Lawrence M. Hinman:

Esse est indicato in Google: Ethical and Political Issues in Search Engines

Abstract:

Search engines play an increasingly pivotal role in the distribution and eventual construction of knowledge, yet they are largely unnoticed, their procedures are opaque, and they are almost completely devoid of independent oversight. In this paper the author examines three areas in which we encounter difficult and persistent ethical issues in search engine technology: The problem of algorithm and the lack of transparency of the search process, the problem of privacy with regards of the possibility to monitor search histories, and the problem of local censorship. The given findings lead to the conclusion that the development of structures of accountability for search engines is an important task for the near future.

Agenda

Author:

Lawrence M. Hinman, Ph.D.:

- Organization and contact address: Department of Philosophy University of San Diego 5998 Alcalá Park San Diego, CA 92110-2492 USA
- Telephone, email and personal homepage: ☎ 619-260-4787 , ✉ hinman@sandiego.edu

www.i-r-i-e.net   **19**
  ISSN 1614-1687

Introduction

In the final months of 2004, rumors began to circulate on the Internet that the infamous prison abuse photographs from Abu Ghraib were no longer available on a Google image search, although they continued to show up on other search engines.[1] The implication was that political considerations might have been influencing the search engine results, and implication that Google denies.[2] When I emailed Google directly about this issue, Nate Tyler, a spokesman for Google, wrote: "Basically, Google did show these images but only for a limited period of time, as our index (collection of web images) cycles through every so often to update itself. New images replace the old. At no point did we filter these images." This explanation seems implausible, given the large number of old photos that seem to stay in the Google database and the high level of importance (and back-links) of these particular photos.

This was not the first instance of ethical issues being raised about search engines. In the early years of search engines, the line had not always been clearly drawn between "sponsored sites" (i.e., sites that pay the search company to put their sites on the top of the list) and regular, non-paying sites. This has in large measure been worked out, and search results typically label those sites that have paid to be listed. This strikes a nice balance between the demands of honesty and those of business. Search engines are understandably heavily dependent on advertising revenues, so it was important to provide a solution that permitted that to continue; at the same time, it was important that users find themselves directed toward the most relevant sites.

Subtle variations upon this theme, however, are now pervasive. Search engine companies sell certain keywords to advertisers in such a way that, when searches enter that term, certain advertising results are displayed in the results page. The advertiser then pays the search engine company a fixed amount per click. This has given rise to "click fraud," generated by the lure of an estimated 3.8 billion dollars annually in advertising revenues.[3] Competitors may repeatedly click on the ads, thereby driving up the advertising costs paid by their competitors. The average price-per-click for popular keywords is $1.70, and can range in rare cases as high as $50 per click. It's easy to see how an unscrupulous competitor could drive the advertising budget of another company into the ground.

Other issues have proved more troublesome. In a typical Google search on the word "Jew," several of the first ten sites that come up are virulently anti-Semitic, including "Jew Watch" and "The International Jew: The World's Foremost Problem." Comparable searches on "Christian" or "Muslim" or "Hindu" do not yield critical sites among the top-ranked entries. In a note from Google on "Offensive Search Results,"[4] The Google Team points out that anti-Semitic sites do not typically appear in a search for "Jewish people," "Jews," or "Judaism," only in a search for the singular word "Jew."

---

[1] When I did a search on "Abu Ghraib" in December 2004 on Alta Vista (http://www.altavista.com/image/results?q=abu+ghraib&mik=photo&mik=graphic&mip=all&mis=all&miwxh=all), I came across a number of the infamous photos on the first page of results; the research listed a total number of 2,579 results. However, when I did a comparable search on Google (with SafeSearch turned off) (http://images.google.com/images?q=abu+ghraib&hl=en&lr=&safe=off&start=0&sa=N), I got 137 results, but almost none of them were the prison abuse photos that from Abu Ghraib that so electrified the world. The same search, repeated in February 2005, yielded far more images in Google, although still some of the original infamous photos seemed not to be present.

[2] Email from Mr. Tyler to me on 1/4/05.

[3] Michael Liedtke, "Click Fraud Looms As Search-Enging Threat," Associated Press, Feb. 11, 2005; http://www.miami.com/mld/miamiherald/business/national/10876986.htm?1c. Also see Jessie C. Stricchiola, "Click Fraud—An Overview." Alchemist Media, Inc http://www.alchemistmedia.com/CPC_Click_Fraud.htm .

[4] http://www.google.com/explanation.html . They write, in part, that "If you use Google to search for "Judaism," "Jewish" or "Jewish people," the results are informative and relevant. So why is a search for "Jew" different? One reason is that the word "Jew" is often used in an anti-Semitic context. Jewish organizations are more likely to use the word "Jewish" when talking about members of their faith."

In an international counterpart to the United States emphasis on local standards for judging pornography, international search engines encounter the problem that such anti-Semitic websites are illegal in some countries. Responding to the legal requirements of their home countries, Google.de and Google.fr do not list those anti-Semitic sites. A search for "Juden" (the plural—the singular in German, "Jude," returns many entries on Jude Law) on Google.de yields over 2M entries, but the first page contains no critical entries; nor does a search on "Juif" on Google.fr yield anti-Semitic sites.

Google's official policy on this issue is clearly stated in the note on offensive entries:

> Our search results are generated completely objectively and are independent of the beliefs and preferences of those who work at Google. Some people concerned about this issue have created online petitions to encourage us to remove particular links or otherwise adjust search results. Because of our objective and automated ranking system, Google cannot be influenced by these petitions. The only sites we omit are those we are legally compelled to remove or those maliciously attempting to manipulate our results.[5]

Several of the first page sites that appear in a search on the "Klu Klux Klan" are highly critical of the Klan; no note appears in that search about offensive results.

These cases raise interesting and extremely important ethical issues about access to information on the Web and the role of search engines. Let me begin by commenting on the public function and responsibility of search engines.

The Public Function and Responsibility of Search Engines

Search engines occupy a privileged place in the world of information technology. They are like windows onto the web—and, like windows, tend to be largely unnoticed because our gaze focuses on what is visible through them. With windows, however, it is easy to detect when they are cloudy or distorted. With search engines, however, it is much more difficult to tell when they are providing dis-

torted or incomplete pictures. Several points should be noted here.

First, the vast amount of information available on the Web would be almost useless without search engines. They play an absolutely crucial role in the access to information.[6] In the world of the Web, *esse est indicato in Google*: to exist is to be indexed on Google. The challenge in information retrieval is not simply to find the right piece of information, but also to avoid listing all the pieces of extraneous information. (The success of Google was precisely in its ability to help users find exactly the information they were seeking and to avoid irrelevant sites.) Search engines are the gatekeepers of the web,[7] helping people to reach their desired destinations. Without them, much of the web would simply be inaccessible to us.

Second, access to information is crucial for responsible citizenship.[8] Citizens in a democracy, and

---

[5] Ibid.

[6] In March 2005, Google was ranked fourth in most accessed U.S. sites by Nielsen, with a unique audience that month of 60M viewers, which equaled an audience reach of 43%.

http://www.netratings.com/news.jsp?section=dat_to&country=us The other principal mode of access to the Web has been guides done by individuals. In the early stages of the Web, these flourished. More recently, with increasing accuracy of search engines, they have declined in importance.

[7] On the gatekeeper metaphor, see Baye, M. R. and Morgan, J (2001). Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets, *American Economic Review* 91(3): 454-474.

[8] On the political dangers associated with search engines, see Introna, Lucas D. and Helen Nissenbaum (2000) "Shaping the Web: Why the Politics of Search Engines Matters", *The Information Society*, Vol. 16, No.3, 1-17; available at http://www.indiana.edu/~tisj/readers/full-text/16-3%20Introna.html. On government surveillance, see "The Nature and Scope of Governmental Electronic Surveillance Activity," Center for Democracy and Technology (2004), at http://www.cdt.org/wiretap/wiretap_overview.html; for current standards, see "CURRENT LEGAL STANDARDS FOR ACCESS TO PAPERS, RECORDS, AND COMMUNICATIONS: What Information Can the

indeed members of the international community in general, cannot make informed decisions without access to accurate and complete information. Within a few years, the Web has become the favored source of information retrieval. When we want to find more information about a topic, whether it be torture or tsunamis, we turn first—and often only—to the Web. The Web has become the principal source of research information for most Americans who do casual research. Typically, users turn first to Google for searches; Machill et al. estimated that 74% of users turn to Google first.[9]

Third, search engines have become central to education.   Students today perform countless web searches in an average day. They search Google far more often than they go to the library, undoubtedly more often than they look in a book for information. Search engines play a role analogous to the card catalogue in traditional libraries and the indices, such as the *Reader's Guide to Periodical Literature*, that were so important to students of the previous generation. Imagine a library without a card catalogue; that would be a close analogy to the Web without search engines, but with one important difference. Books would still be written without card catalogues, but without search engines, many persons and groups would probably not develop their websites.
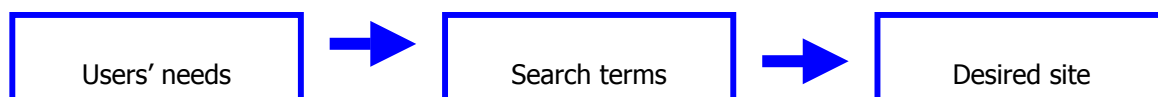
Fourth, search engines are owned by private corporations, businesses that are quite properly seeking to make a profit. These companies, especially Google since it has become the search engine of choice for so many millions, have a crucial public responsibility but are accountable to shareholders, not the general public. This sets up a tension between the public role of search engines and their corporate accountability.

Government Get About You, and How Can They Get It?"   at http://www.cdt.org/wiretap/govaccess/govaccesschart.html

[9] Machill, M., Neuberger, C., Schweiger, W. and Wirth, W, "Wegweiser im Netz" Qualität und Nutzung von Suchmaschinen," in *Wegweiser im Netz: Qualität und Nutzung von Suchmaschinen,* Verlag Bertelsman Stiftung, Bielefeld, p. 397.

Let's now examine three areas in which we encounter difficult and persistent ethical issues in search engine technology.

The Problem of the Algorithm

The key to the success of Google was an important conceptual shift in the understanding of searches. Initially search engines used fairly elementary algorithms to determine page rank such as the number of visits to a page, the number of other pages which link to a given page. What is common to these initial approaches to user searches was that they depended on objective criteria such as the number of page views. A given search engine could certainly get it wrong, but that did not diminish the fact that there was an objective fact of the matter to be gotten wrong. These initial searches were at least intended to rank the most popular sites, where "popularity" would have a technical and objective meaning.

The shift in what we could call second-generation search engines involved looking much more closely at what users wanted to find, which was not always the most popular site, but the site that most closely meets their needs. The remarkable success of Google depends in part on its ability to offer users what they are looking for, based on the search terms that are entered. Thus we have the following relationship:

This is conceptually very different from a ranking of page popularity alone; what the user wants becomes an integral part of the formula, as does the set of search terms most commonly used to express what the user wants.

The situation described above is complicated by the fact that the search algorithms that govern searches are well-kept secrets, and properly so. Not only do these algorithms give some companies a competitive edge, but potential spammers can manipulate search engine results much more easily if they know the details of the algorithms used to rank search results. Consequently, the search process is not transparent, that is, we do not know why certain sites have been included or excluded and we do not know what some sites are ranked above others.

The Politics of Searching: Privacy and Liberty

In the aftermath of the September 11th attacks, the Federal Bureau of Investigation in the United States proposed to develop an email intercept system that could sniff out possible terrorist threats, getting right to the "meat" of the message and disregarding the inessential. Carnivore, as it came to be known,[10] was designed to monitor email traffic, but it is easy to see the way in which the same argument could justify monitoring internet searches. Carnivore, like most FBI computer projects, was a technical failure and abandoned, after an expenditure of $6-15M, in favor of commercial software.[11] After all, if the government is entitled by the Patriot Act of 2001 to see what books we have been taking out from the library,[12] wouldn't the same logic mandate access to search requests?

The potentially chilling effects of such a situation are clear. The technical difficulties are significant but surmountable. Certainly it is virtually impossible to check who is doing searches from a public computer. From office or home machines, it's at least possible to obtain ip addresses, and sometimes more if, for example, someone has cookies enabled. Most recently, Google has offered a voluntary search history, "My Search History," that allows users to store and retrieve their searches. It "lets you easily view and manage your search history from any computer."[13] Google stresses the benefits for end users, building on the fact that most of us have at one time or another been unable to retrieve a reference we originally found in a Google search but cannot find again. However, there is obviously an economic motive behind this helpful attitude: Google can provide advertisers with far more sophisticated consumer profiles if it maintains a comprehensive database of search histories that can be sorted by individual user. To some extent, this is already

possible with cookies and with individuals signed in with a Gmail account, but the new "My Search History" feature increases accuracy dramatically and tracks users across multiple machines.

Economics is driving these technological developments in tracking search engine users, but the truly frightening aspect of this is political rather than economic. We all leave countless virtual footprints as we move through the day, using credit cards, making cell phone calls, accessing ATM machines, etc. These already provide a surprisingly detailed picture of an individual's daily life at least in terms of external activities. Search histories, however, go one step further: they provide an excellent source of insight into what someone is *thinking*, not just what that person is doing.

The danger, at least in the United States, is that such monitoring may be used increasingly to monitor and eventually suppress political dissent. The terrorist attacks of September 11th were ironically effective in strengthening public support for the erosion of personal liberty in the United States, and one can easily imagine government monitoring of search engine activity justified as a counter-terrorism measure.[14]

If such a scenario seems too implausible, and if it seems unthinkable that major search engine companies would cooperate with such an undertaking, one only has to look at Internet filtering in China today to see what the future may hold.

Local Standards in a Global Village

Perhaps the most frightening aspect of the power of search engines has occurred recently in China, which has made massive and highly effective efforts to prevent average Chinese citizens from accessing certain sites on the Internet. The accepted wisdom has been that the Internet is an unstoppable force for democratization, a force for liberation that cannot be tamed by local governments.

---

[10] Later, it was called DCS-1000.

[11] "FBI cuts Carnivore Internet probe," CNN website. Tuesday, January 18, 2005 Posted: 9:59 PM EST (0259 GMT) Tuesday, January 18, 2005.

[12] "FBI monitoring library records in terror probe," Associated Press, June 25, 2002 (http://www.freedomforum.org/templates/document.asp?documentID=16468; last accessed 5/3/05).

[13] https://www.google.com/searchhistory/login

[14] For an insightful discussion of this issue in the European context, including a discussion of the differences between the American and European contexts, see Michael Nagenborg, "Privacy and Terror: Some Remarks from Historical Perspective, IJIE International Journal of Information Ethics, Vol. 2 (11/2004), 1-5.

This assumption has been proved false in the case of Internet censorship in China. The Chinese government has succeeded in blocking the access of the average Chinese computer user to political sites dealing with the Dalai Lama and free Tibet, the Falun Gong, Tiananmen Square and—most recently—the Chinese demonstrations against Japan's most recent attempts at revisionist history.[15] The report of the ONI on "Internet Filtering in China 2004-2005" indicates that China has been far more successful in preventing its citizens from accessing certain websites than previously imagined. China's approach has been multi-pronged. Much of it occurs at the backbone level, which is highly effective, but this is supplemented by restrictions on internet service providers and even down to the level of cybercafés, which are required to track customer usage.[16] Email appears to be filtered at the service provider level, not at the backbone level, and increasingly sophisticated anti-spam filtering software can also be modified for use in political filtering. Blog provides are carefully monitored through keyword filtering, and politically incorrect bloggers are typically removed quickly from the servers. Within China, when one looks for Google, one often reaches alternative search engines such as Openfind, Globepage, chinaren.com, search.online.sh.cn, and fm365.com.[17] These

search engines are easily manipulated to carry out the kind of filtering that the Chinese government mandates.[18]

It is important to realize here the degree of cooperation that China has gotten from the West in its Internet filtering programs. Certainly much of the backbone of China's Internet has been supplied by American manufacturers. According to the ONI Country Study on China, Cisco Systems has played a pivotal role in providing the infrastructure that enables the Chinese government to filter the Internet so effectively.[19] Without the technical expertise and physical infrastructure provided by American companies, China's Internet filtering endeavors would be far less successful.

The role of Google in this situation, at least what we know of that role, does little to quell fears about the ways in which Google may be subject to political pressure. In 2004, the Chinese government began intermittently to shut down access from within China to the China Edition of Google News. Eventually, Google decided to shape its search results within China to the expectations of the Chinese government. A Google statement describes the situation in the following terms.

> There has been controversy about our new Google News China edition, specifically regard-

---

[15] Jonathan Krim, "Web Censors In China Find Success," *Washington Post*, Thursday, April 14, 2005; Page A20. Also see Jonathan Zittrain and Benjamin Edelman, "Empirical Analysis of Internet Filtering in China," Berkman Center for Internet & Society, Harvard Law School: http://cyber.law.harvard.edu/filtering/china/ ; last accessed 5/2/05; this includes a complete list of the 18,931 sites blocked by the Chinese government.

[16] OpenNet Initiative (ONI), "Internet Filtering in China 2004-2005: A Country Study," April 14, 2005. http://opennetinitiative.net/studies/ **china**/ONI_China_Country_Study.pdf Also see Jonathan Zittrain and Benjamin Edelman, "Internet Filtering in China," 2003. http://unpan1.un.org/intradoc/groups/public/documents/apcity/unpan011043.pdf

[17] Berkman Center for Internet & Society, Harvard Law School, "Replacement of Google with Alternative Search Systems in China: Documentation and Screen Shots,"

http://cyber.law.harvard.edu/filtering/china/google-replacements/

[18] OpenNet Initiative: Bulletin 005, "Probing Chinese search engine filtering," August 19, 2004 http://www.opennetinitiative.net/bulletins/005/

[19] "There has been considerable debate about the complicity of Western corporations in the development and maintenance of China's filtering system. China's Internet infrastructure includes equipment and software from U.S. companies, including Cisco Systems, Nortel Networks, Sun Microsystems, and 3COM.28 Cisco Systems in particular has been integral to China's Internet development. The core of China's Internet relies on Cisco technology; Cisco specifically implemented the backbone networks for ChinaNet29 and CERNet30, China's nation-wide educational network. Cisco's involvement continues to this day with the company's role in the development of China's "Next-Generation Network," known as CN2.31." "Internet Filtering in china 2004-2005," pp. 6-7.

ing which news sources we include. For users inside the People's Republic of China, we have chosen not to include sources that are inaccessible from within that country.[20]

In other words, Google decided to respect the Chinese political censorship rather than allow it to be shut down once again.

Although China is a vast potential market, it currently has little economic influence over Google, and presumably no political power over it. Nevertheless, Google seems to have accommodated itself to the wishes of the Chinese government. If this is the case, one cannot help but worry that Google could eventually be much more strongly influenced by the United States government, which has far greater economic and political impact on Google than does the government of China.

Conclusion

Search engines play an increasingly pivotal role in the distribution and eventual construction of knowledge, yet they are largely unnoticed, their procedures are opaque, and they are almost completely devoid of independent oversight: powerful, cloaked in secrecy, and not subject to external control. Insofar as the flourishing of deliberative democracy is dependent on the free and undistorted access to information, and insofar as search engines are increasingly the principal gatekeepers of knowledge,

we find ourselves moving in a politically dangerous direction. We risk having our access to information controlled by ever-powerful, increasingly opaque, and almost completely unregulated search engines that could shape and distort our future largely without our knowledge. For the sake of a free society, we must pursue the development of structures of accountability for search engines. Based on the cases discussed above, there is little reason to think that search engines will remain impervious to external political and economic pressures.

Lawrence M. Hinman, Director of the Values Institute and Professor of Philosophy at the University of San Diego, writes widely in the area of applied ethics. He is the founder of *Ethics Updates* (http://ethics.sandiego.edu) and *Ethics Videos* (http://ethics.sandiego.edu/video/).

---

[20]

http://www.google.com/googleblog/2004/09/china-google-news-and-source-inclusion.html Google concludes, "On balance we believe that having a service with links that work and omits a fractional number is better than having a service that is not available at all. It was a difficult tradeoff for us to make, but the one we felt ultimately serves the best interests of our users located in China. We appreciate your feedback on this issue." Also see the links at http://www.google-watch.org/china.html .