, *et al.* (2014) have investigated the propagation of false political news stories on social media, pointing out their persistence. In the medical area, vaccine-related conspiracy theories have attracted notable interest for their tangible damage in terms of health policies. In her study of the rhetoric of the anti-vaccination movement, Kata (2012) noted how post-modern relativism and a dubious degree from the “University of Google” can become the basis to make crucial decisions on children’s vaccination [3].

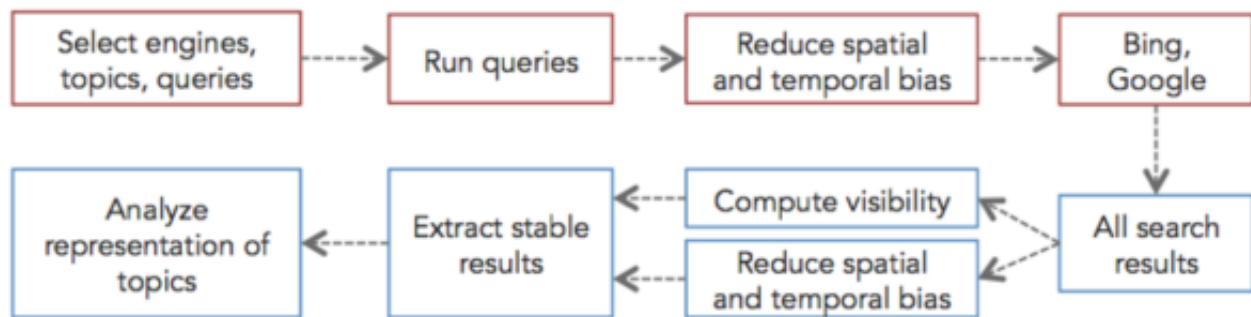
#### 4. Methodology

This new methodology was devised to investigate the representation of topics in search engines, treating the search results as editorial products with embedded biases. To illustrate the methodology,

the representation of 15 conspiracy theories was investigated with respect to the types of content and its ideological bias, providing a detailed description of methodological steps on real data. The work flow of the methodology is summarized in [Figure 2](#), starting from the selection of search engines, topics (*i.e.*, conspiracy theories), and text-based queries, to the data collection and analysis. The methodology follows these steps:

1. Select a sample of search engines and topics (in this case, Google, Bing, and 15 conspiracy theories).
2. For each topic, select a sample of text-based queries (in this case, six queries from Google Trends for each conspiracy theory).
3. Execute all the queries at different times through a decentralized proxy (in this case, the Tor network), and collect the search results (*i.e.*, ranked URLs).
4. For each URL, compute visibility score  $v$ . Extract stable results that maintain high visibility over time.

These stable results were then subsequently classified regarding the content type in five classes (*academic, blog, news, wikipedia, or misc*) and the content bias in five classes (*conspiratorial, neutral, debunking, related, unrelated*). For each engine, for each Web site, and for each topic, two indices — the Conspiracy Index (CI) and Polarization Index (PI) — are computed and analyzed. The remainder of this section describes each step in detail, discussing its advantages as well as limitations.



**Figure 2:** Workflow of the methodology, from the design phase to the implementation and analysis.

#### 4.1. Selection of search engines

The selection of a sample of search engines, restricting the inquiry to the Anglo-American world, did not present particular challenges. A 2014 report from comScore [\[4\]](#) clearly indicates that the vast majority of text-based searches in the U.S. are performed on Google (67 percent), Bing (18 percent), and Yahoo! (10 percent). In the U.K., a study by theEword [\[5\]](#) shows an even stronger dominance of Google (88 percent), followed by Bing (six percent), and Yahoo! (three percent). Hence, these search engines constitute an ideal sample for this study, representing about the 95 percent of the search engine market in the English-speaking world. As Yahoo! currently relies on the Bing index for its search product, only the two most influential engines, Google Search [\[6\]](#) and Microsoft Bing [\[7\]](#), were included in the study.

#### 4.2. Selection of topics

Given the definitional problems surrounding conspiracy theories, more caution was needed to select a suitable sample. Thousands of ephemeral narratives, rumors, urban legends, hoaxes, and fringe beliefs are relentlessly generated and flow through blogs, forums, and social media, leaving traces in search engines. Hence, to restrict the scope of the study to a set of salient case studies, precise criteria were designed and adopted. The sample includes 15 conspiratorial narratives that (i) are considered implausible by a wide scientific consensus; (ii) appeared at least a decade ago; (iii) occur in many variants but have stable core claims; (iv) are extensively present online; (v) currently have active supporters; and (vi) belong to diverse ideological beliefs (see [Table 1](#)).

The sample covers diverse categories of narratives, including political conspiracies (the assassination of John F. Kennedy and 9/11 as an “inside job”), historical conspiracies (Holocaust revisionism), and technological conspiracies (free energy suppression). While all of these conspiracy theories have stable core beliefs, they tend to appear in uncountable variants. For example, secret societies are often believed to carry out many plans, including depopulation, free energy suppression, and others. Although a detailed critique of these narratives is beyond the scope of this study, their plausibility

varies widely. Whilst the implausibility of the grotesque Reptilian Elite conspiracy by David Icke is self-evident, other cases, such as JFK's assassination, are more complex. Moreover, the boundaries between genuine scientific controversies about global warming and uninformed, naive, or malicious conspiracy theorizing by populist right-wing American media can be rather difficult to establish. Similarly, Holocaust denialism is deeply intermingled with legitimate academic historical scholarship. In this sense, the sample covers a wide range of conspiracy theories, both in terms of ideological leanings on the left-right spectrum, and in terms of epistemic plausibility.

#### 4.3. Selection of queries

Search engine users explore a topic by typing text-based queries. As users can enter any keyword or phrase in the input field, a set of highly representative queries had to be chosen for each conspiracy theory. To achieve this goal, it was necessary to access the most common queries submitted to the search engines. Google provides information on the most popular queries entered by its users. The Google Trends tool [8] enables the exploration of the topics searched by the engine's users. A topic (e.g., World War I) is presented with the most frequent semantically related queries (e.g., "world war," "ww1," "wwi," etc.).

For each of the 15 conspiracy theories, Google Trends was therefore utilized to obtain the most popular queries. For example, the top queries for chemtrails include "chemtrails," "what are chemtrails," and "contrails chemtrails." The manual inspection of the top queries revealed that the popularity of top queries decreases rapidly after the fifth or sixth query, with the Google Trends relative popularity measure decaying from 100 to 10, indicating that topics are searched predominantly with few queries. To reduce the bias in the query selection process, the task was carried out by the author with a collaborator familiar with the methodology and the topics. As a result, six queries were selected for each topic, for a total of 90 queries, listed in [Table 1](#).

<b>Table 1: Sample of 15 conspiracy theories with main claims, estimated origin, and six text-based queries that represent them, based on Google Trends data.</b>			
<b>Conspiracy theory</b>	<b>Core claims</b>	<b>Origin</b>	<b>Top text-based queries</b>
9/11	9/11 attacks were planned/helped by the U.S. Government/army/CIA/Jews.	2002	9 11 conspiracy; 9 11 truth; 9 11 theories; government did 9 11; 9 11 inside job; 9 11 government planned
AIDS-HIV	HIV does not cause AIDS. AIDS does not exist. AIDS was manufactured to attack minorities.	1980s	aids conspiracy; conspiracy of aids; hiv aids conspiracy; does aids exist; aids government conspiracy; aids conspiracy theory
Chemtrails	Airplanes controlled by a secret society spread chemicals to carry out geo-engineering or depopulation plans.	1996	chemtrails; chemtrails haarp; what are chemtrails; contrails chemtrails; chemtrail spraying;

			chemtrail planes
<i>Depopulation plan</i>	Secret organizations plan to reduce the world population.	1970s	depopulation conspiracy; depopulation agenda; illuminati symbols; bill gates depopulation; illuminati depopulation; agenda 21 depopulation
<i>Fake moon landings</i>	Moon landings were filmed in a studio by the U.S. government and NASA to win the space race against the Soviet Union.	1974	conspiracy moon landing; moon hoax; moon landing hoax; fake moon landing; moon landing conspiracy; nasa moon hoax
<i>Free energy suppression</i>	Technologies that produce unlimited and free energy are suppressed by energy corporations.	1850s	free energy suppression; free energy conspiracy; energy for free; free energy generator; nikola tesla conspiracy; tesla conspiracy theories
<i>Global warming denialism</i>	Global warming is a fake theory disseminated by liberals and left-wingers to make profits from environmental regulations.	1990s	climate hoax; global warming myth; climate change hoax; global warming fake; global climate hoax; climate warming hoax
<i>HAARP secret weapon</i>	U.S. project High Frequency Active Auroral Research Program (HAARP) is used as a weapon to cause tsunamis and earthquakes.	1990s	haarp; haarp conspiracy; haarp weather; haarp earthquakes; alaska haarp; haarp machine
	The Jewish Holocaust		holocaust denial; holocaust fake; holocaust never

<i>Holocaust revisionism</i>	perpetrated by the Nazis is a fabrication of Allied propaganda.	1960s	happened; denial of holocaust; holocaust deniers; irving holocaust denial
<i>Jewish conspiracy</i>	Financial crises, wars, and major disasters are caused by an international Jewish elite.	nineteenth century	jewish conspiracy; jewish world conspiracy; elders of zion; the rothschild conspiracy; international jewish conspiracy; world jewish conspiracy
<i>JFK</i>	Secret motives and actors behind John F. Kennedy's assassination, including CIA, Mafia, and Fidel Castro.	1960s	jfk assassination conspiracy; kennedy cia conspiracy; jfk conspiracy theories; kennedy conspiracy; who killed jfk; jfk assassination mafia
<i>Reptilian elite</i>	A species of extraterrestrial, lizard-like mutants secretly controls the world.	1990s	reptilians; icke reptilians; aliens reptilians; reptilians illuminati; greys reptilians; david icke reptilians
<i>Secret societies</i>	Freemasons, Illuminati, or other secret societies steer the world towards a totalitarian global government (the "new world order").	nineteenth century	secret societies; freemasons; freemasons conspiracy; illuminati; bilderberg group; new world order
<i>UFO cover up</i>	Governments have secret contacts with alien species.	1950s	ufos; ufo sightings; ufo cover up; alien cover up; roswell ufo; ufo area 51
<i>autism</i>	Vaccines can cause autism in children. Vaccines are	1990s	vaccines autism; vaccines cause autism; flu vaccine



	harmful and are promoted to damage the population.	conspiracy; swine flu conspiracy; mmr vaccines; autism mmr
--	----------------------------------------------------	---------------------------------------------------------------------------

#### 4.4. Data collection

To collect the data from Google and Bing, the technical complexity and the opaque workings of these search engines had to be taken into account. Each engine has several advanced personalization options, and has different versions for different markets (e.g., *google.com* for the U.S., *google.co.uk* for the U.K.), which return different results. To extract stable, representative results, the most complex aspect to tackle was the geographic and history-based personalization, present in both engines, i.e., the dynamic adaptation of search results based on the user's search history, settings, and current geographic location — typically based on the IP address of the request.

Initially, the personalization was disabled manually, with the goal of obtaining results as close to the product's default results as possible. A preliminary experiment was run from two IP addresses, one located in the U.S. and one in the U.K. Specific versions of each engine were selected explicitly for each experiment (U.S. and U.K. versions of Google and Bing). A small sample of 10 queries was executed on both machines, on the four versions of the search engines on three separate days. By comparing the results, it became clear that the engines completely ignored the manual settings and applied geographic personalization to the results, introducing a strong spatial bias. Since search engines are highly dynamic products whose algorithms are constantly re-engineered, the results also changed over time, returning different URLs in a different order. Hence, the extraction of representative results required three steps: (i) reduction of spatial bias, (ii) reduction of temporal bias, and, (iii) extraction of stable results.

#### 4.5. Reduction of spatial bias

To reduce the spatial bias, instead of executing the queries from the same IP address, an anonymization technique was used. The requests to the search engines were carried out through Tor [9], which provides a dynamic, highly distributed network of machines to obfuscate the routing path of a request, therefore changing the IP of the caller. To increase the randomness of the IP addresses, the IP of the machine was refreshed for each query, querying the search engines from machines located around the globe, rather than from a single network location [10].

Another decision concerned how many results were to be included in the study, considering that search engines might return a very high volume of pages for each query. Studies from Web marketing and related fields consistently show that the vast majority of links are obtained on the first page of results returned by the search engine (92 percent of all clicks), and analogously, the top 10 links attract 91 percent of links [11]. Therefore, only the results in the first page were included in the study.

#### 4.6. Reduction of temporal bias

To tackle the temporal variation in the results, instead of relying on results collected at any particular time, the queries were executed every day, at randomized times. The results were collected 34 times, once a day from 7 May 2014 to 9 June 2014 through the Tor network. Overall, 45,900 queries were executed from the U.S. and U.K. versions of Google and Bing, for a total of 111,214 URLs (~3,300 per day). To extract a subset of stable results from this dataset, the temporal change in the results was analyzed by defining a measure of change in the URLs  $CU$  in the URLs between observations:

$$CU_i = \left(1 - \frac{|U_{i-1} \cap U_i|}{|U_{i-1} \cup U_i|}\right) 100$$

Where  $U_i$  is the set of URLs returned for a given query at time  $i$ ,  $CU$  is a percentage that ranges from 0 percent (no change) to 100 percent (total change).  $CU$  was computed across the 34 observations, grouping the results by search engines, by topic, and by the version of search engines. As shown in [Table 2](#), over 34 days, about 25 percent of URLs changed (*cumulative CU*), with a mean daily  $CU$  of



12 percent, suggesting that about 75 percent of URLs remained stable over a month. No statistically significant difference was found between the versions of search engines (U.S. or U.K., t-test  $p=.86$ ). By contrast, Bing and Google presented significant differences in their  $CU$  ( $p<.001$ ). Bing results change every day by  $20.2\pm12.8$  percent, peaking at 44 percent, for a cumulative change equal to 33.3 percent. Google results appear considerably more stable (daily  $CU$   $6.9\pm1.9$  percent), for a cumulative change of 24.3 percent.

<b>Table 2: Variation in search results over time (34 days) in terms of cumulative <math>CU</math>, daily <math>CU</math> (mean, standard deviation, and maximum).</b>					
<b>Parameter</b>	<b>Value</b>	<b>Cumulative <math>CU</math></b>	<b>Mean daily <math>CU</math></b>	<b>Standard deviation daily <math>CU</math></b>	<b>Maximum daily <math>CU</math></b>
<i>Overall</i>	—	24.8	11.9	6.1	23.8
<i>Engine</i>	Bing	33.3*	20.2	12.8	44.0
	Google	24.3*	6.9	1.9	12.4
<i>Version</i>	U.K.	26.6	16.6	9.6	37.3
	U.S.	26.9	16.6	10.4	38.6
<i>Topic</i>	Holocaust denial	39.6	11.8	8.4	33.2
	Depopulation	32.4	11.7	7.9	30.3
	HAARP	29.9	12.2	5.2	24.9
	Free energy suppression	28.3	12.1	7.6	27.0
	Vaccine-autism	27.6	12.3	6.7	24.8
	Secret societies	27.2	16.8	8.1	33.9
	9/11 inside job	26.9	13.5	10.0	36.5
	UFO cover up	21.9	15.4	7.9	31.7
	Chemtrails	21.5	11.4	8.1	31.2
	Reptilian elite	20.7	9.2	5.6	20.9
	Fake moon landing	20.6	8.4	6.7	23.3
	Global warming denial	18.9	11.7	8.4	35.2
	AIDS-HIV	18.2	7.5	6.1	22.4
	Jewish conspiracy	18.1	11.5	9.6	39.0
	JFK	17	10.5	7.2	30.9

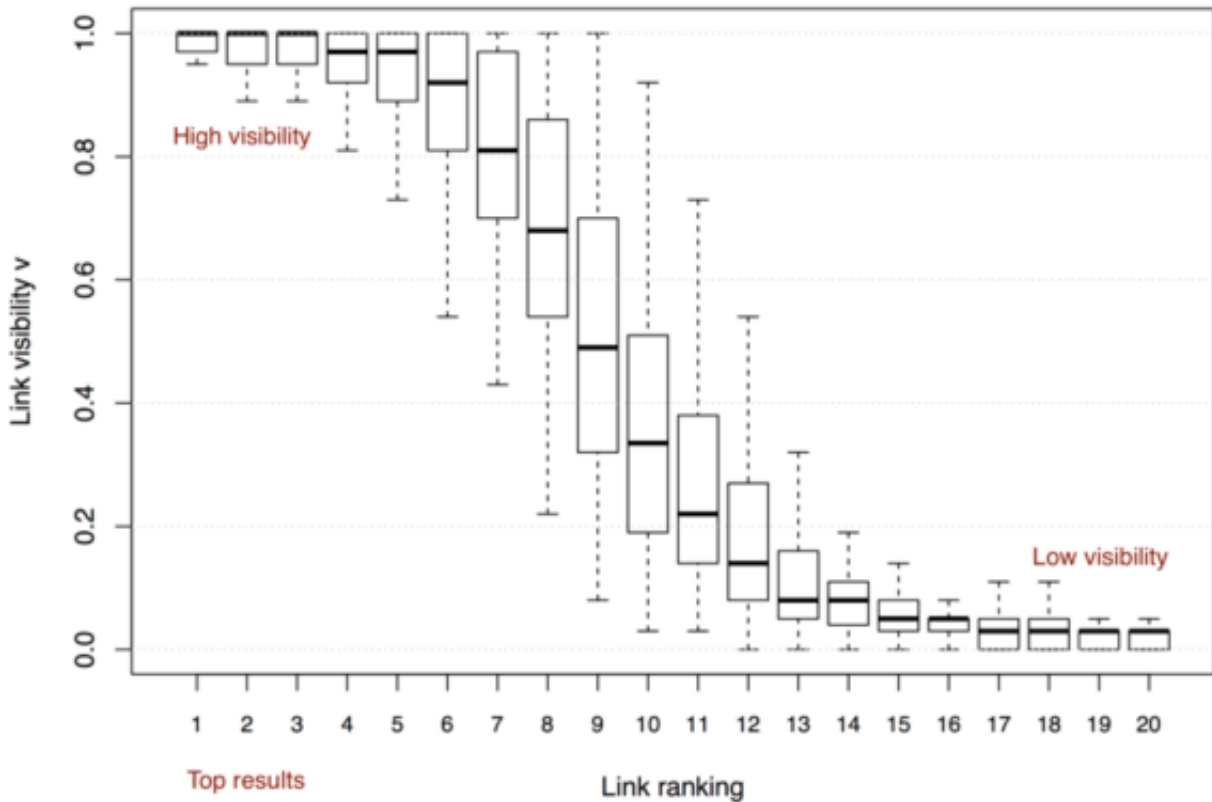
Significant differences were also visible in different topics, whose cumulative  $CU$  ranges between 17 percent and 40 percent. Some conspiracy theories, such as Holocaust denialism, showed an above-average cumulative change (39.9 percent), whilst others remain well below average (e.g., AIDS and JFK-related conspiracies <20 percent). These differences can be interpreted as resulting from the interplay between the engines' internal workings and the Web activity regarding a given topic, highlighting the rate of change in the indexed web content. To identify stable, core content that did not change in the results, a measure of visibility  $V$  of each URLs was computed. The visibility  $V$  of URL  $u$ , where  $RS$  is the set of all results  $U$  over time, is computed as follows:

$$V_u = \frac{|u \in RS|}{|RS|} \quad RS = \{U_1 \cdots U_n\}$$

$V$  ranges from 0 (URL  $u$  is never shown) to 1 ( $u$  is present in all results, every day).  $V$  was subsequently computed to rank all the URLs in each topic, search engine, and time, from the most visible (ranking equal to 1) to the least visible. [Figure 3](#) shows the overall distribution of  $v$  in the entire datasets, aggregating the URLs ranked by  $V$ , highlighting how some URLs are consistently visible, some fluctuate in an average position, and others appear sporadically and disappear:

- *High visibility*: rankings 1 to 7,  $V \in (.7, 1]$
- *Average visibility*: rankings 8 to 11,  $V \in [.2, .7]$
- *Low visibility*: rankings 12 to 20,  $V \in (0, .2)$

For example, for the query “9 11 conspiracy” on the American version of Google (*google.com*), the conspiratorial Web site *www.911truth.org* was present at all times in the results ( $V=1$ ), while a related news story on *www.dailymail.co.uk* was shown only in three percent of the times ( $V=.03$ ). Hence, it is possible to exclude from the analysis the unstable content that obtained low visibility, weighting the URLs with respect to their  $V$ , under the assumption that  $V$  is proportional to the URL’s representativeness of the search engine content. The index can also be used to filter out unstable results with  $V$  below a suitable threshold, for example, discarding the *Daily Mail* article.



**Figure 3:** Distribution of visibility index  $V$  for all URLs in the search results.

The 34 result sets were merged into one, including the top 10 results for each query, resulting in 8,208 URLs across the different queries, engines, and engine versions. To further reduce the weight of unstable results, the minimum threshold for  $V$  was set to .05, excluding the tail of the least visible 2,473 links (30.1 percent of the total number of links), while the sum of  $V$  for these excluded links indicates that they represent the least visible of the dataset (2.3 percent of  $V$ ). The resulting dataset of stable results contained 5,734 URLs.

#### 4.7. Methodological limitations

Despite the advantages of the presented approach, several limitations need to be borne in mind when drawing conclusions from the search results. Search engines like Google Search and Microsoft Bing are products in permanent flux and, as discussed, results do change over time. While the proposed methodology reduces spatio-temporal bias in the results, it cannot remove it altogether. For example, particular media events can create spikes of activity around certain conspiracy theories for short periods of time, altering the type and ideological composition of the results.

More specifically, while the proposed methodology aims at extracting the default representation of a given topic, modern search engines utilize advanced personalization techniques to tailor the search results to specific users based on a wide number of indicators (previous searches, click behavior, social networks, etc.). It is therefore expected that a user searching and exploring conspiracy theories through a search engine will receive increasingly divergent results in “filter bubbles” (O’Hara, 2014) that are difficult to study systematically. Finally, this approach focuses on search engine results themselves, leaving the actual behavior of the users outside the scope. Other, complementary methods from the social sciences must be employed in this enterprise (Wouters and Gerbec, 2003).



## 5. Analysis of search engine representation

Once the search results have been weighted, aggregated, and filtered, having reduced spatial and temporal biases in the data, it is possible to analyze the representation of the selected topics in search engines. In this case study focused on conspiracy theories, the URLs were classified into five classes of content type and five classes of ideological bias (see [Table 3](#)). Based on these categories, the classification was performed manually by inspecting each of 5,734 pages. As expected, because of the complexity and nuances of the subject matter and the heterogeneity of the content returned by search engines, some pages defied the type classification scheme.

<b>Table 3: Classes of content type and ideological bias.</b>	
<i>Content type</i>	
<i>Academic (a)</i>	Content produced by educational institutions such as universities, and official government Web sites. This content often belongs to domains .edu, .gov, .ac.uk, and includes reputable scientific publishers, libraries, and official Web sites of public institutions. This content tends to be of high quality, and is often peer-reviewed. Examples: www.cdc.gov, thelancet.com, eoearth.org, epa.gov, stanford.edu.
<i>News (n)</i>	Content produced by newspapers, magazines, and news agencies that have editorial control, such as the <i>New York Times</i> , <i>Guardian</i> , and <i>Wall Street Journal</i> . Content farms and automated news aggregators are excluded, but it includes tabloids. Examples: telegraph.co.uk, theatlantic.com, thestar.com, timesofisrael.com
<i>Blogs (b)</i>	Content generated on Web 2.0 blogging and social networking platforms, without editorial control. Examples: ufosightingsdaily.com, www.reptilians-exposed.com, elderofziyon.blogspot.com
	Pages of Wikipedia or related projects.

<i>Wikipedia (w)</i>	Because of its popularity, its collective authorship, and its consistently high ranking on search engines, Wikipedia deserves a dedicated category. Examples: en.wikipedia.org, en.wikiquote.org
<i>Misc. (m)</i>	All content that does not clearly fall into the preceding categories, such as online stores.
<b><i>Ideological bias</i></b>	
<i>Conspiratorial (c)</i>	Content that openly supports one or more conspiracy theories. This includes critiques of a conspiracy theory that suggest comparable theories. Examples: 911truth.eu, agenda21conspiracy.com, rense.com
<i>Neutral (n)</i>	Content that describes a conspiracy theory without expressing clearly positive or negative value judgments. This includes reference articles and news stories.
<i>Debunking (d)</i>	Content that openly attacks the conspiracy theory as unsound, unsubstantiated, implausible, illogical, or ridiculous. Examples: snopes.com, rationalwiki.org
<i>Related (r)</i>	Content that is thematically related to a conspiracy theory, but does not mention it explicitly.
<i>Unrelated (r)</i>	Content that is not related to a conspiracy theory. This content is noise in the search results.

The editorial control of online newspapers varies widely, and the boundary between news and blogs can be unclear, such as in the case of the *Huffington Post*. Similarly, a writer's attitude toward a conspiracy theory can be unintelligible or sarcastic, such as the parodies by the Mad Revisionist [12]. To reduce the classification bias of a single coder, a collaborator performed the classification separately on a random subset of 50 URLs. The second, independent classification of content type showed agreement on 94 percent on the cases, while the agreement on the ideological bias was lower (88 percent). The disagreements were mostly found between indirect or direct discourses in the text.

<b>Table 4: Content type and bias (%).</b> Note: The total sum of the cells is 100%. <i>N</i> Web pages=5,734. Top four cells are in bold, for a total of about 74 percent of results. The results are not weighted by visibility <i>V</i> .						
<b>Bias/Type (%)</b>	<b>Academic</b>	<b>Blogs</b>	<b>News</b>	<b>Wikipedia</b>	<b>Misc.</b>	
<i>Conspiratorial</i>	.2	<b>47.5</b>	3.0	.1	1.0	51.8
<i>Debunking</i>	.2	<b>6.0</b>	3.4	4.2	.9	26.8
<i>Neutral</i>	3.1	<b>12.8</b>	<b>7.5</b>	3.0	.4	14.7
<i>Related</i>	1.0	.9	.8	.5	.5	3.7
<i>Unrelated</i>	0	1.5	.3	.1	1.1	3.0
Type total	4.5	68.7	15.0	7.9	3.9	—

The outcome of this classification is summarized in Table 4. Out of 5,734 Web pages, only 4.5 percent were classified as academic, while the majority of the results belonged to the blog category (68.7 percent). Overall, 51.8 percent of the results were conspiratorial, while only 26.8 percent were neutral and 14.7 percent debunking. By observing the relationship between content type and ideological bias, it appears that the bulk of the content consists of conspiratorial blogs (47.5 percent), and debunking blogs (12.8 percent). Newspapers perform some debunking (7.5 percent), while Wikipedia pages provide a substantial part of results (7.9 percent), for the most part neutral or debunking. Two

observations emerge from the analysis: (i) academic results have extremely low visibility; and, (ii) conspiratorial material is much more visible than debunking or neutral material.

### 5.1. Conspiracy and polarization indexes

To analyze the representation of each conspiracy theory, two domain-specific indexes were defined and computed for all cases. This approach is built on the tradition of the study of media bias, providing tools to study political representations (Entman, 2007), adapted to the context of search engine results. The conspiracy index ( $CI$ ) summarizes the ideological bias of a set of results  $U$ , as the difference between the conspiratorial and the debunking results:

$$CI_U = \frac{|U_{consp}| - |U_{debunk}|}{|U| - |U_{rel}| - |U_{unrel}|} \quad CI \in [-1,1]$$

This index ranges from -1 (all results are openly against the conspiracy theory) to 1 (all results support the conspiracy theory). 0 indicates either a balance between conspiratorial and debunking results, or the dominance of neutral results. A complementary index to  $CI$  is the polarization index ( $PI$ ) of set of results  $U$ , defined as the ratio between neutral results and relevant results:

$$PI_U = \frac{|U_{neutral}|}{|U| - |U_{rel}| - |U_{unrel}|} \quad PI \in [0,1]$$

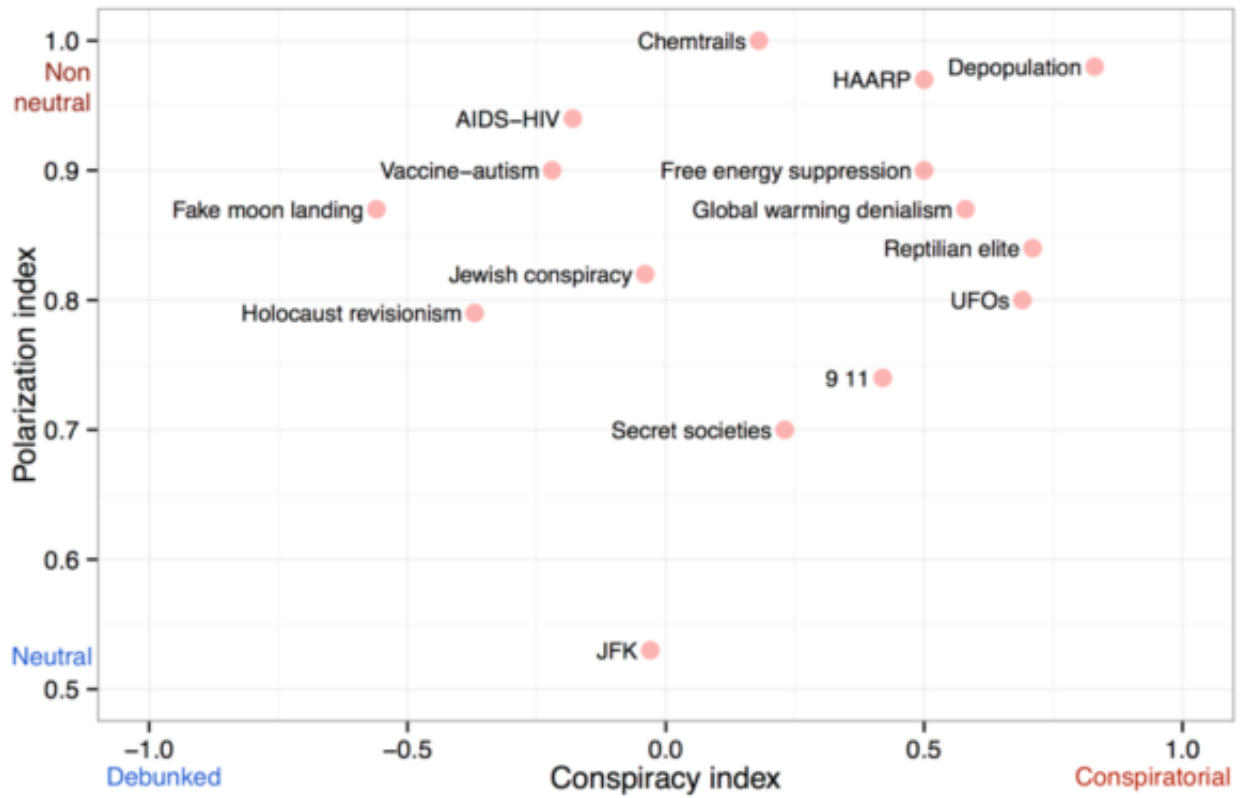
This index quantifies the ratio between neutral and non-neutral results, ranging between 1 (all results are openly pro or against the conspiracy theory, *i.e.*, totally polarized) and 0 (all results are neutral, *i.e.*, non-polarized). [Table 5](#) shows the results of the classification, including the two indexes, grouped by search engine and conspiracy theory. For each conspiracy theory, it is possible to see the representation returned by the search engines in terms of ideological bias and content type. These results show the high variability between the search engines. Overall, Bing ( $CI=.3$ ) provides more conspiratorial and less debunking material than Google ( $CI=.16$ ). By contrast, the weight of neutral results is comparable ( $PI\approx.84$ ). While Bing and Google return a roughly comparable amount of academic results ( $\sim 4$ -5 percent), blogs ( $\sim 65$ -68 percent), and Wikipedia ( $\sim 10$ -13 percent), Google gives considerably more visibility to news content (19.1 percent) than Bing (10.2 percent). As the ideological bias of most news content is debunking or neutral (see [Table 4](#)), this variation accounts for the difference between the two engines.

Value	Indexes		N results	Ideological bias (%)					Content type (%)				
	CI	PI		Conspir.	Neutral	Debunk.	Related	Unrel.	Acad.	Wiki.	Blog	News	Misc.
All	.22	.84	5,734	50.7	14.9	29.7	2.9	1.8	4.7	11.1	66.6	15.2	2.4
Bing	.30	.83	3,089	53.7	15.9	25.5	3.0	1.9	5.4	12.9	68.1	10.2	3.4
Google	.16	.85	2,645	48.4	14.1	33.1	2.8	1.7	4.1	9.7	65.5	19.1	1.6
Depopulation	.83	.98	389	90.1	1.9	7.4	-	.7	-	2.9	96.5	.6	-
Reptilian elite	.71	.84	398	76.6	16.0	6.3	.1	.9	-	14.6	80.5	3.8	1.0
UFOs	.69	.80	399	70.8	18.6	5.2	4.5	.9	1.9	9.8	70.4	16.3	1.7
Global warming denial.	.58	.87	397	71.8	13.2	14.4	0.6	-	2.1	6.7	65.8	25.5	0.1
Free energy suppr.	.50	.90	378	65.8	9.2	19.0	1.2	4.8	.1	10.7	82.3	1.3	5.7
HAARP	.50	.97	359	67.1	2.3	21.7	1.0	7.8	.5	10.9	76.5	9.8	2.2
9/11 inside job	.42	.74	387	56.7	25.9	15.5	1.7	.3	.1	12.3	69.5	16.8	1.3
Secret societies	.23	.70	385	40.1	25.8	20.1	10.5	3.5	1.0	14.4	63.8	12.4	8.4
Chemtrails	.18	.99	371	59.1	.3	40.6	-	-	.1	10.8	82.1	7.0	-
JFK	-.03	.53	361	24.1	46.3	27.1	2.4	.1	11.2	13.1	30.1	43.2	2.4
Jewish conspiracy	-.04	.82	360	37.9	17.6	42.3	.9	1.3	2.6	14.5	68.3	9.5	5.2
AIDS-HIV	-.18	.94	381	37.4	5.6	55.4	1.6	-	5.6	12.9	55.1	25.5	.9
Vaccine-autism	-.22	.90	377	29.1	8.4	47.6	14.8	-	19.8	8.4	56.4	15.4	-
Holocaust revision.	-.37	.79	409	19.1	19.5	53.1	3.7	4.6	16.8	12.3	48.5	16.4	6.0
Fake moon landing	-.56	.87	383	14.7	13.1	69.9	.8	1.6	8.5	12.7	53.4	24.4	1.1

**Table 5:** Overview of results grouped by search engines (Bing and Google) and topics (rest of the table). *CI*= conspiracy index; *PI*=polarization index. The results are order by *CI*, and weighted by visibility *V*.

These indices enable the quantitative mapping of topic representations along specific dimensions. When observing the representations of the 15 conspiracy theories, the *CI* has a striking variability, ranging from .83 for the depopulation conspiracy to -.56 for the fake moon landing. The values of the indexes of the conspiracy theories are also mapped in [Figure 4](#). The scatterplot shows that the engines portray some conspiracies with many supporting results ( $CI > .5$ ), providing limited debunking or neutral viewpoints. Extreme conspiracy theories involving alien species and global elites tend to fall in this area (depopulation, HAARP, reptilian elite, and UFOs), suggesting low debunking efforts. The chemtrail conspiracy appears less biased ( $CI = .18$ ), and is more polarized, having no neutral results ( $PI = 1$ ). Secret societies, 9/11 as inside job fall in the same region ( $CI \in [.2, .5]$ ). More balanced representations are found for the JFK and Jewish conspiracies ( $CI \approx .03$ ).

Other conspiracies, in the left-hand side of the figure with negative *CI*, are portrayed primarily through debunking results. Medical conspiracies (AIDS- and vaccine-related) have  $CI \approx -.2$ . The fake moon landing and Holocaust denialism are the most debunked ( $CI < -.37$ ), indicating a systematic effort by highly visible online data sources. The variability of the polarization index (*PI*) is lower than that of *CI*, falling in range .5 and 1. Most conspiracy theories exhibit very strong polarization ( $PI > .7$ ), indicating that these topics tend to be represented either in a positive or negative light, with relatively few results portraying them in a balanced way. HAARP, depopulation and chemtrail conspiracies, all thematically related, obtain totally polarized results ( $PI > .95$ ). A notable exception to this general trend is the JFK-related conspiracies, which obtains a relatively low  $PI = .53$ , because of an extraordinarily high proportion of neutral results (46.3 percent).



**Figure 4:** Conspiracy index (CI) w.r.t. polarization index (PI). As no results had  $PI < .5$ , the axis limits are  $[.5, 1]$ . Each dot represents a conspiracy theory.

### 5.2. Dominant Web sites

After having analyzed the search results grouped by search engine and conspiracy theory, the stable results extracted through the presented methodology can be used to identify the Web sites that dominate the results. In many cases, material originating from the same Web sites was returned for different conspiracy theories. Hence, it is useful to observe the most visible Web sites across the dataset, rather than in any specific query or topic. The domain was extracted from each of the 5,734 URLs, and a cumulative visibility measure  $V$ , and the number of topics covered by the Web site was computed. For example, the aggregation of all URLs from the conspiratorial site *beforeitsnews.com* obtained  $V=37.6$ , covering eight of the 15 topics, for a  $CI=1$  and  $PI=1$ .

This analysis resulted in 972 unique Web sites, in which the most visible sites are *en.wikipedia.org* ( $V=350.5$ ), *youtube.com* (189.4), *rationalwiki.org* (70), *rense.com* (51.2), and *time.com* (47.8). The distribution of the cumulative visibility  $V$  reflects closely the scale-free nature of online content: the top 20 percent of domains are the source of 79 percent of the content, suggesting a Pareto distribution. [Table 6](#) shows the top 15 sites grouped by ideological bias (conspiratorial, neutral, and debunking). These sites are therefore the most likely to be clicked, and provide the core content returned by search engines.



Conspiratorial ( $CI = 1$ )			Neutral (variable $CI$ )				Debunking ( $CI = -1$ )		
Website	V	T	Website	V	T	CI	Website	V	T
rense.com	51.2	11	en.wikipedia.org	350.5	15	-.39	rationalwiki.org	70.0	11
beforeitsnews.com	37.6	8	answers.yahoo.com	35.7	14	0	content.time.com	47.8	5
newsone.com	26.6	2	listverse.com	34.9	4	0	howstuffworks.com	41.1	5
wanttoknow.info	24.2	5	telegraph.co.uk	27.3	8	-.26	ushmm.org	33.9	2
davidicke.com	23.9	3	amazon.com	24.8	13	.29	jewishvirtuallibrary.org	28.9	2
bibliotecapleyades.net	23.7	6	huffingtonpost.com	20.8	6	-.21	nationalgeographic.com	23.3	3
bariumblues.com	20.3	3	abcnews.go.com	18.6	4	0	cnn.com	18.7	3
arcturi.com	19.3	1	cdc.gov	18.2	1	0	archive.adl.org	16.8	2
collective-evolution.com	19.1	5	dailymail.co.uk	18.2	9	.04	slate.com	15.5	5
911truth.org	18.6	1	usatoday.com	15.4	2	-.07	mcadams.posc.mu.edu	15.0	1
globalresearch.ca	18.6	5	uk.answers.yahoo.com	15.0	6	0	thetalkingdrum.com	13.9	1
biblebelievers.org.au	17.9	4	webmd.com	13.5	1	-.44	csicop.org	13.6	5
globalclimatescam.com	17.4	1	theweek.co.uk	10.2	2	-.04	theonion.com	10.4	2
trutv.com	16.3	8	hdot.org	9.3	1	0	snopes.com	10.2	2
cracked.com	16.0	2	mirror.co.uk	9.1	2	-.19	contrailscience.com	9.8	1

**Table 6:** Top 15 conspiratorial, neutral, and debunking Web sites in the whole dataset, with cumulative visibility index  $V$  and number of topics  $T$  (out of 15).

## 6. Findings from the case study

The methodology described in this article provides an empirical tool to investigate new hypotheses in the area of search engine media research. In the selected case study, 15 popular conspiracy theories were investigated. The findings of this study are summarized as follows.

*Scarcity of academic sources.* A striking finding is the lack of academic sources in the search results. Only 4.5 percent of results come from a reputable source, while almost 70 percent of the content originates from Web 2.0 platforms. Substantially more academic sources are returned in the case of vaccine-related conspiracies (19.8 percent), and of Holocaust revisionism (16.8 percent). Debunking is performed mainly by blogs (48 percent of debunking material) and news (28 percent).

*Ideological diversity and relativism.* Search engines are often accused of reinforcing mainstream positions (e.g., Vaidhyanathan, 2011) and of suppressing controversies (Gerhart, 2004). The results of this study show that, in the case of conspiracy theories, search engines return a combination of pro- and anti-conspiracy material. Nine conspiracy theories out of 15 obtained primarily conspiratorial material ( $CI > .15$ ), while the fake moon landing, Holocaust denialism, and medical conspiracies (AIDS- and vaccine-related) appear predominantly debunked in the results ( $CI < -.15$ ). Only two conspiracies (JFK-related and Jewish conspiracies) present balanced results ( $CI \approx 0$ ). Overall, the search engines returned material that was biased in favor of conspiracy theories ( $CI = .22$ ) and highly polarized ( $PI = .84$ ). Based on these results, it seems that fringe and non-mainstream views are heavily represented by search engines, displaying high variability in the type of content and ideological leaning for different topics. These results indicate that the search engines offer ideologically diverse viewpoints, to the point of epitomizing post-modern relativism, “flattening truth” in a homogenous collection of links (Kata, 2012). No evidence was found of suppression of controversies. Given the high visibility of conspiratorial blogs (47.5 percent of all content), it seems

hard to argue that search engines reinforce and perpetuate mainstream views. Even in highly sensitive contexts (Holocaust and vaccine related conspiracies), blogs on free Web 2.0 platforms are more visible than well-funded, respected academic sources.


*Search engines as mirror of society.* In the interplay between search engines and cultural processes, it is plausible to interpret search results as a biased, imperfect, and yet useful mirror of society, highlighting the differences between conspiracy theories. The conspiracy and polarization index can inform hypotheses about underlying cultural and political mediated conversations. The results indicate that UFO-related and depopulation conspiracies do not elicit extensive debunking, and are expressed primarily on self-referential blogs and forums. By contrast, global warming denialism is heavily present in mainstream news media such as Fox News and *Forbes* (25.5 percent), resulting in very high *CI* (.58). Conspiracy theories triggering some degree of public outrage, such as the fake moon landing, Holocaust denialism, and anti-vaccine narratives have resulted in wide debunking efforts by scientific and medical communities, reflected in the search engine results.



## 7. Conclusion

This paper presented a methodology to study the representation of topics in popular search engines, extracting stable, highly visible results from a large number of volatile search results. Treating search engine results as editorial products, the proposed technique reduces spatio-temporal bias in the results. The methodology is applied to a set of conspiracy theories, observing their representations in Google and Bing in terms of content type and ideological bias. Given the increasing influence of search engines and their gatekeeping role in the circulation of information, conspiracy theories constitute an ideal arena to study the impact of search technologies on opinion formation and on broader political, cultural, and social matters.

Two indexes — conspiracy (*CI*) and polarization index (*PI*) — were used to extract patterns from the data, showing the general trends as well as detailed aspects of the representations. Based on a dataset of 5,734 URLs returned by Bing and Google, the analysis revealed underlying features of these representations, illustrating the new possibilities enabled by the proposed methodology. For instance, the results indicate that, far from suppressing minority views, search engines such as Google and Bing give easy access to diverse viewpoints, including fringe beliefs.

The study of search engines as editorial products can provide insights about a wide variety of social, political, and cultural phenomena, such as filter bubbles and extremism (O'Callaghan, *et al.*, 2013), and possible solutions, such as search engine regulations (Grimmelmann, 2013), and technological approaches to measure the credibility of online material (Lewandowski, 2012). Further research on search engines' cultural and social impacts can benefit from the availability of heterogeneous content made visible by such unobtrusive and yet hegemonic search technologies. 

## About the author

Andrea Ballatore is a postdoctoral researcher and the research coordinator at the Center for Spatial Studies, University of California, Santa Barbara. In 2013, he received a Ph.D. in geographic information science from University College, Dublin. He has worked as a lecturer in the Department of Computer Science at the National University of Ireland, Maynooth, and as a software engineer in Italy and Ireland. His interdisciplinary research focuses on the digital representations of place, crowdsourcing, and the technological imaginary at the intersection between computer science, geography, and media studies.  
E-mail: andrea [at] ballatore [at] gmail [dot] com

## Notes

[1.](#) Grimmelmann, 2010, p. 436.

[2.](#) Grimmelmann, 2013, p. 29.

3. Kata, 2012, p. 3,780.

4. <http://www.comscore.com/Insights/Press-Releases/2014/4/comScore-Releases-March-2014-U.S.-Search-Engine-Rankings>, accessed 2 November 2014.

5. [http://theeword.co.uk/info/search\\_engine\\_market.html](http://theeword.co.uk/info/search_engine_market.html), accessed 2 November 2014.

6. <http://www.google.com>, accessed 2 November 2014.

7. <http://www.bing.com>, accessed 2 November 2014.

8. Google Trends analyzes a percentage of Google Web searches to determine how many searches have been done for the terms you've entered compared to the total number of Google searches done during that time. See <http://www.google.com/trends>, accessed 2 November 2014.

9. <http://www.torproject.org>, accessed 2 November 2014.

10. It is important to note that the distribution of machines in the Tor network is not genuinely random, but it nonetheless greatly reduces the spatial bias in the results, providing a wide range of IP addresses instead of one.

11. See, for example, <http://chitika.com/google-positioning-value>, accessed 2 November 2014.

12. See the parodies of conspiracy theories at <http://www.revisionism.nl>, accessed 2 November 2014.

## References

D. Brossard and D.A. Scheufele, 2013. "Science, new media, and the public," *Science*, volume 339, number 6115 (4 January), pp. 40–41.

doi: <http://dx.doi.org/10.1126/science.1232329>, accessed 18 June 2015.

S. Clarke, 2002. "Conspiracy theories and conspiracy theorizing," *Philosophy of the Social Sciences*, volume 32, number 2, pp. 131–150.

doi: <http://dx.doi.org/10.1177/004931032002001>, accessed 18 June 2015.

R. M. Entman, 2007. "Framing bias: Media in the distribution of power," *Journal of Communication*, volume 57, number 1, pp. 163–173.

doi: <http://dx.doi.org/10.1111/j.1460-2466.2006.00336.x>, accessed 18 June 2015.

G. Eysenbach and C. Köhler, 2002. "How do consumers search for and appraise health information on the World Wide Web? Qualitative study using focus groups, usability tests, and in-depth interviews," *British Medical Journal*, volume 324, number 7337, pp. 573–577.

doi: <http://dx.doi.org/10.1136/bmj.324.7337.573>, accessed 18 June 2015.

B.J. Fogg, C. Soohoo, D.R. Danielson, L. Marable, J. Stanford, and E.R. Tauber, 2003. "How do users evaluate the credibility of Web sites? A study with over 2,500 participants," *DUX '03: Proceedings of the 2003 Conference on Designing for User Experiences*, pp. 1–15.

doi: <http://dx.doi.org/10.1145/997078.997097>, accessed 18 June 2015.

S.L. Gerhart, 2004. "Do Web search engines suppress controversy?" *First Monday*, volume 9, number 1, at <http://firstmonday.org/article/view/1111/1031>, accessed 18 June 2015.

E. Goldman, 2008. "Search engine bias and the demise of search engine utopianism," In: A. Spink and M. Zimmer (editors). *Web search: Multidisciplinary perspectives. Information Science and Knowledge Management*, volume 14. Berlin: Springer, pp. 121–133.

doi: [http://dx.doi.org/10.1007/978-3-540-75829-7\\_8](http://dx.doi.org/10.1007/978-3-540-75829-7_8), accessed 18 June 2015.

M. Graham, R. Schroeder, and G. Taylor, 2014. "Re: Search," *New Media & Society*, volume 16, number 2, pp. 187–194.

doi: <http://dx.doi.org/10.1177/1461444814523872>, accessed 18 June 2015.

J. Grimmelmann, 2013. "What to do about Google?" *Communications of the ACM*, volume 56, number 9, pp. 28–30.

J. Grimmelmann, 2010. "Some skepticism about search neutrality," In: B. Szoka and A. Marcus (editors). *The next digital decade: Essays on the future of the Internet*. Washington, D.C.: TechFreedom, pp. 435–459.

A. Halavais, 2009. *Search engine society*. Cambridge: Polity.

E. Hargittai, 2007. "The social, political, economic, and cultural dimensions of search engines: An

introduction," *Journal of Computer-Mediated Communication*, volume 12, number 3, pp. 769–777.  
doi: <http://dx.doi.org/10.1111/j.1083-6101.2007.00349.x>, accessed 18 June 2015.

K. Hillis, M. Petit, and K. Jarrett, 2013. *Google and the culture of search*. New York: Routledge.

R. Hofstadter, 1964. "The paranoid style in American politics," *Harper's* (November), pp. 77–86, and at <http://harpers.org/archive/1964/11/the-paranoid-style-in-american-politics/>, accessed 18 June 2015.

L.D. Introna and H. Nissenbaum, 2000. "Shaping the Web: Why the politics of search engines matters," *Information Society*, volume 16, number 3, pp. 169–185.  
doi: <http://dx.doi.org/10.1080/01972240050133634>, accessed 18 June 2015.

A. Kata, 2012. "Anti-vaccine activists, Web 2.0, and the postmodern paradigm — An overview of tactics and tropes used online by the anti-vaccination movement," *Vaccine*, volume 30, number 25, pp. 3,778–3,789.  
doi: <http://dx.doi.org/10.1016/j.vaccine.2011.11.112>, accessed 18 June 2015.

M.T. Keane, M. O'Brien, and B. Smyth, 2008. "Are people biased in their use of search engines?" *Communications of the ACM*, volume 51, number 2, pp. 49–52.  
doi: <http://dx.doi.org/10.1145/1314215.1314224>, accessed 18 June 2015.

P. Knight, 2000. *Conspiracy culture: From Kennedy to the X-Files*. New York: Routledge.

R. König and M. Rasch (editors), 2014. *Society of the query reader: Reflections on Web search*. Amsterdam: Institute for Network Cultures.

D. Lewandowski, 2012. "Credibility in Web search engines," In: M. Folk and S. Apostel (editors). *Online credibility and digital ethos: Evaluating computer-mediated communication*. Hershey, Pa.: IGI Global, pp. 131–146.  
doi: <http://dx.doi.org/10.4018/978-1-4666-2663-8.ch008>, accessed 18 June 2015.

D. Mocanu, L. Rossi, Q. Zhang, M. Karsai, and W. Quattrociocchi, 2014. "Collective attention in the age of (mis)information," *arXiv:1403.3344*, at <http://arxiv.org/abs/1403.3344>, accessed 18 June 2015.

D. O'Callaghan, D. Greene, M. Conway, J. Carthy, and P. Cunningham, 2013. "The extreme right filter bubble," *arXiv:1308.6149*, at <http://arxiv.org/abs/1308.6149>, accessed 18 June 2015.

K. O'Hara, 2014. "In worship of an echo," *IEEE Internet Computing*, volume 18, number 4, pp. 79–83.  
doi: <http://dx.doi.org/10.1109/MIC.2014.71>, accessed 18 June 2015.

B. Pan, H. Hembrooke, T. Joachims, L. Lorigo, G. Gay, and L. Granka, 2007. "In Google we trust: Users' decisions on rank, position, and relevance," *Journal of Computer-Mediated Communication*, volume 12, number 3, pp. 801–823.  
doi: <http://dx.doi.org/10.1111/j.1083-6101.2007.00351.x>, accessed 18 June 2015.

P. Reilly, 2008. "'Googling' terrorists: Are Northern Irish terrorists visible on Internet search engines?" In: A. Spink and M. Zimmer (editors). *Web search: Multidisciplinary perspectives. Information Science and Knowledge Management*, volume 14. Berlin: Springer, pp. 151–175.  
doi: [http://dx.doi.org/10.1007/978-3-540-75829-7\\_10](http://dx.doi.org/10.1007/978-3-540-75829-7_10), accessed 18 June 2015.

A. Spink and M. Zimmer (editors), 2008. *Web search: Multidisciplinary perspectives. Information Science and Knowledge Management*, volume 14. Berlin: Springer.

C. Stempel, T. Hargrove, and G.H. Stempel, 2007. "Media use, social structure, and belief in 9/11 conspiracy theories," *Journalism & Mass Communication Quarterly*, volume 84, number 2, pp. 353–372.  
doi: <http://dx.doi.org/10.1177/107769900708400210>, accessed 18 June 2015.

C.R. Sunstein and A. Vermeule, 2009. "Conspiracy theories: Causes and cures," *Journal of Political Philosophy*, volume 17, number 2, pp. 202–227.  
doi: <http://dx.doi.org/10.1111/j.1467-9760.2008.00325.x>, accessed 18 June 2015.

S. Vaidhyanathan, 2011. *The Googlization of everything (and why we should worry)*. Berkeley: University of California Press.

L. Vaughan and M. Thelwall, 2004. "Search engine coverage bias: evidence and possible causes," *Information Processing & Management*, volume 40, number 4, pp. 693–707.  
doi: [http://dx.doi.org/10.1016/S0306-4573\(03\)00063-3](http://dx.doi.org/10.1016/S0306-4573(03)00063-3), accessed 18 June 2015.

M. Wood, 2013. "Has the Internet been good for conspiracy theorising?" *Psychology Postgraduate Affairs Group (PsyPAG) Quarterly*, number 88, pp. 31–33, and at <http://www.psypag.co.uk/wp->

[content/uploads/2013/09/Issue-88.pdf](#), accessed 18 June 2015.

M.J. Wood, K.M. Douglas, and R.M. Sutton, 2012. "Dead and alive: Beliefs in contradictory conspiracy theories," *Social Psychological & Personality Science*, volume 3, number 6, pp. 767–773.  
doi: <http://dx.doi.org/10.1177/1948550611434786>, accessed 18 June 2015.

P. Wouters and D. Gerbec, 2003. "Interactive Internet? Studying mediated interaction with publicly available search engines," *Journal of Computer-Mediated Communication*, volume 8, number 4, at <http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2003.tb00221.x/full>, accessed 18 June 2015.  
doi: <http://dx.doi.org/10.1111/j.1083-6101.2003.tb00221.x>, accessed 18 June 2015.

---

## Editorial history

Received 20 November 2014; accepted 16 June 2015.

---



This paper is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](#).

Google chemtrails: A methodology to analyze topic representation in search engine results  
by Andrea Ballatore.

*First Monday*, Volume 20, Number 7 - 6 July 2015

<http://firstmonday.org/ojs/index.php/fm/rt/prINTERfriendly/5597/4652>

doi: <http://dx.doi.org/10.5210/fm.v20i7>.