MARQUETTE
UNIVERSITY

LAW SCHOOL

SEARCH ENGINE BIAS AND
THE DEMISE OF SEARCH ENGINE UTOPIANISM

Eric Goldman

(March 2006)

Eric Goldman
Assistant Professor of Law
Marquette University Law School
Sensenbrenner Hall
P.O. Box 1881
Milwaukee, Wisconsin 53201-1881
eric.goldman@marquette.edu

Search Engine Bias and the Demise of Search Engine Utopianism

By Eric Goldman[*]

Abstract: Due to search engines' automated operations, people often assume that search engines display search results neutrally and without bias. However, this perception is mistaken. Like any other media company, search engines affirmatively control their users' experiences, which has the consequence of skewing search results (a phenomenon called "search engine bias"). Some commentators believe that search engine bias is a defect requiring legislative correction. Instead, this Essay argues that search engine bias is the beneficial consequence of search engines optimizing content for their users. The Essay further argues that the most problematic aspect of search engine bias, the "winner-take-all" effect caused by top placement in search results, will be mooted by emerging personalized search technology.

In the past few years, search engines have emerged as a major force in our information economy, helping searchers perform hundreds of millions (or even billions) of searches per day.[1] With this broad reach, search engines have significant power to shape searcher behavior and perceptions. In turn, the choices that search engines make about how to collect and present data can have significant social implications.

Typically, search engine automate their core operations, including the processes that search engines use to aggregate their databases and then sort/rank the data for presentation to searchers. This automation gives search engines a

---

[*] Assistant Professor, Marquette University Law School, Milwaukee, Wisconsin. Home page: http://www.ericgoldman.org. Email: egoldman@gmail.com. I appreciate the comments of participants at the Yale Law School Regulating Search Symposium and the 2005 Association of Internet Researchers (AoIR) Annual Meeting in Chicago.

[1] In 2003, search engines performed over a half-billion searches a day. *See* Danny Sullivan, *Searches Per Day*, Search Engine Watch, Feb. 25, 2003, http://searchenginewatch.com/reports/article.php/2156461.

veneer of objectivity and credibility.[2] Machines, not humans, appear to make the crucial judgments, creating the impression that search engines bypass the structural biases and skewed data presentations inherent in any human-edited media.[3]

Unfortunately, this romanticized view of search engines does not match reality. Search engines are media companies. Like other media companies, search engines make editorial choices designed to satisfy their audience.[4] These choices systematically favor certain types of content over others, producing a phenomenon called "search engine bias."

Search engine *bias* sounds scary, but this Essay explains why search engine bias is both necessary and desirable. The Essay also explains how emerging personalization technology will soon ameliorate many concerns about search engine bias.

I.   SEARCH ENGINES MAKE EDITORIAL CHOICES

Search engines frequently claim that their core operations are completely automated and free from human intervention,[5] but this characterization is false.

---

[2] *See* Jason Lee Miller, *Left, Right, or Center? Can a Search Engine Be Biased?*, WebProNews.com, May 10, 2005, http://www.webpronews.com/insidesearch/insidesearch/wpn-56-20050510LeftRightorCenterCanaSearchEngineBeBiased.html.

[3] There is a broad perception that search engines present search results passively and neutrally. *See* Leslie Marable, *False Oracles: Consumer Reaction to Learning the Truth About How Search Engines Work*, June 30, 2003, http://www.consumerwebwatch.org/news/searchengines/ContextReport.pdf; Maureen O'Rourke, *Defining the Limits of Free-Riding in Cyberspace: Trademark Liability for Metatagging*, 33 GONZ. L. REV. 277 (1998).

[4] *See*, *e.g.*, C. EDWIN BAKER, ADVERTISING AND A DEMOCRATIC PRESS (1994).

[5] *See*, *e.g.*, *Does Google Ever Manipulate Its Search Results?*, Google.com, http://www.google.com/support/bin/answer.py?answer=4115&topic=368 ("The order and contents of Google search results are completely automated. No one hand picks a particular result for a given search query, nor does Google ever insert jokes or send messages by changing the order of results.") (last visited Dec. 16, 2005); *Does Google Censor Search Results?*, Google.com, http://www.google.com/support/bin/answer.py?answer=17795&topic=368 ("Google does not censor results for any search terms. The order and content of our results are completely automated; we do not manipulate our search results by hand.") (last visited Jan. 3, 2006); *Technology Overview*, Google.com, http://www.google.com/corporate/tech.html ("There is no human involvement or manipulation of results….") (last visited Mar. 22, 2006); *see also* Complaint at ¶¶ 37-38, 52-56, KinderStart.com LLC v. Google, Inc., Case No. C 06-2057 (N.D. Cal. Mar. 17, 2006)

Instead, humans make numerous editorial judgments about what data to collect and how to present that data.

Indexing. Search engines do not index every scrap of data available on the Internet. Search engines omit (deliberately or accidentally) some web pages entirely[6] or may incorporate only part of a web page.[7]

During indexing, search engines are designed to associate third party "metadata" (data about data) with the indexed web page. For example, search engines may use and display third party descriptions of the website in the search results.[8] Search engines may also index "anchor text" (the text that third parties use in hyperlinking to a website),[9] which can cause a website to appear in a search results for a term the website never used (and may object to).[10]

(giving other examples of Google's claims to be passive). Note that Google has subsequently revised some of these cited pages after its censorship controversy in China.

[6] *See* Judit Bar-Ilan, *Expectations Versus Reality – Search Engine Features Needed for Web Research at Mid-2005*, 9 CYBERMETRICS 2 (2005), http://www.cindoc.csic.es/cybermetrics/articles/v9i1p2.html.

[7] For example, many search engines ignore metatags. *See* Eric Goldman, *Deregulating Relevancy in Internet Trademark Law*, 54 EMORY L.J. 507, 567-68 (2005). Search engines also incorporate only portions of very large files. *See* Bar-Ilan, *supra* note 6; *Why Doesn't My Site Have a Cached Copy or a Description?*, Google.com, http://www.google.com/support/bin/answer.py?answer=515&topic=365 (last visited Nov. 25, 2005) (describing how some pages are "partially indexed"); *Has Google Dropped Their 101K Cache Limit?*, ResearchBuzz!, Jan. 31, 2005, http://www.researchbuzz.org/2005/01/has_google_dropped_their_101k.shtml (discussing how historically Google indexed only the first 101k of a document).

[8] *See My Site's Listing Is Incorrect and I Need it Changed*, Google.com, http://www.google.com/webmasters/3.html (last visited Nov. 25, 2005). Google's automated descriptions have spawned at least one lawsuit by a web publisher who believed the compilation created a false characterization. *See* Seth Fineberg, *Calif. CPA Sues Google Over "Misleading" Search Results*, ACCT. TODAY, Apr. 19, 2004, at 5.

[9] *See* Jagdeep S. Pannu, *Anchor Text Optimization*, WebProNews.com, Apr. 8, 2004, http://www.webpronews.com/ebusiness/seo/wpn-4-20040408AnchorTextOptimization.html.

[10] For example, the first search result in Google and Yahoo! for the keyword "miserable failure" is President George W. Bush's home page because so many websites have linked to the biography using the term "miserable failure." *See* Tom McNichol, *Your Message Here*, N.Y. TIMES, Jan. 22, 2004, at G1. This algorithmic vulnerability has spawned a phenomenon called "Google bombing," where websites coordinate an anchor text attack to intentionally distort search results. *See* John Hiler, *Google Time Bomb*, Microcontent News, Mar. 3, 2002, http://www.microcontentnews.com/articles/googlebombs.htm.

Finally, once indexed, search engines may choose to exclude web pages from their indexes for a variety of reasons, ranging from violations of quasi-objective search engine technical requirements[11] to simple capriciousness.[12]

Ranking. To determine the order of search results, search engines use complex proprietary "ranking algorithms." Ranking algorithms obviate the need for humans to make individualized ranking decisions for the millions of search terms used by searchers, but they do not lessen the role of human editorial judgment in the process. Instead, the choice of what factors to include in the ranking algorithm, and how to weight them, reflects the search engine operator's editorial judgments about what makes content valuable. Indeed, to ensure that these judgments are produce desired results, search engines manually inspect search results[13] and make adjustments accordingly.

Additionally, search engines claim they do not modify algorithmically-generated search results, but there is some evidence to the contrary. Search engines allegedly make manual adjustments of a web publisher's overall ranking,[14] and search engines occasionally modify search results presented in response to particular keyword searches.[15]

---

[11] *See*, *e.g.*, Stefanie Olsen, *Search Engines Delete Adware Company*, CNET News.com, May 13, 2004, http://news.com.com/2102-1024_3-5212479.html?tag=st.util.print (Google and Yahoo kicked WhenU.com out of their indexes for allegedly displaying different web pages to searchers and search engine robots, a process called "cloaking").

[12] This is the heart of KinderStart's allegations against Google. *See* Complaint, KinderStart.com LLC v. Google, Inc., Case No. C 06-2057 (N.D. Cal. Mar. 17, 2006). Although the complaint's allegations about Google's core algorithmic search may not be proven, Google does liberally excise sources from Google News. For example, Google claims that "news sources are selected without regard to political viewpoint or ideology," *see Google News (Beta)*, Google.com, http://news.google.com/intl/en_us/about_google_news.html#25 (last visited Dec. 16, 2005), but Google dropped a white supremacist news source from Google News because it allegedly promulgated "hate content." *See* Susan Kuchinskas, *Google Axes Hate News*, Internetnews.com, Mar. 23, 2005, http://www.internetnews.com/xSP/article.php/3492361.

[13] *See* Posting of Eric Goldman to Technology & Marketing Law Blog, *Google's Human Algorithm*, http://blog.ericgoldman.org/archives/2005/06/googles_human_a.htm (June 5, 2005) (Google hires students to manually review search results for quality purposes).

[14] *See* Search King Inc. v. Google Tech., Inc., No. CIV-02-1457-M, at 4 (W.D. Okla. Jan. 13, 2003) ("Google knowingly and intentionally decreased the PageRanks assigned to both SearchKing and PRAN."). This manual adjustment has also been alleged in the recent KinderStart lawsuit. *See* Complaint, KinderStart.com LLC v. Google, Inc., Case No. C 06-2057 (N.D. Cal. Mar. 17, 2006).

[15] *See MSN Blockades phpBB Searchers*, trimMail's Email Battles, Jan. 18, 2006, http://www.emailbattles.com/archive/battles/vuln_aacgfbgdcb_jd/ (some search engines blocked

Conclusion. Search engines have some duality in their self-perceptions, and this duality creates a lot of confusion. Search engines perceive themselves as objective and neutral because they let automated technology do most of the hard work. However, in practice, search engines make editorial judgments just like any other media company. Principally, these editorial judgments are instantiated in the parameters set for the automated operations, but search engines also make individualized judgments about what data to collect and how to present it. These manual interventions may be the exception and not the rule, but these exceptions only reinforce that search engines play an active role in shaping their users experiences when necessary to accomplish their editorial goals.

II.   SEARCH ENGINE EDITORIAL CHOICES CREATE BIASES

Search results ordering has a significant effect on searchers and web publishers. Searchers usually consider only the top few search results; the top-ranked search result gets a high percentage of searcher clicks, and clickthrough rates quickly decline from there.[16] Therefore, even if a search engine delivers

---

certain search terms containing the keyword "phpBB"). Other examples of search engines manually adjusting algorithmically-generated results:

- In response to the search term "Jew," Google displays a special result in the sponsored link, saying "Offensive Search Results: We're disturbed about these results as well. Please read our note here." The link leads to a page explaining the results (see http://www.google.com/explanation.html).

- Amazon's book search functionality offered the prompt "did you mean adoption?" in response to searches for the keyword "abortion." When pointed out to Amazon, it manually changed the database to disable that prompt. *See* Laurie J. Flynn, *Amazon Says Technology, Not Ideology, Skewed Results*, N.Y. TIMES, Mar. 20, 2006.

- Google removed some websites from its index in response to a 512(c)(3) take-down demand from the Church of Scientology. However, Google displayed the following legend at the bottom of affected search results pages (such as search results for "scientology site:xenu.net"): "In response to a complaint we received under the US Digital Millennium Copyright Act, we have removed 2 result(s) from this page. If you wish, you may read the DMCA complaint that caused the removal(s) at ChillingEffects.org." *See* http://www.google.com/search?sourceid=navclient&ie=UTF-8&rls=GGLD,GGLD:2005-09,GGLD:en&q=scientology+site%3Axenu%2Enet   (last visited Nov. 25, 2005); *see generally* Chris Sherman, *Google Makes Scientology Infringement Demand Public*, Search Engine Watch, Apr. 15, 2002, http://searchenginewatch.com/searchday/article.php/2159691.

[16] *See* Jakob Nielsen, *The Power of Defaults*, Jakob Nielsen's Alertbox, Sept. 26, 2005, http://www.useit.com/alertbox/defaults.html (citing a study by Cornell professor Thorsten Joachims that the first search result gets 42% of clicks and the second search result gets 8%; further, when

hundreds or even thousands of search results in response to a searcher's query, searchers effectively ignore the vast majority of those search results. Accordingly, web publishers desperately want to be listed among the top few search results.[17]

For search engines, results placement determines how the searcher perceives the search experience. If the top few search results do not satisfy the searcher's objectives, the searcher may deem the search a failure. Therefore, to maximize searcher perceptions of search success, search engines generally tune their ranking algorithms to support majority interests.[18] In turn, minority interests (and the websites catering to them) often receive marginal exposure in search results.

To gauge majority interests, search engines frequently include a popularity metric in their ranking algorithm. Google's popularity metric, PageRank, treats inbound links to a website as popularity votes, but votes do not count equally; links from more popular websites count more than links from lesser-known websites.[19]

Beyond promoting search results designed to satisfy majority interests, PageRank's non-egalitarian voting structure causes search results to be biased towards websites with economic power[20] because these websites get lots of links due to their marketing expenditures and general prominence.

---

the first two search results are switched, the first search result gets 34%—meaning that positioning dictated searcher behavior); Nico Brooks, *The Atlas Rank Report: How Search Engine Rank Impacts Traffic*, Atlas Institute Digital Marketing Insights, June 2004, http://www.atlasdmt.com/media/pdfs/insights/RankReport.pdf (the first ranked search result may get ten times the quantity of clicks as the tenth ranked search result).

[17] *See* Michael Totty & Mylene Mangalindan, *Web Sites Try Everything To Climb Google Rankings*, WALL ST. J., Feb. 26, 2003.

[18] *See* Lucas D. Introna & Helen Nissenbaum, *Shaping the Web: Why the Politics of Search Engines Matters*, INFO. SOC'Y, July-Sept. 2000, at 169.

[19] *See Our Search: Google Technology*, Google.com, http://www.google.com/technology/ (last visited Mar. 10, 2006).

[20] *See* Nina Elkin-Koren, *Let the Crawlers Crawl: On Virtual Gatekeepers and the Right to Exclude Indexing*, 26 U. DAYTON L. REV. 179, 188 (2001); Frank Pasquale, *Rankings, Reductionism, and Responsibility*, Seton Hall Public Law Research Paper No. 888327, Feb. 25, 2006, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=888327, at 25; Trystan Upstill et al., *Predicting Fame and Fortune: PageRank or Indegree?*, Proceedings of the 8th Australasian Document Computing Symposium, Dec. 15, 2003, http://research.microsoft.com/users/nickcr/pubs/upstill_adcs03.pdf (showing that BusinessWeek Top Brand, Fortune 500 and Fortune Most Admired companies get disproportionately high PageRank).

Indeed, popularity-based ranking algorithms may reinforce and perpetuate existing power structures.[21] Websites that are part of the current power elite get better search result placement, which leads to greater consideration of their messages and views. Furthermore, the increased exposure attributable to better placement means that these websites are likely to get more votes in the future, leading to a self-reinforcing process.[22] In contrast, minority-interest and disenfranchised websites have a difficult time cracking through the popularity contest, potentially leaving them perpetually relegated to the search results hinterlands.[23]

A number of commentators have lamented these effects and offered some proposals in response:

Improve Search Engine Transparency. Search engines keep their ranking algorithms secret.[24] This secrecy hinders search engine spammers from gaining more prominence than search engines want them to have, but the secrecy also prevents searchers and commentators from accurately assessing any bias. To enlighten searchers, search engines could be required to disclose more about their practices and their algorithms.[25] This additional information has two putative benefits. First, it may improve market mechanisms by helping searchers choose among search engine competitors. Second, it may help searchers determine the appropriate level of cognitive authority to assign to their search results.

Publicly Fund Search Engines. Arguably, search engines have "public good"-like attributes, such as reducing the social costs of search behavior. If so, private actors will not incorporate these social benefits into their decision-

---

[21] *See* Introna & Nissenbaum, *supra* note 18.

[22] *See Egalitarian Engines*, ECONOMIST, Nov. 17, 2005 ("there is a widespread belief among computer, social and political scientists that search engines create a vicious circle that amplifies the dominance of established and already popular websites"); *see also* Junghoo Cho & Sourashis Roy, *Impact of Search Engines on Page Popularity*, WWW 2004, May 2004, http://oak.cs.ucla.edu/~cho/papers/cho-bias.pdf; Upstill, *supra* note 20. *But see* Santo Fortunato et al., *The Egalitarian Effect of Search Engines*, WWW2006, Nov. 2005, http://arxiv.org/pdf/cs.CY/0511005 (questioning this "rich-gets-richer" effect).

[23] *See* Cho & Roy, *supra* note 22.

[24] *See* Search King Inc. v. Google Tech., Inc., No. CIV-02-1457-M, at 3 n.2 (W.D. Okla. Jan. 13, 2003) ("Google's mathematical algorithm is a trade secret, and it has been characterized by the company as 'one of Google's most valuable assets.'"); Stefanie Olsen, *Project Searches for Open-Source Niche*, CNET News.com, Aug. 18, 2003, http://news.com.com/2102-1032_3-5064913.html?tag=st_util_print.

[25] *See* Introna & Nissenbaum, *supra* note 18.

making. In that case, public funding of search engines may be required to produce socially optimal search results.[26] Indeed, there have been several proposals to create government-funded search engines.[27]

Mandate Changes to Ranking/Sorting Practices. Search engines could be forced to increase the exposure of otherwise-marginalized websites. For example, Pandey et al. advocate a "randomized rank promotion" scheme where obscure websites randomly should get extra credit in ranking algorithms, appearing higher in the search results on occasion and getting additional exposure to searchers accordingly.[28] As another example, Pasquale proposes that, when people think the search engines are providing false or misleading information, search engines should be forced to include a link to corrective information.[29]

III.   SEARCH ENGINE BIAS IS NECESSARY AND DESIRABLE

Before trying to solve the problem of search engine bias, we should be clear that search engine bias actually is a problem requiring correcting. From my perspective, search engine bias is the unavoidable consequence of search engines exercising editorial control over their databases. Like any other media company, search engines simply cannot passively and neutrally redistribute third party content (in this case, web publisher content). If a search engine does not attempt to organize web content, its system quickly and inevitably will be overtaken by

---

[26] *See id.*; Eszter Hargittai, *Open Portals or Closed Gates? Channeling Content on the World Wide Web*, 27 POETICS 233 (2000); *cf.* CASS SUNSTEIN, REPUBLIC.COM 170-72 (2001) (advocating publicly funded "deliberative domains").

[27] *See* Kevin J. O'Brien, *Europeans Weigh Plan on Google Challenge*, INT'L HERALD TRIB., Jan. 18, 2006 (discussing a European initiative called Quaero, which is intended to break the American hegemony implicit in Google's dominant market position); Graeme Wearden, *Japan May Create Its Own Search Engine*, CNET News.com, Dec. 21, 2005, http://news.com.com/Japan+may+create+its+own+search+engine/2100-1025_3-6004037.html.

[28] *See* Sandeep Pandey et al., *Shuffling a Stacked Deck: the Case for Partially Randomized Ranking of Search Engine Results*, http://www.cs.cmu.edu/~olston/publications/randomRanking.pdf; *cf.* SUNSTEIN, *supra* note 26 (websites should be forced to link to contrary views as a way of increasing exposure to alternative viewpoints).

[29] *See* Pasquale, *supra* note 20, at 28-30 (proposing that the link be displayed as an asterisk to the search results).

spammers, fraudsters and malcontents.[30] At that point, the search engine become worthless to searchers.

To prevent anarchy and preserve credibility, search engines unavoidably must exercise some editorial control over their systems. In turn, this editorial control will create some bias.

Fortunately, market forces limit the scope of search engine bias. Searchers have high expectations for search engines: they expect search engines to read their minds[31] and infer their intent based solely on a small number of search keywords.[32] Search engines that disappoint (either by failing to deliver relevant results, or by burying relevant results under too many unhelpful results) are accountable to fickle searchers.[33] There are multiple search engines available to searchers,[34] and few barriers to switching between them.[35]

---

[30] Every Internet venue accepting user-submitted content inevitably gets attacked by unwanted content. If left untended, the venue inexorably degrades into anarchy. *See*, *e.g.*, Alorie Gilbert, *Google Fixes Glitch That Unleashed Flood of Porn*, CNET News.com, Nov. 23, 2005, http://news.com.com/2102-1025_3-5969799.html?tag=st.util.print (describing how Google Base, a venue for user-submitted content, was overtaken by pornographers: "the amount of adult content on Google Base was staggering considering Google only launched the tool a week ago"); Josh Quittner, *The War Between alt.tasteless and rec.pets.cats*, WIRED, May 1994, at 46 (describing how a group of anarchists took over a USENET newsgroup about pets for fun).

[31] *See Our Philosophy*, Google.com, http://www.google.com/corporate/tenthings.html (last visited Mar. 26, 2006) ("The perfect search engine…would understand exactly what you mean and give back exactly what you want."); Chris Sherman, *If Search Engines Could Read Your Mind*, Search Engine Watch, May 11, 2005, http://searchenginewatch.com/searchday/article.php/3503931.

[32] Searchers routinely use a very small number of keywords to express their search interests. *See* iProspect.com, Inc*., iProspect Natural SEO Keyword Length Study*, Nov. 2004, http://www.iprospect.com/premiumPDFs/keyword_length_study.pdf (88% of search engine referrals are based on only one or two keywords); *see also* Declan Butler, *Souped-Up Search Engines*, NATURE, May 11, 2000, at 112, 115 (citing an NEC Research Institute study showing that up to 70% of searchers use only a single keyword as a search term); Bernard J. Jansen et al., *Real Life Information Retrieval: A Study of User Queries on the Web*, 32 SIGIR FORUM 5, 15 (1998) (average keyword length was 2.35 words; one-third of searches used one keyword and 80% used three keywords or fewer); Jakob Nielsen, *Search: Visible and Simple*, May 13, 2001, http://www.useit.com/alertbox/20010513.html (average keyword length was 2.0 words).

[33] *See* Kim Peterson, *MSN Learns to Crawl*, SEATTLE TIMES, May 2, 2005 (MSN Search "learned that the arcane searches were the make-or-break moments for Web searchers. People weren't just happy when a search engine could find answers to their most bizarre, obscure and difficult queries. They would switch loyalties.").

[34] In addition to the recent launch of major new search engines by providers like MSN, the open-source software community is developing Nutch to allow anyone to build and customize his or her

As a result, searchers will shop around if they do not get the results they want,[36] and this competitive pressure constrains search engine bias. If a search engine's bias degrades the relevancy of search results, searchers will explore alternatives even if searchers do not realize that the results are biased. Meanwhile, search engine proliferation means that niche search engines can segment the market and cater to underserved minority interests.[37] Admittedly, these market forces are incomplete—searchers may never consider what results they are not seeing—but they are powerful nonetheless.

In contrast, it is hard to imagine how regulatory intervention will improve the situation. First, regulatory solutions become a vehicle for normative views about what searchers should see—or should *want* to see.[38] How should we select among these normative views? What makes one bias better than the other?

Second, regulatory intervention that promotes some search results over others does not ensure that searchers will find the promoted search results useful. Instead, government regulation rarely can do better than market forces at delivering results that searchers find relevant, so searchers likely will find some of the promoted results irrelevant.

---

own web search engine. *See* http://lucene.apache.org/nutch/ (last visited Nov. 25, 2005); *see also* Olsen, *Open-Source Niche*, *supra* note 24.

[35] *See* Rahul Telang et al., *An Empirical Analysis of Internet Search Engine Choice*, Aug. 2002. For example, search engines use the same basic interface (a white search box), and searchers rarely use advanced search features that might require additional learning time at other search engines.

[36] *See* Press Release, Vividence, Inc., *Google Wins Users' Hearts, But Not Their Ad Clicks* (May 25, 2004), http://www.vividence.com/public/company/news+and+events/press+releases/2004-05-25+ce+rankings+search.htm (up to 47% of searchers try another search engine when their search expectations are not met).

[37] *See* Rahul Telang et al., *The Market Structure for Internet Search Engines*, 21 J. MGMT. INFO. SYS. 137 (2004), *available at* http://www.heinz.cmu.edu/~rtelang/engine_jmis_final.pdf (describing how searchers sample heterogeneous ranking algorithms, which supports a diversity of search engines); Mário J. Silva, *The Case for a Portuguese Web Search Engine*, http://xldb.fc.ul.pt/data/Publications_attach/tumba-icwi2003-final.pdf (describing the value of a Portuguese-oriented search engine); *cf.* Jakob Nielsen, *Diversity is Power for Specialized Sites*, Jakob Nielsen's Alertbox, June 16, 2003, http://www.useit.com/alertbox/20030616.html (describing how specialized sites will flourish on the Internet).

[38] For a good example, *see* Susan L. Gerhart, *Do Web Search Engines Suppress Controversy?*, First Monday, Jan. 2004, http://www.firstmonday.org/issues/issue9_1/gerhart/. Gerhart argues that search engines do not adequately prioritize search results that expose controversies about the search topic. However, her argument assumes that controversy-related information has value to consumers, an assumption that deserves careful evaluation.

The clutter of unhelpful result may hinder searchers' ability to satisfy their search objectives, undermining searchers' confidence in search engines' mind-reading abilities.[39]   In this case, regulatory intervention could counterproductively degrade search engines' value to searchers. Whatever the adverse consequences of search engine bias, the consequences of regulatory correction are probably worse.[40]

IV.   TECHNOLOGICAL EVOLUTION WILL MOOT SEARCH ENGINE BIAS

Currently, search engines use "one-size-fits-all" ranking algorithms to deliver homogeneous search results to searchers with heterogeneous search objectives.[41] One-size-fits-all algorithms exacerbate the consequences of search engine bias in two ways: (1) it creates winners (websites listed high in the search results) and losers (those with marginal placement), and (2) it delivers suboptimal results for searchers with minority interests.[42]

These consequences will abate when search engines migrate away from one-size-fits-all algorithms towards "personalized" ranking algorithms.[43] Personalized algorithms produce search results that are custom-tailored to the searcher's interests, so searchers will see different results in response to the same search query. For example, Google offers searchers an option that "orders your search results based on your past searches, as well as the search results and news headlines you've clicked on."[44]

---

[39] *See* Eric Goldman, *A Coasean Analysis of Marketing* (forthcoming 2006).

[40] *Cf.* Susan P. Crawford, *Shortness of Vision: Regulatory Ambition in the Digital Age*, 74 FORDHAM L. REV. 695 (2005) (discussing the shortcomings of regulatory intervention in organic information systems).

[41] *See* James Pitkow et al., *Personalized Search*, COMM. ACM, Sept. 2002, at 50, 51.

[42] *See* Michael Kanellos, *Microsoft Aims for Search on Its Own Terms*, CNET News.com, Nov. 24, 2003, http://news.com.com/2102-1008_3-5110910.html?tag=st.util.print (quoting a Microsoft researcher as saying "If the two of us type a query [into a search engine], we get the same thing back, and that is just brain dead. There is no way an intelligent human being would tell us the same thing about the same topic."); Personalization of Placed Content Ordering in Search Results, U.S. Patent App. 0050240580 (filed July 13, 2004).

[43] *See* Pitkow, *supra* note 41, at 50.

[44]   *What's   Personalized   Search?*,   Google.com, http://www.google.com/support/bin/answer.py?answer=26651&topic=1593 (last visited Nov. 25, 2005).

Personalized ranking algorithms represent the next major advance in search relevancy. One-size-fits-all ranking algorithms have inherent limits on their maximum relevancy potential, and further improvements in one-size-fits algorithms will yield progressively smaller relevancy benefits. Personalized algorithms transcend those limits, optimizing relevancy for each searcher and thus implicitly doing a better job of searcher mind-reading.[45]

Personalized ranking algorithms also reduce the effects of search engine bias. Personalized algorithms mean that there are multiple "top" search results for a particular search term instead of a single "winner,"[46] so web publishers will not compete against each other in a zero-sum game. Also, personalized algorithms necessarily will diminish the weight given to popularity-based metrics (to give more weight for searcher-specific factors), reducing the structural biases due to popularity. Personalized ranking algorithms are not a panacea—any process where humans select and weight algorithmic factors will produce some bias[47]— but personalized algorithms will eliminate many of the current concerns about search engine bias.

V.   CONCLUSION

Complaints about search engine bias implicitly reflect some disappointed expectations. In theory, search engines can transcend the deficiencies of predecessor media. In practice, search engines are just like every other medium—heavily reliant on editorial control and susceptible to human biases. This fact shatters any illusions of search engine utopianism.

Fortunately, search engine bias may be largely temporal. In this respect, I see strong parallels between search engine bias and the late 1990s keyword metatag "problem."[48] Web publishers used keyword metatags to distort search

---

[45] *See* Jaime Teevan et al., *Personalizing Search via Automated Analysis of Interests and Activities*, SIGIR '05, http://haystack.lcs.mit.edu/papers/teevan.sigir05.pdf; Terry McCarthy, *On the Frontier of Search*, TIME, Aug. 28, 2005 ("Search will ultimately be as good as having 1,000 human experts who know your tastes scanning billions of documents within a split second.") (quoting Gary Flake, Microsoft Distinguished Engineer).

[46] *See* Kevin Lee, *Search Personalization and PPC Search Marketing*, Clickz News, July 15, 2005, http://www.clickz.com/experts/search/strat/print.php/3519876.

[47] Personalized algorithms have other potentially adverse consequences, such as creating self-reinforcing information flows. *See* SUNSTEIN, *supra* note 26. For a critique of these consequences, *see* Goldman, *Coasean Analysis*, *supra* note 39.

[48] *See generally* Goldman, *Deregulating Relevancy*, *supra* note 7.

results, but these techniques worked only so long as search engines considered keyword metatags in their ranking algorithms. When search engines recognized the distortive effects of keyword metatags, they changed their algorithms to ignore keyword metatags.[49]   Search result relevancy improved, and the problem was solved without regulatory intervention.

Similarly, search engines naturally will continue to evolve their ranking algorithms and improve search result relevancy—a process that, organically, will cause the most problematic aspects of search engine bias to largely disappear. To avoid undercutting search engines' quest for relevance, this effort should proceed without regulatory distortion.

---

[49] *See* Danny Sullivan, *Death of a Meta Tag*, Search Engine Watch, Oct. 1, 2002, http://www.searchenginewatch.com/sereport/print.php/34721_2165061.