2
Through the Google Goggles: Sociopolitical Bias in Search Engine Design

A. Diaz

Summary Search engines like Google are essential to navigating the Web's endless supply of news, political information, and citizen discourse. The mechanisms and conditions under which search results are selected should therefore be of considerable interest to media scholars, political theorists, and citizens alike. In this chapter, I adopt a "deliberative" ideal for search engines and examine whether Google exhibits the "same old" media biases of mainstreaming, hypercommercialism, and industry consolidation. In the end, serious objections to Google are raised: Google may favor popularity over richness; it provides advertising that competes directly with "editorial" content; it so overwhelmingly dominates the industry that users seldom get a second opinion, and this is unlikely to change. Ultimately, however, the results of this analysis may speak less about Google than about contradictions in the deliberative ideal and the so-called "inherently democratic" nature of the Web.

2.1 Introduction

As knowledge, commerce, and politics continue move online and to the Web in particular, search engines have quickly become the "gatekeepers" of cyberspace. What's more, a *single* search engine – Google – now handles the majority of Web queries. Google directs hundreds of millions of users towards some content and not others, towards some sources and not others. As with all media gatekeepers, if we believe in the principles of deliberative democracy – and especially if we believe that the Web is an open, "democratic" medium – then we should expect our search engines to disseminate a broad spectrum of information on any given topic.

In the first section of this chapter, I describe how a "deliberative media" ideal can be used to evaluate search engine and why, as media critics have done with prior innovations, we should examine Google's content biases, its advertising policies, and consolidation in the industry as a whole. Subsequent sections will dive into each of these areas: first, we will look at the deliberative implications of the PageRank algorithm Google uses to crawl and order Web content; next,

we will critically examine the role advertising plays in Google's search results; lastly, we discuss the implications of a highly concentrated and commercial competitive search landscape. It is hoped that through this investigation, we might start to uncover the sociopolitics of search.

2.2   In Search of a Democratic Medium

The Supreme Court once observed that "the dissemination of the widest possible information from diverse and antagonistic sources is essential to the welfare of the public" (Associated Press v. United States 1945: 20). This goes to the heart of 'deliberative democracy,' a concept that has in recent years gained considerable currency among political scientists and media critics (Benhabib 1996; Elster 1998; Fishkin 1991; Sunstein 1997). For Benjamin Page, "In order that the public as a whole can collectively control what its government does, the public, collectively, must be well informed. Some kind of *public* deliberation is required" (Page 1996: 5). Individuals' exposure to "diverse and antagonistic views" is central to such debate, as John Stuart Mill (1859) once argued:

> He who knows only his own side of the case knows little of that. His reasons may be good … but if he is unable to refute the reasons of the opposite side, if he does not so much as know what they are, he has no ground for preferring either opinion … [H]e should hear the arguments … from the persons who actually believe them, who defend them in earnest and do their very utmost for them. He must know them in their most plausible and persuasive form. (p. 35)

For Mill, it does not matter whether arguments are popular or unpopular, correct or incorrect, offensive or pleasing; what matters is that public opinion is given the *opportunity* to "be set right when it is wrong" (p. 19). This is why "streets and parks," according to Justice Roberts, "have immemorially been held in trust for the use of the public and…have been used for purposes of assembly, communicating thoughts between citizens, and discussing public questions" (Hague et al. v. CIO et al., 1939, 515).

The deliberative model appears to capture what we usually mean by "democratic media": forums in which every corner of society is represented fairly – spaces where the debate isn't dominated by corporations, politicians, or privileged groups. Given that we are a nation too large and too distributed to engage in a singular, Habermasian debate (1990, 89) the media have an important role to play in ensuring that speakers have access to heterogeneous citizens. As Justice Kennedy observed, "minds are not exchanged in streets and parks as they once were. To an increasing degree, the more significant exchanges … occur in mass and electronic media" (DAETC et al. v. FCC 1996, 132). Given the enormous reach of radio, television, and newspapers, the media could allow citizens to access a range of perspectives they might not otherwise encounter.

2.2.1   The Traditional Media and a New Hope

But as countless critics have argued, the mass media have fallen far short of these aspirations. "[M]arket forces," writes Cooper (2003), "provide neither adequate incentives to produce the high quality media product, nor adequate incentives to distribute sufficient amounts of diverse content necessary to meet consumer and citizen needs" (p. 43). The economics of dissemination and the politics of deregulation, rather than encourage the formation of alternative outlets, have concentrated the media in fewer and fewer hands (p. 141). The scarcity of alternative channels has allowed media companies to pursue ever-greater profit margins through advertising, sponsorship, and product placement with little fear of consumer retaliation (McChesney 2000: 39–42). "The media," according to Bagdikian (1992), "have become partners in achieving the social and economic goals of their patrons" (p. 151). The value of large audiences has tended to yield "middle-of-the road," nonpolitical, mainstream content that creates a "buying mood" but fails to represent unpopular or diverse opinions. The result is a media landscape characterized by sameness, by a suppression of controversy, and by hypercommercialism.

But a new medium has recently emerged, and it has promised to change all this. Decentralized and distributed, the global Internet – and, in particular, the Web – allows anyone and everyone to make their views accessible, and to access anyone's views. It is arguably more like the printing press than radio and, indeed, information on the network is not constrained by the limits of printed matter, by delivery distances, or seemingly by time, space, and matter itself. With a click of the mouse, you can read information and opinions that have not been "filtered" by profiteering corporations or corrupt governments. At the same time, underrepresented and unheard groups can cheaply bypass the "monolithic media empire" to have a voice. The Internet is many-to-many, all-to-all, and it has for many restored faith mediated deliberation. Aspirations are expressed repeatedly, and with understandable excitement:

> The Web…breaks the traditional publishing model. … [It] says instead, "You have something to say? Say it. You want to respond to something that's been said? Say it and link to it. … And you never have to ask anyone's permission." (Weinberger 2002: vii–ix)

> You don't have to be writing for an organization to have a credible voice. The Net elevates those voices. What the large media were about was distribution capacity to communicate with hundreds of thousands of people. Now the Net does that. (Barlow, qtd. in Lasica 1996)

The prospects seemed so exhilarating that some jumped to label the Internet "inherently" democratic (Gilder, qtd. in Schuler 2003: 72).

And indeed, the Web has had many tangible, positive effects for diverse, democratic discussion. Access and content on the new medium has exploded; the majority of Americans now have Internet connections in their homes (Wellman and Haythornthwaite 2002: 13). Anyone with basic computer competence can now publish a Web site viewable around the globe. Activists have grown their own "grassroots" communities to pursue particular policy objectives while bloggers – self-made 'journalists' who report their findings and solicit comments in a

sort of "deliberative diary" – have gained loyal followings and the attention of the mainstream media (Rodzvilla 2002). Real-world community projects have sprung up online, "evidence of an overdue renewal of interest in democracy" (Schuler 2003: 73).

And yet, over the last ten years, user traffic on the Web has gravitated around a few, large, and increasingly commercial sites. In a fascinating book, Notre Dame physicist Albert-Lásló Barabási (2002) recounts how his team of scientists mapped the Web's structure to reveal disturbing evidence about the supposed "egalitarianism" of the network. He found that a small number of pages – what he called "hubs" – are linked to by a great many other pages, while the vast majority of documents are linked to by few or no sites at all. Hubs are very easy to "come across" from anywhere on the Web; they are therefore more likely to be linked to, which further increases their discoverability (the so-called "rich get richer" phenomenon). Meanwhile, a typical page – one pointed to by only couple documents – remains almost impossible to find. It's no wonder that, by 2001, over half of users' online time was being spent at four sites; one third of the total time was spent at AOL-Time Warner properties (CNN.com 2001). On the political Web – the set of sites dealing with democratically urgent issues such the death penalty, Congress, and gun control – Hindman et al. (2003) found "strong and consistent" patterns consistent with Barabási's research: "the number of highly visible sites is small" and "almost all prominent sites are run by long-established interest groups, by government entities, by corporations, or by traditional media outlets" (p. 26). The link structure of the Web suggests the medium exhibits the same old problems: "it is hard for all but a few 'ordinary citizens' to post their views prominently – and conversely, to read the views of other ordinary citizens, unless they are highlighted by a small number of prominent sites" (p. 30). Or as Barabási put it, "The hubs are the strongest argument against the utopian vision of an egalitarian cyberspace. Yes, we all have the right to put anything we wish on the Web. But will anybody notice?" (p. 58).

2.2.2   Search Engines as Intermediaries

That so many accessible pages go unseen suggests that the Internet has done away with "spectrum scarcity" but not with *attention scarcity* (Kottke 2003). Sure, there are literally billions of pages ("channels") available on the Web. But there is a rather fixed limit to how many we, as individuals, can consume. With television, radio, and the print media, we rely on the mass media to condense the available opinions and make them easily accessible through newspapers, the evening news, radio broadcasts, and so on. And the same sort of intermediation is required online.

The key "general interest intermediaries" of the Web, I argue, are the search engines. These sites are the primary means by which Internet users are directed towards particular sources of information and are among the first and most frequently accessed pages for the vast majority of users. Consider: each one of the top 5 sites

is either a portal or search engine (Burns 2007); by 2004, 84% of online Americans had used search engines, and a majority of these used them at least once a day (Fallows and Rainie 2004); search engines are the most popular way to locate medical, governmental, and religious information on the Web (Fallows 2005); fully 79% of those seeking online election information began their journeys at portals and search engines (Cornfield and Rainie 2003, p. 25).

So when Steven Levy (1995) said that "instead of a gatekeeper, users get an open invitation to the electronic world and can choose whatever they want" (p. 59), he was being less than accurate. Internet users *do* get a gatekeeper – the search engine – and they choose primarily among the sites it offers to them. As with all such intermediaries, we expect search engines to present the available information in a fair and diverse manner; we expect them, in other words, to be "democratic." We should ask about search engines like Google the same questions scholars have asked about the traditional media: *Can underrepresented voices and diverse viewpoints be heard through the filter of search engines? What role does advertising play in the returned results? Do a few players dominate the industry?* Only by answering these questions – as we will do in turn – can one assess the true "deliberativeness" of the Web itself.

2.3 The Politics of PageRank

Just as the mass media have the power and responsibility to disseminate unusual and heterodox views, so too do search engines have the capability to highlight those high-quality, out-of-the-mainstream sites that would otherwise be lost in the deafening din of the Web. Automated crawling and ranking can do what we, as individuals, cannot: find, catalog, and consider millions of poorly-linked and underrepresented pages – what Chris Anderson (2004) has called "the long tail" – and ultimately break through the link inequality that calls into question the egalitarian ideal of the Web.

2.3.1 The Mathematics of PageRank

So does Google actually promote those dissident and minority views so critical for a "well-functioning democracy"? Given the complex and propriety nature of Google's search technology its software looks at over 100 features of a page to ascertain "relevance" (Mayer 2005) – answering this question is exceedingly difficult. But we can start with what Google (2004) calls "the heart" of its software: the PageRank algorithm. PageRank estimates the "importance" of an arbitrary page by looking at how many *other* "important" pages link *to* it. Mathematically, the PageRank of your page is the weighted count of links pointing to it, with links from high-PageRank documents contributing more to your score than links from low-PageRank documents.

An analogy may be useful: an academic paper is "important" if many other papers cite it – and especially if it is referenced by other, highly cited works ("canons").

With PageRank, then, being "important" means being "popular" or "visible." PageRank actually turns out to be the precise probability that a "random surfer" clicking links from page to page will come across a given document. Thus the highly-referenced "hubs" Barabási worried about have the highest PageRanks; these tend to be the sites of large, famous, technology-oriented companies such as Amazon and eBay (Upstill et al. 2003). In contrast, the millions of "typical" pages – those we are already unlikely to "randomly" stumble across – have among the lowest PageRank values. Google apparently uses PageRank to guide its crawlers such that popular sites have a better chance of being indexed (Cho et al. 1998). Sites with high PageRank also tend to be more prominent among the search results (Diaz 2005: 81–85).

2.3.2   PageRank as a Voting Mechanism

According to Google's public relations literature (2004), PageRank is not only consonant with democratic principles; it in fact embodies the very process of democracy itself: "Google interprets a link from page A to page B as a vote, by page A, for page B." Princeton computer science professor and cyberactivist Ed Felten (2004) puts it more colorfully:

> Google is a voting scheme … not a mysterious Oracle of Truth. … It's a form of democracy – call it Googlocracy. Web authors vote by creating hyperlinks, and Google counts the votes. If we want to understand Google we need to see democracy as Google's very nature, and not as an aberration.

But what Ed Felten and other PageRank proponents fail to recognize is the important distinction between the ideal process of "democratic" *governance* and that of "democratic" *discourse*. Sure, a political democracy generally requires that the aggregated preferences of the majority be put into practice. But this does not imply that only the majority's views should be heard during deliberation, nor does it suggest that popular opinions should be preferred *ipso facto*. To the contrary, the validity of voting – of aggregating preferences – depends precisely on the dissemination of a broad spectrum of opinions, especially those put forth by unpopular or minority groups (Mill 1859: 16).

From the perspective of deliberative democrats, then, PageRank is highly problematic. Unpopular but nevertheless democratically critical voices face a double bind: search engines like Google are "biased against [these pages], ignoring them as they crawl the Web" (Barabási 2002: 58) and, even if the pages make it to the index, they may find themselves buried among the results. To the degree that Google adopts a PageRank bias, it *mirrors* rather than mitigates the Web's link inequality.

Indeed, some scholars have argued that the use of PageRank actually *magnifies* the Web's skewed distribution of links, making it increasingly difficult for new sites to be discovered (Fortunato et al. 2006; Hindman et al. 2003; Pandey et al. 2005). The problem is this: a well-linked page appears prominently on search engines like Google; this page therefore enjoys greater traffic; and, as users become even more aware of the site, they link to it on their own pages, increasing the document's PageRank and visibility even further. The result is a "vicious cycle," "entrenchment bias," or "googlearchy" wherein popular pages are, over time, increasingly likely to maintain their prominence while new pages become more difficult to discover. Cho and Roy's (2004) computer simulation indicated that "it takes 66 times longer" for a new page to become popular by means of highly PageRank-biased search engines than by pure "random surfing."

2.3.3 The "Common Case" and Majoritarian Interests

PageRank therefore seems to reproduce the same sort of "antideliberative" bias typically associated with the traditional media. To recall Cooper's (2003) remarks about big media: "In the commercial model, popular, mainstream, and middle of the road ideas will almost certainly find a voice, one that is likely to be very loud. However, the unpopular, unique, and minority points of view will not" (p. 16). Similarly, "search engines wishing to achieve greatest popularity … tend to cater to majority interests" (Introna and Nissenaum 2000: 176). According to Google's founders, this bias was by design:

> One of the design goals of PageRank was to handle the common case for queries well. … [T]he goal of finding a site that contains a great deal of information … is a very different task … There is an interesting system that attempts to find sites that discuss a topic in detail … this results in good results for queries like "flower"; the system will return good navigation pages from sites that deal with the topic of flowers in detail. Contrast that with the common case approach which might simply return a commonly used commercial site that had little information except how to buy flowers … [W]e are concentrating only on the common case approach. (Page et al. 1999: 10–11).

PageRank, in other words, abandons the goals of actually reflecting a page's "importance" or "authoritativeness" on a given subject, and instead aims to mirror the "common" wishes of users. This, as the creators' own example illustrates, can have the problematic effect of promoting popular, commercial pages over more detailed, noncommercial sources of information.

To be sure, these problems are more or less typical of commercial search engines in general. In their groundbreaking overview of search engine bias, Introna and Nissenbaum (2000) observed that "while markets undoubtedly would force a degree of comprehensiveness and objectivity in listings, there is unlikely to be much market incentive to list sites of interest to small groups of individuals … or, for that matter, individuals of lesser economic power" (p. 177). PageRank's

"one size fits all" approach does little for the atypical, outside-the-mainstream individuals that might actually wish to see or communicate controversial content.

2.3.4   Suppression of Controversy

Susan Gerhart (2004) makes a similar point in a unique content analysis of Web search results. Gerhart queried Google, Teoma, and AllTheWeb for information on five broad topics, each of which she knew to contain some controversial subtopic that was well documented on the Web. Gerhart then recorded, in painstaking detail, whether and how such disputed perspectives were raised within the search results. She looked, for example, at whether a search for "distance learning" would return sites that shared David Noble and other academics' concern about "the loss of control over their intellectual products, as well as contact with students" and the tendency of these programs to act as "digital diploma mills." Similarly, she looked at whether the results for "Einstein" mentioned the debate over whether his first wife received appropriate credit for contributions to his work.

Her findings indicate that when a controversy was frequently discussed within a topic and widely recognized as important (e.g., the effectiveness of St. John's Wort) the disputed matters were, indeed, represented among the results. When searching for female astronauts or St. John's Wort, for example, it was possible for a user to "definitely recognize the existence of controversy, which [a result] explains in some detail." But for three of the topics – distance learning, Albert Einstein, and Belize – the respective disputes were to a great extent "suppressed," such that most surfers would not "be exposed to the controversies by [a general] search…alone." In these cases, the controversies were overrun by "organizational clout" (e.g., official Belizean tourism sites or distance learning programs) or by pages that reflected what users "wanted to see" (e.g., Einstein quotations, 'bland' biographies for term papers, etc.). In the end, the controversial viewpoints that perhaps matter most from a deliberative point of view – those antagonistic perspectives that haven't garnered widespread attention – are precisely those that are left out of the search engine's results. Gerhart concludes that

> Search technology tends to present the 'sunny side' of a topic. This bias reflects authors' links and searchers' choices. A few organizations often exert strong commercial (or nonprofit) influence through Web site investments and accrue high link counts through their off-Web prominence. ('Conclusion').

If we really believe that through "democratic media" like the Web individuals "must have the freedom to communicate radical and unpopular ideas and opinions" – and, what's more, that citizens should be exposed to what "they don't want to hear" – then search engines fall short of these aspirations when they fail to disseminate those dark, uncomfortable views on a given topic.

Of course, "Web search engines do not conspire to suppress controversy." Rather, this is direct consequence of the seemingly laudable attempt to please its

users. As Gerhart suggests, "On the simplest query for a topic, a searcher expects to see the most influential organizations appear, not a bundle of dirty laundry or diatribes attacking the topic's leaders or ideas … Searchers user a particular engine because its biases give them the results they usually want." The deliberative model may ask of too much from users: pushing for them to see what they *don't* want to see because, really, it's "good for them" (Rostbøll 2005). To this extent, it conflicts with intuitive and reasonable ideas for how search engines should work.

2.3.5   Small Players (Still) Matter

Even if PageRank does, in theory, encode an antideliberative, antidemocratic bias, a few caveats are in order. First, as Dan Bricklin (2002) has pointed out, even if popular sites do get a sizeable boost for some queries, rarely do the same corporate megasites pop up across different search topics. As a result, "small players [still] matter," especially when we are conducting 'typical' searches for specialized information not easily found in the traditional media. Although it is difficult for a page to gain visibility on established topics – Microsoft," "abortion," or "flowers" – an unprecedented number of "ordinary citizens" may still be reaching sizeable publics through the Google search engine.

Second, PageRank is only one element of Google's ranking algorithm; consider, for example, that PageRank is completely *query-independent*, capturing the "importance" of a page irrespective of the user's stated interest. In practice, Google takes many other factors into account when ordering search results: whether the query appears in the page's title, what words people use to link to the page, and so forth. While it is true that PageRank predicts rank position *in the aggregate*, individual result sets exhibit at best a weak correlation (Diaz 2005: 84). For this reason, Cho and Roy's simulation – which assumed search results were strictly ordered by PageRank – may be unrealistic. Indeed, a more recent study suggests that search engines' query-dependent heuristics actually "smear out the traffic attraction of high-degree pages…counteracting the skewed distribution of links in the Web [by] directing some traffic toward sites that users would never visit otherwise" (Fortunato et al. 2006: 6). Clearly, there is a need for continued and systematic research into the many; sometimes counteracting biases of today's advanced search technology.

2.4   Commercialism, Advertising, and "Mixed Motives"

Advertising is, by and large, how the commercialized media make money. Newspapers, magazines, radio, and television outlets provide free or inexpensive content to their readers, listeners, and viewers; in exchange, they sell advertisers access to these audiences. Advertising is, however, a mixed blessing. On the one hand, it makes it viable to disseminate information to a broad audience at a low

cost; on the other hand, there is the persistent threat that the wishes of sponsors will subtly work their way into the content itself, narrowing the range of opinions that can be profitably and widely expressed.

These competing forces come strongly into play in the arena of search engines. As the primary gatekeepers of the Web, search engines not only direct users to particular pages but can also direct consumers towards particular services and products. This presents an enormous opportunity for targeted advertising: search engines can "sell" access to highly segmented audiences while marketers can target individuals who are actively expressing interest in a topic or product. The money-making potential is enormous and, indeed, one industry report predicted as early as 2003 that "worldwide search revenue estimates of $7B by 2007 are conservative" (Raschtchy and Avilio 2003). By 2005, advertisements on Google *alone* brought in over $6 billion – or over 99% of the company's yearly revenue (Google 2006).

But by selling advertising, Google and its competitors have an enormous financial incentive to direct users away from the "free," "organic" results and towards the sites of its sponsors. These "mixed motives" are stated eloquently by none other than Google's founders themselves, in an appendix to their 1998 Stanford research paper:

> The goals of the advertising business model do not always correspond to providing quality search to users. For example, in our prototype search engine one of the top results for cellular phone is … a study which explains [the] risk associated with conversing on a cell phone while driving … It is clear that a search engine which was taking money for showing cellular phone ads would have difficulty justifying the page … For this type of reason and historical experience with other search engines we … expect that advertising funded search engines will be inherently biased towards the advertisers and away from the needs of the consumers. … Since it is very difficult even for experts to evaluate search engines, search engine bias is particularly insidious … [and] less blatant bias are likely to be tolerated by the market. (Brin and Page 1998: 17–18)

2.4.1   A Brief History of Search Advertising

When Google's founders wrote those words, the predominant form of search advertising was the so-called "banner" ad. As it turns out, these ads tended not to work well in the context of search. For one, only a few banners can reasonably be placed on each page, and searchers would often click their result before the image had finished loading (Sullivan 2003a). More importantly, users quickly developed an ability to unconsciously spot and ignore banners, focusing – with "laser beam accuracy" – on what they perceived to be the actual search results (Pagendarm and Schaumburg 2001). If sponsors wished to be noticed, their solicitations must look like, and appear amongst, the actual results. As the CEO of one search engine company put it, "The money is in the search results themselves, not the billboards on the site of that road. The question is how do you profit from the search results, when they have been given away for free" (Thornley qtd. in Pagendarm and Schaumburg 2001).

The way many of Web search engines have gone about "profiting from their results" is by offering various kinds of "paid listings." The most common scheme, called *paid placement,* allows sponsors to purchase search-result-like text ads that appear above, below, or alongside the "organic" results for their chosen keywords. Sometimes these paid results are marked as "sponsored" listings; other times, "it may be hard for the average user to distinguish" (Crowell 2003). Unsurprisingly, paid placement proved vastly more effective than previous methods at drawing users' attention towards sponsors' sites. To the degree that these "matches" walk, talk, and act like relevant results, users click them. As *Business Week* puts it, paid placements have become "the Holy Grail of Internet advertising, and no wonder" (Reinhardt 2003). These ads have caught on, in some form or another, among virtually all of Web's most popular search engines (Google, Altavista, AOL, AskJeeves, Hotbot, Google, Lycos, MSN, and Yahoo! have similar offerings). The demand for paid listings quickly became so great that, according to *The Economist* (2004), they "lead the recovery in advertising expenditure on the Internet."

While paid listings may be a bonanza for search companies, investors, and advertisers alike, their implications for online, egalitarian discourse are depressingly obvious:

> [The] concept that Web sites should be able to buy their way to the top of search listings is being copied in one way or another by every major search and portal site. As they do, the search engines, which are still the most popular gateways to the Web, are transforming themselves from infinite electronic encyclopedias to the more prosaic, if profitable, role of universal commercial directories. (Hansell 2001)

To the extent that the commercial interests of the rich dominate the results of even noncommercial queries, the practice of selling prominence can seriously distort what the Web consists of for millions of users.

But just as market forces drive search engines to paid placement, so too do market forces push back. If, as commercial listings become more numerous, the relevancy of a search engine's results decline, dissatisfied users may switch to a competitor, resulting in an overall decline in advertising revenues. From this angle, the amount of paid listings to include is a straightforward optimization problem. Economists Bhargava and Feng (2002) respond to it by proposing "a mathematical model for optimal design of a paid placement strategy" that would "give a search engine the best balance between revenues from content providers and revenues based on user base" (p. 122).

2.4.2   'Clearly Labeled'?

For search engine critics, however, such economic models are not very comforting. We might reasonably wonder – as Brin and Page did in 1998 – whether users will actually see what's missing from their search results. Although Bhargava and Feng assume "that search engines cannot hide the fact that they perform paid placement" (p. 118), it appears that many Internet users remain unaware of such practices.

In 2002, a study commissioned by Consumers Union found that *fewer than one in four* Internet users had ever heard of search engines "taking fees to list some sites more prominently than others" (Princeton Survey Research Associates 2002: 17). After being told that, in fact, most search engines do exactly this, "a solid majority (80%) say it is important for search engines to tell users about their fee details, including 44% who say it is *very* important" (p. 17). At the time of the study, several search engine companies were using remarkably vague and misleading terminology to demarcate their paid listings (e.g., "Featured Sites," "Products and Services"). So in 2001 the watchdog group Commercial Alert filed a complaint with the FTC alleging that seven search companies were engaging in "deceptive advertising" practices (Miller 2001). When FTC responded in June 2002, it did not call for immediate action against the search engines named in the complaint (Gallagher 2002), but it did agree that there was a "need for clear and conspicuous disclosures of paid placement … to advise consumers as to when they are being solicited, as opposed to being impartially informed" (Hippsley 2002).

Google has largely avoided criticism for its AdWords paid placement program and the company was noticeably absent from the Commercial Alert complaint. While other search engines were happily crowding their search results with "Featured Links," Google insisted on drawing a line – quite literally – between "paid" and "organic" results. Algorithmically, advertising was to have no effect on the selection and ordering of the free results, and ads were "clearly marked" as "Sponsored Links." These results initially appeared only to the right of the "organic" results, but today Google includes up to three sponsored links directly above the top result. Even though these are also labeled as "sponsored," selected by relevance not price, and appear over a colored background (Sullivan 2002, AdwordsRep 2004), the fact remains: a considerable portion of Google's revenue comes from moving ads to the most prominent positions *above* the "first hit." It is unclear whether, in practice, users perceive these as ads; Google, after all, has an enormous interest in blurring that line.

In any case, disclosure alone does not solve the problems of paid listings. If we really wish to promote ideals of democratic discourse, then we should worry about any policy that allows those with money to be featured prominently among results for a given topic. This concern, it should be emphasized, is not with advertising in general. It is with a particular type of advertising competes with "organic," relevant content; it is with advertising that supplants, rather than complements, the pages individuals might otherwise see. Despite what Brin and Page say today, paid listings, even if disclosed, are not "just like" advertising in the traditional media. Industry reporter Danny Sullivan (2003a), however, disagrees:

> Think newspapers. Newspapers have both "editorial" copy, which is not supposed to be influenced by advertising, as well as ads themselves. You may read the paper primarily for the articles, but there are certainly times when you may find the advertisements useful, as well … In "old" media … most people can readily identify ads because they look or act so very different from "content."

But there's the rub. In the new media of search engines, paid listings (as opposed to banner ads) *don't* "look or act so very different" from normal results. Search

engines with paid listings are hardly like a newspaper with lots of informative, unbiased content and obvious, product-oriented ads sprinkled here and there. They are, perhaps, more like a newspaper in which half articles on the front page are written and paid for by commercial groups and discretely labeled as "sponsored articles." Or like network television if half the primetime spots were allotted to infomercials. No wonder, despite Sullivan's claim that users will eventually "learn" to distinguish paid and unpaid content, a 2005 study continued to find that "While most consumers could easily identify the difference between TV's regular programming and its infomercials…only a little more than a third of search engine users are aware of the analogous sets of content commonly presented by search engines" (Fallows 2005: 3).

2.4.3 'Wine but not Beer' and Other Ad Policies

Although the relationship between Google's editorial and paid listings has largely escaped criticism, much controversy *has* surrounded the company's 'bias' with respect to the selection of advertisements. As Brin admits, "We don't try to put our sense of ethics into the search results, but we do when it comes to advertising" (Sheff 2004). The resulting scheme is a patchwork of proscriptions: the search engine doesn't accept ads for beer, but it does for wine (Sheff 2004); ads for pornography are fine, but ads for guns are not (Johnson 2003); you can promote T-shirts depicting the cannabis leaf and drug paraphernalia, but you may not advertise water pipes (Kopytoff 2004). Most worrisome, perhaps, is that ads have been rejected because the sponsoring site – or even a page it links to – advocates *against* an individual or group. When the nonprofit environmental advocacy group Oceana tried to run ads on Google, they were rejected because the organization's site was critical of Royal Caribbean Cruise Lines, a Google advertiser (Liedtke 2004). In August 2004, the *San Francisco Chronicle* obtained internal documents detailing the company's advertising policies (Kopytoff 2004). These policies prohibited ads for sites that bashed politicians, gave special scrutiny to ads by the Church of Scientology, and allowed sites to advertise on the keyword 'abortion' only if they made no reference to religion. Yahoo! and MSN, by many accounts, impose far fewer restrictions on the content of the ads they run.

It is not immediately clear what this bias means from the perspective of democratic discourse. Media scholars like McChesney and Bagdikian are, after all, not so much worried about biased advertising standards as they are about the dissolution of the boundary between editorial content and advertising. Google has seemingly adopted a similar position, steadfastly reminding the press, its users, and advertisers that its advertising biases in "no way affect the results [they] deliver" – as if that puts an end to the matter (Google 2005). But it is not too difficult see how advertising selectivity may have political and deliberative implications, as Lawrence Lessig (2004) suggests in his latest book:

> Say you want to run a series of ads that try to demonstrate the extraordinary collateral harm that comes from the drug war. Can you do it? Well, obviously, these ads cost lots of money. Assume you raise the money … Can you be sure your message will be heard then?
>
> No. You cannot. Television stations have a general policy of avoiding "controversial" ads. Ads sponsored by the government are deemed uncontroversial; ads disagreeing with the government are controversial … Thus, the major channels of commercial media will refuse one side of a crucial debate the opportunity to present its case. (p. 167)

By recognizing that advertising may be used as a tool not only for promoting products, services, political candidates but also as medium to voice antagonistic opinions about these subjects, Lessig and other scholars have argued that advertising too may serve as a kind of deliberative forum (though one largely confined to the well-heeled).

Ultimately, in newspapers or television, radio or the Internet, whenever editorial content is interspersed with paid content that deals with similar topics, the spectrum of views put forth on that subject encompasses both types of material, for better or for worse. And so, the more advertisements Google includes under the constraints of this policy the more likely it is that users will find mainstream, commercial sites promoting a particular position, product, or service, and the less likely it is that they will hit noncommercial, antagonistic, 'controversial' voices. These voices, so critical for deliberative discourse, are multiply penalized: they are less likely to appear in the 'organic' listings that tend to suppress controversy; they are less likely to have the financial means to buy a prominent advertising spot; and, even if they had the money, their message may not conform to Google's content standards. Brin and Page knew all this back in 1998; indeed, they were so uncomfortable with the "inherent bias" of commercial search engines that they declared it "crucial" to keep Google "a competitive search engine that is transparent and in the academic realm" (p. 19).

2.5   Towards 'Coke and Pepsi'? Search Engine Consolidation

Over the past few decades, concerns over media consolidation have reached a fever pitch (e.g., McChesney 1999). Media concentration allows companies to increase profit margins by leveraging economies of scale (via horizontal integration) and by developing mutually-reinforcing, cross-promotional "synergies" (through vertical conglomeration). Unfortunately, this raises the barriers of entry for newcomers who may not be able to effectively engage in wage price wars or gain access to cross-promotional outlets (Compaine and Gomery 2000: 521). To the extent that a few firms succeed in amassing control of the media, the dissemination of diverse and antagonistic views is potentially undermined.

Applying these concerns to the field of search engines, we might suppose, as Kawaguchi and Mowshowitz (2002) do in their study of variance among the search engines, that "too few intermediaries spells trouble":

> The only real way to counter the ill effects of search engine bias on the ever-expanding Web is to make sure a number of alternative search engines are available. Elimination of competition in the search engine business is just as problematic for a democratic society as consolidation in the news media. (p. 60)

Surprisingly, the issue of concentration in the search engine industry has received relatively little attention (exceptions include Sheu and Carley 2001, and Elizabeth van Couvering's chapter in this volume). Perhaps this is because the relatively nascent, "transitioning" state of the industry makes it difficult to distinguish long-term patterns from the normal wax and wane of competitors in new markets.

But as the dot-com dust settles, and as consistent, comparative market data become available, a fairly clear pattern of consolidation starts to emerge. Users are, first of all, increasingly converging on a smaller set of search engines. In 1998, each of the top 8 search engines was used by at least 10% of the online audience and, on average, reached about 23% of all Web users (Sullivan 1998). But today, the top three sites – Yahoo!, Microsoft, and Google – handle over four-fifths of all search traffic (Sullivan 2005), and almost half of Web users frequent a single search site (Fallows and Rainie 2004). Thus, whereas users were once distributed across many portals and individually relied on several different search engines, today they stick to a few, overwhelmingly popular sites (Diaz 2005: 130).

In addition, ownership of the various search sites has been consolidated into the hands of a decidedly smaller number of companies. These developments were predicted as early as 1996, when Jupiter Communications, an industry research firm, forecast an imminent "shake-out" in the sector. "There are simply too many players," they warned investors, "offering similar functionality and features, competing for a limited number of advertising dollars and users" (qtd. in Sullivan 2001a). Even Excite's CEO, George Bell, was pessimistic about the chances for survival: "There are a lot of 'two' examples out there … There's Pepsi and Coke, Time and Newsweek … the third always tends to struggle, the fourth tends to get bought. I think [Yahoo and Excite] will make it" (qtd. in Sullivan 2001a). Excite, of course, did not make it. After a steady decline in profitability and traffic, it was ultimately acquired by Ask Jeeves, which also gobbled up DirectHit, Teoma, iWon, MyWay, and MyWebSearch (Waters and Lee 2003). Yahoo! was relatively lucky, keeping a significant market share while acquiring Inktomi, AllTheWeb, Altavista, Del.icio. us, and paid listings pioneer Overture (Ostrom 2003). While Robin Kellet of MSN UK believes that the "period of consolidation is probably almost over" (qtd. in van Vark 2004), the sector is already dominated by a few, relatively large corporations, not a multitude of independent startups.

In light of these developments, it is not surprising to find many referring to the emerging 'search oligopoly' (Arnold 2003). And, economically speaking, that characterization seems apt. Under Kaysen and Turner's determination of oligopoly, for example, "type I" oligopoly is achieved when "the eight largest firms have 50% of receipts and the 20 largest at least 75%" (Compaine and Gomery 2000: 555–556) — a threshold easily exceeded regardless of whether we look at ad revenues or traffic share (Diaz 2005: 131–132). Applying the more complex Herfiendhal-Hirshman Index (HHI), which "reflects … the number and size distribution of firms

in a market, as well as concentration of the output" (Rhodes qtd. in Compaine and Gomery 2000: 558–559) we see that search achieves a level of concentration exceeding those of the much-ballyhooed traditional media sectors (Sheu and Carley 2001 14; Cooper and Cooper 2003). For Sheu and Carley, these indicators suggest that "the industry looks close to being plagued by anticompetitive practices" (p. 22). But unlike the oil, film, and music oligopolies – which "work together to … restrict the game of profit maximizing to themselves" (Compaine and Gomery 2000: 275) – there is little evidence today of a "search engine cartel," and new competitors do occasionally spring up (Grossman 2003, Hansell 2005). The situation may be in this respect more like the "loose and open oligopoly" of book publishing industry, in which "one can properly lament some concentration…[is] nowhere near as tightly controlled as movies and music" (Compaine and Gomery 2000: 517–518).

Oligopoly or not, Google is perched firm and tall at the top of the search industry. While it was certainly a latecomer – only 5% of Web users had accessed the site as of December 2000 (Sullivan 2001b) – it now handles half of all U.S. Web searches and its users are far and away the most loyal, with 56% of them using nothing else (Fallows 2005). Amazingly, Google rose to the top without the aid of mergers, acquisitions, or even a large advertising budget, and it did so amidst a stock market crash that was decimating the dot-coms. Its success was simply attributed to the quality of its results, a product of its unique and groundbreaking technologies like as PageRank. Google's algorithmic superiority clearly caught they eye of users, but it also garnered the attention of "portal" operators such as Yahoo! and AOL, who had previously ignored the importance of search and were now clamoring to license Google's technology to power its own sites (Prather 2002; Rosenberg 1998). And so, by May 2003, after Yahoo! and AOL outsourced their search technology to Google, the Web's top three search destinations were all powered by the Mountain View startup. Taken together, this meant that it was fielding a whopping 76% of all Web search queries performed in the United States (Sullivan 2003b). Although Yahoo! has since switched to in-house search technology, today Google continues to field twice as many queries as its Sunnyvale rival (Nielsen 2006).

Google's technology thus has an enormous influence on virtually all online discourse and communication. In what is perhaps the "ultimate measure of impact" on the public consciousness, its name has become a verb: potential mates "google" each other before a date, recruiters "google" job applicants, citizens "google" information on Iraq, and schoolchildren "google" for everything from encyclopedia articles to games for their graphing calculators. As Jonathan Zittrain of the Harvard Law School explains, Google has quickly become the "the traffic cop at the main intersection of the information society" (qtd. in Markoff and Zachary 2003).

If decades of media criticism are any guide, this should be a cause for concern. A dominant intermediary like Google has both the opportunity and the incentives to hypercommercialize content and to bias results in a self-interested manner. The effects of any such negative, 'antidemocratic' bias are intensified in proportion to how widely Google's search technology is diffused (McGinn 1990, p. 99). For instance, if users find Websites primarily through search engines (they do), if Google

handles the vast majority of these search queries (it does), and if the use of PageRank does result in popular, mainstream opinions dominating the search results, then Google's monopoly could make it considerably more difficult for 'ordinary' sites to be seen by a significant population of Web users.

But concern over Google's dominance need not hinge on whether or not the company has illicit motives, or on whether its results are 'democratically' selected. All intermediaries, even the fairest ones, must have biases; they must all somehow choose to elevate some issues, opinions, and voices and to ignore others (Goldman 2006). But when many intermediaries can reach a sizable chunk of the public each encoding its own opinions about what is interesting, relevant, or valid – these biases can counteract each other and, taken as a whole, a broad array of opinions can be disseminated (in part, through a second step of interpersonal communication as discussed in Katz 1973). In contrast, when only one or a few outlets have any significant reach, there is enormous inequality in what is transmitted: some views garner lots attention, and those left out are not heard at all. Consequently, for Bagdikian (1992), it is consolidation – irrespective of commercialization – that is the real enemy:

> The threat does not lie in the commercial operation of the mass media. It is the best method there is and, with all its faults, it is not inherently bad. But narrow control, whether by government or corporations, is inherently bad. In the end, no small group, certainly no group with as much uniformity of outlook and as concentrated in power as the current media corporations can be sufficiently open and flexible to reflect the full richness and variety of society's values and needs. The answer is not elimination of private enterprise in the media, but the opposite. It is the restoration of genuine competition and diversity. (p. 223–224)

We should therefore worry when we hear one writer conclude, "so powerful has Google become that many … view it as the Web itself: if you're not listed on its indexes, they say, you might as well not exist" (Olsen 2002). While the opposite extreme – a highly balkanized audience with little "common ground" – has its own problems (Sunstein 2001: 91–99), a World Wide Web that consists only of what appears at the top of Google's results is, frankly, a very attenuated sort of deliberative public forum.

2.6 Letting the Market Decide: Barriers of Entry to Search

There is an obvious retort to all this: if Google does betray the values of the Web and the needs of its users, the quality of its product will decline, and its users will just switch to another search engine. Search engines are highly substitutable commodities: users can just type in a different URL or can change their browser's start page. This is how David Zetland (2005) dismisses "naïve claims" that Google "reduces our access to dissident, minority or heterodox views": "The objection … is groundless" because "[e]ntry is easy, and Google has major rivals" (p. 7). Or as Eric Goldman (2006) puts it, "market forces limit the scope of search engine bias" (p. 196).

Sadly, it's not so simple. First, because it is so hard to see what's "missing" from search results; users are unlikely to have the necessary knowledge to even consider a switch (Telang et al.1999). Second, bundled services such as email tend to "lock" users to a search engine. Google, for example, is building brand loyalty is through tight integration into the major browsers, through its release of the Google Toolbar (Miller 2005), and through its countless "portal-like" services including GMail, Google Calendar, and Personalized Search. These are clear – and logical – ways of increasing both switching costs for users and barriers to entry for competitors (Sheu and Carley 2001: 17–18).

Of course, our romantic visions of tech innovation – billion-dollar companies sprouting from Silicon Valley garages – suggest that the next great search engine may be just around the corner. Google, after all, began its ascent to the top, and eventually overtook multi-billion-dollar giant Yahoo!, on a 'mere' $25 million dollar investment (Marshall 2005). Since then competitors have continued to appear, though very few been able to obtain a significant market share (Hansell 2005). Thus, we may conclude that the market is, as Compaine argues, "oligopoly proof" (Compaine and Gomery: 476).

But already, companies are finding it hard to keep up with the exponential growth of the Web, which demands highly complex technical systems and enormous expertise to manage:

> Today, the wholesale search market has significant barriers to entry. Economies of scale have asserted themselves, secondary competitors have folded, and the creation of new search engines by startups is becoming prohibitively expensive. Consider: to crawl, index, and search more than eight billion pages still only a fraction of the Web – Google now operates a global infrastructure of more than 250,000 Linux-based servers of its own design … and is becoming a major consumer of electrical power, computer hardware, and telecommunications bandwidth. (Ferguson 2005)

These economic hurdles, according to an executive at Ask Jeeves, are "likely to lead to more consolidation rather than competition from new entrants" (Cox qtd. in Glover 2005). Already, the "search engine wars" are between Yahoo!, Google, and Microsoft; independent general interest search engines are few and far between, unlikely to raise the necessary capital. Even niche, "vertical" search engines seldom gain significant market share, and profitability is likely to continue inducing search engines to focus on majoritarian interests.

What about the "competitive," "academic" search engine Brin and Page promised us? Is such a search engine still possible? Probably not. As one search engine manager points out, making such a system available would "cost you a ton of money":

> This is why ever since 2000, 2001, most of the search research done at the universities is what I call Metacrawler-esque, which is people not building a search engine but doing something on top of a search engine, because they just can't afford to build their own. Which is a shame, because you're not getting these big engines coming out of academia any more. (qtd. in van Couvering 2004: 10)

The virtually insurmountable regulatory and economic challenges plaguing existing proposals for more "egalitarian" search engines suggest that the market

mechanism – despite all its problems – may be, at least for now, the only practical means of getting a viable search engine off the ground. Google's founders may have abandoned their original vision for a search engine that is "competitive" and "in the academic realm" not because they sold out, but because they had come to see this as a contradiction.

2.7 Conclusion: Is Google 'Evil'?

Given the critical-analytic lens with which we have approached the sociopolitics of search, it may seem that I am promoting the view that Google is, in fact, "evil." But this is certainly not the case. It is my view that the observed tensions between the search engine and democratic aspirations are, for the most part, not the product of malicious or even profiteering intent. Instead, they stem from both the high demands of the democratic model and the inherent limitations of commercialized search. It is hard to imagine a search company staying afloat, after all, if it does not present what its users want; it is difficult to make money if it does not display advertising; it is unprofitable to operate a competitive search engine without a very significant market share.

It would be quite difficult to suggest that we are better off without Google or, for that matter, without any of the other search engines. Awash in a sea of bits, we may be tempted to look at 'democraticness' and 'bias' as binaries, as things you either have or you don't. It makes more sense to take a step back, and to think of Google as *one more way* in which people can get information. Only the most hardened cynic would think that the success of Google has resulted in a *net loss* of sources to which we are exposed. And so, as Compaine reminds,

> the questions to ask yourself are: Are there more or fewer voices available today than 15, 25 or more years ago? And, is it easier or harder, are the regulatory barriers higher or lower, is it more expensive or less expensive, to gain access, in whatever format, to a large audience … than in 1900? in 1950? in 1990? (Compaine and Gomery: 576).

I believe that the answers to all these questions are emphatically positive, and that Google – certainly more than the traditional broadcast media – is making it possible for more people to hear and contribute to a broader spectrum of opinions.

But there are a number of ways we might think about improving the deliberativeness of search engines like Google. For many, the answer is in technology – not regulation or subsidization. Cho and his colleagues, for example, have proposed two alternatives to PageRank – random selection (Pandey et. al 2005) and popularity increase rate (Cho and Adams 2003) – that arguably surface high-quality content while mitigating popularity bias and entrenchment effects. Echoing Brin and Page's (1998) footnoted musings about user-seeded PageRank computation (p. 15), Goldman (2006) has put his faith in personalized search: "Personalized algorithms mean that there are multiple 'top' search results for a particular search term … so Web publishers will not compete against each other in a zero-sum game … reducing

structural biases" (p. 199). At least one search engine has already "spent a lot of R&D" on algorithms that attempt to distinguish between commercial and noncommercial searches – ensuring that paid results "only show up under paid queries" – and Yahoo! Mindset allows users to interactively bias their results according to whether they are "shopping" or "researching" (Raschtchy and Avilio 2003). Such innovative solutions are what we need to pursue if search engines are to serve the needs of both citizens and of consumers. It is what we need if Web search engines are to serve democracy, while remaining economically viable.

The purpose of this chapter has been to take a hard look at the search engine we rely on. But it is also intended to reveal the many, difficult entailments of utopian, democratic ideals associated with the Web. The deliberative standard is quite clearly an extremely difficult – some might say impossible – one to meet. And so, it's not that haven't moved forward. It's just that we aren't quite "there" yet.

References

AdWordsRep: Top position? Impossible? In: Search Engine Watch Forums. http://forums.searchenginewatch.com/showthread.php?t=2952 (2004). Cited 9 Apr 2007

Anderson C The long tail. *Wired* 12(10). http://www.wired.com/wired/archive/12.10/tail.html (2004). Cited 9 Apr 2007

Arnold S In search of clicks that make cash: three search companies are all chasing the same pool of advertising spend. World Information Review (1 Apr 2003)

Associated Press v. United States, 326 U.S. 1, US Supreme Court (1945)

Bagdikian B (1992) The media monopoly, 4th edn Beacon Press, Boston

Barabási (2002) A Linked: the new science of networks. Perseus Books, Cambridge

Benhabib S (1996) Toward a deliberative model of democratic legitimacy. In: Democracy and difference: contesting the boundaries of the political. ed by Benhabib, S. Princeton University Press, Princeton, pp 67–94

Bhargava H, Feng J (2002) Paid placement strategies for Internet search engines. In: Proceedings of the WWW2002 Conference, Honolulu, Hawaii. http://portal.acm.org/citation.cfm?id=511462 (2002). Cited 8 Apr 2007. pp 117–123

Bricklin D (2002) Why small players matter. Dan Bricklin's Web site. http://www.bricklin.com/smallplayers.htm (2002). Cited 10 Apr 2005

Brin S, Page L (1998) The anatomy of a large-scale hypertextual web search engine. Seventh International WWW Conference, Brisbane, Australia. http://www-db.stanford.edu/pub/papers/google.pdf (April 1998). Cited 1 May 2005

Burns E (2007) Top U.S. parent companies and stickiest brands on the Web, February 2007. Clickz Network. http://www.clickz.com/showPage.html?page=3625269 (2007). Cited 9 Apr 2007

Cho J, Adams R (2003) Page quality: in search of an unbiased web ranking. Technical report, UCLA Computer Science Department. http://oak.cs.ucla.edu/~cho/papers/cho-quality-long.pdf (2003). Cited 9 Apr 2007

Cho J, Roy S (2004) Impact of search engines on page popularity. In: Proceedings of the WWW2004 Conference, New York. http://oak.cs.ucla.edu/~cho/papers/cho-bias.pdf (May 2004). Cited 6 Aug 2006

Cho J, Garcia-Molina H, Page L (1998) Efficient crawling through URL ordering. In: Proceedings of 7th World Wide Web Conference, http://dbpubs.stanford.edu/pub/1998-51 (2000). Cited 9 Apr 2007

CNN.com: Four sites account for half of Web surfing. http://www.cnn.com/2001/TECH/ Internet/06/05/Internet.consolidation/ (5 June 2001). Cited 27 Apr 2004

Compaine B, Gomery D (2000) Who owns the media? competition and concentration in the mass media industry, 3rd edn, Lawrence Erlbaum Associates, Mahwah

Cooper M (2003) Media ownership and democracy in the digital information age. http://cyberlaw. stanford.edu/blogs/cooper/archives/mediabooke.pdf (2003). Cited 15 May 2005

Cooper M, Cooper S (2003) Hope and hype v. reality: The role of the commercial Internet in democratic discourse and prospects for institutional change. Stanford Law School Center for Internet and Society. http://cyberlaw.stanford.edu/blogs/cooper/archives/HOPEALL.pdf (2003). Cited 7 May 2005

Cornfield M, Rainie L (2003) Untuned keyboards: online campaigners, citizens, and portals in the 2002 elections. In: Pew Internet and American Life Project. http://www.pewInternet.org/pdfs/ PIP_IPDI_Politics_Report.pdf (2004). Cited 11 May 2005

Crowell G (2003) The "secret system" of search engine advertising. In: Search Engine Watch. http://searchenginewatch.com/_subscribers/articles/article.php/3289361 (2003). Cited 28 Apr 2004

Cunningham F (2002) Theories of democracy: a critical introduction. Routledge, New York

Denver Area Educational Telecommunications Consortium, Inc., et al. v. Federal Communications Commission et al., 518 U.S. 727, US Supreme Court(1996).

Diaz A (2005) Through the google goggles: sociopolitical bias in search engine design. Undergraduate honors thesis, Stanford University.

The Economist (2004) Spiders in the web (15 May) pp 16

Elster J (ed) (1998) Deliberative democracy. Cambridge University Press, Cambridge

Fallows D (2005) Search engine users: Internet searchers are confident, satisfied, and trusting –but they are also unaware and naïve. In: Pew Internet and American Life Project. http://www. pewInternet.org/pdfs/PIP_Searchengine_users.pdf (2005). Cited 8 Apr 2007

Fallows D, Rainie L (2004) The popularity and importance of search engines. In: Pew Internet and American Life Project. http://www.pewInternet.org/pdfs/PIP_Data_Memo_Searchengines.pdf (2004). Cited 10 May 2005

Felten E (2004) Googleocracy in action. In: Freedom to Tinker. http://freedom-to-tinker.com/ archives/000509.html (2004). Cited 15 May 2005

Ferguson C (2005) What's next for google? The search firm wants to organize all digital information; that means war with Microsoft. Technology Review: MIT's Magazine of Innovation 108: 38–46

Fishkin J (1991) Democracy and deliberation: new directions for democratic reform. Yale University Press, New Haven

Fortunato S, Flammini A, Menczer F, Vespignani A (2006) The egalitarian effect of search engines. In: Proceedings of the WWW2006 Conference, Edinburgh, United Kingdom. http:// arxiv.org/pdf/cs.CY/0511005 (May 2006). Cited 6 Aug 2006

Gallagher D (2002) U.S. Warns web sites to label sponsorships. The New York Times (2 July 2002)

Gerhart S (2004) Do web search engines suppress controversy? First Monday 9(1). http://firstmonday. org/issues/issue9_1/gerhart/index.html (2004). Cited 1 May 2005

Glover T (2005) Search engines power new dotcom boom. Sunday Business (27 March 2005)

Goldman E (2006) Search engine bias and the demise of search engine utopianism. Yale Journal of Law and Technology 8: 188–200

Google (2004) Our search: google technology. http://www.google.com/technology (2004). Cited 13 May 2005

Google (2005) Google adwords content policy. http://adwords.google.com/select/contentpolicy. html (2005). Cited 21 Mar 2005

Google (2006) Financial data. google investor relations. http://investor.google.com/fin_data.html (2006). Cited 6 Aug 2006

Grossman L (2003) Search and destroy: a gang of web search companies is gunning for Google. Time (22 Dec 2003) pp 46–50

Habermas J (1990) Moral consciousness and communicative action. MIT, Cambridge

Hague, Mayor et al. v. Committee for Industrial Organization et al., 307 U.S. 496, US Supreme Court (1939)

Hansell S (2001) Clicks for sale: paid placement is catching on in web searches. The New York Times (4 June 2001)

Hansell S (2005) Search sites play a game of constant catch-up. The New York Times (31 Jan 2005)

Hindman M, Tsioutsiouliklis K, Johnson J (2003) Googlearchy: how a few heavily-linked sites dominate politics on the Web. In: Annual Meeting of the Midwest Political Science Association. http://citeseer.ist.psu.edu/hindman03googlearchy.html (2003). Cited 8 Apr 2007

Hippsley H (2002) Re complaint requesting investigation of various Internet search engine companies for paid placement and paid inclusion programs. U.S. Federal Trade Commission. http://www.ftc.gov/os/closings/staff/commercialalertletter.htm (2002). Cited 20 May 2004

Introna L, Nissenbaum H (2000) Shaping the web: why the politics of search engines matters. The Information Society 16: 1–17

Johnson J (2003) Google accepts porn ads but refuses those for guns. CNSNews. http://www.cnsnews.com/Culture/archive/200310/CUL20031007c.html (2003). Cited 23 July 2003

Katz E (1973) The two-step flow of communication: an up-to-date report of a hypothesis. In: Marketing classics: a selection of influential articles, ed by Enis, B., Cox, K. Allyn and Bacon, Boston

Kopytoff V (2004) Google ad rules complex, controversial: documents reveal details about what the popular search engine accepts, rejects. San Francisco Chronicle (9 Aug 2004)

Kottke J (2003) Weblogs and power laws. http://kottke.org/03/02/weblogs-and-power-laws. Cited 1 Aug 2006

Lasica JD (1996) Interview with John Perry Barlow. http://www.jdlasica.com/interviews/barlow.html (1996). Cited 25 April 2004

Lessig L (2004) Free culture: how big media uses technology and the law to lock down culture and control creativity. Penguin, New York

Levy S (1995) How the propeller heads store the electronic future. New York Times Magazine (Sep 24) pp 58–59

Liedtke M (2004) Google bans environmental group's ads. USA Today Online. http://www.usatoday.com/tech/news/2004-02-12-google-bans-ad_x.htm (12 Feb 2004). Cited 10 June 2004

Markoff J, Zachary G (2003) In searching the web, google finds riches. The New York Times (13 Apr)

Marshall M (2005) Google founders' brashness sparks debate. San Jose Mercury News (18 Aug)

Mayer M (2005) Progress in research and ads (talk). In: Google factory tour. (Google.com, 2005), http://www.google.com/intl/en/press/factorytour.html (2005). Cited 20 May 2005

McChesney R (1999) Oligopoly: the big media game has fewer and fewer players. The Progressive (November) 1: 20–24

McChesney R (2000) Rich media, poor democracy: communication politics in dubious times. New Press, New York

McGinn R (1990) Science, technology and society. Prentice Hall, Englewood Cliffs

Mill JS (1978) On liberty, 1859, ed by Rapaport, E. Hackett, Indianapolis

Miller L (2001) How to net results in search-site seas. USA Today (13 Aug)

Miller M (2005) Broadband everywhere: Internet explorer still dominates the web, but firefox is growing fast and igniting innovation. PC Magazine. http://www.pcmag.com/article2/0,1759,1770267,00.asp (22 March 2005). Cited 9 Apr 2007

Mowshowitz A, Kawaguchi A (2002) Bias on the web. Communications of the ACM 45: 56–60

Nielsen//Netratings: Google accounts for half of all U.S. searches in April. http://www.nielsen-netratings.com/pr/pr_060525 (25 May). Cited 31 July 2006

Olsen S (2002) The google gods: does search engine's power threaten web's independence? CNet News. http://news.com.com/2009-1023-963618.html (31 Oct 2002). Cited 9 Apr 2007

Ostrom M (2003) Pasadena, calif: commercial search firm to buy web search properties. San Jose Mercury News (26 Feb)

Page B (1996) Who deliberates? University of Chicago Press, Chicago

Page L, Brin S, Motwani R, Winograd T (2007) The pagerank citation ranking: bringing order to the web. http://dbpubs.stanford.edu/pub/1999-66 (1999). Cited 8 Apr 2007

Pagendarm M, Schaumurg H (2001) Why are users banner-blind? The impact of navigation style on the perception of web banners. Journal of Digital Information 2 http://journals.tdl.org/jodi/article/view/jodi-37/38

Pandey S, Roy S, Olston C, Cho J (2005) Shuffling a stacked deck: The case for partially randomizing ranking of search engine results. http://reports-archive.adm.cs.cmu.edu/anon/2005/CMU-CS-05-116.pdf (2005). Cited 8 Apr 2007

Prather M. (2002) Ga-ga for google. Entrepreneur Magazine. http://www.entrepreneur.com/article/0,4621,297807,00.html. Cited 9 Apr 2007

Princeton Survey Research Associates: A Matter of Trust: What Users want From Web Sites. Results of a National Survey of Internet Users for Consumer WebWatch. http://www.pewtrusts.org/pdf/vf_web_watch_trust_0402.pdf. Cited 8 Apr 2007

Raschtchy S, Avilio J (2003) Industry note: search symposium shows bigger role for search in advertising. U.S. Bancorp Piper Jaffray. http://marketwatch.multexinvestor.com/download.asp?docid=29692599&sid=26. Cited 5 Feb 2005

Reinhardt A (2003) And you thought the web ad market was dead: sales of so-called 'Paid placement' listings are soaring. In: BusinessWeek Online. http://www.businessweek.com/magazine/content/03_18/b3831134_mz034.htm (2003). Cited 8 Apr 2007

Rodzvilla J (2002) (ed) We've got blog: how weblogs are changing our culture. Perseus Books, Cambridge

Rosenberg S (1998) Yes there is a better search engine: while the portal sites fiddle, Google catches fire. Salon.com. http://archive.salon.com/21st/rose/1998/12/21straight.html (1998). Cited 5 Mar 2005

Rostbøll C (2005) Preferences and paternalism: on freedom and deliberative democracy. Political Theory 33: 370–396

Schuler D (2003) Reports of the close relationship between democracy and the Internet may have been exaggerated. In: Democracy and New Media, ed by Jenkins, H., Thorbum, D. MIT, Cambridge

Sheff D (2004) Playboy interview: google guys. Playboy 51: 55–60, 142–145

Sheu T, Carley K (2001) Monopoly power on the web: a preliminary investigation of search engines. In: Proceedings of the 29th Telecommunications Policy Research Conference, Alexandria, Virginia. http://arxiv.org/as/cs.CY/0109054. Cited 7 May 2005

Sullivan D (1998) NetRatings search engine ratings, June 1998. Search Engine Watch. http://searchenginewatch.com/mhts/9806-10-netratings.mht. Cited 20 Dec 2004

Sullivan D (2001a) Nielsen NetRatings search engine Ratings, December 2000. Search Engine Watch. http://searchenginewatch.com/mhts/9902-0012-netratings.mht. Cited 21 Dec 2004

Sullivan D (2001b) The end for search engines? Clickz Experts. http://www.clickz.com/experts/search/opt/article.php/837281. Cited 23 Dec 2004

Sullivan D (2002) FTC recommends disclosure to search engines. Search Engine Watch. http://searchenginewatch.com/sereport/article.php/2164891. Cited 20 May 2004

Sullivan D (2003a) Buying your way in search engine advertising chart. In: Search Engine Watch. http://searchenginewatch.com/wemasters/article.php/2167941. Cited 20 May 2004

Sullivan D (2003b) comScore Media Metrix search engine ratings, August 2003. Search Engine Watch. http://searchenginewatch.com/mhts/0305-mediametrix.mht. Cited 1 May 2005

Sullivan D (2005) ComScore Media Metrix search engine ratings, December 2004. Search Engine Watch. http://searchenginewatch.com/reports/article.php/2156431. Cited 8 May 2005

Sunstein C (1997) Deliberation, democracy, and disagreement. In: Justice and democracy: cross-cultural perspectives, ed by Bontekoe, R., Stepaniants, M. University of Hawaii Press, Honolulu, pp 92–117

Sunstein C (2001) Republic.com. Princeton University Press, Princeton

Telang R, Mukhopadhyay T, Wilcox R (1999) An empirical analysis of the antecedents of Internet search engine choice. Workshop on Information Systems and Economics, Charlotte, North Carolina

Upstill T, Craswell N, Hawking D (2003) Predicting fame and fortune: pagerank or indegree? In: Eighth Australasian Document Computing Symposium, Canberra, Australia. http://cs.anu.edu.au/~Trystan.Upstill/pubs/upstill_adcs03.pdf. Cited 10 May 2005

van Couvering E (2004) New media? The political economy of Internet search engines. Presented at the Annual Conference of the International Association of Media and Communications Researchers, Porto Alegre, Brazil. http://www.media.uio.no/prosjekter/ctp/papers/IAMCR-CTP04_S1-1_vanCouver.pdf. Cited 9 Apr 2007

van Vark C (2004) Search engines: search still sets the pace. Revolution (21 Apr)

Walters R, Lee A (2003) Ask jeeves to join excite Internet. The Financial Times (5 March)

Weinberger D (2002) Small pieces loosely joined: a unified theory of the web. Perseus, Cambridge

Wellman B, Haythornthwaite C (eds) The Internet in everyday life. Blackwells, Oxford

Zetland D. (2005) Is Google evil? Knowledge commodification, community and innovation. Presented at the Tenth International Kal Polanyi Conference, Istanbul (14 Oct). http://www.kysq.org/pubs/IsGoogleEvil.pdf. Cited 12 Nov 2007