

2006

## **Rankings, Reductionism, and Responsibility**

Frank Pasquale

Follow this and additional works at: <https://brooklynworks.brooklaw.edu/faculty>



Part of the [Intellectual Property Law Commons](#), and the [Internet Law Commons](#)

---

# RANKINGS, REDUCTIONISM, AND RESPONSIBILITY

FRANK PASQUALE\*

I. THE BASICS OF RANKING.....	117
A. <i>Keeping Material From Being Highly Ranked</i> .....	119
B. <i>Getting Material Highly Ranked</i> .....	123
II. TWIN GOALS FOR SEARCH LAW: RESPONSIBLE AND AUTHORITATIVE METADATA .....	125
III. ENCOURAGING DOMINANT SEARCH ENGINES TO BE RESPONSIBLE .....	128
A. <i>Negative Externalities of Ranking</i> .....	130
B. <i>Reducing Negative Externalities of Ranking</i> .....	135
C. <i>Addressing Harm from Inclusion in Highly-Ranked Results</i> .....	135
D. <i>Addressing Harm from Exclusion from Results</i> .....	137
VI. CONCLUSION.....	139

An invitation to engage with the work of Margaret Jane Radin is a privilege. Scholars of intellectual property (IP) have particularly benefited as she has applied her insights on law, philosophy, and property to the novel problems in this field. In her contribution to this symposium, Professor Radin has proposed that those making intellectual property policy and interpreting IP laws pay closer attention to the legal “neighborhood” of these rights—particularly, antitrust laws governing *competition* and free speech laws related to *expression*. Radin’s provocative and constructive work suggests that courts would establish more reasonable and balanced IP doctrines if they took into account the policy behind competition and free speech law.

For the most part, I agree with this perspective on the ameliorative potential of competition law. Antitrust can play an important role in tailoring or limiting IP protections in many concentrated IP industries, where monopolists or oligopolists consistently try to leverage control over one product into control over all means of

---

\*Associate Professor of Law, Seton Hall Law School. I would like to thank David Barnhizer, the Cleveland State Law Review, and the organizers of the symposium for inviting my contribution. I would also like to thank the organizers of the Seton Hall Faculty Retreat for permitting me to present my works in progress on search engines there. I greatly appreciated the extensive comments offered by Gaia Bernstein, Marina Lao, James Grimmelmann, R. Erik Lillquist, Eric Goldman and Thomas Healy, and wish I had more time to address them in this piece. Thanks to Charles Sullivan and Richard Murphy for their incisive and constructive conversations on the topic. And I am grateful to research assistants Mohammed Azeez and Dean Murray for their extraordinary helpfulness and diligence.

selling, distributing, or even referring to it.<sup>1</sup> The competition “neighborhood” promises to bring some balance and perspective to IP law, just as environmental and bankruptcy law symbiotically influence much CERCLA litigation, or ERISA and state insurance regulation lead to a nuanced (if complex) regime of pension and benefit law.<sup>2</sup>

Yet I am not as confident as Radin of the salutary influence of the “free speech” neighborhood. We can certainly applaud *Eldred* for acknowledging that First Amendment principles require fair use and the idea/expression dichotomy in copyright law.<sup>3</sup> Important trademark “fair use” and “nominative use” cases rely even more directly on First Amendment law, and as Radin suggests, these principles should influence the evolving doctrine of “initial interest confusion.”<sup>4</sup> Yet as Owen Fiss has observed, First Amendment “absolutism” can often be as formalistic as the Lochneresque propertization movement now dominating copyright policy.<sup>5</sup> Rights to express oneself, regardless of the consequences, are not only tools for those marginalized by the current propertization trend, but can also be exploited by those at its core.

Given the brevity of a symposium contribution, I cannot explore in detail how IP and First Amendment law mutually reinforce one another at the expense of competition policy. However, I can comment on a suggestive recent case. In *SearchKing v. Google*, a federal district court dismissed various unfair competition complaints on the grounds that search engine results are protected under the First Amendment.<sup>6</sup> In *SearchKing*, the court simultaneously called the *automation* of search results an important reason not to hold search engines accountable for them,

---

<sup>1</sup>See, e.g., Michael A. Carrier, *Resolving The Patent-Antitrust Paradox Through Tripartite Innovation*, 56 VAND. L. REV. 1047 (2003); Matthew Fagin, Frank Pasquale, and Kim Weatherall, *Beyond Napster: Using Antitrust to Advance and Enhance Online Music Distribution*, 8 B.U. J. SCI. & TECH. L. 451 (2002).

<sup>2</sup>See, e.g., Joel Gross, *The Effect of Bankruptcy on Obligations to Clean Up Contaminated Properties: Recent Developments and Open Issues Two Decades After Kovacs and Midlantic* (2003), [http://www.arnoldporter.com/pubs/files/Effect\\_Bankruptcy\\_Contaminated\\_Properties.pdf](http://www.arnoldporter.com/pubs/files/Effect_Bankruptcy_Contaminated_Properties.pdf).

<sup>3</sup>*Eldred v. Ashcroft*, 538 U.S. 186 (2003) (characterizing as “the Framers’ view” “that copyright’s limited monopolies are compatible with free speech principles” and claiming that the “‘fair use’ defense . . . allows the public to use not only facts and ideas contained in a copyrighted work, but also expression itself for limited purposes . . . [and] thereby affords considerable latitude for scholarship and comment.”).

<sup>4</sup>See David W. Barnes & Teresa A. Laky, *Classic Fair Use in Trademark: Confusion about Defenses*, 20 SANTA CLARA COMPUTER & HIGH TECH L.J. 833 (2004); Eric Goldman, *Deregulating Relevancy in Internet Trademark Law*, 54 EMORY L.J. 507 (2005).

<sup>5</sup>See, e.g., OWEN FISS, *THE IRONY OF FREE SPEECH* (1996) (concentrating on First Amendment victories for those opposing limits on campaign expenditures, hate speech, and discriminatory associations); Julie Cohen, *Lochner in Cyberspace: The New Economic Orthodoxy of “Rights Management,”* 97 MICH. L. REV. 462 (1998) (critiquing formalism of the new Lochnerism).

<sup>6</sup>Dahlia Lithwick, *Google-opoloy, The Game Only Google Can Play*, SLATE, Jan. 29, 2003, <http://www.slate.com/id/2D77875> (discussing the case and its implications for search policy).

and yet reached out to immunize Google's acknowledged *human intervention* to bury results referring to a potential competitor. Admittedly, the plaintiff in the case was not terribly sympathetic, and the court may just have been trying to avoid the chaotic consequences of exposing search engines to all manner of state laws governing business torts. Nevertheless, one need not reject the court's conclusion that a business tort occurred here in order to agree that *some* accountability for search engine results is increasingly necessary as they become the primary portal for net users.

After discussing how search engines operate in Part I below, and setting forth a normative basis for regulation of their results in Part II, this piece proposes (in Part III) some minor, non-intrusive legal remedies for those who claim that they are harmed by search engine results. Such harms include unwanted high-ranking results relating to them, or exclusion from a page they claim it is their "due" to appear on. In the first case (deemed "inclusion harm"), I propose a right not to suppress the results, but merely to add an asterisk to the hyperlink directing web users to them, which would lead to the complainant's own comment on the objectionable result.<sup>7</sup> In the latter case (deemed "exclusion harm"), complainants should have some right to a limited explanation of why they did not appear in highly ranked results. Both these rights are based on the Fair Credit Reporting Act's guarantee to consumers that they both get to correct and comment on negative information in their credit reports, and that they receive some account of why their credit reports might have led to a denial of credit.

Given extraordinary advances in the annotation software of "wiki's," automation itself should make these basic prerogatives relatively easy to implement. But even if these particular proposals are deemed implausible, they do focus attention on matters of principle that will have increasing importance in coming years: the copyrightability and First Amendment status of search engine rankings and other "machine speech" resulting from computerized algorithms. Given the rapidly growing importance of rankers and other aggregators of information, law should not lightly permit machine expression to garner these protections. Rather, they are merited to the extent that rankers are *responsible*, reflecting actual human judgment and providing due process to those harmed by inclusion or exclusion in relevant results.

### I. THE BASICS OF RANKING

Search engines work by gathering a great deal of information and then providing responses to search queries.<sup>8</sup> Once a search engine has indexed a set of web pages

---

<sup>7</sup>For now, I would limit this remedy to searches a) of a person's name; and b) of a trademarked good. For example, if a search for "American Blinds" yielded links to several competitors of the owner of the mark American Blinds, the owner of the mark could use the asterisk to indicate that such rivals were not authorized sellers of its products. Trying to expand rights of control beyond these narrow categories might prove unworkable given the number of potential claimants to an asterisk. For more on the parallels between names and trademarks, see Greg Lastowka, *The Trademark Function of Authorship*, 85 B.U. L. REV. 1171 (2005).

<sup>8</sup>See Frank Pasquale, *The Law and Economics of Information Overload Externalities*, 56 VAND. L. REV. (forthcoming 2007), available at <http://papers.ssrn.com/sol3/papers.cfm?>

and other information resources, it can generate responses to searchers' queries.<sup>9</sup> For example, if I were to type in "plastic Christmas trees" to the Yahoo! search engine, I would receive a page with several pieces of information.<sup>10</sup> First, the most prominently displayed materials are two "paid listings" from major retailers who sell such trees. Paid listings border the page, and are permitted a variety of attention-getting design strategies. Below them appear the search results: several web pages commenting on Christmas trees, and some from sellers of such trees. The top right hand corner indicates that this initial page is displaying merely ten of the 3,540,000 potentially relevant pages. Theoretically, an obsessive searcher could review all of them (ten sites at a time) merely by clicking the next button on the bottom right hand corner of the page 354,000 times; in practice, the number of results initially "claimed" tends to decline as one clicks toward lower-ranked results.

As that fanciful idea suggests, actual human inspection plays a very small role in search engine's sorting and ranking of webpages. Rather, software does nearly all the work.<sup>11</sup> Although they have long tried to keep their ranking algorithms a "black box" to preserve competitive advantage, search engines have begun to reveal the basics of how their searches work.<sup>12</sup> The basic strategy is to index pages on two axes: 1) relevance to the query; and 2) overall importance. Step one, relevance, compares the searcher's query with the text on the webpage, the "metatags" embedded in the coding of the page,<sup>13</sup> and other aspects of the page.<sup>14</sup> Step two, importance, relies on a number of heuristics, the most important being the number of other pages that link to the page, the number of pages that link to the linking pages, and so on, recursively.<sup>15</sup>

---

abstract\_id=888410 (discussing this process); Niva Elkin-Koren, *Let the Crawlers Crawl: On Virtual Gatekeepers and the Right to Exclude Indexing*, 26 DAYTON L. REV., 180 (2001).

<sup>9</sup>For a taxonomy of the types of searches now prevalent, see LOUIS ROSENFELD & PETER MORVILLE, *INFORMATION ARCHITECTURE FOR THE WORLD WIDE WEB* § 6.2 (1998).

<sup>10</sup>See <http://www.yahoo.com> (results from typing "plastic Christmas trees" in the Search box and clicking "Web Search.").

<sup>11</sup>Google claims that "There is no human involvement or manipulation of results, which is why users have come to trust Google as a source of objective information untainted by paid placement," available at [http://www.reedusdesign.com/google\\_pagerank.htm](http://www.reedusdesign.com/google_pagerank.htm). But see *Google, Inc. v. Am. Blinds & Wallpaper Factory, Inc.*, No. C 03-05340 JF, 2005 WL 832398 (N.D. Cal. Mar. 30, 2005); *Search King, Inc. v. Google Tech., Inc.*, No. CIV-02-1457-M, 2003 WL 21464568 (W.D. Okla. May 27, 2003). Eric Goldman also discusses human "quality assurance" personnel at Google. See Goldman, *Google's Human Algorithm*, TECH. & MARKETING BLOG, June 5, 2005, available at [http://blog.ericgoldman.org/archives/2005/06/googles\\_human\\_a.htm](http://blog.ericgoldman.org/archives/2005/06/googles_human_a.htm) (last visited May 15, 2006).

<sup>12</sup>See, e.g., Google, *Technology Overview*, <http://www.google.com/corporate/tech.html> (last visited Mar. 31, 2006) [hereinafter *Technology Overview*].

<sup>13</sup>To view this source code, just right-click on any webpage and then click on "View Source." This should reveal the underlying structure of and metadata in the page.

<sup>14</sup>There can be many other aspects; "Google's order of results is automatically determined by more than 100 factors, including our PageRank algorithm." Google, *Google Information for Webmasters*, <http://www.google.com/webmasters/4.html> (last visited Mar. 29, 2006).

<sup>15</sup>*Technology Overview*, *supra* note 12 ("Google uses PageRank™ to examine the entire link structure of the web and determine which pages are most important. It then conducts

Over the past decade, the main legal controversies sparked by web rankings have centered on trademark law. Mark owners have sued search engines for permitting competitors to appear as highly ranked results in response to a query based on the mark owner's name.<sup>16</sup> A growing amount of secondary literature has offered many interesting perspectives on these trademark disputes.<sup>17</sup> However, it has not adequately anticipated the next generation of "ranking" controversies that will likely arise around two contrasting demands: 1) by those seeking privacy (who want search engines to remove links to sensitive, misleading, or false information); and 2) by those seeking publicity (who want to become more highly ranked in response to relevant queries). Currently the law does little to hold search engines accountable on either score.

#### *A. Keeping Material From Being Highly Ranked*

The "privacy" or "reputational" challenges to keep materials off search results come from individuals, corporations, and even countries.<sup>18</sup> Some are amusing--one New York Times reporter recently recounted her unsuccessful effort to keep a webpage featuring an unflattering photograph of herself from appearing as a high-ranked web result when her name was queried on Google.<sup>19</sup> More troubling was

---

hypertext-matching analysis to determine which pages are relevant to the specific search being conducted. By combining overall importance and query-specific relevance, Google is able to put the most relevant and reliable results first."); *see also* PageRank patent U.S. Patent No. 6,285,999 (filed Jan.9, 1998) (discussing Google's patented method of ranking pages, but not revealing company trade secrets that also contribute to the algorithm).

<sup>16</sup>For example, Starbucks might object if a search query for "Starbucks" ends up with rival coffee shops as highly ranked results, especially if they come up as paid results. A good discussion of current disputes in this area can be found in Susan Kuchinskas, *Google Takes Hit in U.S. Trademark Case*, <http://www.internetnews.com/bus-news/article.php/3494406> (last visited May 15, 2006). Important current cases include *Government Employees Insurance Company v. Google, Inc.*, 330 F.Supp.2d 700, (E.D.Va. 2004) (granting Google's motion to dismiss trademark suit based on Google's sale of "GEICO" as an adword to GEICO's competitors in the car insurance business); *American Blind & Wallpaper Factory, Inc.*, 2005 WL 832398 (refusing to dismiss key claims in American Blind's suit to enjoin the sale of "American Blind" as an "adword" to competitors of American Blind).

<sup>17</sup>The leading article here is probably Goldman, *supra* note 4 (proposing common law and statutory safe harbors for search providers accused of trademark infringement). I generally agree with Goldman's approach, although the "asterisk" proposal in Part V below would likely give trademark holders more rights than his proposal would envision.

<sup>18</sup>Privacy might seem an odd term here, given the wide array of concerns expressed by those who want to keep certain sites off highly ranked result pages. I am using the term in the general sense of the "power to share information discriminately." Helen Nissenbaum, *Privacy as Contextual Integrity*, 79 WASH. L. REV. 119, 121 (2004) (citing James Rachels, *Why Privacy is Important*, in *PHILOSOPHICAL DIMENSIONS OF PRIVACY: AN ANTHOLOGY* 290, 294 (Ferdinand David Schoeman ed., 1984)). Though we have not traditionally thought of groups or corporations as privacy "rightsholders," the examples given in this section should suggest both normatively attractive and suspect examples of situations where they might reasonably be considered such.

<sup>19</sup>Stephanie Rosenbloom, *Loosing Google's Lock on the Past*, N.Y. TIMES, June 2, 2005, G1. ("In the winter of 1996, back when I was a brunette who wore sensible shoes, a photographer snapped my picture during a rehearsal for a college musical. . . . The resulting

Yahoo's failure to hide a webpage that featured the pictures, name, phone number, and address of a woman whose ex-boyfriend had posted them in order to spite and humiliate her.<sup>20</sup> Corporations have also tried to control their presentation on the web, trying to keep critical websites from appearing atop a websearch including their company names.

Search engines' general policy is to refuse to intervene in situations like this.<sup>21</sup> For example, in response to the query "Can you remove my information from Google's search results?," Google answers, "We'd like to assist you, but information in our search results is actually located on third-party publicly available webpages. In order to remove your information from our search results, you'll need to contact the webmaster of this third-party site."<sup>22</sup> In its "terms of service," Google gives a slightly lengthier rationale for its refusal to intervene:

Google's indices consist of information that has been identified, indexed and compiled through an automated process with no advance review by human beings. Given the enormous volume of web site information added, deleted, and changed on a frequent basis, Google cannot and does not screen anything made available through its indices. For each web site reflected in Google's indices, if either (i) a site owner restricts access to his or her web site or (ii) a site is taken down from the web, then, upon receipt of a request by the site owner or a third party in the second

---

portrait showed a pasty, gaunt girl being swallowed by a XXX-large T-shirt. The only thing more unfortunate than the photo is that nearly a decade after it was taken—a decade in which I became a blonde and graduated to stilettos—it is still the definitive image of me on the World Wide Web, the one that pops up every time my name is entered in a Google search.”). Rosenbloom's story was sparked by what is known as a “vanity search”: the entry of one's own name into Google (usually in quotation marks) in order to see how many, and which, links it brings up.

<sup>20</sup>An Oregon woman “filed suit against Yahoo for \$3 million, charging that the company had not removed . . . photographs of her from the Web, as it had promised it would. [The plaintiff's] suit claims that an ex-boyfriend posted the pictures, her e-mail address, and her work phone number without permission.” *Id.*

<sup>21</sup>*See, e.g.,* Ted Frank, *Mark Maughan Sues Google Over Search Results*, OVERLAWYERED, Mar. 28, 2004, [https://www.Overlawyered.com/2004/03/mark\\_maughan\\_sues\\_google\\_over.html](https://www.Overlawyered.com/2004/03/mark_maughan_sues_google_over.html) (“CPA Mark Maughan is upset that a 2003 Google search for his name leads to a web page that, he says, accuses him incorrectly of wrongdoing. So he's suing Google, AOL, Time Warner, and Yahoo.”). Commenting on the case, John Battelle observes: “[I]n the minds of others, you are what the index says you are.” JOHN BATTELLE, *THE SEARCH: HOW GOOGLE AND ITS RIVALS REWROTE THE RULES OF BUSINESS AND TRANSFORMED OUR CULTURE* 191 (2005). In this way, the “digital person” that Dan Solove describes becomes accessible not merely to credit bureaus, the government, and other data collectors, but to the general public. *See* DAN SOLOVE, *THE DIGITAL PERSON: TECHNOLOGY AND PRIVACY IN THE INFORMATION AGE* (2004).

<sup>22</sup>*See* Google Help Center, *Can You Remove My Information From Google's Search Results?*, <http://www.google.com/support/bin/answer.py?answer=9111&topic=360> (last visited Mar. 26, 2006).

instance, Google would consider on a case-by-case basis requests to remove the link to that site from its indices.<sup>23</sup>

Given the immensity of the World Wide Web, it is easy to sympathize with Google's plight. It is certainly in no position to review individually all new content put on the web. That is one reason why the DMCA extended cybertort liability originally designed for internet service providers like Verizon to search engines like Google and Yahoo.<sup>24</sup>

Nevertheless, the pleas of impotence here are a bit disingenuous. When confronted by important enough entities, Google does intervene in search results. For example, due to a number of anti-Semites' efforts to manipulate search rankings (a process known colloquially as "Google-bombing")<sup>25</sup>, a Holocaust-denial site routinely appeared in the top ten results for the query "Jew."<sup>26</sup> In response to a number of complaints from the Anti-Defamation League,<sup>27</sup> Google added a headline titled, "An explanation of our search results" to the top of the page.<sup>28</sup> The linked

---

<sup>23</sup> Google Privacy Center, *Google Terms of Use for Your Personal Use*, [http://www.google.com/terms\\_of\\_service.html](http://www.google.com/terms_of_service.html)

<sup>24</sup> 17 U.S.C. § 512(k)(1)(B) (2000) (defining "service providers" as applied to DMCA safe harbor categories). These protections cover not only internet service providers like Verizon or Comcast, but also search engines like Google and Yahoo.

<sup>25</sup> Google bombing refers to "a technique through which a group of bloggers working together can make a webpage come up when someone searches Google for certain keywords." Microcontent News, Glossary, "Google bombing," <http://www.microcontentnews.com/resources/glossary/googlebomb.htm> (last visited Mar. 30, 2006). Google bombing is often considered a type of "spamdexing." "[Spamdexing is defined as those techniques] employed by some Web marketers and site designers in order to fool a search engine's spider and indexing programs. The objective is to ensure that their Web site always appears at or near the top of the list of search engine results." PCMag.com, Encyclopedia, "Spamdexing," [http://www.pcmag.com/encyclopedia\\_term/0,2542,t=spamdexing&i=51796,00.asp](http://www.pcmag.com/encyclopedia_term/0,2542,t=spamdexing&i=51796,00.asp) (last visited Mar. 26, 2006); see also Ira S. Nathanson, *Internet Infoglut and Invisible Ink: Spamdexing Search Engines with Metatags*, 12 HARV. J.L. & TECH. 43 (1998).

<sup>26</sup> The site, <http://www.jewwatch.com/>, is an obvious hate site.

<sup>27</sup> See Anti-Defamation League, *Google Search Ranking of Hate Sites Not Intentional*, [http://www.adl.org/rumors/google\\_search\\_rumors.asp](http://www.adl.org/rumors/google_search_rumors.asp) ("The ranking of Jewwatch and other hate sites is in no way due to a conscious choice by Google, but solely is a result of this automated system of ranking. . . . The longevity of ownership, the way articles are posted to it, the links to and from the site, and the structure of the site itself all increase the ranking of [the anti-Semitic site] within the Google formula.").

<sup>28</sup> Google, *An Explanation of Our Search Results*, <http://www.google.com/explanation.html> (last visited Mar. 31, 2006) [hereinafter *Explanation*] (explaining that "If you recently used Google to search for the word 'Jew,' you may have seen results that were very disturbing. We assure you that the views expressed by the sites in your results are not in any way endorsed by Google. We'd like to explain why you're seeing these results when you conduct this search."). But see Goldman, *supra* note 4, at 533 (discussing the many ways in which search engines make what would traditionally be deemed editorial decisions; for example, they "make editorial choices about which websites to include in their database[, removing] websites based on a publisher's overzealous efforts to game the search engine's



webpage explains the reasons why the anti-Semitic site appeared so high in the relevant ranking and distances Google from the result.<sup>29</sup>

The proper degree of toleration for the intolerant in a democratic society has vexed liberal political philosophy<sup>30</sup> and American constitutional law<sup>31</sup> for some time. It is not the place of this paper to assess whether Google's response in this particular controversy was adequate. However, the very fact that there was a response at all raises some interesting questions. Do search engines have an obligation to assure that material appearing on the web fairly, accurately, or otherwise constructively relates to the query at issue? Are certain types or sites or particularly sensitive information *by their very nature* undeserving of the type of publicity high-ranked results provide? Consider, for instance, a hapless man whose name generates a highly-ranked result linked to a page on the site "Don't Date Him, Girl," (which permits disgruntled ex-wives and ex-girlfriends to post in great detail the shortcomings of their former husbands or boyfriends).<sup>32</sup> If it is true, or "merely" an opinion, he can't sue the poster for defamation.<sup>33</sup> But should this be the end of the story—his web presence forever compromised by a virtual "scarlet letter" attached to his name?<sup>34</sup>

---

system or due to liability concerns. In other cases, search engines may choose not to index certain websites based on the costs and benefits of catering to esoteric interests.").

<sup>29</sup>Explanation, *supra* note 28.

<sup>30</sup>See JOHN RAWLS, POLITICAL LIBERALISM xiii (1993).

<sup>31</sup>See, e.g., DONALD ALEXANDER DOWNS, NAZIS IN SKOKIE: FREEDOM, COMMUNITY AND THE FIRST AMENDMENT (1985); PHILIPPA STRUM, WHEN THE NAZIS CAME TO SKOKIE: FREEDOM FOR SPEECH WE HATE (1998).

<sup>32</sup>See The Bitch, *Refuse, Rebuff, Reject, Repel, Repulse*, MIAMI NEW TIMES, Sept. 22, 2005 ("Tasha Joseph, a furiously scorned Miami-based publicist . . . decided to take action. 'You can do a background check for an employee. Why can't you do one for a potential boyfriend?' she says. Joseph launched the Website dontdatehimgirl.com this past summer. Since July, more than a hundred community-minded women have posted profiles and photos of men caught cheating . . . . Joseph says her site has been receiving about 2500 hits per day."). Dan Solove has written extensively, and persuasively, about the dangers of untrammelled rights to publish even true information. See, e.g., Daniel J. Solove, *The Virtues of Knowing Less: Justifying Privacy Protections Against Disclosure*, 53 DUKE L.J. 967 (2003).

<sup>33</sup> Dontdatehimgirl.com is not a mere hobbyist's entertainment, but rather aspires to be the "universal dating background check tool." Kimberly Palmer, *Should Cheaters be Outted Online?*, CHICAGO TRIBUNE REDEYE, Oct. 20, 2005, at 2 ("Joseph, who at 33 is in what she calls a 'committed' relationship, envisions the site as an empowering tool for women. If every cheating guy is eventually posted, it will become the universal dating background check tool, she says. And she's not worried about getting sued, mostly because she had a team of lawyers help her set up the site and create a terms of use form that women have to agree to before posting cheaters. 'We put the onus on women who use the site to make sure she's telling us the truth. . . . We have no liability,' she says. Plus, she adds, 'Truth is the absolute defense to libel.'").

<sup>34</sup> Law seems to take an interest in preventing such embarrassment by banning blackmail—which, as James Boyle notes, is a rather odd exception to information economics' general endorsement of contractual control over information. See James Boyle, *A Theory of Law and Information: Copyright, Spleens, Blackmail, and Insider Trading*, 80 CAL. L. REV. 1413 (1992). (discussing several leading scholars' effort to resolve the "paradox" of

Search engines like to advise individuals and companies who don't like the current highly-ranked results associated with them to engage in self-help—by trying to persuade the offending poster to take down or modify their site, by creating their own highly ranked information, and by purchasing “adwords” to assure highly ranked “paid results” above and to the side of unpaid results.<sup>35</sup> For example, if Wal-Mart is deeply unhappy that two critical sites come up in the top five responses to the query “Walmart,” it could eventually create a group of well-connected “supportive” sites that eventually displace “Walmart Watch” and the “Walmart Movie.” It could also purchase several relevant “adwords” (such as “sex discrimination” or “low wages”) to assure that such sites appear in “paid results” listings near results listings generated by queries for critical sites.<sup>36</sup> Search engines give a number of tips on how those concerned with search results can attempt to secure better placement for their sites<sup>37</sup>--and, in turn, displace unflattering material or contextualize it with purchased links.<sup>38</sup>

### B. Getting Material Highly Ranked

So far, search engines' *laissez-faire* approach to complaints about ranking results may appear well-founded. They claim to be merely directories of websites and not to generate content themselves.<sup>39</sup> Those dissatisfied with results can put up their own rival content or pay for an advertisement connected to key queries. Given these opportunities, search engines claim that they are just one more forum in the “marketplace of ideas” which give the disgruntled more than adequate opportunity to respond to or reconfigure search results they dislike.

---

blackmail law); see also *Cosby v. Jackson*, 741 F. Supp. 740 (N.D. Ill. 1970). Dontdatehimgirl.com does offer some measure of self-help, a chance for the accused to present “his side of the story.” Nevertheless, the “asterisk” solution proposed in Part V below provides a more systematic approach.

<sup>35</sup>See, e.g., Google, *Google Can Help Your Business Make More Money*, <http://services.google.com/marketing/links/US-HA-CMBNINE2> (last visited Apr. 1, 2006).

<sup>36</sup>For an account of adword auctions, see DAVID VISE & MARK MALSEED, *THE GOOGLE STORY* 116-18 (2005) (“Companies of all sizes were participating in these keyword auctions, spending anywhere from hundreds of dollars to many millions each quarter on Google.”).

<sup>37</sup>See, e.g., Google, *Google Information for Webmasters: Webmaster Guidelines*, <http://www.google.com/intl/en/webmasters/guidelines.html> (last visited Apr. 1, 2006).

<sup>38</sup>See Rosenbloom, *supra* note 19, at G1 (“‘The trick . . . is if you can get lots of people that have a great PageRank to link to you, you're going to be driven up very high.’ [Therefore], the secret to burying unflattering Web details about yourself is to create a preferred version of the facts on a home page or a blog of your own, then devise a strategy to get high-ranking Web sites to link to you. Many people assume that a Google ranking has something to do with Web traffic, but that is incorrect, as is the notion that the more links a site has, the higher its PageRank. A PageRank can be high even if a Web site is linked to only once, Ms. Mayer of Google said, provided that the one link is itself a highly ranked page.”).

<sup>39</sup>Goldman provides a number of reasons to dispute this claim. See Goldman, *supra* note 4, at 532-33 (“[I]t is 100% wrong to treat search engines as passive agents for publisher content. Search engines are media companies, not neutral providers of information. Thus, search engines often view their search database as a proprietary asset that they tightly control to provide a quality user experience and, not coincidentally, maximize profits.”).

However, efforts to displace unfavorable information may ultimately increase its visibility, as “important” websites may be drawn to comment on the efforts themselves.<sup>40</sup> Search engines do not want website publishers to get *too* adept at manipulating the rankings, and routinely punish those who they say unfairly influence their results. Although Google gives webmasters a number of general suggestions on how to increase the rank of their websites, they warn against hiring certain unsavory “Search Engine Optimizers” (“SEO’s”) companies which attempt to increase one’s rank in response to relevant (and sometimes not-so-relevant) queries.<sup>41</sup> As litigation recently revealed, at least one search engine “buried” the website of a Search Engine Optimizer under dozens of other listings in order to punish the company for manipulating page rankings.<sup>42</sup> After becoming aware of the SEO SearchKing’s strategy to increase its clients’ ranking, “Google knowingly and intentionally decreased the PageRanks assigned to . . . SearchKing,” greatly diminishing potential clients’ ability to find SearchKing via a Google search.<sup>43</sup>

At first glance, this seems like perfectly reasonable behavior—the search engine’s unpaid rankings are both a public service and the foundation of its commercial enterprises, and it has a right to maintain its integrity. Moreover, the unpaid rankings are only commercially feasible given the paid rankings, and not many website owners would pay for rankings if they could easily assure top placement via other means. Yet search engines’ secrecy regarding the way they rank pages makes the whole idea of “fair” or “unfair” ranking manipulation exceptionally vague.<sup>44</sup> If site owners have only a vague idea of what counts as a legitimate factor in webpage ranking, they have little idea of whether their tactics are a brilliant maneuver that will lead to cyberfame or a censured ploy that will consign them to anonymity.<sup>45</sup>

---

<sup>40</sup>For example, when a California accountant sued search engines for reporting contested professional disciplinary action against him as the first result attached to his name, the popular website Overlawyered increased the prominence of his disgrace by mocking his litigiousness. See BATTELLE, *supra* note 21, at 191. It is possible that other self-help measures would be similarly self-defeating, enhancing the notoriety of the very site they are designed to occlude.

<sup>41</sup>Google, *Google Information For Webmasters: Search Engine Optimizers*, <http://www.google.com/intl/en/webmasters/guidelines.html> (last visited Apr. 1, 2006) (“Many SEOs provide useful services for website owners, from writing copy to giving advice on site architecture and helping to find relevant directories to which a site can be submitted. However, a few unethical SEOs have given the industry a black eye through their overly aggressive marketing efforts and their attempts to unfairly manipulate search engine results.”).

<sup>42</sup>*Search King, Inc.*, 2003 WL 21464568, at \*2. After offending Google with its tactics for elevating its clients rankings, Search King would not show up on a search for “Search King” until many pages after the initial results page—though it obviously appeared to be the most relevant response to such a search. *Id.*

<sup>43</sup>*Id.*

<sup>44</sup>Google’s mathematical algorithm is a trade secret, and it has been characterized as one of its most valuable assets. See Frederick Townes, *SEO [Search Engine Optimization] vs. Marketing: The Fine Art of Copywriting*, SEARCH GUILD, available at <http://www.searchguild.com/article772.html>.

<sup>45</sup>For example, if I pay an “important” site to link to my site, in order to enhance its ranking, I have little idea whether this approach will be rewarded or condemned.

Despite these paradoxes, the *SearchKing* controversy ultimately gave search engines little to worry about. Following earlier decisions that deemed bond rating agencies' assessment of firms' creditworthiness protected speech, the court found that Google's page rankings were an opinion protected under the First Amendment.<sup>46</sup> Google, and by implication other search engines, had no obligation to assure that their published results were actually generated by the "objective," unmanipulated algorithm that they claim produces them.<sup>47</sup>

There are several reasons to object to such blanket protection for search results. The court may call rankings opinions, but the world does not treat them as such; rather, the more dominant a search engine is, the more its ranking is treated as (and becomes) a *fact* about the relevance, quality, and prominence of the ranked. Given that the purported "objectivity" of rankings and lack of human intervention is the main reason why search engines refuse to review or remove links to the material mentioned above,<sup>48</sup> the legal protection of their capacity to alter results at will, and secretly, is curious at best. Search engines serve valuable social functions, but indiscriminate application of First Amendment immunities to their results provides unnecessarily expansive protections.<sup>49</sup>

## II. TWIN GOALS FOR SEARCH LAW: RESPONSIBLE AND AUTHORITATIVE METADATA

As book publishers, news agencies, and others launch copyright lawsuits against Google, information law appears eminently capable of scuttling search engines' quest for an authoritative index of materials. However, outside of a few trademark disputes, plaintiffs appear incapable of supporting minimal demands for responsible publication of search results. This is unacceptable in an age of data proliferation, as *metadata* (the organization and classification of data) becomes an ever more

---

<sup>46</sup>*Search King, Inc.*, 2003 WL 21464568, at \*3 ("Two questions remain. First, are PageRanks constitutionally protected opinions? Second, if PageRanks fall within the scope of protection afforded by the First Amendment, is the publication of PageRanks *per se* lawful under Oklahoma law, thereby precluding tort liability premised on the intentional and even malicious manipulation of PageRanks by Google? The Court answers both questions in the affirmative.").

<sup>47</sup>*Id.* at \*4 ("[T]he Court finds that under Oklahoma law, protected speech-in this case, PageRanks-cannot give rise to a claim for tortious interference with contractual relations because it cannot be considered wrongful, even if the speech is motivated by hatred or ill will.").

<sup>48</sup>As Google avers in response to queries about objectionable rankings, "The order and contents of Google search results are completely automated. No one hand picks a particular result for a given search query . . . ." Google Help Center, *Principles*, <http://www.google.com/support/bin/topic.py?topic=368> (last visited Apr. 1, 2006) [hereinafter *Principles*]; see also Stepforth, *The SEO Blog: Are Google SERPs Entirely Organic?*, <http://news.stepforth.com/blog/2006/02/are-google-serps-entirely-organic.php> (last visited Apr. 2, 2006) [hereinafter *SEO Blog*].

<sup>49</sup>See also Michael L. Rustad & Thomas H. Koenig, *Rebooting Cybertort Law*, 80 WASH. L. REV. 335 (criticizing CDA immunities for ISP's (such as search engines) as overbroad); Eugene Volokh, *Crime-Facilitating Speech*, 57 STAN. L. REV. 1095 (2005) (cataloging limits to First Amendment protection for speech that induces or enables certain crimes).

important resource for navigating oceans of expression.<sup>50</sup> Some balance is needed before copyright law turns the Internet into a pay-per-use anticommons<sup>51</sup> and misguided First Amendment concerns render search engines entirely unaccountable for the results they provide. In an increasingly fragmented, strident, and sensationalistic media environment, there is ever more need for *authoritative and responsible providers of metadata*. Each of these aspirational terms is described in a bit more detail below.

*Authoritativeness:* Librarians, archivists, and collectors have always dreamed of a comprehensive source of data—one that includes all relevant material in a single index. With the advance of digitization and interconnection, there is no *technical* obstacle to such a “celestial jukebox,” “new library of Alexandria,” or consolidated collection of all types of expression.<sup>52</sup> As technology governed by Moore’s Law advances, storage and search costs continue to decline.<sup>53</sup> However, *legal* and *business* obstacles to authoritativeness appear to arise as quickly as technical barriers come down. Certainly some of these projects, like digital rights management<sup>54</sup> and the development of commercial “deep webs”<sup>55</sup> available only to those licensees, may be necessary to secure compensation to copyright holders and other entrepreneurs.<sup>56</sup>

---

<sup>50</sup>As David Weinberger comments on the Google Print controversy, “[D]espite the present focus on who owns the digitized content of books, the more critical battle for readers will be over how we manage the information about that content—information that’s known technically as metadata.” David Weinberger, *Crunching the Metadata: What Google Print Really Tells Us About the Future of Books*, BOSTON GLOBE, Nov. 13, 2005, at E3.

<sup>51</sup>An anti-commons develops when fragmented ownership causes high transaction costs that stunt the development of a resource. Michael A. Heller, *The Tragedy of the Anticommons: Property in the Transition from Marx to Markets*, 111 HARV. L. REV. 621 (1998). Just as Garrett Hardin observed that insufficient propertization could lead to overuse of resources, anticommons theorists show how excessive propertization can lead to underuse of resources. *Id.* at 624 n.10.

<sup>52</sup>See PAUL GOLDSTEIN, COPYRIGHT’S HIGHWAY: FROM GUTENBERG TO THE CELESTIAL JUKEBOX (2003); Charles C. Mann, *The Heavenly Jukebox*, ATLANTIC, Sep. 2000, available at <http://www.theatlantic.com/issues/2000/09/mann.htm>; Raymond Shih Ray Ku, *The Creative Destruction of Copyright: Napster and the New Economics of Digital Technology*, 69 U. CHI. L. REV. 263 (2002).

<sup>53</sup>Raymond Shih Ray Ku, *The Creative Destruction of Copyright: Napster and the New Economics of Digital Technology*, 69 U. CHI. L. REV. 263, 270-71 (2002).

<sup>54</sup>Digital Rights Management (DRM) is defined as a “system for authorizing the viewing or playback of copyrighted material on a user’s computer or digital music player. DRM has centered around copyrighted music, with Apple’s FairPlay and Microsoft’s Windows Digital Rights Manager being the two predominant DRM systems. Video DRM is on the horizon as broadband Internet and more highly compressed video formats take hold.” The Free Dictionary, “DRM,” <http://computing-dictionary.thefreedictionary.com/DRM>.

<sup>55</sup>“The ‘Deep Web’ refers to content available online that is not indexed by search engines. This content may be inaccessible to search robots because the content is presented on dynamically generated pages, which the robots cannot access.” Goldman, *supra* note 4, at 533 n.76 (citing Michael K. Bergman, *The Deep Web: Surfacing Hidden Value*, 7 J. ELECTRONIC PUB. 1, Aug. 2001, available at <http://www.press.umich.edu/jep/07-01/bergman.html>).

<sup>56</sup>Admittedly, some predict that untrammelled “deep webs” will ultimately make general purpose search engines’ business model untenable. See Jack Shafer, *The Great Google*

But the mere indexing and archiving of works—the core of the Google Print project—has little if any negative commercial impact on information creators. Holdouts should not be permitted to stop this project in the same way that permission culture has crippled innovation in the music and film industries.<sup>57</sup>

*Responsibility:* Nevertheless, as Socrates reminds us, the best physician is also often the best poisoner.<sup>58</sup> As search engines become more authoritative, encompassing more and more sources of data, they are also likely to become more important sources of information in our daily lives. If an individual has some rights to control personally identifiable information in the transfer of bank, medical, and video store records, shouldn't similar protections apply to search results? Moreover, in a world where dates, employers, and even casual acquaintances "google" the individuals they meet and companies they do business with, shouldn't the "googled" have some opportunity to present their side of the story in response to potentially misleading, biased, or merely cruel links? Finally, if search engines can evade responsibility in the foregoing scenarios by pointing to the automation of their services, shouldn't they also have some legal obligation to actually follow the algorithms they claim to adhere to?<sup>59</sup> Internet policymakers should address each of these concerns in coming years; I propose some small steps toward accountability in Part V below.

---

*Wipeout: Chronicle of a Corporate Death Foretold*, SLATE, Nov. 21, 2005, <http://www.slate.com/id/2130795> (predicting that a growing proprietary "deep web" kept out of search engine indexes may eventually render search engine results trivial as the most important and relevant sites opt out of their indices). To the extent the "deep webs" accomplish this isolation via technical measures, there may be little law can do to stop their development.

<sup>57</sup>Examples are legion; several participants at the Harvard *Signal or Noise* conference described spending dozens of hours attempting to hunt down the owner of tiny snippets of music, only to be confronted with near-extortionate demands for a license. *Signal or Noise? The Future of Music on the Net*, Briefing Book, available at [http://cyber.law.harvard.edu/events/netmusic\\_brbook.html](http://cyber.law.harvard.edu/events/netmusic_brbook.html). The owners of copyrights in the television show *The Simpsons* demanded \$10,000 for a ten-second clip that a documentary filmmaker wanted to insert into his work. See LAWRENCE LESSIG, *FREE CULTURE* 96 (2004). See also Center for Social Media, Copyright and Fair Use, [http://www.centerforsocialmedia.org/resources/fair\\_use/](http://www.centerforsocialmedia.org/resources/fair_use/) (last visited May 15, 2006) (discussing controversies); KEMBREW MCLEOD, *FREEDOM OF EXPRESSION@: OVERZEALOUS COPYRIGHT BOZOS AND OTHER ENEMIES OF CREATIVITY* (2005).

<sup>58</sup>"As Socrates points out in Book I of the Republic, the outstanding doctor, by virtue of his technical knowledge of medicine and the human body, also makes the best poisoner." Peter Berkowitz, *Other People's Mothers*, NEW REPUBLIC, Jan. 10, 2000, at 27, available at <http://www.peterberkowitz.com/otherpeoplesmothers.htm> (reviewing PETER SINGER, *PRACTICAL ETHICS* (1993)).

<sup>59</sup>The FTC has warned sites to disclose which are paid and which are unpaid rankings. A similar proviso regarding "humanly manipulated" ranks may be in order. Danny Sullivan at Search Engine Watch has also proposed that search engines should at least indicate that a government-censored site may be relevant, even if it is legally required not to show the site. Danny Sullivan, *Got To Censor Search Listings? Why Not Disclose?*, Search Engine Watch Blog, Sep. 27, 2004, <http://blog.searchenginewatch.com/blog/040927-093434>. Google apparently already does this in response to European restrictions on content (such as body modification sites), but it is not clear if Google.cn will follow such a policy in China. *Id.* Alan Davidson, Presentation at Yale Search Engine Conference, Dec. 5, 2005, available at <http://islandia.law.yale.edu/isp/regulatingsearch.html> (last visited May 15, 2006).

Authority and responsibility can be mutually reinforcing. As the costs of comprehensiveness decline, the costs of entry into the search market decline. As more search services emerge, the stakes of inclusion or exclusion in any particular directory decline. Lower stakes make responsibility an easier task, as the process due is should be proportional to the harm inflicted by any given inclusion in or exclusion from search engine results.<sup>60</sup>

As Helen Nissenbaum and Lucas Introna argued in their seminal piece on the development of search engines, these information providers are not merely one among many voices on the World Wide Web.<sup>61</sup> They organize and rank all the rest, acting as a “public good” as indispensable as phone books, traffic signals, or lighthouses.<sup>62</sup> We all gain when authoritative, comprehensive search engines can refer us to *all* the webpages, books, and other media relating to the topics that interest us. Yet we all stand to lose if increasingly authoritative search engines do not act responsibly when publishing and disseminating the information they gather.

Unfortunately, information law appears to aid the development of partial and irresponsible search engines. If things are to change, the negative consequences of absolute rights granted pursuant to copyright and the First Amendment must be quantified, or at least subject to economic analysis. I examine the copyright side of the equation in another piece, proposing special fair use protections for categorizers and indexers.<sup>63</sup> Here, I concentrate on the untoward consequences of irresponsible search engine results in a concentrated search engine market and propose some legal responses.

### III. ENCOURAGING DOMINANT SEARCH ENGINES TO BE RESPONSIBLE

As Google, Yahoo, MSN, and other leading search engines become extraordinarily powerful research tools, society’s interest in *comprehensive* collections of information is advanced. But is there a downside to such authoritative search engines? To the degree that the search engine market is not merely authoritative, but also *concentrated*, the answer is yes. All those harmed mentioned in Part II above—ranging from those unable to keep misleading or embarrassing information from being associated with given queries, to those who felt they deserved more prominent placement—would find their troubles exacerbated to the extent that any given search engine, or small group of search engines, became

---

<sup>60</sup>American due process jurisprudence has generally scaled the amount of “process due” to the gravity of deprivation caused by government action. See *Mathews v. Eldridge*, 424 U.S. 319 (1976).

<sup>61</sup>Lucas D. Introna & Helen Nissenbaum, *Shaping the Web: Why the Politics of Search Engines Matters*, 16 INFO. SOC’Y 3 (2000), available at <http://www.nyu.edu/projects/nissenbaum/papers/searchengines.pdf>.

<sup>62</sup>See Carol Rose, *Comedy of the Commons: Custom, Commerce and Inherently Public Property*, 53 U. CHI. L. REV. 711 (1986) (discussing the treatment of libraries and lighthouses).

<sup>63</sup>See Pasquale, *supra* note 8. Focusing on the “authoritativeness” goal, that piece argues that, in an era of information overload, comprehensive collections and categorizations of information are becoming ever more important to our economic and psychological well-being. *Id.* Building on my earlier work on the fourth fair use factor, I argue for favorable treatment for categorizers and organizers of information under the first of the statutory fair use factors.

authoritative information sources. Moreover, there are other negative externalities arising from dominant rankings systems generally, which Section IV(A). below addresses.

Individuals' limited capacity (or willingness) to process information is the main reason why high-ranking search engine results are so important. Though literally thousands or millions of results can appear in response to a query, only about ten to fifteen can appear on the first page. Of these, the first unpaid result is likely to get ten times the traffic as the tenth, and twice that of the second. The resulting competition has created various strategies to influence rankings, which in turn drive search engines to make their ranking algorithms more opaque. Rankings can also generate self-fulfilling prophecies, whereby the top-ranked site may become the most popular and successful one, regardless of its merits.<sup>64</sup>

Though some might argue that strong copyright protection would ameliorate such problems, by reducing the power of search engines, in fact the reverse is true: policies that cheapen the data necessary for search engine *comprehensiveness* also advance the ultimate solution to negative externalities from a concentrated search market: namely, a *diversity* of search engines as ranking resources. In a world where search engines have to bargain with every rightsholder, only the wealthiest companies could maintain the legal staff necessary for such an endeavor—which would probably be Sisyphean to boot.<sup>65</sup> If search engines eventually gain fair use protections for indexing and providing samples of content, far more entrants could compete in this market.

Nevertheless, we cannot expect such a competitive market to arise organically, even if search engines manage to succeed in copyright litigation. The economics of search predict that, in any given market for information, there are enormous returns to scale—individuals are going to seek out the most comprehensive and authoritative source of information, and the very fact that this occurs gives the leading source enormous leverage to assure that information sources will want to appear (and be highly ranked) on its search results.<sup>66</sup> Growing numbers of searches on a given

---

<sup>64</sup>A commenter on John Battelle's blog discusses this in the context of searches for sites about the Civil War. "Thus, a person unschooled in search might start with a broad term, such as 'civil war'—because that's the way we used to search for things in books and catalogs—she already knows to think a bit before hitting send to identify much more specific search terms likely to approach unique for what she needs: e.g., 'bloodiest battle civil war.'" Tim Windsor, *Comment*, John Battelle's Search Blog, available at <http://battellemedia.com/archives/001863.php>

<sup>65</sup>See Gary Wolf, *The Great Library of Amazonia*, WIRED, Oct. 24, 2003 (describing Amazon's painstaking efforts to convince, coax, or coerce the owners of 100,000 book titles to participate in its "Look Inside The Book" program).

<sup>66</sup>See DAVID A. VISE & MARK MALSEED, *THE GOOGLE STORY* 215 (2005):

[T]he potent Googleware blend of software and hardware gave the company more computing power than anyone else. It would not be easy for a newcomer to match its raw scale, said Peter Norvig, director of search quality. "We're like Dell," he said, referring to the way Google assembled and customized each of the personal computers in its vast network. Moreover, prior to each product rollout, Brin and Page scrutinized the product's potential for rapid scaling up so that they could maintain a competitive advantage through size and distribution even as other mimicked and matched the ideas.



service give that service ever more data to refine and improve its index.<sup>67</sup> The “rich get richer,” making the search and rankings field a very difficult one to enter.

To the extent that search is a natural monopoly or oligopoly, government must try to assure that search engines are responsible for their results. The final section of this part tentatively suggests some potential reforms based on prior interventions in parallel information industries. Not only would such regulation check the worst aspects of dominant search engines, it might also tend to deconcentrate the market if it were only applied to “major” search engines.

#### *A. Negative Externalities of Ranking*

Creative *organization* of data (and concomitant generation of metadata) can counteract the negative effects of information overload. Unfortunately, ranking, a powerful and almost inevitable method of organizing data, has its own negative externalities. Unlike *rating*, which usually permits a number of rated entities to occupy some terrace of status, ranking creates a zero-sum game in which any one advance leads to other displacements. The resulting struggle is particularly acute on search engines, where the top-ranked site is likely to get twice the traffic of the second-ranked site and ten times the traffic of the tenth ranked site.<sup>68</sup>

The importance of ranking is regularly discussed by leading authors in the field of web-site promotion:

There is competition for those top ten seats. There is serious competition. People are trying to take away the top spots every day. They are always trying to fine tune and tweak their HTML code and learn the next little trick. The best players even know dirty ways to ‘bump off’ their competition while protecting their own sites.<sup>69</sup>

Economists have explored how positional dynamics in a number of different markets (including housing, test preparation, and automobiles) have led to socially wasteful “arms races” for positional advantage. In ordinary markets, the presence of high-spending consumers will draw more producers so that, eventually, supply will approach demand. However, there can only be one “top-ranked” site.<sup>70</sup> Tactics to

---

*Id.* For economic theory predicting such returns to scale in information industries, see HAL R. VARIAN & CARL SHAPIRO, *INFORMATION RULES: A STRATEGIC GUIDE TO THE NETWORK ECONOMY* (2000); Hal R. Varian, *Economics and Search*, Aug. 1999, available at <http://www.sims.berkeley.edu/~hal/Papers/sigir/sigir.html>.

<sup>67</sup>For example, if 100 people search for “skis” on a search engine on a Tuesday, and all pick the third-ranked result, the search algorithm may adjust and put that result as the first result the next day. The most used search engine will have more data to tweak its algorithms than its less-used rivals. We should then expect its algorithms to better reflect mass taste, and in turn to draw in more of the data that permits it to do so.

<sup>68</sup>See Jakob Nielsen, *The Power of Defaults*, USEIT.COM, Sept. 26, 2005, <http://www.useit.com/alertbox/defaults.html>; see also Nico Brooks, *The Atlas Rank Report: How Search EngineRank Impacts Traffic*, ATLAS INSTITUTE DIGITAL MARKETING INSIGHTS, June 2004, available at <http://app.atlasonepoint.com/pdf/AtlasRankReport.pdf>

<sup>69</sup>Introna & Nissenbaum, *supra* note 61, at 14.

<sup>70</sup>It is easy to see how spending could spiral out of control for such spots. There are only twenty universities ranked in the “Top 20” by *U.S. News & World Report*, regardless of the level of demand for slots at Top 20 schools. Even if we go from a society where 50%, instead

influence unpaid listings and prices for paid listings are sure to escalate, but it is not clear that this competition creates much utility.<sup>71</sup>

As Robert H. Frank and Cass Sunstein have observed, "In many contexts, consumers find themselves on a positional treadmill, in which their choices do not really make them happier or better off, but instead serve largely to keep them in the same spot in the hierarchy."<sup>72</sup> For example, if one applicant buys a designer suit before an interview, he may well look better than all the other applicants; but if *all* applicants do the same, no one gains any advantage, and each is out the cost of the suit.<sup>73</sup> As search engines become more important guides to products and services and advertising budgets continue to migrate toward them, we can expect more wasteful jockeying for "top spots" in response to search queries.<sup>74</sup>

---

of 5%, of high schoolers aspire to be in such a school, or where potential students have \$200,000, as opposed to \$20,000, in resources to apply to education, there can only be twenty Top 20 schools. Perhaps this is why top universities continue to fundraise while barely spending returns on extant endowments.

<sup>71</sup>Newspapers and other traditional venues for advertising also worry that, to the extent all (or the most desirable) consumers begin to use search engines as one-stop sources of information on goods and services, vital sources of revenue may be diverted from them. Given the extraordinary profitability of such media over the past twenty years, it is hard to take these objections seriously currently. However, it is important to note that even as search engines may cross-subsidize socially useful functions (such as the indexing of "research space") with revenues from advertising, they may be undercutting the position of other media to do so.

<sup>72</sup>Robert H. Frank, *The Demand for Unobservable and Other Nonpositional Goods*, 75 AM. ECON. REV. 101, 102-03 (1985) (asserting that workers will work under less optimal conditions to maintain relative position compared to coworkers).

<sup>73</sup>Robert H. Frank, H. J. Louis Professor of Econ., Johnson Graduate School of Mgmt., Cornell Univ., Address Before Australian Economics Association: Are Positional Externalities Different? (Oct. 28, 2003), *available at* <http://www.bu.edu/econ/seminars/micro/pdf/fall03/Frank.pdf>.

<sup>74</sup>*See* Robert H. Frank and Philip J. Cook, *The Winner-Take-All Society* 130 (1995) (discussing Richard Thaler, *The Winner's Curse: Paradoxes and Anomalies in Economic Life* (1992)). The winner's curse has been a widely acknowledged phenomenon among investors; as one popular business site notes, it is "A tendency for the winning bid in an auction to exceed the intrinsic value of the item purchased. Because of incomplete information, emotions or any other number of factors regarding the item being auctioned, bidders can have a difficult time determining the item's intrinsic value. As a result, the largest overestimation of an item's value ends up winning the auction." *See* Investopedia, "Winner's Curse," <http://www.investopedia.com/terms/w/winnerscurse.asp> (last visited May 19, 2006); *see also* Barry Schwartz, *The Paradox of Choice: Why More is Less* 194 (2004) (citing Thomas C. Schelling, *Micromotives and Macrobehavior* (1978)):

It's like being in a crowded football stadium, watching the crucial play. A spectator several rows in front stands up to get a better view, and a chain reaction follows. Soon everyone is standing, just to be able to see as well as before. Everyone is on their feet rather than sitting, but no one's position has improved. And if someone, unilaterally and resolutely, refuses to stand, he might just as well not be at the game at all. When people pursue goods that are positional, they can't help being in the rat race. To choose not to run is to lose.

Competition for rank also tends to lead to a destructive dynamic of *gaming* and *opacity*. The search engine optimizer (SEO) industry has long attempted to give advice to webmasters on how to increase their visibility. The line between legitimate and illegitimate tactics is blurry, and search engines are quick to alter their algorithms in order to thwart the latter and encourage the former. They consider these algorithms their most important asset, protecting them with a forbidding wall of trade secret, copyright, patent, and paracopyright protection. As algorithm transparency gives way to opacity, SEO's engage in ever more desperate tactics, leading to new rounds of surveillance and secrecy.<sup>75</sup>

Perhaps if we could be sure that all searches were running according to these algorithms, without human intervention, we could accept this opacity, assured that skilled "reverse engineers" could eventually make sense of the rating system.<sup>76</sup> However, self-serving intervention has been documented. Google has not merely punished "bad" search engine optimizers by making them "disappear" from relevant rankings, but it reportedly manipulated search results immediately before trademark litigation in order to increase its chances of winning.<sup>77</sup> This results in an untenable situation where searchers cannot fully understand the algorithm behind searches but also cannot be sure that algorithm is actually being used in any given case.<sup>78</sup>

Lack of transparency in search algorithms also exacerbates the problem of "self-fulfilling prophecies," which arise when a top-ranked site ends up being the most prestigious or profitable, not on account of its intrinsic qualities, but only because it was top-ranked to begin with. Few people are better acquainted with this phenomenon than law school faculty, whose schools are largely ranked according to "reputational" scores in *U.S. News & World Report* which are themselves an artifact of past rankings.<sup>79</sup> Nearly all the participants at a recent symposium on law school rankings criticized *U.S. News & World Report's* annual effort to ordinally rank law

---

<sup>75</sup>See Orin Kerr, *SSRN as a Measure of Scholarly Performance*, The Volokh Conspiracy, Aug. 16, 2005, available at [http://volokh.com/archives/archive\\_2005\\_08\\_14-2005\\_08\\_20.shtml#1124241868](http://volokh.com/archives/archive_2005_08_14-2005_08_20.shtml#1124241868) (discussing potential "gaming" of downloads by professors who download their own papers (or develop perl script to do so)).

<sup>76</sup>See Suzanne Scotchmer and Pamela Samuelson, *The Law and Economics of Reverse Engineering*, 111 YALE L.J. 1575 (2002).

<sup>77</sup>See *Search King, Inc.*, 2003 WL 21464568, at \*2; BATTELLE, *supra* note 21, at 183-86.

<sup>78</sup>Similar problems afflict the credit bureaus, whose leaders' methods are also protected by trade secret protection. Concerned about possibly racist elements of credit ratings, several states have banned the use of consumer credit scores in the setting of insurance premia.

<sup>79</sup>See Jeffery Evans Stake, *The Interplay Between Law School Rankings, Reputations, and Resource Allocation*, 81 IND. L.J. 229, 250 (2006) ("The schools ranked highly by U.S. News moved up in faculty reputation score and tended to stay there, and the schools that got low initial rankings moved down. In other words, the rankings were reflexive; there was an echo effect.") (citing Brian Leiter, Commentary, *How to Rank Law Schools*, 81 IND. L.J. 47, 51 (2006)). But see Richard Smalbeck, *The Durability of Law School Reputation*, 48 J. LEGAL EDUC. 568, 585 n.21, 586 (1998) (disputing the self-fulfilling prophecy characterization). Stake offers several reasons to doubt Schmalbeck's complacency about the self-fulfilling prophecy effect of the U.S. News rankings, including effects on lawyer reputation scores "significant at the 0.001 level." Stake, *supra*, at 253.

schools—including those usually enamored of untrammelled competitive struggle.<sup>80</sup> Educational rankings remain entirely unregulated, despite bitter complaints from some of the ranked and the growing sense that they have catalyzed a self-destructive arms race among schools. If any one search engine becomes as dominant in the general ranking field as *U.S. News* is in the law school field, similar dynamics may be unleashed in many other spheres.<sup>81</sup>

Unless search engines use their revenues to cross-subsidize socially useful projects, adword bidding wars may prove extraordinarily wasteful. To the extent that such information is used as a tool for competing in zero-sum games, it is unclear whether the production of more information is either efficient or equitable. For instance, imagine the development of an innovative method of moving up in search results, which sells for \$10,000. Though the number of users of such a method may be small, they will nevertheless manage to leverage financial advantage into notoriety. Moreover, even if the price of the method goes down, inequity persists *until it is universally accessible*. This inequity persists because the more accessible the method is, the greater the pressure to use it (spreading its cost more generally), and the greater the disadvantage suffered by those who don't use it (increasing the penalty for non-participation). But perversely, once the method becomes universally accessible, it will afford *no one* a positional advantage, because one can only rise in rank *relative to others*. Therefore, the innovative method either generates some inequity or inefficiency, since universal accessibility destroys the only value of the method—the ability to raise one's score relative to others.

This example suggests a more general theory regarding the development of primarily position-enhancing information (PPEI) like adwords and search engine gaming techniques. First, the development of such information immediately divides society into two groups—those who can afford to buy or use the information and those who cannot. Thus the innovation commodifies notoriety. Moreover, even

---

<sup>80</sup>See, e.g., Leiter, *supra* note 79; Richard A. Posner, *The Next Generation of Law School Rankings: Framing the Rankings Debate: Law School Rankings*, 81 IND. L.J. 13 (2006); Stake, *supra* note 79. But see Russell Korobkin, *In Praise of Law School Rankings: Solutions to Coordination and Collective Action Problems*, 77 TEX. L. REV. 403 (1998) (“Students read and value rankings because they know that attending a highly-ranked school signals their quality to desirable employers, who also study the rankings in order to interpret these signals. This coordination function is served whether or not the rankings accurately measure the quality of law schools, however defined.”).

<sup>81</sup>This is not merely a parochial concern of legal educators; recall the Pentagon's quick decision to scrap a prediction market in terrorism. Both rankings and prediction markets are dangerous as self-fulfilling prophecies. For example, in American elections, a candidate will often pull ahead of another candidate on a site called Tradesports, which gives “odds” on electoral and other competitions. This putative advantage can quickly become real, as investing donors (especially those seeking help from the government) tip to the more likely winner. For this reason, Roger Shattuck suggests that this is a type of information that is best kept private. ROGER SHATTUCK, *FORBIDDEN KNOWLEDGE* (1996). But see Cass Sunstein, *Group Judgments: Statistical Means, Deliberation, and Information Markets*, 80 N.Y.U. L. REV. 962, 1048 (2005) (arguing that “information markets [like Tradesports] . . . tend to correct, rather than to amplify, the effects of individual errors. . . . [and] create powerful incentives to disclose, rather than to conceal, privately held information.”); Tom Bell, *The Prediction Exchange: Progress in Promoting the Sciences and Useful Arts*, <http://www.tomwbell.com/writings/PredEx.pdf>.

though the group that can afford access to the PPEI is advantaged relative to those who cannot, the development of PPEI may be inefficient for that group as well. The benefit of distancing itself from those who cannot afford the information may well be outweighed by the cost of access. And precisely to the extent that this cost is lowered, the positional advantage afforded by the innovation dissipates. In other words, PPEI creates inequity when it first arises and inefficiency as it becomes more universally accessible.

Even if such commercial competition is deemed an inevitable accompaniment of markets, pervasive ranking also threatens to coarsen culture by encouraging a false sense of commensurability. Consider Richard Posner's recent effort to "rank" public intellectuals, a laborious project made ever easier by the powerful computing systems driving search engines.<sup>82</sup> There is no way to make a ranking process "fair" unless we subscribe to the bizarre idea that the quality of the work of economists, historians, philosophers, scientists, et al. can be ordinally ranked with some commensurating metric.<sup>83</sup> The danger is particularly acute in social sciences and literature, where quality is hard to measure and notoriety often becomes a substitute for it.<sup>84</sup> Only positivistic delusion could lead us to ignore the contestability of the basic assumptions and aspirations of authors in such fields--for example, how closely tied they are to visions of the good society and ideological commitments. Yet the temptation to assess the importance and even quality of a person's intellectual work with reference to the number of "hits" they generate on a search engine database is sure to advance as these services play a greater role in our lives.

---

<sup>82</sup>RICHARD POSNER, *PUBLIC INTELLECTUALS: A STUDY IN DECLINE* (2001). Posner ranked the "importance" of public intellectuals by counting the number of times they were quote in Lexis databases. See *id.* at 209-11, table 5.3 ("Top 100 public intellectuals by media mentions (1995-2000)"); *id.* at 212-14, table 5.4 ("Top 100 public intellectuals by scholarly citations (1995-2000)"). There is an excellent collection of critiques of Posner's project at *Reactions to Posner's Public Intellectuals, Complete Review*, COMPLETE REVIEW QUARTERLY, May, 2002, <http://www.complete-review.com/quarterly/vol3/issue2/posner.htm> (last visited May 25, 2006).

<sup>83</sup>For a normative defense of incommensurability, see MICHAEL WALZER, *SPHERES OF JUSTICE* (1978); Michael Walzer, *Response to Critics*, in *PLURALISM, JUSTICE, AND EQUALITY* 281 (David Miller, ed., 1995); Brian Leiter, *Incommensurability: Truth or Consequences?* 146 U. PENN. L. REV. 1723 (1998)

<sup>84</sup>See, e.g., LEO BRAUDY, *THE FRENZY OF RENOWN: FAME AND ITS HISTORY* (1997); JAMES F. ENGLISH, *THE ECONOMY OF PRESTIGE: PRIZES, AWARDS AND THE CIRCULATION OF CULTURAL VALUE* (2005) (describing the proliferation of prizes for expression); Tim Harford, *What Really Counts*, FIN. TIMES Jan. 28, 2006, at 12 ("Chicago-based economist David Galenson . . . recently demonstrated that Picasso was by far the greatest artist of the 20th century. Galenson's method is simplicity itself: round up every art history textbook of the past 15 years, and see whose art is reproduced most frequently. Picasso, with 395 illustrations in 33 textbooks, scores nearly as many as his three closest rivals (Matisse, Duchamp and Mondrian) put together."). Although such developments may portend the extension of some of the worst aspects of celebrity culture into academia, they also do create some positive pressures. For example, academics can no longer complacently permit their work to be locked into restrictive databases or books, for fear that their notoriety will decline as other, accessible work (approximating a good enough substitute for it) becomes a standard reference. See Dan Hunter, *Walled Gardens*, 62 WASH. & LEE L. REV. 607, 608-9 (discussing the advantages of open-access publishing).

### *B. Reducing Negative Externalities of Ranking*

None of the negative dynamics mentioned above is easy to measure. Once a self-fulfilling prophecy begins, it is nearly impossible to distinguish between legitimate and “echo chamber” aspects of a ranking. Moreover, there is no direct solution to the problem of arms races for position, or the destructive dynamic of gaming and opacity in ranking algorithms. Creeping misperceptions of commensurability may be more of a cultural than a legal problem. Nevertheless, all of these problems may be ameliorated if the stakes of any particular ranking diminish. This can probably only occur if more search engines enter a currently concentrated market, or if government intervenes to make extant search engines present their results more responsibly.

A diversity of search engines would make the results of any particular ranking far less consequential. It would also make the search field less vulnerable to gaming. Some search engines could continue to make their ranking algorithms opaque, while others might compete with an “open standard.” Search engines might concentrate on particular fields, developing algorithms that are particularly sensitive to the concerns of particular searchers or web publishers.

Unfortunately, in the rough and tumble world of online competition and consolidation, a “peaceable kingdom” of diverse search engines is unlikely to develop. This reality makes a baseline of accountability for dominant search engine results all the more important. The Federal Trade Commission (FTC) took an important first step in this direction by warning search engines to clearly distinguish between paid and unpaid search listings, lest the former constitute a form of false advertising.<sup>85</sup> The FTC could further advance fair competition in the search market by requiring large search engines to put in place basic procedural protections for those potentially harmed by query results.<sup>86</sup>

### *C. Addressing Harm from Inclusion in Highly-Ranked Results*

Just as major credit bureaus must respond to consumers’ allegations that a piece of information on their credit report is false or misleading, complaints about false or misleading search results on major search engines should lead to more than polite advice about self-help or a price list for adwords. When meritorious, such complaints should result in the right to give one’s own side of the story.<sup>87</sup> A small

---

<sup>85</sup>See Reuters, *Search Engine Providers Address FTC Complaints*, USA TODAY, July 2, 2002, available at <http://www.usatoday.com/tech/news/2002/07/02/ftc-search-engines.htm> (discussing the FTC response to Public Citizen complaint letter about Alta Vista).

<sup>86</sup>In future work, I hope to work out a comparative institutional analysis in order to develop some potential legal reforms to respond to these concerns. See, e.g., NEIL KOMESAR, *IMPERFECT ALTERNATIVES: CHOOSING INSTITUTIONS IN LAW, ECONOMICS, AND PUBLIC POLICY* 3-4 (1997) (advocating comparative institutional analysis before policy prescriptions).

<sup>87</sup>For individuals, this might amount to an asterisk leading to a refuting or clarifying site. For companies, the remedy might include an asterisk next to unfair competitor’s entry. If, for example, I own the trademark “Oh-So-Con” for convenience stores, and someone else’s store comes up first in an Google search for “Oh-So-Con,” I could have an asterisk next to the hyperlink to that entry, and the asterisk would lead to a page that would permit to explain that I own the mark.

asterisk next to the offending result, linked to the complainant's own website, would accomplish this at minimal cost to the affected search engine.<sup>88</sup>

Search engines will likely resist such regulation by pointing to *SearchKing's* characterization of ranking as opinion protected by the First Amendment. Wouldn't such an asterisk constitute precisely the type of "forced speech" ruled unconstitutional by the Supreme Court in *GLIB v. Hurley*?<sup>89</sup> When California voters attempted to place three dollar signs ("\$\$\$") next to the ballot entries of candidates who had engaged in certain disfavored campaign funding practices, this "compelled speech" was quickly ruled unconstitutional.<sup>90</sup> What is to prevent similar First Amendment absolutism from protecting search engines' right to produce whatever rankings they want?

Setting aside the intricacies of First Amendment doctrine for now, the Fair Credit Reporting Act (FCRA) appears to provide a clear precedent for limiting metadata providers' capacity to provide whatever accounts of data they deem fit. The FCRA requires credit bureaus to permit individuals to dispute negative information on their credit reports and to give their own side of the story on reports generated for potential creditors, insurers, and employers. Applied to search engines, such protections would merely permit those dogged by negative information in search engines to put an asterisk next to the metadata indexing the information which would be hyperlinked to "their side of the story."<sup>91</sup> The asterisk is not speech itself, but merely a way of restructuring a forum with deep roots in our First Amendment tradition.

---

<sup>88</sup>Scholars have only recently realized the power of asterisks to send all manner of signals. See, e.g., ANTHONY GRAFTON, *THE FOOTNOTE: A CURIOUS HISTORY* (2003); Charles Sullivan, *The Undertheorized Asterisk Footnote*, 93 GEO. L.J. 1093, 1094 (2005) (discussing evolution of "asterisk footnote" in law review articles.).

<sup>89</sup>See *Hurley v. Irish-Am. Gay, Lesbian & Bisexual Group*, 515 U.S. 557 (1995); *Boy Scouts v. Dale*, 530 U.S. 640, 656 (2000) (ruling that civil rights laws could not force institutions to send a "pro-gay" message by including gays in their activities). The Supreme Court will likely limit this doctrine in *Forum for Academic & Institutional Rights, Inc. v. Rumsfeld*, in 2006 by reversing a Third Circuit opinion that relied on *Dale* in order to overturn the Solomon Amendment (which had effectively prevented law schools from banning military recruiters) on the basis of expressive and associational autonomy. 390 F.3d 219, 230 (3d Cir. 2004), *rev'd*, 126 S. Ct. 1297 (2006).

<sup>90</sup>Cal. Prolife Council Political Action Comm. v. Scully, No. CIV.S-96-1965 at 16 (E.D. Cal. Mar. 1, 2001), available at <http://www.law.ucla.edu/faculty/bios/lowenste/slatemailorder.pdf> ("Statutes compelling speech, like those forbidding speech, address matter well within the protection of the First Amendment."); Shanto Iyengar, Daniel H. Lowenstein, & Seth Masket, *The Stealth Campaign: Experimental Studies of Slate Mail in California*, 17 J.L. & POL. 295 (2001) (noting that "the requirement for dollar signs instead of asterisks, and the requirement that \$50,000 contributors to ballot measure campaigns be disclosed on the face of a slate mailer" were ruled unconstitutional). Full background on the case is available at *California Prolife Council Political Action Committee v. Scully*, 989 F.Supp. 1282 (E.D.Cal.1998).

<sup>91</sup>The asterisk might also provide a good way to settle disputes over trademarked adwords. For example, if a user types in "American Blinds" as a query, and a number of competitors appear as results, perhaps American Blinds should have the right to put an asterisk next to each result in order to indicate that the competitor is not actually the source legally designated by the trademarked words American Blinds.

Federal regulation along the lines of FCRA would also offer some comfort to dominant search engines if it resulted in a preemption of state law liability. The SearchKing court was correct to worry that state unfair competition, defamation, and false light claims could lead to de facto judicial regulation of search and an unwieldy expansion of secondary liability. Moreover, permitting such lawsuits would probably only entrench dominant search engines, as they are the only players with the resources to consistently defend against them.

#### *D. Addressing Harm from Exclusion from Results*

As the stories in Part IV demonstrate, search engine results can be arbitrary and capricious, suddenly dropping individuals who built their livelihoods and reputations on high rankings.<sup>92</sup> Sometimes drops in ranking result from understandable efforts to assure the integrity of ranking algorithms.<sup>93</sup> However, more suspect manipulations have also been documented. John Battelle relates the following stunning turn of events in the *American Blinds* case:

September 17, 2004 was the day the San Jose District Court was to hear arguments in the American Blinds case [regarding a dispute over the trademarked adwords “American Blinds”]. . . . Google had filed a motion to dismiss. . . . [When a] member of American Blinds’ legal team . . . [attempted to] test the system, he brought up Google and entered what had become a habitual search query: “American Blinds.” . . . . Every [other] time someone entered ‘American Blinds’ into Google’s search field, competitors to American Blinds came up on the screen.

Only this morning, for some reason, they did not.

The lawyer suspected Google had changed its results, and called colleagues in other parts of the country. Sure enough, searches in other regions returned different results, including the potentially infringing advertisements. . . . The lawyer quickly documented his findings.<sup>94</sup>

Despite this evidence, Google’s PR team claimed that “‘Google would certainly never do such a thing’” and attributed the discrepancy to a technical glitch.<sup>95</sup>

As long as ranking algorithms are a “black box,” such potentially troubling practices are likely to continue—and perhaps become more undetectable. Given the strength of IP protection of search algorithms,<sup>96</sup> it is unlikely that proposals to make

---

<sup>92</sup>See, for example, the story of Neil Moncrief, the proprietor of 2bigfeet.com (a seller of large-sized men’s shoes), whose site was knocked off the first page of Google’s rankings by a sudden algorithm shift in November, 2003, right before the Christmas buying season. BATTELLE, *supra* note 21, at 157. Moncrief attempted to contact Google several times, but “never got a response.” *Id.*

<sup>93</sup>*Id.* at 157-59 (describing a particular intervention, dubbed “Florida,” designed to thwart illicit “search engine optimization” techniques).

<sup>94</sup>*Id.* at 184.

<sup>95</sup>*Id.*

<sup>96</sup>Such algorithms may be protected by copyright, trade secret, and patent protections. Moreover, some courts have suggested that *revisions* to such laws might constitute



ranking methods transparent can succeed.<sup>97</sup> However, those hurt by sudden changes in ranking should have some opportunity to learn about what caused the change. Just as credit bureaus must report to consumers the *nature* of “adverse information” that makes their credit reports less than perfect, search engines should have some responsibility to reveal the reasons for their re-rankings to those adversely affected. Though such information may prove anodyne at first, scrutiny might eventually reveal any overly self-serving or unfair manipulations of rank.<sup>98</sup> Moreover, search engines cannot complain that these minor revelations would overly burden their business—the credit rating agencies still manages to keep the exact nature of their most important scoring mechanisms (such as the FICO score) a tightly-held secret.

## VI. CONCLUSION

Automation has increased the amount and sped the dissemination of information exponentially over the past few decades. Research tasks that would have taken a team of assistants days or weeks can now be completed by a computer in seconds.<sup>99</sup> Search engine databases index information equal (by some measures) to all that was

---

compensable takings. See, e.g., *Philip Morris v. Reilly*, 312 F.3d 24 (1st Cir. 2002); *Monsanto v. Ruckelshaus*, 753 F.2d 649, 651 (8th Cir. 1985). For an effort to clarify and extend this caselaw, see Adam Mossoff and Eric Claeys, *Patents, Prices, and Property: Drug Patents and the Takings Clause*, Lecture at Benjamin Cardozo School of Law Intellectual Property Conference (August 2005) (setting forth rationales for applying regulatory takings doctrine to IP).

<sup>97</sup>Nissenbaum and Introna proposed such a “transparency” rule as a *sine qua non* of search engine reform. Introna & Nissenbaum, *supra* note 61, at 32:

As a first step we would demand full and truthful disclosure of the underlying rules (or algorithms) governing indexing, searching, and prioritizing, stated in a way that is meaningful to the majority of Web users. Obviously, this might help spammers. However, we would argue that the impact of these unethical practices would be severely dampened if both seekers and those wishing to be found are aware of the particular biases inherent in any given search engine. We believe informing users, on the whole, will be better than the status quo, in spite of the difficulties. Those who favor a market mechanism would perhaps be pleased to note that disclosure would move us closer to fulfilling the criteria of an ideal competitive market in search engines. Disclosure is a step in the right direction because it would lead to a clearer grasp of what is at stake in selecting among the various search engines, which in turn should help seekers make more informed decisions about which search engines to use and trust.

*Id.*

<sup>98</sup>Particularly if the secret ranking algorithm is held “in escrow” by an independent agency, which can verify the validity of reasons given for changes. See, e.g., Dan L. Burk and Julie E. Cohen, *Fair Use Infrastructure for Rights Management Systems*, 15 HARV. J. L. & TECH. 41 (2001) (proposing keeping digital “keys” to digital rights management software “in escrow” for government authorities who could “unlock” the material in response to legitimate request for “fair use” of it).

<sup>99</sup>For example, compare the process of “shephardizing” a case in “the books” to the “keycite” feature on Westlaw, or the computerized “Shephard’s” offered by Lexis.

thought and written before their advent.<sup>100</sup> Advanced algorithms can now organize and categorize this data in a remarkable number of ways.

Nevertheless, as search engines become more authoritative, encompassing more and more sources of data, they become capable of harms commensurate with their benefits. So far public attention has focused on privacy concerns raised by these data aggregators' storage of all search requests connected to given users.<sup>101</sup> Major players in the "creative industries" have also claimed that search engines will "Napsterize" their content or lead to a digital "Wal-Mart" capable of squeezing content providers as effectively as the brick-and-mortar retail behemoth drives down the prices it pays its suppliers.

Though both these concerns are important, they focus on likely *effects* of search engines, and not on the ranking process itself. Courts have largely failed to hold that process up to normative scrutiny, preferring instead to view search engine results as *both* 1) unregulable speech that is but one more contribution to the marketplace of ideas; *and* 2) an automated result not susceptible to the kind of moral judgment we usually reserve for humans.

This piece argues that neither justification for leaving search engine results unregulated is convincing. Search engines are not merely one more voice in a pluralistic public dialogue, but are poised to become the chief *organizer* and *forum* for research, public discussion, and commercial competition among internet users. Moreover, the same automated processes that make search engine results so rapid, useful, and flexible can also be deployed to make them more fair and responsible. Rather than hiding behind the mechanical application of First Amendment protections, new public fora like search engines should promote the First Amendment goal of open public forums. Only an ongoing concern for the *policy* behind freedom of expression will make that legal "neighborhood" an ameliorative influence on the gated privatopia of IP law it borders.<sup>102</sup>

---

<sup>100</sup>See Peter Lyman & Hal R. Varian, *How Much Information* 2003, <http://www.sims.berkeley.edu:8000/research/projects/how-much-info-2003/>. Note that the title of the survey is not *How Much Wisdom?*.

<sup>101</sup>Though long a concern of groups like the Electronic Privacy Information Center, this controversy only received sustained attention with Google's resistance of the federal government's subpoena of search requests. Google, however, did not resist the subpoena on the basis of the privacy rights of its users, but rather because (it claimed) the information could disclose company trade secrets.

<sup>102</sup>See, e.g., EVAN MCKENZIE, *PRIVATOPIA: HOMEOWNERS' ASSOCIATIONS AND THE RISE OF RESIDENTIAL PRIVATE GOVERNMENT* 177 (1994); Lee Ann Fennell, *Contracting Communities*, 2004 U. ILL. L. REV. 829, 832 (2004). Though the influence of free speech and competition law on IP law holds great promise, contracts of adhesion and paracopyright-like protections may ultimately render it a hermetically "gated community," incapable of absorbing the norms of other, more progressive milieux.

