THE STRUCTURE OF SEARCH ENGINE LAW

James Grimmelmann[*]

[*] Resident Fellow, Information Society Project, Yale Law School. My thanks to Jack Balkin, Yochai Benkler, Shyam Balganesh, Aislinn Black, Michael Carroll, Anne Huang, Dan Hunter, David Johnson, Thomas Lee, Beth Noveck, Frank Pasquale, Guy Pessach, Chris Riley, Steven Wu, Tal Zarsky, and the participants in the Yale Regulating Search conference and in Jack Balkin's Knowledge and Information Policy seminar for their comments.

INTRODUCTION

Search engines are the new linchpins of the Internet.[1] A large and growing fraction of the Internet's large and growing volume of traffic flows through them. They are librarians, bringing order to the chaotic online accumulation of human knowledge and creativity. They are messengers, creating new information flows and reorienting others. They are critics, wielding the power to elevate content to prominence or consign it to obscurity. They are inventors, devising new technologies and business models in their relentless drive to better describe complex online realities. And they are spies, asked to carry out investigations with dispatch and discretion.

Lawyers and the law have taken notice of search engines. Governments around the world are casting an increasingly skeptical eye on search engines, questioning whether their actions have always been in the interests of society. More and more parties are presenting themselves at the courthouse door with plausible stories of how they have been injured by search engines. Only a few foresighted legal scholars have recognized the growing importance of search engines.[2]

This Article will provide a road map to the legal issues posed by search. It will indicate what questions we must consider when thinking about search engines, and it will detail the interconnections among those questions. It will not endorse any particular normative framework for search.[3] Nor will it recommend who should regulate search.[4] Instead, it will provide the necessary foundation for informed decision-making, by whatever regulator and whatever its normative approach.

What gives the diverse questions of law to be discussed in this Article their coherence is that they all lay claim to regulate the same few core information flows. The essence of a search engine is the combination of information about content with user queries to provide recommendations to users who can then find the content. Every doctrine or policy value discussed herein relates directly to this single core process. Plenty of other law affects search engines—Google's well-publicized IPO, for example raised substantial issues of securities law,[5]

---

[1] *See* JOHN BATTELLE, THE SEARCH: HOW GOOGLE AND ITS RIVALS REWROTE THE RULES OF BUSINESS AND TRANSFORMED OUR CULTURE (2005); DAVID VISE WITH MARK MALSEED, THE GOOGLE STORY (2005).

[2] *See* Urs Gasser, *Regulating Search Engines: Taking Stock and Looking Ahead*, 9 YALE J.L. & TECH. 124 (2006); Frank Pasquale, *Rankings, Reductionism, and Responsibility*; 54 CLEV. ST. L. REV. 115 (2006); Eric Goldman, *Search Engine Bias and the Demise of Search Engine Utopianism*, 8 YALE J.L & TECH. 188 (2006); Niva Elkin-Koren, *Let the Crawlers Crawl: On Virtual Gatekeepers and the Right to Exclude Indexing*, 26 DAYTON LAW REVIEW, 180 (2001).

[3] In a companion piece, I will propose that the guiding principle for regulating search should be *maximizing broad user access to a diversity of effective search tools*. *See* James Grimmelmann, *A Theory of Search Engine Policy* (draft on file with author). I have split these two articles, because while an accurate description of search law as it is today is an essential prerequisite to that proposal, others may draw different lessons from the description than I do.

[4] *Cf.* Oren Bracha & Frank Pasquale, *Federal Search Commission? Access, Fairness, and Accountability in the Law of Search* (draft outline on file with author) (comparing institutional forms for search regulation).

[5] *See* Google, Inc., *Form S-1 Registration Statement*, available at http://www.sec.gov/Archives/edgar/data/1288776/000119312504073639/ds1.htm.

and it has been sued for employment discrimination[6]—but these other issues are separable. They can be resolved on their own merits, in isolation, without much affecting the core disputes about the information flows making up search, and without needing to be balanced with other concerns about those flows.

Part I will explain how modern search engines function and describe the business environment within which they operate. Search engine operations can be understood in terms of the *information flows* among four principal *actors*: *search engines* themselves, their *users*, information *providers*, and *third parties* (such as copyright holders and censorious governments) with interests in particular content flows. There are, in turn, four significant information flows: the *indexing* by which a search engine leans what content providers are making available, user *queries* to the search engine for information about particular topics, the *results* returned by the search engine to users, and finally, the *content* that providers send to users who have found them through searching. Because so many major search engines are funded through advertising, this Part will also include a brief survey of how search engine advertising works and the distinctive fraud problems confronting search engines and their advertisers.

Part II, the heart of the Article, will present a descriptive analysis of the legal struggles over search, showing how questions of search policy, many of which have long been latent in different fields of Internet law, are increasingly confronting lawyers, courts, and regulators. It will describe those struggles in terms of the legitimate *interests* that each of these actors brings to debates over search. Users want *high-quality results* without too great a sacrifice of *privacy*. Content providers want *favorable placement in search results* without paying more than their fair share of the *costs of supporting search* and without facing *unfair competition* from search engines. Third parties want to prevent unauthorized distribution of *copyrighted content*, to preserve their own *privacy*, to protect their *reputation*, and to preserve *user virtue*. And finally, search engines want to preserve their ability to *innovate*, to protect themselves from *fraud*, and to ensure that the search market remains open to *competition*. Each entry in this list of a dozen interests has its own associated legal theories; this systematic taxonomy allows us to recognize how any given legal theory affects the search ecology.

Part III will then show, with five examples, how taking a broad view of search yields otherwise unavailable insights into pressing controversies. This is not to say that the end result must be a body of search-specific law, [7] only to note that failing to consider the larger forces at work in search is antithetical to sensible policy-making. First, the broad, systematic view illustrates how various claims in search engine disputes can serve as functional substitutes for each other. Second, it shows that the degree of transparency of the search process is a highly contested variable, which some concerns pressing for greater transparency and some pressing for

---

[6] *See* Elwell v. Google, Inc., No. 05 Civ. 6487, 2006 U.S. Dist. LEXIS 3114 (S.D.N.Y. Jan 31, 2006).

[7] *Cf.* Frank Easterbrook, *Cyberspace and the Law of the Horse*, 1996 U. Chi. L. F. 207 (arguing that there neither is nor should be a distinct body of cyberlaw). Lawrence Lessig has responded that "more than law alone enables legal values, and law alone cannot guarantee them," Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, 113 Harv. L. Rev. 501 (1999), and argued that cyberlaw (or "Internet law," depending on one's view of the subject) provides a broader view of law itself. Although I have generally sided with Lessig's view of cyberlaw's significance in other work, *see* James Grimmelmann, *Regulation by Software*, 114 YALE L.J. 1719 (2005), I do not make the same claims for search engine law. Search engines are more important in the consideration of what law should do than in the consideration of what law is.

less. Third, it illustrates that user privacy is a deeply knotty problem, and that preserving reasonable user expectations will involve difficult trade-offs with other interests—including some of users' own. Fourth, it shows that we require a theory of search engine speech; the most well-developed theory of search engine results as speech so far articulated by a court is insufficiently complex. And fifth, it illustrates the richness of debates over search engines' relationship to providers' trademarks.

Finally, a brief Conclusion will take note of some of the many open issues facing search engine law and scholarship.

I. SEARCH ENGINE TECHNOLOGY AND BUSINESS

In what follows, I will use "search engine" in a fairly narrow sense: to refer to *a service that helps its users locate publicly accessible content on the Internet*. At the core of this definition are the Web search sites that most people think of when they hear the term "search engine," such as Google, MSN Search, and Yahoo! Many search engines, however, work with content beyond simply Web pages. Google Scholar searches journal articles, Yahoo! Local searches businesses near the user, the Internet Movie Database searches lists of film casts and crews, and Amazon.com searches books and other products that it sells. For this reason, it is better to say that search engines help users find "content" rather than "pages" or "sites."[8]

Just outside of this definition, but close enough to be worth noting, are search services that work with content not available to the public. LexisNexis, for a fee, allows users to search a large proprietary database of legal and news documents. Similarly, peer-to-peer file sharing systems such as Gnutella[9] and Grokster[10] generally include search functionality; the content available through them is only accessible through the peer-to-peer service itself.

The exact dividing line is not of great importance for present purposes. Because so many applications on the Internet include search functionality, little is likely to turn on whether a technology is characterized as a "search engine" or as something else. We shall see that much of what I refer to as "search engine law" is drawn from many other technologies and will in turn be applicable to them.

A. *Technology*

A search engine in isolation is useless. It becomes valuable only through its interactions with *content providers* and with *users*. By aggregating its knowledge of what content providers have to offer and organizing that knowledge in a form useful to users, a search engine can match users with appropriate content providers, to the benefit of both. This matching, however, can antagonize *third parties* who would rather that certain connections not be made. (As discussed in more detail below, such third parties include copyright holders, targets of libel, stalker-fearing privacy lovers, and censorious governments.) Visualizing the information flows between search engines and these three groups illustrates how search works.

At its core, search consists of four flows of information:

1. The search engine gathers content.
2. A user queries the search engine.
3. The search engine provides the user with results.

---

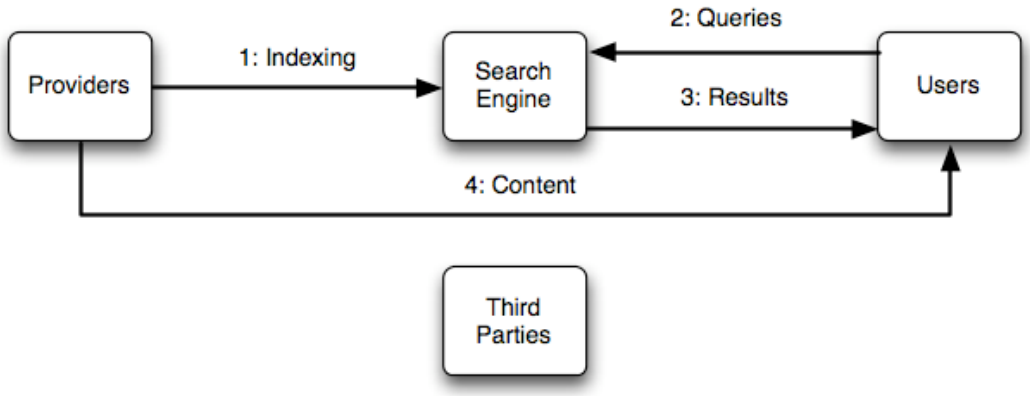
[8] I say "content" rather than "information" to distinguish it from significant information flows connected with search. It is also worth noting that not all search engines are themselves Web sites.

[9] *See Gnutella*, Wikipedia, at http://en.wikipedia.org/wiki/Gnutella.

[10] *See Grokster*, Wikipedia, at http://en.wikipedia.org/wiki/Grokster

4. The user obtains the content.

These four steps describe traditional Web search engines, but they also describe eBay, Grokster, Wikipedia, and any number of other Internet-enabled search applications. They are diagrammed in Figure 1:



1. The Search Engine Gathers Content

While users use a search engine to search for content, a search engine itself must search out the content it is to recommend to them. It must therefore work with content providers to learn what they have to offer. With Web search, the process is normally driven by the search engine, which uses automated software agents—"robots", "spiders," or "crawlers"—to explore the Web and find content. It generally does so in the same manner that a user would, requesting Web pages from content providers' servers and seeing what those pages contain.[11]

Other forms of search involve different forms of information gathering. Some search engines simply take existing collections of information and organize them more effectively—a nationwide phone number search (now built into some computer operating systems) aggregates information at one time only available in a shelf of phone books.[12] Others rely on content providers to come to them. Under paid search inclusion, once practiced by a number of Web search engines, content providers pay a search engine and supply it with content; the engine promises to index any content given to it with the appropriate fee.[13]

The line between search engine and provider can be quite indistinct. Online merchants (e.g. Amazon.com) typically provide search engines for their own sites, as do most sites that aggregate content supplied by users (e.g. Wikipedia). One does not have to use the site-specific search engine to reach the content, but one can. Decentralized peer-to-peer systems collaboratively use the same computers both to provide files and to index them.

---

[11] *See* Google, *How Google crawls my site*, at http://www.google.com/support/webmasters/bin/topic.py?topic=8843.

[12] *See* Apple, *White Pages – Dashboard – Reference*, at http://www.apple.com/downloads/dashboard/reference/whitepages.html.

[13] *See* Danny Sullivan, *The Evolution of Paid Inclusion*, Search Engine Watch (July 2, 2001), at http://searchenginewatch.com/showPage.html?page=2163971.

2. The User Queries the Search Engine

Content is only one of two inputs to search. The other is the *search query*, a request by a user for information on a particular topic. Most Web search engines have queries made up of a few keywords or short phrases. Some search engines are non-textual, allowing users to issue queries in more exotic forms, such hummed tunes,[14] pictures of celebrities,[15] or even color (e.g. to find clothing matching a given shade[16]). In whatever form, the user provides the search engine with some criteria by which to narrow the vast universe of possible results.

A query is typically only an approximation of the user's intention, and the same query may express many different intentions. Common intentions include navigational queries (the user wishes to find a specific site or datum), informational queries (the user wishes to find information on a topic), and transactional queries (the user wishes to perform an activity, such as purchasing a good).[17] Because words have multiple meanings, the same query could reasonably be directed at many different possibilities, even within one of these categories. Further, the user may not have in her head a clear idea of what she is searching for. What she is searching for may not even exist. By way of example, consider a search for "apple." The user might have intended any of the following:

- To find the home page for Apple Computer (navigational);
- To learn about apples, the fruit (informational);
- To purchase an Apple MacBook computer (transactional);
- To purchase apples, the fruit (transactional, but different);
- To learn about oranges (informational, but confused);
- To find the home page for Apple Records, the Beatles' record label (navigational, but no such page exists[18]); or
- To test whether her connection to the Internet is working.

Along with the query itself may come various user information, such as past searches, geographic location, preferred types of results, operating system and browser, and preferences among search results revealed through clicks on past search results. Some search engines, such as Google and Yahoo!, keep extensive histories on the searches and preferences of registered users. Personalized search engines may customize their results on each query for the particular user, showing different results to users with different geographic locations or announced interest.

3. The Search Engine Provides the User With Results

In the defining step of search, the search engine integrates information about available content and the user's question to return to the user information about content relevant to her query. Web search engines typically show results in descending order of perceived relevance, grouped into pages of ten at a time. Each result normally contains the name of the identified

---

[14] *See* Midomi Video Tour (2007), http://www.midomi.com/index.php?action=main.video.

[15] *See* Bob Tedeschi, *Shopping Site Offers a Way to Raid a Celebrity's Closet*, New York Times (Nov. 13, 2006); Like Visual Search, http://www.like.com.

[16] *See* Become, http://www.become.com/.

[17] *See* Andrei Broder, *A Taxonomy of Web Search*, 36 ACM SIGIR Forum, no. 2 (2002). Broder's taxonomy is obviously not exhaustive nor are its three categories entirely distinct from each other, but it provides a good first approximation.

[18] The closest substitutes are probably the home page for Apple Corps, Ltd. (the conglomerate parent of Apple Records), http://www.applecorps.com/ and *The Complete Apple Records* (an unaffiliated site with a complete Apple Records discography), http://www.schomakers.com/.

piece of content, its location, and a very short preview, summary, or excerpt of the content. This excerpt sometimes indicates where, if at all, the query appears within the content.

Note also that as a user inspects a set of search results, she may also refine her query, choosing slightly different keywords in an attempt to better convey her intention to the search engine. The engine, in turn, will supply her a different set of results. This process is iterative and thus, even within a single search session, query and results flows may repeat any number of times.[19]

This step typically involves extensive organization and processing of the information the search engine has at its disposal. The algorithms that search engines use have been growing more complicated and resource-intensive. At one time, Web search engines simply scanned the text of Web pages to determine which topics the pages discussed. That technique was later augmented by analysis of additional information about pages (called "metadata"), such as their age, the number of links they contain, or the keywords used by their authors to describe them. More recently, powerful link structure techniques involving study of which Web pages link to which other pages—from which search engines can determine how "popular" a page is with other page creators—have become the dominant Web search paradigm.[20]

To be useful, the search engine must also integrate the particular query with its indices of content. Different search engines structure this integration in different ways. Some are almost entirely non-interactive. They have pre-generated lists of results and merely allow users to choose. in effect, from a predetermined list of queries. Such search engines function like the index of a book, allowing a user to find quickly anything the search engine has specifically chosen to show. Yahoo's directory of the Web was perhaps the best-known example, but there are still many examples of static directories with less comprehensive ambitions.[21] Most search engines today are more interactive and generate different lists of results for every possible query. Some preprocess the information they have gathered from content providers before analyzing a query; others flip the order and only gather information in response to a query.

Search engines are also increasingly learning from the large volumes of query data they have accumulated. Divining user intent from a search query is a notoriously difficult problem and the same query may indicate a different intent in different contexts. Search engines use various analytical technologies to transform queries into a hypothesized intent of the user. Query histories also provide valuable information. Search engines often use them to understand their own failings—thus, for example, repeated searches on closely related terms may indicate that a user is having trouble locating relevant information.

---

[19] *See* Goldman, *supra* note 2. As Goldman explains, this refinement of results is a critical process by which users correct for ambiguities in their initial queries—and the interactivity of the process argues against trying to infer a single fixed meaning for a query.

[20] *See* U.S Patent No. 7,058,628 (June 6, 2006) ("Method for node ranking in a linked database"); Lawrence Page et al., *The PageRank citation ranking: Bringing order to the Web* (1999); Jon M. Kleinberg. *Authoritative sources in a hyperlinked environment*, Proc. Ninth Ann. ACM-SIAM Symp. Discrete Algorithms 668 (1998). A similar growth in sophistication has been evident in other search domains. More recent peer-to-peer systems determine the location, bandwidth, and likely quality of the files they index. Sites like Flickr (for photographs) and Del.icio.us (for user-bookmarked Web pages) use user-supplied "tags" to categorize content in a flexible way not tied to any particular taxonomy.

[21] The Open Directory Project, http://dmoz.org/, still provides a static directory, although it now seems to wish to include "only the best content." *About the Open Directory Project*, at http://dmoz.org/about.html.

4. The User Obtains the Content

At the end of the day, what matters to the user is the content. Most often, she takes the location information given her by the search engine in its results and uses that information to contact the relevant provider, with whom she can then negotiate to obtain the content. She may purchase goods and services (or decide not to, after considering her options). Where the goods are information goods, the she may consume them directly, downloading them from the relevant content provider. Typically, this matching is valued both by the user and the content provider, since this sort of exchange is why the content provider is online and why the user turned to a search engine.

Some search engines themselves provide content to users. Where providers create their own search engines, the process of providing results and the process of providing the content itself may merge. A successful search on Wikipedia, for example, simply returns the desired Wikipedia entry directly. Independent search engines sometimes also provide content to users. They may cache content, storing copies to make it easier for users to receive it quickly, or archive content, enabling users to receive it even when the original provider cannot be reached. They may also allow users to preview content, offering smaller (e.g. thumbnails for pictures) or excerpted versions of it. At the extreme end of this trend, search engines become true middlemen, simply purchasing the content from willing sellers and retailing it to users.

B. *Business*

It is impossible to understand the legal controversies over search without some understanding of the most common search business models. The overwhelmingly predominant model for Web search today is contextual advertising, in which, the search engine, in addition to showing its users results, shows them advertisements, most commonly textual ones. Almost all are for other Web sites: Content providers have learned that they can be found by users either by being listed as search results or by advertising on search engines.

The line between these two forms is not always clear. Some search engines simply mix advertisements with so-called "organic" search results, in a process known as paid placement.[22] The theory behind this approach is to create an explicit market in search ranking, so that users can count on finding the providers who most value them—and are most eager for their business. Most search engines segregate paid and organic results, however, showing the ads in a different position on the screen, in a different color, and with a descriptive header identifying them as ads.[23]

---

[22] *See* Benjamin Edelman & Michael Ostrovsky, *Strategic Bidder Behavior in Sponsored Search Auctions*, Decision Support Systems (forthcoming), available at http://www.benedelman.org/publications/cycling-060703.pdf; F. Gregory Lastowka, *Search Engines Under Siege: Do Paid Placement Listings Infringe Trademarks?*, 13 Intell. Prop. & Tech. L.J. 6 (2002). The process is closely related to paid inclusion; in each, a content provider has paid the search engine in the hope of being seen by users entering particular queries. *But see* Thomas A. Weber & Zhiqiang (Eric) Zhang, *A Model of Search Intermediaries and Paid Referrals*, Wharton School OPIM Working Paper No. 02-12-01, available at http://ssrn.com/abstract=601903 (arguing that ranking based on provider bids decreases overall social welfare).

[23] *But see* Deborah Fallows, *Search Engine Users*, Pew/Internet (Jan. 23, 2005), http://www.pewinternet.org/pdfs/PIP_Searchengine_users.pdf (finding 62% of search engine users unaware of distinction between organic and sponsored results).

Even these ads, however, are still often search "results" in the sense that the search engine's decision of which ads to show is driven by the user's query. The search engine uses its access to query data to pick particular ads to show to particular users. Search engines have developed sophisticated bidding algorithms that balance advertisers' willingness to pay with the popularity of their ads in choosing which so-called "sponsored links" to show.

Search engines use three common billing techniques to sell ads. An advertiser using pay-per-impression pays a given fee to the engine each time a user sees the ad;[24] under pay-per-click, the advertiser pays each time a user not only sees the ad but clicks on it; under pay-per-conversion (also known as pay-per-action or pay-per-performance), the advertiser pays only when the user makes a purchase or takes some similar action that indicates serious interest in the advertiser's site. Pay-per-impression is the least correlated with sales; pay-per-conversion is the most, but requires that advertisers turn over significant information to the search engine so that billing can be properly calculated. Pay-per-click currently strikes the most popular overall compromise between accuracy and convenience.[25]

These advertising models are not confined to search engine sites. Web-wide advertising middlemen, such as DoubleClick, match Web sites and advertisers. By tracking users across multiple sites, these middlemen have tried to target ads to particularly responsive users. Search engines have more recently entered this business, using their sophisticated content analysis algorithms to pick the most promising ads for the particular Web site on which they will appear. In all of these affiliate network models, the advertiser pays for each event, with the middleman and the site on which the ad appears splitting the revenue.[26]

A few Web search engines do not provide advertising. Most commonly, providers who establish their own site-specific search engines can monetize them directly because every result will be from their own site. Other directories and search engines are maintained out of altruism;

---

[24] For non-interactive media, pay-per-impression is generally the only feasible model. Magazines, newspapers, and television networks, for example, typically compete for advertisers based on their circulation or viewership, which provides a rough proxy for the number of impressions.

[25] A full discussion of the business issues involved in search engine advertising is beyond the scope of this paper. As Tal Zarsky has explained in correspondence, the different billing models have significantly different implications for the incentives of search engines and advertisers and for the privacy of search engine users. Thus, for example, under a pay-per-click system, search engines will try to favor not just those ads for which advertisers will pay the most per click, but also those ads that will generate the most clicks. Advertisers, in turn, will therefore be able to purchase prominent advertising placement more cheaply if their ads are well designed to encourage clicks. The process of monitoring clicks, however, is both vulnerable to fraud and requires close observation of user behavior, two facts that implicate the general tension between openness and transparency in search advertising and search engine operations. *See also* Ben Elgin, *The Vanishing Click-Fraud Case*, Business Week (Dec. 4, 2006), available at http://www.businessweek.com/print/technology/content/dec2006/tc20061204_923336.htm.

[26] Affiliate marketing is actually substantially more general than this brief description would suggest. Amazon is credited with initiating the practice on the Web, by offering site owners a commission on any sales generated by referrals from their sites. The referral system was not new; the commission and concomitant fraud potential were. *See* John Schwartz & Bob Tedeschi, *New Software Quietly Diverts Sales Commissions*, New York Times (Sept. 27, 2002).

the search facility is provided ad-free as a public service. A few search engines have attempted to make money by analyzing user search patterns and selling aggregate information.[27]

These advertising techniques have generated their own forms of fraud. First, those upset at a particular advertiser or in competition with it may engage in click fraud, repeatedly viewing and clicking on ads to run up the advertiser's bills.[28] More insidiously, some engage in affiliate fraud, registering as affiliates and then clicking on ads they themselves have served. Some of the money goes to the affiliate, while the bill goes to the advertisers.

The great demand for high placement[29]—when combined with the zero-sum nature of ranking decisions—leads to search engine optimization: the business of redesigning content (or creating it) to attract search engines and convince them to rank content highly.[30] Some "white hat" SEO techniques are generally considered desirable, because they make the content easier for search engines and users to access. Other, "black hat" techniques involve mimicking the superficial features that search engines use as proxies for quality content. When Web search engines scanned page text and keywords, optimizers would hide popular keywords in invisible tiny text on a page, or show a search engine a different page (one larded with thousands of keywords) than the one shown to users. As search engines shifted to analyzing link structure, optimizers switched to creating link farms: sets of thousands of sites and pages pointing to each other, mimicking a community of real users and hoping to trick search engines into treating them as authoritative, popular sources of information.[31] Search engines and black hat SEOs are locked in a technical arms race that pits increasingly sophisticated algorithms to distinguish fraudulent from authentic content against increasingly subtle forms of mimicry.[32] As might be expected,

---

[27] The Alexa Toolbar uses this strategy. Once installed in a user's browser, it displays extra information about the pages that the user visits, including related links. In exchange, it can track which pages users visit. This aggregate user data is valuable even just as a Nielsen-type rating for Web pages.

[28] *See* Charles C. Mann, *How Click Fraud Could Swallow the Internet*, Wired (Jan. 2006), available at http://www.wired.com/wired/archive/14.01/fraud_pr.html.

[29] *See* Nico Brooks, *The Atlas Rank Report* (parts I and II) (Atlas Institute 2004), available at http://www.atlassolutions.com/institute/insights.aspx. Users are far likelier to click on the first result in a list shown them than on any other result; if they are shown a page of ten results, only a small fraction of them will click through to see even the second page. Results after the hundredth or so may as well not exist.

[30] *See* David Kesmodel, '*Optimize' Rankings At Your Own Risk*, Wall Street Journal (Sept. 23, 2005); Google, *What's an SEO?*, at http://www.google.com/support/webmasters/bin/answer.py?answer=35291.

[31] Link farms have cropped up in some surprising places. *See* Barry Schwartz, United Press International Selling PageRank, Search Engine Watch (Oct. 19, 2006), at http://blog.searchenginewatch.com/blog/061019-091933 (United Press International); Andy Baio, *Wordpress Website's Search Engine Spam*, Waxy.org (Mar. 30, 2005), at http://www.waxy.org/archive/2005/03/30/wordpres.shtml (homepage for "popular open-source blogging software"); Blake Ross, *Stanford Daily link spam harms the web and students*, Blakeross.com May 27, 2005), at http://blakeross.com/index.php?p=136 (college newspapers).

[32] *See, e.g.*, *Wikipedia:Send in the Clones*, at http://en.wikipedia.org/wiki/Wikipedia:Send_in_the_clones (copying content available elsewhere); Lee Gomes, *Our Columnist Creates Web 'Original Content' But Is in for a Surprise*, Wall St. Journal (Mar. 1, 2006), available at http://online.wsj.com/public/article/SB114116587424585798-0qH9qUYuUug__vRSFKGvxIEwLGw_20070301.html?mod=blogs (creating low-quality but semantically "original" content). *Cf.* Stephen Baker, *Asbestos and the art of blogging for money*, Blogspotting, BusinessWeek (May 27, 2005) at

black hat SEO techniques are highly controversial, and the line between black and white hat techniques is both unclear and contested, as is the line between authentic and fraudulent content.

These two realities of the search engine business—advertising and SEO—come together in an unsavory mix of parasitic business practices. Both seek to turn user attention into profit. Affiliate network advertisers who misdirect users to their own sites make the users into agents of affiliate fraud. SEO provides an essential tool in this effort by causing the search engine to participate in the misdirection. These forms of fraud are also linked to many other practices that degrade the Internet experience for all, including fraudulent domain registration[33] spyware installations,[34] and straightforward consumer fraud.[35] Black-hat SEOs have also used all sorts of techniques to create ranking-boosting links, in effect turning media besides email into vectors for hyperlink spam. They have posted link-filled comments to blogs and discussion boards,[36] created fake Web sites and blogs,[37] and sent hyperlinks in email, instant messages, and even in requests for Web pages.[38] The possibility of hijacking a highly-ranked site's reputation through fraudulent links also gives a powerful incentive to discover and exploit security holes.

II. THE STRUCTURE OF SEARCH ENGINE LAW

The array of legal theories asserted by, against, and in relation to search engines may seem bewildering. It has, however, a recurring deep structure that becomes evident if we focus on four concepts: the actors involved, the information flows among them, the interests that they bring to search, and the legal theories that they use to vindicate their interests. We have already met the key actors: search engines themselves, content providers, users, and concerned third parties. We have also discussed the relevant information flows: indexing, queries, results, and content. This Part will detail the actors' interests in these information flows, and examine the legal theories various associated with those interests.

Every dispute involving a search engine is, at a deeper level, a dispute between or within the other three actors. They are attempting both to capture the gains created by search and to use search engines to gain greater leverage over each other. Both goals could at times be considered legitimate, or illegitimate, depending on the particular details. Recognizing when one, the other, or both are at work illuminates what is really at stake in any given legal battle over search.

More specifically, on the one hand, search engines create enormous social benefits. They allow willing users and content providers to find each other, reducing transaction costs and

---

http://www.businessweek.com/the_thread/blogspotting/archives/2005/05/asbestos_and_th.html (describing experiment in which Michael Buffington created a topical blog on asbestos specifically to capture some of the $15 and upwards that lawyers would pay per click on ads triggered by keywords such as "mseothelioma.) There is no clear line between such experiments and "real" professional blogs.

[33] *See, e.g.*, Declan McCullagh, *Dotster named in massive cybersquatting suit*, C|Net News.com (June 2, 2006), at http://news.com.com/2102-1032_3-6079567.html.

[34] *See* Ben Edelman & Hannah Rosenbaum, *The Safety of Internet Search Engines* (May 12, 2006), at http://www.siteadvisor.com/studies/search_safety_may2006.html ("Despite search engines' efforts, we see too many sites trying to deceive unsuspecting users.").

[35] *See* Ben Edelman, *False and Deceptive Pay-per-Click Ads* (Oct. 10, 2006), at http://www.benedelman.org/ppc-scams/.

[36] *See Six Apart Guide to Combatting Comment Spam*, at http://www.sixapart.com/pronet/comment_spam.html

[37] *See* Christopher Heun, *Invasion of the Splogs*, InformationWeek (May 15, 2006).

[38] *See* Michelle Delio, *When the Spam Hits the Blogs*, Wired News (Oct. 26, 2002), at http://www.wired.com/news/culture/1,56017-0.html.

enabling mutually beneficial exchanges. These benefits depend on the contributions of users, providers, and search engines, in the form of queries, content, and ranking algorithms, respectively. Good search engine policy would therefore give each group appropriate incentives to maximize its productive contributions while avoiding rent-seeking behavior.

On the other hand, search engines can also cause enormous harms to particular parties. By controlling the matching process between users and content providers, they create winners and losers within these communities. Both users and providers entrust search engines with valuable information and may be upset at the terms on which search engines reveal that information. Third parties who would prefer that certain content not flow from providers to users also are injured when search engines enable such flows. Good search engine policy will prevent search engines from inflicting serious harms on others.

Complicating matters, these two features of search are inextricably intertwined. Users value search engines precisely because the search engines pick and choose among possible providers. Third parties are most upset at new content flows precisely when users and providers value those flows the most. Search engines do not generally cause harms out of their own inherent malice. They cause harms in the process of serving their other constituencies. Attempts to remediate particular hams, then, almost invariably involve a contest between the interests of these constituencies. The law's choice to intervene or not, therefore, is a choice among their interests.

The central position that search engines occupy also creates problems even when the balance between their various constituencies seems appropriate. If the balance comes with too much deference to search engines, the risk is that they will behave unaccountably and upset the balance by aligning themselves with one group against another. But if the balance comes with too many restrictions on the actions of search engines, the risk is that those restrictions will squander the innovative potential of search engines to benefit all.

With these general principles in mind, let us take up the specific interests that each constituency brings to any discussion of search law:

Users are the most obvious constituency served by search engines. When they make search queries, they are revealing some potentially private information; thus, they may be harmed if a search engine misuses or reveals their private information. They also desire useful search results, so a search engine can harm them by providing low-quality or deliberately biased results.

Providers have three interlocking interests in search. They may wish not to shoulder more than what they perceive as their fair share of the costs involved in providing search. Some such costs are contractual; they do not wish to pay more to search engines for advertising than necessary. Other such costs are technical, imposed on them by the search engine's spidering or by large numbers of users following a search engine's recommendation. Second, on a closely related note, by opening up new ways to view content, or simply by delivering content directly to users, search engines can compete with content providers, to the providers' great displeasure. Third, prominent placement in search results can be economically valuable. Providers have therefore gone to significant technical and legal lengths to achieve such placement

Third parties can be harmed principally by the content flows that search engines enable. Some have a copyright interest in the material now flowing freely; some may be the subjects of that material and feel that it defames them or invades their privacy; some (most often governments) may simply wish to suppress the flow of content they consider objectively harmful.

And finally, search engines themselves have at least three significant interests worth discussing. The first is operational—keeping their activities safe from SEO and click fraud. The second is innovative—retaining sufficient incentives and freedom to develop and deploy new forms of search technology. And the third is competitive—keeping other search engines from unfairly dominating the market for search.

These controversies are summarized in Table 1:

Constituency	Interest	Information Flows	Sample Legal Theories
Users	Privacy	Queries	Information privacy, contract
	Quality	Results	Consumer protection
Providers	Costs	Indexing, Content	Trespass, contract
	Unfair competition	Content	Copyright, trademark, contract
	Placement	Results	Trademark, business torts
Third Parties	IP	Content	Copyright (incl. DMCA)
	Reputation	Content	Defamation
	Privacy	Content	Information privacy
	Virtue	Content	Direct regulation
Search Engines	SEO/Click fraud	All	Fraud, contract
	Innovation	All	Intellectual Property
	Competition	All	Antitrust

The remainder of this Part will take up these issues in order.

A. *Users' Interests*

1. Query Privacy

Effective search requires that users disclose information about their interests and intentions. Whether they click on a topic heading in an index or craft a complex query with various exclusions and inclusions, the very fact that they are curious about something will be evident in their queries. If they are repeat users, the search engine may be able to construct an extensive history of their queries. It may also be able to correlate this curiosity with users' actual behavior in obtaining content, or with other information it has about them from their use of other applications and features it provides. In addition to collecting in-depth data on each user, a search engine also has broad access to information about many users.

Much of this data is personal or sensitive. When AOL publicly released the search queries of some 650,000 search users, the logs included queries such as "can you adopt after a suicide attempt,"[39] "cocaine in urine,"[40] and "How to deal with mental abuse in a Christian

[39] Declan McCullagh, *AOL's disturbing glimpse into users' lives*, C|Net News.com (Aug. 9, 2006).

[40] *Id.*

marriage."[41] Even though the search logs identified users only by pseudonymous numbers, reporters had little trouble tracking down people behind some of the searches.[42] Given the sensitivity of this information and the ease of linking it back to particular individuals, users have an evident privacy interest that their queries not be misused. Even where the information is genuinely not personally identifiable, it can still be used in ways that cause privacy harms.

The most obvious potential harms come from the release of user and query data to providers and third parties. In the terminology of Daniel Solove's *A Taxonomy of Privacy*,[43] the releases constitute *disclosures* and possibly a breach of promises of *confidentiality*. A release to a provider could enable the provider to engage in *decisional interference* against users who arrive at the provider's site. The provider could, for example, quote a higher price to a user whose query history reveals her not to be an informed purchaser.[44] Releases to third parties can cause *exposure* of embarrassing personal searches and the *identification* of particular searchers for further investigation, possibly leading to prosecution or other sanctions.

Even without external releases, however, users may feel uneasy about the treatment of their queries. Some may be upset by any *secondary use* for a purpose other than answering the specific query. (Even the display of advertising keyed to a user's query is arguably a purpose other than that intended by the user; it can also be viewed as a form of mild decisional interference.) Especially when the engine engages in *aggregation* of a user's entire search history, users often feel uncomfortable at how much an engine seems to know about them.[45] Whether based on disclosure, resulting harm, or the search engine's internal use, users' sense of *surveillance* is itself a harm—one that exerts a chilling effect on their searches[46] and may harm the search engine by deterring searching.

Caselaw dealing with collection of personal information online suggests that users will not enjoy significant legal recourse against search engines for the misuse of their queries. Multiple courts have held that users fail to state a claim when they allege that Web advertising services (with the cooperation of the Web sites on which the ads appear) have captured their browsing habits and tracked them through time.[47] Once in possession of the information, an engine would be free to disclose it to others unless it has undertaken not to.[48] A search engine with carefully drafted terms of service, therefore, can largely immunize itself from user privacy suits. Even a violation of its own stated privacy policy may not expose it to significant legal

---

[41] Lee Gomes, *What Are Web Surfers Seeking? Well, It's Just What You' Think*, Wall St. Journal (Aug. 16, 2006).

[42] Michael Barbaro and Tom Zeller Jr., A Face Is Exposed for AOL Searcher No. 4417749, New York Times (Aug. 9, 2006)

[43] 154 U. Penn. L. Rev. 477 (2006).

[44] *See generally* Joseph Turow, Lauren Feldman, & Kimberly Meltzer, *Open to Exploitation: American Shoppers Online and Offline* (Annenberg Public Policy Center 2005) (discussing consumer anger at online retailers practicing price discrimination by profiling customers); Tal Zarsky, *Mine Your Own Business*, 5 Yale J. L. &Tech. (2002–03) (explaining "autonomy trap" caused by such practices).

[45] *See* Jeffrey Zaslow, *If TiVo Thinks You Are Gay, Here's How to Set It Straight*, Wall St. Journal (Nov. 26, 2002); *Gmail Is Too Creepy*, http://www.gmail-is-too-creepy.com/.

[46] *See* Julie Cohen, *A Right to Read Anonymously: A Closer Look at "Copyright Management" in Cyberspace*, 28 Conn. L. Rev. 981 (1996).

[47] *See, e.g.*, Chance v. Avenue A, Inc., 165 F. Supp. 2d 1153 (W.D. Wash. 2001); In re DoubleClick Inc. Privacy Litigation, 154 F. Supp. 2d 497 (S.D.N.Y. 2001).

[48] *See, e.g.*, In re JetBlue Airways Corp. Privacy Litigation, 379 F. Supp. 2d 299 (E.D.N.Y. 2005).

risk.[49] (Legislation to change this baseline and require the deletion of stored queries was introduced in the 109th Congress, but died in committee.[50])

Even though a search engine may not be *obliged* to keep queries private, is it *permitted* to do so against the legal demands of others? Answers vary. The Fourth Amendment and various statutes allow several procedures by which computers can be searched, balancing the government's showing of relevance to an investigation against user expectations of privacy. [51] A full search warrant, properly supported by an affidavit showing probable cause, will trump any expectation of privacy. Internationally, search engine operators (albeit not in their roles as search engines) have shown a willingness to comply with government demands for identifying data, even when the consequences for the identified users are severe.[52] Some have been proactive about cooperation with law enforcement.[53]

Where the demand comes from a private-sector third party, users may have more leverage. A traditional *subpoena duces tecum* issued to an intermediary, such as a search engine, affords the intermediary and the adversary in the underlying litigation an opportunity to object.[54] The relevance of the information to the litigation will be balanced against the burden on the recipient of the subpoena. Google recently successfully resisted a subpoena for a random sample of user queries, with users' loss of trust in a search engine that releases their queries constituting the significant portion of the burden.[55] Courts considering the use of subpoenas to learn the identity of particular users have developed tests that depend on the speech interests of the users, [56] although they have yet to determine what speech interests users have in anonymous search. Outside of litigation, courts have held that the statutory subpoena process of the Digital Millennium Copyright Act[57] does not apply to ISPs that do not themselves host allegedly

---

[49] *See, e.g.*, In re Northwest Airlines Privacy Litigation, No. 04-126, 2004 U.S. Dist LEXIS 10580 (D. Minn. 2004) (dismissing breach of contract claim); In re Geocities, 127 F.T.C. 94 (1999) (entering FTC consent order without fine or punishment).

[50] *See* Eliminate Warehousing of Consumer Internet Data Act of 2006, H.R. 4731, 109th Cong. (introduced Feb. 8, 2006, referred to committee Feb. 17, 2006).

[51] *See generally* Computer Crime and Intellectual Property Section, Criminal Division, United States Department of Justice, *Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations* (2002).

[52] *See, e.g.*, Philippe Naughton, *Yahoo blamed for jailing of Chinese reporter*, Times (London) Online (Sept. 7, 2005); Anne Broache, *Google to hand over Brazilian user data*, C|Net News.com (Sept. 5, 2006), http://news.com.com/2061-10812_3-6112176.html.

[53] *See, e.g.*, Jonah Engle, *Buyer Beware: eBay Security Chief Turns Website Into Arm of the Law*, The Nation (June 20, 2003), available at http://www.thenation.com/doc/20030707/engle.

[54] *See* Fed. R. Civ. P. 45(c).

[55] Gonzales v. Google, 234 F.R.D. 674 (N.D. Cal. 2006). AOL, Yahoo!, and Microsoft did not resist subpoenas served on them for similar data.

[56] *See, e.g.*, O'Grady v. Superior Court, 139 Cal. App. 4th 1423 (Ct. App. 2006) (granting reporters' motion to quash subpoena to ISP to learn identity of confidential source); Sony Music Entertainment Inc. v. Does 1-40, 326 F. Supp. 2d 556 (S.D.N.Y. 2004) (requiring ISP to disclose identities of copyright infringement defendants); Doe v.2TheMart.com, 140 F. Supp. 2d 1088 (W.D. Wash. 2001) (requiring higher showing of need to learn identity of non-party to underlying litigation).

[57] 17 U.S.C. § 512(h).

infringing files,[58] but the applicability of these holdings to search engines has not yet been determined.

2. Unbiased Results

Users turn to search engines to help them find useful information. They therefore care that the search engine return them the highest-quality information it can. This desire gives them a first-order interest that the search engine do its best to return them relevant results, and a second-order interest in improving the "best" options open to search engines. Scholars have described the gap between the optimal results for a given search and the result actually returned as "bias."

The first significant challenge here is distinguishing bias from simple failure by the search engine to do as well as it could have. Several problems make it difficult to set a proper baseline of "unbiased" results. As noted above, different users may have different needs and desires, a given query may reflect any number of different intentions, and even the user may not know what she is searching for when she queries a search engine. Some have even questioned whether users are the proper judges of what search results would be best for them.[59]

Even without a precise definition of "unbiased" results, however, some authors have articulated a concern with certain sorts of bias. A search engine might consciously bias its results by favoring one provider or viewpoint over another. In China, major search engines remove from their indices content disfavored by the government, such as information on the banned Falun Gong movement.[60] Both liberal and conservative groups have accused Google of bias towards the other in its advertising policies. [61] The concern is commercial as well as political: some have claimed that search engines systematically favor their own advertisers or providers corporately affiliated with them.[62]

---

[58] In re: Charter Communications, Inc., Subpoena Enforcement Matter, 393 F.3d 771 (8th Cir. 2005); Recording Industry Association of America, Inc. v. Verizon Internet Services, Inc., 351 F.3d 1229 (D.C. Cir. 2003).

[59] *SEE* CASS SUNSTEIN, REPUBLIC.COM (2001) (arguing that users will choose to see only content confirming their preexisting biases if they are given total control over their information inputs). *But see* NICHOLAS NEGROPONTE, BEING DIGITAL (1996) (positing that users have tastes both for narrowly personalized information and for the common information seen by many others).

[60] *See* Clive Thompson, *Google's China Problem (and China's Google Problem)*, New York Times Magazine (Apr. 23, 2006).

[61] *See, e.g.*, PERRspectives, *Google's Gag Order: An Internet Giant Threatens Fee Speech* (June 20, 2004), at http://www.perrspectives.com/articles/art_gagorder01.htm (claiming conservative bias); Rightmarch.com, *Google says NO to Conservative Ads!*, at http://www.rightmarch.com/google.htm (claiming liberal bias); *but see* Eric Ulken, *A Question of Balance: Are Google News search results politically biased?* (May 5, 2005), *available at* http://ulken.com/thesis/googlenews-bias-study.pdf (claming no bias towards either side in Google News selection of articles).

[62] *See, e.g.*, Sergey Brin and Lawrence Page, *The Anatomy of a Large-Scale Hypertextual Web Search Engine*, available at http://www-db.stanford.edu/~backrub/google.html (stating "we expect that advertising funded search engines will be inherently biased towards the advertisers and away from the needs of the consumers"); *but see* Google, *Why we sell advertising, not search results*, at http://www.google.com/honestresults.html.

Technical design features of search engines can also introduce unconscious structural biases in their coverage and ranking of content.[63] Some have claimed that the link-heavy nature of weblogs leads to their overrepresentation by search engines using PageRank-type link-analysis algorithms.[64] Studies of relative traffic and links to Web sites have also caused some to discern a "Googlearchy," in which the most popular content receives more attention from users and therefore becomes even more popular, effectively preventing new providers from entering because they can never hope to catch up with established content in this vicious circle.[65] Mathematical models and empirical studies both support and undercut this theory.[66] Separating cause from effect in the wildly uneven popularity of content has proven difficult, as have attempts to show whether these differences are undesirable or not.[67]

If users or regulators are concerned about bias, what might they do? Because users are often not in a position to evaluate the search engine's performance directly, there is a substantial agency concern that bias may be undetectable.[68] If my query is navigational, I can usually tell whether my intended destination is among the search results; if my query is transactional, however, I may be less able to tell whether I have really been directed to the best sources for me. Some of this concern may be alleviated by users' ability to compare results from different search engines, although other search engines may suffer from the same systematic biases (think of the Chinese censorship and Googlearchy problems). Users might respond by demanding additional

---

[63] *See* Batya Friedman & Helen Nissenbaum, *Bias in computer systems*, 14 ACM Transactions on Information Systems 330 (1996); Lucas D. Introna & Helen Nissenbaum, *Shaping the Web: Why the politics of search engines matters*, 16 The Information Society, iss. 3, at 1 (2000).

[64] *See* John Hiler, *Google ♥ Blogs*, Microcontent News (Feb. 26, 2002), at http://www.microcontentnews.com/articles/googleblogs.htm; *see also* Ulken, *supra* note 61 (claiming inclusion of "non-traditional news sources" in Google News creates increased prominence for extreme viewpoints). *Cf.* Anil Dash, *Nigritude Ultramarine*, Dashes.com, at http://www.dashes.com/anil/2004/06/04/nigritude_ultra (describing how collaboration among blog authors to link Dash's entry using keywords "negritude ultramarine" nearly won an SEO contest).

[65] Matthew Hindman et al., *"Googlearchy": How a Few Heavily-Linked Sites Dominate Politics on the Web* (July 28, 2003), available at http://www.princeton.edu/~mhindman/googlearchy--hindman.pdf**.**

[66] *Compare* Junghoo Cho & Sourashis Roy, *Impact of Search Engines on Page Popularity*, *in* Proceedings of the World-Wide Web Conference (2004), *available at* http://oak.cs.ucla.edu/~cho/papers/cho-bias.pdf ("New and valuable pages are ignored just because they have not been given a chance to be noticed by people.") *with* Santo Fortunato et al.*, The egalitarian effect of search engines*, *in* Proceedings of the World-Wide Web Conference (2006), *available at* http://arxiv.org/abs/cs.CY/0511005 ("[T]he use of search engines actually has an egalitarian effect.").

[67] *See* Clay Shirky, *Power Laws, Weblogs, and Inequality*, at http://www.shirky.com/writings/powerlaw_weblog.html (Feb. 8, 2003) ("Inequality . . . is a reliable property that emerges from the normal functioning of the system."). Notably, even where scholars agree on the overall distribution of attention, they disagree on its implications. *Compare* Hindman, *supra* note 65 (x^{-1} powerlaw distribution sign of lack of diversity), *with* CHRIS ANDERSON, THE LONG TAIL (2006) (x^{-1} powerlaw distribution sign of diversity).

[68] *See, e.g.*, Alejandro M. Diaz, *Through the Google Goggles: Sociopolitical Bias in Search Engine Design* 147 (2005) ("The complexity and opacity of search technology makes it almost impossible for users to notice what is 'missing' from their search results."), available at http://www.stanford.edu/~amd/download/thesis_final.pdf.

information from a search engine about its ranking algorithms, to understand why it has made the choices it has.[69]

Legally, search engines have a strong first line of defense against user suits for bias in their browsewrap terms of service.[70] Further, few business tort theories provide users with enforceable rights. One exception may be the FTC's jurisdiction over misleading business practices. Although it has not taken direct action against any search engines, the FTC has communicated to search engines its belief that any paid placement results should be clearly disclosed and distinguished from organic search results.[71] In cases of conscious, human-directed manipulation of results, search engines may also have opened themselves to claims of fraud on consumers by emphasizing the mechanical and supposedly "objective" nature of their algorithms.[72]

Those who are concerned about systematic biases have also proposed various forms of forced ranking or inclusions. One proposal would have search engines be required randomly to intermix new content that has not yet had the time to establish itself with older and already popular content.[73] Others would require search engines to show users more diverse content to break down their biases towards the familiar and towards their own viewpoint.[74] There is a strong counterargument, however, that regulators would be grossly incompetent (and even more biased) at the task of dictating search results in general, a claim that would place a significant upper limit on the ambition of any anti-bias proposal.[75]

B. Providers' Interests

1. Minimizing Costs

Providers do not want to pay more to attract users than absolutely necessary. If they pay for search engine advertising, the money is money diverted from some other cause, whether advertising in another medium, other operations, or profits. Such advertising campaigns are increasingly managed like any other advertising campaign, often as part of a comprehensive

---

[69] *But see* James Grimmelmann, Note: *Regulation by Software*, 114 Yale L.J. 1719 (2005) (observing that some algorithms are so complicated that it may not be possible to say "why" a computer made a particular decision).

[70] *See, e.g.*, Hubbert v. Dell Corp., 359 Ill. App. 3d 976 (Aug. 12, 2005) (holding arbitration clause in "Terms and Conditions" on Web site enforceable).

[71] *See* Letter from Heather Hippsley, Acting Associate Director, F.T.C. Division of Advertising Practices, to [search engine company[(June 27, 2002), available at http://www.ftc.gov/os/closings/staff/commercialalertattatch.htm.

[72] *See* Rebecca Tushnet, *KinderStart: The Return*, 43(B)log, ("If Google continues to tell searchers one thing about how search results are generated and tell webmasters another, it might behoove the FTC – the only entity with a realistic chance of affecting Google – to look into the matter."), at http://tushnet.blogspot.com/2006/09/kinderstart-return.html. *But see Search King*, 2003 U.S. Dist. LEXIS 27193, at *11 ("Google and Page's statements as to the purported objectivity of the PageRank system cannot transform a subjective representation into an objectively verifiable fact.").

[73] *See* Sandeep Pandey et al., *Shuffling a Stacked Deck: The Case for Partially Randomized Ranking of Search Engine Results*, *in* Proceedings of 31st International Conference on Very Large Databases (VLDB) (2005), available at http://oak.cs.ucla.edu/~cho/papers/cho-shuffle.pdf.

[74] *See* Sunstein, *supra* note 59; Susan Gerhart, *Do Web search engine suppress controversy?*, First Monday (Jan. 2004).

[75] *See* Eric Goldman, *A Coasean Analysis of Marketing*, Wisconsin L. Rev. (forthcoming 2006).

marketing strategy—and with the same advertiser pressure for lower rates.[76] Since users could also be directed to a provider through organic search results, search advertising is a partial substitute for search rankings. There have been occasional accusations that search engines churn their rankings or deliberately demote some providers to spur them to purchase search advertising.

Even non-advertiser providers must bear some of the technical costs of search. Every time a search engine asks to index content from a provider, the provider must use a little server time and network bandwidth to respond to the request. Every time a user requests content from the provider, the provider must similarly expend a little server time and bandwidth. Especially for providers with extensive collections, who deal with multiple search engines, and who must respond to requests from many users, the burden can be substantial. [77] Most of the time, providers willingly cooperate. Being easily searchable brings more users, something so valuable that some providers create their own search engines and others pay for inclusion, placement, or advertising on search engines. It is a rare provider who complains about being searchable or about the costs of indexing, but there are some. [78]

In addition to the technical costs of being indexed, there are technical costs associated with providing content to users. It might seem paradoxical that a provider should complain about having too many users, but the fear is real. The so-called "Slashdot effect" takes its name from a popular computer news site, Slashdot.org[79]; the site's large readership of heavy Internet users has meant that a link from a Slashdot news story to a Web site can reliably be expected to produce a sudden and huge influx of traffic.[80] The linked site can expect to be hit with a potentially huge bandwidth bill[81] and to face a significant risk of seeing its servers crash under the load.[82]

Generalizing from this example, the reality of bandwidth and hosting bills emphasizes that Internet speech can be surprisingly—even unexpectedly—expensive. Moreover, a flood of new traffic may be less attentive or in some other way a different audience from the one the provider hopes to reach, making the attention into a sort of denial of service attack.[83] When the

---

[76] *See generally* Search Engine Marketing Professional Organization (SEMPO), *About Sempo*, http://www.sempo.org/about/.

[77] In 2006, 46% of all requests for pages from the author's Web site came from the Yahoo! robot.

[78] This is not the only reason a provider might wish to be findable by users but not searchable. Individuals may be more comfortable revealing personal information when they do not expect it to be easily searchable. *See generally* danah boyd, *Facebook's "Privacy Trainwreck": Exposure, Invasion, and Drama* (Sept. 8, 2006), at http://www.danah.org/papers/FacebookAndPrivacy.html. Providers may also have unfair competition objections to being searched.

[79] http://www.slashdot.org/. *See also Slashdot effect*, Wikipedia, at http://en.wikipedia.org/wiki/Slashdot_effect.

[80] *See* Jason Kottke, *Digg vs. Slashdot (or, traffic vs. influence)*, Kottke.org (Jan. 12, 2006), at http://www.kottke.org/06/01/digg-vs-slashdot.

[81] *Mark Ward, The dangers of having a good idea*, BBC News (May 3, 2003), at http://news.bbc.co.uk/1/hi/technology/2995343.stm.

[82] *But see* Mirrordot FAQ, http://www.mirrordot.org/faq/ ("system to automatically mirror any Slashdot-linked pages and ensure the content would remain available, even if the original site got clobbered").

[83] *Cf.* Christian Zappone, Help! YouTube is kililng my business!, CNNMoney.com (Oct 12, 2006), at http://money.cnn.com/2006/10/12/news/companies/utube/index.htm (describing overloads at the Web site of Universal Tube and Rollerform, utube.com, caused by misdirected users looking for youtube.com).

flood is too severe, the provider may reach *none* of the users; its server is too busy trying to keep track of the incoming requests for it to answer any of them.[84] While technologies are being developed to respond to these problems, none of them are fully effective.[85]

Turning to the law, providers have attempted to defend their servers from unnecessary indexing burdens on three principal grounds. First, they have used the common-law tort of trespass to chattels. Despite some successes for this theory,[86] the 2003 California Supreme Court's holding in *Intel v. Hamidi*[87] that the tort would not lie without a showing of harm to the chattel or the owner's ability to use it would largely (if followed elsewhere) preclude a trespass to chattels claim unless the search engine was egregiously burdensome.[88]

Providers have had more luck under *sui generis* state and federal computer intrusion statutes, which generally prohibit "access" to a computer system without "authorization." Here, courts have been willing to say that any use of a server—including spidering—that the owner does not wish to have take place is *ipso facto* "unauthorized." Despite academic criticism[89] and at least one court's discomfort with the theory,[90] courts considering claims against search engines under the federal Computer Fraud and Abuse Act's civil provisions[91] have held for the providers.[92]

Third, providers have at times alleged the existence of a contract prohibiting a search engine from spidering their content. In one notable 2004 case, the Second Circuit found that displaying a notice on information returned by a server forbidding certain uses of that information was sufficient to bind even a company accessing the information purely through a spidering program.[93] In reasoning directly applicable to search engines, the court held that the repeated access meant that knowledge of the purported contractual terms should be imputed to the spider's operator. The leading case on consumer interactions with contracts presented on the Web, *Specht v. Netscape*,[94] is in accord. While the consumer there was not bound by the contract, it was only because the contract was not clearly displayed on the web page in a way forcing the user to acknowledge it or to see it before clicking on a download link. The court left little doubt that a provider with sufficient willingness could craft terms and an interface for displaying them that would bind users who clicked through.

---

[84] *See* CERT Coordination Center, *Denial of Service Attacks*, at http://www.cert.org/tech_tips/denial_of_service.html (1997) (defining "denial of service attack").

[85] *See generally* Dave Dittrich, *Distributed Denial of Service (DDoS) Attacks/tools*, at http://staff.washington.edu/dittrich/misc/ddos/ (collecting resources on DDoS attacks and defenses).

[86] Register.com v. Verio, 356 F.3d 393 (2d Cir. 2004); Oyster Software, Inc. v. Forms Processing, Inc., No. C-00-0724 (N.D. Cal. Dec. 6, 2001); eBay v. Bidder's Edge, 100 F. Supp. 2d. 1058 (N.D. Cal. 2000).

[87] 30 Cal. 4th 1342; 71 P.3d 296 (2003).

[88] *See also* Ticketmaster Corp. v. Tickets.com, Inc., No. 99-7654 (C.D. Cal. Mar. 7, 2003) (requiring showing of actual harm "pending appellate guidance").

[89] Orin Kerr, *Cybercrime's Scope: Interpreting "Access" and "Authorization" in Computer Misuse Statutes*, 78 N.Y.U. L. Rev. 1596 (2003).

[90] *See Lockheed Martin Corp. v. Speed*, 2006 U.S. Dist. LEXIS 53108 (M.D. Fla. Aug. 1, 2006).

[91] 18 U.S.C. § 1030(g).

[92] *See* Southwest Airlines Co. v. Farechase, Inc., 318 F. Supp. 2d 435 (N.D. Tex 2004); EF Cultural Travel BV v. Explorica, Inc., 274 F.3d 577 (1st Cir. 2001), later proceedings at 318 F.3d 58 (1st Cir. 2003).

[93] *See* Register.com, *supra* note 86.

[94] 306 F.3d 17 (2d Cir. 2002).

None of these theories has been definitively settled. Some of the analyses are cursory at best, and none has yet found liability for a general-purpose search engine, rather than a specialized service that could be characterized as a direct competitor to the aggrieved provider. Nonetheless, the trend seems to be that at the least on a contractual theory, a search engine could be prohibited from indexing an unwilling provider.

On the Web, the matter appears to have settled into a rough equilibrium, with most providers using robot exclusion protocols to inform the public which robots are allowed to spider which portions of the content. Most major search engines as a routine matter of practice agree to respect such exclusions.[95] Whether content providers could demand more fine-grained control over what is later done with the content or whether search engines could demand greater access, neither has shown great desire to unwind this compromise. Note that the compromise does not directly bind users, since the norms of *robot* exclusion protocols apply only to the operators of indexing robots.[96]

Compare, however, the legal regime governing excessive user attention, which draws upon the same doctrinal sources but may work out differently in practice. It seems reasonably well-established that deliberate distributed denial-of-service attacks are both crimes and torts.[97] (Computer viruses that infect millions of computers have been used as launching pads for specific DDoS attacks.[98]) When the distributed agents are acting voluntarily under the control of their individual owners, matters are more ambiguous. Normal principles of co-conspirator liability suggest that an agreement to act in concert with thousands of others to attack a computer system might make each conspirator liable for all resulting harm.[99] "Virtual sit-ins," in which thousands of activists each run a program causing some traffic for a designated target, therefore may be actionable.[100]

Search engines enter this picture as attention lenses. They may direct a crowd to one provider as an expression of their own judgment (whether or not intended to overwhelm the provider's systems). They may also be used as instruments of others' desire to focus attention on a target. Both of these cases fit uneasily into existing categories of DDoS liability. The individual users are acting voluntarily—but they may be unaware until it is too late that their visits are in furtherance of a scheme to overload the provider.

There is at least some tentative case law holding that one who manipulates the inputs to an information location tool so that a provider is incorrectly listed may be liable in trespass to

---

[95] *See* Patricia Bellia, *Defending Cyberproperty*, 79 N.Y.U. L. Rev. 2164 (2004) (arguing that this result is for the most part efficient and just).

[96] *See* Eric J. Feigin, *Architectures of Consent: Internet Protocols and Their Legal Implications*, 56 Stanford L. Rev. 901 (2004) (arguing that Internet norms of access and restriction embedded in technical protocols are entitled to legal respect).

[97] *See, e.g.*, Computer Fraud and Abuse Act, 18 U.S.C. § 1030; *United States v. Ancheta*, No. CR-05-1060 (C.D. Cal. indictment returned Feb. 2005).

[98] *See, e.g.*, SCO, *SCO Offers Reward for Arrest and Conviction of Mydoom Virus Author* (Jan. 27, 2004), at http://ir.sco.com/ReleaseDetail.cfm?ReleaseID=127545.

[99] *See, e.g.*, United States Sentencing Guidelines § 1B1.3(a) ("[I]n the case of a jointly undertaken criminal activity . . . the base offense level . . . shall be determined on the basis of . . . all harm that resulted from . . . all reasonably foreseeable acts and omissions of others in furtherance of the jointly undertaken criminal activity." (quotation reordered)).

[100] *See generally, e.g.*, Stefen Wray, *The Electronic Disturbance Theater and Electronic Civil Disobedience* (June 17, 1998), at http://www.thing.net/~rdom/ecd/EDTECD.html.

chattels.[101] The indirection here creates difficult conceptual puzzles. The potential for unexpected harm is much greater than when the putative trespasser accesses the system directly and can modulate her own degree of use. When she acts indirectly, she may not even intend to access the given system at all, but may have intended a different target than the one ultimately accessed.[102] And, again, the voluntary nature of the intervening accesses is a complicating factor.[103]

The possibility of significant search engine lawsuits for overrecommendations, so far, remain primarily theoretical. But the potential harms are significant. If Google were to start returning the same small business site as its number-one result for every one of the two and a half billion searches performed each month, the traffic would effectively force it offline. The site's owner would seem to have a claim either under one of the access to computer systems theories discussed above, or perhaps in negligence. Drawing a line between malicious or negligent conduct and honest recommendation would require difficult interrogation of a search engine's motivations and the reliability of its processes.

2. Avoiding Unfair Competition

Search engines can cause providers any of a number of potential harms that sound in unfair competition. Here, the search engine offers users the ability to access the provider's content in a different manner than the provider would prefer, to the provider's moral or economic detriment. These harms have an obvious affinity to those caused by unwanted access *qua* access (whether the access is by search engines or by users); they involve disruption of providers' activities induced by search engines.[104] The distinction is that these unfair competition concerns are objections to the loss of properly-focused user interest felt by providers; concerns about access have to do with the burdens borne by providers in supporting search and search-directed users.

Because unfair competition claims generally require a full-cycle analysis of the plaintiff's and defendants' practices to evaluate whether given uses are permissible, it may be somewhat misleading to break down these claims against search engines by the particular activities of which providers complain. Still, because copyright claims in particular hang on particular infringing acts, that is typically how courts have broken down their analyses. A survey of allegedly unfair search practices reveals the extraordinary ingenuity of providers' attorneys and the unsettled state of this area of the law.

A search engine's spidering processes require making at least one initial copy of any content the engine wishes to index; providers have complained that this initial copy is

---

[101] School of Visual Arts v. Kuprewicz, 771 N.Y.S.2d 804, (Sup. Ct. N.Y. Dec. 22, 2003) (holding that defendant who subscribed plaintiff to high-volume email lists could be liable in trespass to chattels);

[102] *See, e.g.*, Dave Plonka, *Flawed Routers Flood University of Wisconsin Internet Time Server* (2003), at http://www.cs.wisc.edu/~plonka/netgear-sntp/.

[103] *See Presley v. Charlottesville*, No. 05-2344 (4th Cir. Sept. 22, 2006) (holding that city that published map erroneously showing public trail across plaintiff's land could be liable under Fourth Amendment for trespasses by private individuals).

[104] Provider complaints about poor search placement, on the other hand, are more fundamentally about providers' competition with other providers, in which the search engine figures when it displays unfair favoritism.

unauthorized, and hence infringing.[105] This is the first doctrinal peg on which the Authors' Guild hangs its copyright suit against the Google Book indexing project.[106] By itself, these initial copies harm only an author's abstract interest in controlling her work; courts have tended to judge their fairness in light of their view of the subsequent purposes to which the engine puts the copies.

If the search engine provides the users directly with content, it may interfere with the providers' proper attribution by severing the link between source and content. This objection appeals to trademark policies,[107] but trademark causes of action may, ironically, be hemmed in by copyright.[108] On the copyright side, unauthorized caching is a strong case for infringement, particularly because it diverts users from authors' preferred business models (whether sale of the content itself or sale of advertising targeted at users).[109] Search engines have successfully required providers to use standard technical measures to signal that they do not wish that their content be cached.[110] (Providers have also begun to argue that even copies retained but not shown to users are troublesome, because of the risk of a security breach exposing the archives to bulk copying.[111] Major archives, therefore, have adhered to the same removal-on-notice policies as have search engine caches.[112])

Issues are more ambiguous when it comes to thumbnailing and other practices that give users only excerpts or summaries of content. Some authority holds that the thumbnails are

---

[105] *See* Kelly, *supra* note 223 (finding fair use); *but see* UMG Recordings v. MP3.com, 92 F. Supp. 2d 349 (S.D.N.Y. 2000) (finding no fair use). In these cases, the line between providers and third parties is at is least distinct. *See also, e.g.*, Sony Computer Entertainment, Inc. v. Connectix Corp., 203 F.3d 596 (9th Cir. 2002) (finding fair use in initial copying for purposes of reverse engineering).

[106] *See* Authors' Guild v. Google, Inc. (S.D.N.Y. complaint filed Sept. 20, 2005).

[107] *See* 15. U.S.C. 1125(a) (making actionable a "false designation of origin . . .likely to cause confusion.")

[108] *See* Dastar Corp. v. Twentieth Century Fox Film Corp., 539 U.S. 23 (2003); *but see* F. Gregory Lastowka, *The Trademark Function of Authorship*, 85 B.U. L. Rev. 1171 (2005).

[109] *See, e.g.*, Twentieth Century Fox Film Corp. v. iCraveTV, No. 00-121, 200 U.S. Dist. LEXIS 1013 (W.D. Pa. 2000) (granting injunction against Internet sites framing TV stations' live feeds with advertisements and rebroadcasting them on the Internet.).

[110] *See* Field v. Google, 412 F. Supp. 2d 1106 (D. Nev. 2006). Providers and search engines have both been developing technical measures to improve the precision of signals in this area. *See, e.g. The Web Robots FAQ*, http://www.robotstxt.org/wc/robots.html; *Publishers aim for some control of search results*, C|Net News.com (Sept. 22, 2006), http://news.com.com/2102-1030_3-6118523.html. *But see* Niva Elkin-Koren, *What Contracts Can't Do: The Limits of Private Ordering in Facilitating a Creative Commons*, 74 Fordham L. Rev. 375 (2005) (arguing that increased precision of permissions increases perception that permission is required for any reuse).

[111] *See* Testimony of Paul Aiken on Behalf of the Authors Guild Before the House Committee on Energy and Commerce Subcommittee on Commerce, Trade, and Consumer Protection (Nov. 16, 2005) ("Since there's no license needed, in Google's view, Google doesn't have to give rightsholders contractual assurances of the security of their database.").

[112] *See, e.g.* Internet Archive, *Removing Documents from the Wayback Machine*, http://www.archive.org/about/exclude.php. *But see* Rebecca Bolin, *Locking Down the Library* (draft on file with author) (arguing that copyright-related removal from online archives threatens preservation of human memory). Failure to understand how archives work has generated some strange lawsuits. *See, e.g.*, Tom Zeller, Jr., *Keeper of Expired Web Pages Is Sued Because Archive Was Used in Another Suit*, N.Y. Times, (July 13, 2005).

protected fair use;[113] this holding is not, however, clearly established and drawing the line between thumbnail and full copy (or derivative work) may require significant case-by-case analysis.[114] Some providers, moreover, have argued that thumbnailing—or even offering search itself—involves the exploitation of value produced by content and properly attributable to that content.[115] Precisely because this related market is valuable, goes the argument, the content owners should have the exclusive right to exploit it.[116]

Even when users obtain the content directly from providers, the search engine may have altered the means to providers' disadvantage. If the engine provides the user with specific technical instructions for obtaining content, the user may experience the content in a context not intended by providers, even though the provider itself supplied the content to the user upon request. Providers have been particularly upset by techniques that bypass the advertising they wish to show users.[117] Courts have not been able to agree on the copyright implications of deep linking, framing, and inlining.[118] Because these techniques distort the providers' preferred presentation, they also raise trademark issues.[119] Further, European courts have also been willing

---

[113] Kelly, *supra* note 223.

[114] *See* Perfect 10 v. Google, Inc., 416 F. Supp. 2d 828 (C.D. Cal 2006) (holing thumbnailing likely not a fair use as against third-party copyright holder; opinion based in part on third party's licensing of reduced-size images for use on cell phones).

[115] *See, e.g.* Nick Taylor, . . . *But Not at Writers' Expense*, Washington Post (Oct. 21, 2005) ("The alphabet ought to be free, most certainly, but the people who painstakingly arrange it into books deserve to be paid for their work."). *But see* Emily Anne Proskine, *Note: Google's Technicolor Dreamcoat: A Copyright Analysis of the Google Book Search Library Project*, 21 Berk. Tech. L.J. 213 (2006) ("None of copyright's exclusive rights suggest that publishers or authors should possess a monopoly over the indexing and searching of their works")

[116] *But see* Rochelle Dreyfus, *Expressive Genericity: Trademarks as Language in the Pepsi Generation*, 65 Notre Dame L. Rev. 397, 405 (1990) ("Furthermore, fallacies in the fundamental assumptions made by courts that have approved this "if value, then right" theory mean that the right lacks a coherent limit.").

[117] *See, e.g.*, Washington Post Co. v. Total News, Inc., No. 97 Civ. 1190 (S.D.N.Y. complaint filed Feb. 20, 1997) ("Yet an advertisement on one of Plaintiffs' sites, when seen through the totalnews.com window, is reduced in size, may even be totally obscured by the totalnews frame, and is forced to compete for the user's attention with the visual clutter of the totalnews.com frame, including other advertising -- possibly including advertising for directly competitive products.")

[118] *See* Kelly v. Arriba Soft, 280 F.3d 934 (9th Cir. 2002), *withdrawn by* 336 F.3d.811 (9th Cir. 2003) (finding that deep linking and framing infringe on exclusive public display right).

[119] *See* Farhad Manjoo, *Public Protests NPR Link Policy*, Wired News (June 20, 2002), at http://www.wired.com/news/business/1,53355-0.html; Digital Equip. Corp. v. AltaVista Tech., 960 F. Supp. 456 (1997) (enjoining defendant from linking to plaintiff and creating false impression of affiliation); Knight-McConnell v. Cummins, 2004 U.S. Dist. LEXIS 14746 (S.D.N.Y. 2004) ("The mere appearance on a website of a hyperlink to another site will not lead a web-user to conclude that the owner of the site he is visiting is associated with the owner of the linked site."). Compare the problems raised by popup advertising and the struggles faced by courts trying to determine whether triggering ads based on the visited site infringers the copyright or trademark interests of the site's provider. *See, e.g.*, 1-800-Contact v. WhenU.com, Inc., 414 F.3d 400 (2d Cir. 2005) (popup advertisements based on URL do not involve "use in commerce" of underlying site's trademarks); Wells Fargo & Co. v. WhenU.com, Inc., 293 F. Supp. 2d 734 (E.D. Mich. 2003) (same); U-Haul Int'l, Inc. v. WhenU.com, Inc., 279 F. Supp. 2d 723 (E.D. Va. 2003) (same).

to find that deep linking to news violates the European Database Directive.[120] An additional set of complications arise when the relevant content is uncopyrightable (e.g. because it consists of unprotectable fact, rather than protectable expression). Here, providers may wish to complain both of direct provision and of altered presentation, but must steer clear of copyright preemption.[121]

Finally, these questions cannot be addressed without consideration of users' rights, because in these contexts search engines are generally acting on behalf of users. Users enjoy significant statutory and fair use rights to make copies for their own caching and archival purposes.[122] Further, there are strong arguments that many of the transformations to which search engines subject content would be wholly legal if carried out by users directly. The role of the search engine as an intermediary carrying out those activities for them raises deeply contentious issues of intellectual property policy.[123]

3. Prominent Placement in Results

Prominent placement in search results can be immensely valuable. Users are likelier to click on the first result than on the second, on the second than on the third, and so on. Each page beyond the first that users must look imposes an even more severe drop-off in interest. Providers whose sites appear beyond the first few pages of results may as well not exist, as far as that particular query is concerned. Given this value, providers seek placement eagerly. Some turn to SEO, seeking to redesign their offerings to better satisfy the search engine's criteria. Others turn to the law.

When a search engine returns a result R in response to query Q, it might be understood in context as asserting that "R is a good source for information about Q," or perhaps that "R is a good place to acquire Q." Assertions of this form implicate the policies of trademark and advertising law, which seek to keep consumers from being confused by misleading claims about the relationship between businesses and the goods they offer.

Trademark law has taken a strong interest in the Internet information location tools used by consumers. First, there were the cybersquatting cases, in which trademark holders sued those using domain names containing or similar to a trademarked term. The sources of law here are many, [124] but the trend has been that one registering a domain name solely to sell it to the

---

[120] Newsbooster v. Danish Newspaper Publishers' Association (Denmark); Copiepress v. Google (Belgium). In a remarkable outgrowth of the last, Google was ordered to post the court order on its Belgian homepage. *But see* Home v. Ofir (Denmark) (finding deep linking legal); Zoekallehuizen.nl v. NVM (Netherlands) (same). *But but see* Bixee.com v. Nahuri.com (India) (finding deep linking illegal).

[121] *See* Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340 (1991); Tickets.com, *supra* note 86 (applying *Feist*).

[122] *See* Sony Corp. of America v. Universal City Studios, Inc., 464 U.S. 417 (1984); Jessica Litman, Lawful Private Use (draft), available at http://www-personal.umich.edu/~jdlitman/papers/LawfulPersonalUse.pdf.

[123] *See, e.g.*, Marvel Enters. v. NCSoft Corp., No CV 04-9253, 2005 U.S. Dist. LEXIS 8448 (C.D. Cal. 2005) (considering copyright and trademark implications of actions by users that would have been lawful private uses offline); Paramount v. ReplayTV (C.D. Cal. complaint filed Oct. 31, 2001) (objecting to feature of personal video recorder allowing consumers to skip commercials automatically).

[124] *See* Internet Corporation for Assigned Names and Numbers, *Uniform Domain-Name Dispute Resolution Policy* (UDRP), at http://www.icannorg/udrp/udrp.html (2001); Anti-cybersquatting Consumer Protection Act (ACPA), 15 U.S.C. §§ 1117 and 1125(d); *e.g.* Lamparello v. Falwell, 420 F.3d 309 (4th Cir. 2005).

trademark holder will often be forced to transfer it, with other cases being resolved under national principles of trademark laws.[125] In this context, United States courts have been willing to find actionable consumer confusion in the initial diversion of a user to the domain, even if the user is not ultimately confused as to the source of the information there.[126]

In a decision that has attracted substantial criticism,[127], the Ninth Circuit adapted trademark reasoning to deal with search engines, holding that the use of trademarks in hidden metatags (a use designed to influence a search engine's ranking decisions) could be actionable as causing "initial interest confusion."[128] Courts have split whether purchasing search advertisements triggered by trademarked keywords can give rise to trademark infringement liability.[129]

Three years later, the Ninth Circuit then extended the doctrine so that liability could run against search engines themselves.[130] The logic was that a search engine's use of a trademarked term as an advertising keyword was use of the trademark capable of causing confusion, whether or not the keywords were used in the advertisements themselves. Other courts have held that adware vendors, whose software displays advertisements based on the Web pages visited by users, do not violate the trademark rights of the providers of the underlying Web pages.[131]

These precedents provide conflicting guidance on the obligations of search engines in triggering advertisements based on search queries containing trademarked terms, and search engines have vacillated in their policies in selling such advertisements.[132] Google, which has been the most vigorous of the major search engines in contesting suits by trademark holders, has both won and lost on nearly identical facts.[133] So far, no search engine has been sued for

---

[125] *See* UDRP, *supra* note 124 (providing expedited remedies only for "[d]isputes alleged to arise from abusive registrations of domain names (for example, cybersquatting)"); 15 U.S.C. § 1125(d) (providing cause of action only for registration of domain name with "bad faith intent to profit").

[126] *See, e.g.*, People for the Ethical Treatment of Animals v. Doughney, 263 F.3d 359 (4th Cir. 2001). *But see* Lamparello, 420 F.3d at 316-17 (distinguishing *Doughney*).

[127] *See* Eric Goldman, *Deregulating Relevancy in Internet Trademark Law*, 54 Emory L.J. 507 (2005); Playboy Enterprises, Inc. v. Netscape Communications Corp., 354 F.3d 1020, 1034-36 (9th Cir. 2004) (Berzon, J., concurring).

[128] Brookfield Communications Corp. v. West Coast Entertainment Corp., 174 F.3d 1036 (9th Cir. 2001).

[129] *Compare* Edina Realty, Inc. v. TheMLSonline.com, No. Civ. 04-4371, 2006 WL 737064 (D. Minn. Mar. 20, 2006) (yes) *with* Merck & Co. v. Mediplan Health Consulting, 425 F.Supp. 2d 402 (S.D.N.Y. Mar. 30, 2006) (no) and Rescuecom Corp. v. Google, Inc., No. 5:04-CV-1055 (N.D.N.Y. Sept. 28, 2006) (no).

[130] Playboy, 354 F.3d at 1024-29.

[131] 1-800-Contacts, Inc. v. WhenU.com, Inc., 414 F.3d 300 (2d Cir 2005); Wells Fargo & Co. v. WhenU.com, Inc., 293 F. Supp. 2d 734 (E.D. Mich. 2003); U-Haul International, Inc. v. WhenU.com, Inc., 279 F. Supp. 2d 723 (E.D. Va. 2003). These cases also raised copyright theories tied to the allegedly altered display of the web pages.

[132] *Compare* Google, *AdWorlds Trademark Complaint Procedure*, at http://www.google.co.uk/tm_complaint_adwords.html (forbidding only "use of the trademark in ad text") *with* Yahoo! Search Marketing, *Trademarks*, at http://searchmarketing.yahoo.com/legal/trademarks.php (forbidding bidding on trademarked terms except under stated conditions).

[133] *See* Government Employees Insurance Co. v. Google, Inc., No. 01:04cv507, 2005 U.S. Dist. LEXIS 18642 (E.D. Va. 2005) (finding for search engine in bench trial), 330 F. Supp. 2d 700 (E.D. Va. 2004) (dismissing theory of trademark infringement for use in advertising keywords but allowing theory of infringement for use in advertising text to proceed); Google Inc. v. American Blind & Wallpaper

returning a competitor's web page as an organic search result in a search on a trademarked term, but such suits cannot be ruled out and some have proposed that a similar result be required by regulation.[134]

Providers aggrieved by their poor placement have also sued search engines for business libel and related theories.[135] The providers have generally argued either that their content was objectively highly relevant and that a search engine ranking it poorly is in effect lying; or that the search engine acted out for malicious reasons in reducing the providers' rankings. These state-law claims have not thus far fared well in court. It may not have helped some providers' legal cases that their pre-litigation actions could arguably be characterized as black-hat SEO.[136]

C. *Third Parties' Interests*

1. Ownership

We turn now to third parties' interests in suppressing certain content flows. We begin with their intellectual property interests in such flows, because it is here that their legal ability to intervene has been at its zenith. This problem also introduces a new political dynamic. The providers of infringing content and the users acquiring it have a common interest in the flow; the copyright holder (or less often, the trademark holder) is their common enemy. The search engine typically has a business interest in serving users and providers, one counterbalanced by any legal pressure the intellectual property holders may be able to bring to bear, and by the possibility of turning them into providers.[137] Note also that the doctrinal analyses available here have much in common with the analyses used when a provider asserts an intellectual property claim against a search engine.

Search engines might trigger direct copyright claims by providing content to users. The substantial weight of authority now holds that neither linking to infringing content nor framing constitutes direct infringement.[138] The initial copies a search engine makes in the process seem

---

Factory, Inc., No. 03-05340, 2005 U.S. Dist LEXIS 6228 (N.D. Cal. 2005) (allowing theory of infringement for use in keywords to proceed); Elinor Mills, *Google loses French trademark lawsuit,* CNet News.com (June 28, 2006) (reporting on similar French decisions in favor of trademark holders).

[134] *See, e.g.*, Frank Pasquale, *Rankings, Reductionism, and Responsibility*, 54 Clev. St. L. Rev. 115 (2006) (proposing that trademark holders be given right to place asterisk next to unsanctioned search results); James A. Rossi, *Protection for Trademark Owners: The Ultimate System of Regulating Search Engine Results*, 42 Santa Clara L. Rev. 295, 347-54 (proposing that search engines be required to offer option of returning results selected by holder of trademarked search term).

[135] Datner v. Yahoo! Inc, No. BC355217 (Cal. Sup. Ct. complaint filed July 11, 2006); KinderStart v. Google, No. 5:06-cv-02057-JF (N.D. Cal. July 13, 2006); Langdon v. Google, 1:06-cv-00319-JJF (D. Delaware complaint filed May 17, 2006); Roberts v. Google, No. 1-06-CV-063047 (Cal. Sup. Ct. May 12, 2006) (voluntary dismissal); CLRB Hanson Industries LLC v. Google, Inc., No. 1-05-CV-046409 (Cal. Sup. Ct. complaint filed August 3, 2005); Search King, Inc., v. Google Technology, Inc., No. CIV-02-1457-M, 2003 U.S. Dist LEXIS 27193 (W.D. Okla. 2003).

[136] *See* Dahlia Lithwick, *Google-Opoly: The Game No One but Google Can Play*, Slate.com (Jan. 29, 2003) ("SearchKing in effect has its clients collude to trick Google into boosting everyone's ratings."). *Cf.* David Kesmodel, *Blogger Faces Lawsuit Over Comments Posted by Readers*, Wall Street Journal (Aug. 31, 2005) (describing defamation lawsuit against blogger whose commenters criticized plaintiff's SEO tools).

[137] *See, e.g.*, Brad Stone, *The Battle Over YouTube*, Newsweek (Oct. 9, 2006) (discussing YouTube's distribution deal with Warner Music).

[138] *See* Perfect 10 and cases cited therein.

to be unambiguous fair uses, at least where the search engine has no actual knowledge that the provider is an infringer.[139] One court has held, however, that thumbnailing is not a fair use, at least as against a third party that has attempted to exploit its exclusive rights by selling thumbnail-sized images.[140] Caching and archiving would seem to raise similar concerns, and most search engines have been behaving particularly cautiously in removing allegedly infringing content from their caches.[141]

Secondary liability for infringement is also a strong possibility. *Grokster* teaches that the makers of infringement-facilitating technologies must both pass the *Sony* staple article of commerce test and steer clear of purposeful, culpable inducement of infringement.[142] Given that Grokster and its brethren consisted of a search application fused with a file-transfer application, this holding applies directly to search engines.[143] The principal Web search engines easily pass the "capable of substantial noninfringing uses" prong of the *Sony* test, but more specialized search engines may not. As for purposeful, culpable inducement, the application of this language to many technologies, not just search engines, remains unclear.[144]

The immunities and subpoena processes detailed in Section 512 of the DMCA[145] will also be significant in search engine copyright litigation. ISP litigation has clarified some of the issues, but Section 512(d), the immunity for "Information Location Tools," has not yet been extensively glossed by courts. That section incorporates by reference the ISP-focused notice-and-takedown procedure for "Information Residing on Systems or Networks At Direction of Users" described in Section 512(c), and search engines have been diligent about removing links for which they receive notices alleging copyright infringement.[146] But the parallel is not exact. On the one hand, because search engines do not have the direct relationship with users that hosting services do, providers are more vulnerable to abuse of the notice-and-takedown process. On the other hand, because the search engine's role in such cases is generally only to link to information, and because the notices must specify the "reference or link, to material or activity claimed to be infringing," search engines can generally undermine a takedown notice by

---

[139] *See* Kelly.

[140] *See* Perfect 10.

[141] *See, e.g.*, *Site Owner Help: Control which of your pages are indexed*, http://search.msn.com.sg/docs/siteowner.aspx?t=SEARCH_WEBMASTER_REF_RestrictAccessToSite.htm. ("Prevent MSNBot from caching a page"). *But see* 17 U.S.C. § 512(b) (providing infringement safe harbor for "system caching"); Parker v. Google, Inc., 422 F. Supp. 2d 492 (E.D. Pa. 2006) (holding that caching by search engines falls within § 512(b) as against third-party copyright holder).

[142] MGM Studios, Inc. v. Grokster, Ltd., 545 U.S. 913 (2005).

[143] *See id.*

[144] *Compare Grokster*, 545 U.S. at ___ ("one who distributes a device with the object of promoting its use to infringe copyright, as shown by *clear expression or other affirmative steps to foster infringement*, is liable") *with Grokster*, No. 01-08541, (C.D. Cal. Sept. 27, 2006) ("need prove only that StreamCast distributed the product with the intent to encourage infringement").

[145] 17 U.S.C. § 512.

[146] Some providers, both in the ISP and search engine contexts, have used such notices to enforce non-copyright-related desires to suppress the material. *See* Online Policy Group v. Diebold, Inc., 337 F. Supp. 2d 1195 (N.D. Cal. 2004) (finding "material misrepresentation" of infringement where posted "embarrassing" content was "not subject to copyright protection"). *Cf.* John Hiler, *Church v. Google*, Microcontent News (Mar. 21, 2002), at http://microcontentnews.com/articles/googlechurch.htm ("[T]he Church may have used a legal complaint to suppress more than just copyrighted material . . . .").

displaying the notice itself.[147] It is also unclear how other provisions of Section 512, such as the termination-of-repeat-infringers requirement of Section 512(i) and the subpoena-to-identify-infringers provision of Section 512(h), apply to search engines.

Cutting across all of these copyright issues are the general problems of what it means for a search engine to have "knowledge" of infringement and the extent to which a search engine profits from infringing activity or can control infringers.[148] Napster was charged with knowledge of particular infringing MP3 files based on notifications from copyright holders.[149] Recently, a copyright holder has won an injunction allowing it serve Google with notice of particular infringing images.[150] Much depends on the scope of such injunctions—whose burden is it to identify whether a given image is so similar to one on the forbidden list that it, too, should be blocked? If the burden is the copyright holder's, the injunction may be nearly useless because infringing providers can make slight tweaks to images; if the burden is the search engine's, image search may need to be shuttered.[151] Identification of infringing providers is also a difficult issue, given that such identifications may effectively be *ex parte*. And the assessment of vicarious liability will also involve some close scrutiny into search business models; it appears that Google's affiliate-network advertising might be decisive in finding vicarious liability under the right circumstances.[152]

Trademark issues are generally simpler and less dangerous for search engines. Direct liability in trademark for trademark use on providers' pages seems unlikely, given the trend in the keyword advertising cases. Secondary liability, based on tests paralleling those in copyright, might be argued by analogy to the offline "swap meet" cases.[153] eBay, the most like a swap meet

---

[147] *See* Google, *Digital Milennium Copyright Act*, at http://www.google.com/dmca.html. ("Please note that a copy of each legal notice we receive is sent to a third-party partner for publication and annotation. . . . A link to your published letter will be displayed in Google's search results in place of the removed content."); Joshua Urist, Who's Feeling Lucky? Skewed Incentives, Lack of Transparency, and Manipulation of Google Search Results Under the DMCA, 1 Brook. J. Corp. Fin. &. Comm. L. 219 (2006) (arguing that such 512(d) takedown requests should be publicly archived).

[148] The question of knowledge is relevant both in the standard for contributory infringement and directly under Section 512; the ability to control infringers is relevant in vicarious infringement and directly under Section 512.

[149] *See* A&M Records v. Napster, 284 F.3d 1091 (9th Cir. 2002).

[150] Perfect 10. In the actual case, the ambiguous initial order was followed by an injunction that instituted a Section 512-style takedown regime, even though Section 512 was not directly applicable.

[151] The difficulty is that while it is easy to test quickly whether one file resembles another, and it is easy to test whether one file is identical to any of a large set of files, no practical general algorithm yet exists to test whether one file resembles any of a large set of others. (The question is the subject of active research efforts.) But such an algorithm is what is necessary to test whether a file a search engine encounters while browsing the Web is just a slightly altered version of one on a list of protected files supplied to it. Thus, in the third-party copyright setting, whichever party's technical burden requires the use of such an algorithm will lose.

[152] *See* Perfect 10 (noting that Google placed affiliate advertisements on sites allegedly hosting infringing images).

[153] *See, e.g.*, Hard Rock Cafe Licensing Corp. v. Concession Services, Inc., 955 F.2d 1143 (7th Cir. 1992) (discussing contributory and vicarious trademark liability for operator of flea market where counterfeit goods are sold).

of the major search engines, has a rigorous trademark protection policy and will take down an auction based on a complaint from a trademark holder.[154]

The case of third-party IP rights, once again, implicates the interests and rights of other constituencies: here, both providers and users (who, for some search systems, such as peer-to-peer ones, may be the same constituency). A familiar concern from other debates over the makers of intermediary technology and services is the degree to which search engines will over- or under-represent the interests of those groups, and the extent to which the search engine should or should not be asked to assert those interests in litigation.

2. Reputation

How different things are when the content flows are defamatory, rather than infringing! Here, search engines are protected by 47 U.S.C. § 230, which gives any "provider . . . of an interactive computer service" blanket immunity from being treated as the "speaker of any information provided by another information content provider or user." Moreover, while search engines need not filter such material, they are also immunized if they voluntarily "in good faith" remove material they believe "obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable."[155] Except in exceptional circumstances, a party defamed or personally harmed by content whose distribution is facilitated by a search engine has no recourse against the engine.[156]

Those circumstances might arise when the engine itself has taken sufficient steps that it could be identified as the provider of the content. First, if it has encouraged the creation of the content and directed the creation, it might be identified with the provider for liability purposes.[157] This scenario is not a significant concern for a pure search engine, but the growing integration of search engines with other applications raises concerns, particularly for search engines associated with creative communities.[158] Second, to the extent that a search engine is viewed as a speaker—something search engines are eager to encourage in the context of defending

---

[154] *See* eBay, *How eBay Protects Intellectual Property (VeRO)*, http://pages.ebay.com/help/tp/programs-vero-ov.html; *But see* Tiffany (NJ) Inc. v. eBay Inc. (S.D.N.Y. complaint filed June 2004) (claiming eBay secondarily liable for sales of trademarked goods). eBay is potentially more liable than most standalone search engines under a vicarious liability theory, since it has the ability to disable any auction on its site. Contributory trademark liability might be more broadly generalizeable beyond eBay.

[155] 47 U.S.C.§ 230.

[156] *See* Chicago Lawyers' Committee for Civil Rights Under the Law, Inc. v. Craigslist, Inc., No. 1:06-CV-00657 (N.D. Ill. Nov. 14, 2006) (applying § 230 to preempt a claim under the Fair Housing Act against an online classified ads site and search engine for the discriminatory housing ads posted by its users). For a remarkable example of the broadness of this immunity, consider Sturm v. eBay, No. 1-06-CV-057926 (Cal. Superior Ct. July 27, 2006), in which eBay was not required to remove defamatory feedback about a user, even when the user and the defamer had entered into a settlement stipulating that the feedback was defamatory and both had written eBay asking that it be removed. *See* Elise Ackerman, *EBay lawsuit reveals foibles of site feedback*. San Jose Mercury News (Aug. 9, 2006), at http://www.mercurynews.com/mld/mercurynews/news/local/15228670.htm.

[157] *See generally* Ken S. Myers, *Wikimmunity: Fitting the Communications Decency Act to Wikipedia*, 20 Harvard J.L. & Tech. (forthcoming 2006) (discussing § 230 cases and the extent to which a service can encourage users to provide defamatory content without losing its immunity).

[158] Google, for example, owns Blogger and YouTube. Yahoo operates Yahoo! Groups, Yahoo! 360, and Flickr. MSN runs Windows Live Spaces.

themselves in suits over rankings—its recommendations of content potentially become endorsements of that content's message. Applicable precedents hold that services have substantial leeway to choose which messages to pass along, but if the search engine itself adds some content of its own to the recommendation (even, for example, a sentence describing the linked-to page[159]), that additional content might fall outside of Section 230's protections.[160]

The presence of search engines, however, also substantially reshapes the dynamics of the struggle between defamer and defamee. Search engines can focus attention on statements that previously would not have spread as far or as fast. Because they are also selective in pairing a query with results, they can firmly link a name to a given piece of information. Indeed, precisely because people may wish to search on others' names, search engines regularly direct them to false and defamatory claims.[161] Arguably, search engines also help diligent third parties discover unflattering information about them before it has spread, allowing them to move directly against the providers.[162] Search engines can rarely be expected to possess information about providers that would be useful in identifying them and facilitating legal action, but there might be unusual cases in which they might facilitate such suits.

One consequence of the broad immunity search engines and other intermediaries enjoy has been to encourage reputational self-help.[163] Companies have been known to engage in substantial SEO tactics to drive unflattering messages about them from search engine prominence.[164] Or consider again the practice of Googlebombing, which can amount to a distributed attempt to tarnish another. In work in progress, Mike Godwin has argued that the increasing democratization of Internet communications technologies means that self-help should be the response of choice; one confronted with an online falsehood should be encouraged first to propagate the truth online.[165] Given the enormous power that search engines wield in shaping which messages are heard and which are not, it is not obvious that the truth will necessarily be able to catch up with the falsehood.[166]

In light of the overall tilt towards defamers and against defames that this technical and legal landscape creates, some scholars have argued that the law of search should be modified (or

---

[159] Yahoo's directory describes each listed site in a sentence. *See* http://dir.yahoo.com. The Open Directory Project does the same, with the taxonomy and descriptions being supplied by volunteers. *See* http://dmoz.org/.

[160] *Cf.* Benjamin Cohen & Helen Nugent, *Cole tackles Google over gay link*, Times (London) (Mar. 7, 2006), available at http://www.timesonline.co.uk/article/0,,2-2073055,00.html (explaining that Google search on "ashley cole" returned "See results for: ashley cole gay," and with Google spokesperson explaining that alternative search suggestions are determined by computer algorithm).

[161] For a domain-specific example, consider Don't Date Him Girl, http://dontdatehimgirl.com/, which allows users to search for personal reports posted by other users about cheating men—by name. At least one alleged cheater has decided to take his chances with Section 230 in a lawsuit against the site. *See* Carl Jones, *Scorned Attorney Sues Kiss-and-Tell Web Site*, Daily Business Review (July 5, 2006), *available at* http://www.law.com/jsp/article.jsp?id=1151658319991.

[162] *See* Daniel Dassey, *A quick Self-Google once a day to guard your reputation*, Sidney Morning Herald (May 23, 2004), available at http://www.smh.com.au/articles/2004/05/22/1085176043551.html.

[163] *See* R. Polk Wagner, *On Software Regulation*, 78 S. Cal. L. Rev. 457 (2005) (arguing that denying online actors legal recourse will encourage them to resort to technical self-help).

[164] *See* Glaser, *supra* note 192.

[165] Mike Godwin, *Libel Law—Time to Let it Die?*, Presentation to Yale Information Society Project (Sept. 19, 2006).

[166] *See* Cho & Roy, *supra note 66*.

interpreted) to provide, at least, some kind of right of reply.[167] By analogy to the Fair Credit Reporting Act, which allows individuals to correct incorrect statements about their credit history (and provides mechanisms to proceed both against original reporters and credit record search agencies), they have suggested that search engines should be required to respond to certain well-specified classes of reputational harms, for example by allowing the subject of an unflattering search result to annotate the result with an asterisked link to her reply.

3. Privacy

Search engines give users remarkable ability to learn about others. They can both root out details that otherwise might have remained obscure and correlate information from many different sources. They democratize investigation, giving anyone the power to develop a broad and deep profile of their chosen target.[168] Even search engines themselves are uncomfortable with the privacy-invading power of their own technologies. Google reacted with petulant anger to a CNet article on privacy concerns raised by search that included an extensive profile of Google CEO Eric Schmidt compiled through use of Google, including his political contributions, wife's name, and hobbies.[169] Google refused to talk to CNet reporters for a year.[170]

The legal baseline when it comes to search subjects' privacy, as with defamation, is that search engines cannot be held liable for the information they pass along.[171] But there are also important structural differences worth noting. First, in many cases, the subject cannot even hold the provider responsible—except in extreme cases or specific subject matters, revealing private information is in general not a tort or a crime. Second, the harm is more typically tied to aggregation than to specific pieces of information. Third, the user who engages in privacy-intrusive searches is typically more actively culpable than the user who is exposed to defamatory content.

And finally, the release of private information is somewhat more tied to the actions of the subject than defamatory information would be—anyone can make up lies about me, but my cooperation is typically needed for anyone else to learn certain true facts about me. That said, the choice of remaining offline is less and less available, and one may have no ability at all to opt out of many data flows. Search engines, in effect, provide commodity database aggregation—whatever has leaked onto the public Internet can be correlated, with some diligence, through clever searching.

---

[167] *See* Pasquale.

[168] *See* Randy Cohen, *The Ethicist: The Way We Live Now*, New York Times (Dec. 15, 2002) (discussing ethics of Googling potential dates). *Cf.* Kevin Poulsen, *MySpace Predator Caught by Code*, Wired News (Oct. 16, 2006), at http://www.wired.com/news/technology/1,71948-0.html (using publicly available datasets to locate registered sex offenders with Myspace profiles).

[169] Elinor Mills, *Google balances privacy, reach*, C|Net News.com (Aug. 3, 2005), at http://news.com.com/2102-1032_3-5787483.html.

[170] *See* Saul Hansell, *Google's Chief is Googled, To the Company's Displeasure*, New York Times C4 (Aug. 8, 2005). The ban was dropped *sub silentio* within a few months. *See* Elinor Mills, *Google to Yahoo: Ours is bigger*, C|Net News.com (Sept. 28, 2005), at http://news.com.com/2102-1038_3-5883345.html (C|Net article based on phone interview with Schmidt).

[171] One significant exception might pertain to the right of publicity, which, if classified as "intellectual property," would survive the Section 230 immunity search engines enjoy. Michael Carroll has noted that this connection between privacy and IP interests is salient in the context of the *Perfect 10* litigation.

The problem of third-party privacy is both closely linked to and different in kind from the problem of user privacy.[172] User queries raise concerns of direct trust with sensitive information; third parties stand in no such relationship to the engine.[173] User queries are valuable in the aggregate for marketing purposes; third parties are more vulnerable to individual targeting (as by stalkers). There is an argument from symmetry that search engines should disclose to search subjects that others are searching for information about them, possibly even including the names of the searchers. At least within closed online communities, social norms often favor some rough balance of privacy reciprocity among searchers and searchees.[174]

4. User Virtue

The most controversial aspect of search engine policy may also be the one least constrained by existing law. When third parties object to certain content flows on the ground that those flows are intrinsically harmful, the debates over whether those flows should be suppressed are inherently political. Most of the arguments have involved government attempts to require search engines to filter results or block certain keywords from being searched.[175] These governments defend their actions on the ground that they are protecting their citizens' virtue. Sometimes, as in attempts to block hate speech or pornography, it is personal virtue. Sometimes, the virtue is political, as in attempts to prevent access to seditious or dissenting points of view.

Here, the policy and legal issues for search engines connect to one of the oldest and most contentious debates in Internet law: whether ISPs must, may, or must not filter out dangerous traffic flowing through their networks. Different national cultures and laws have treated different forms of Internet speech as dangerous. Child pornography touches a nerve in the United States; in Europe it is hate speech; in China it is political speech. Government-mandated filtering and blocking at the network layer is not new; consider the Great Firewall of China[176] and several United States state and federal laws (declared unconstitutional or preempted by Section 230) that would have required intermediary blocking.[177]

---

[172] Provider privacy, while theoretically a possibility, does not raise many substantial issues in practice. Providers who do not wish to be found by searching usually opt off the public Internet entirely. Accidental leaks of information found by search engines raise issues similar to those raised by third-party privacy concerns. That said, a party who accidentally releases information harmful to itself is almost certainly the least-cost avoider as compared with a search engine.

[173] Self-help is not entirely out of the question. Douglas Coupland's novel *JPod* includes a character who creates a juicy but false Web site about herself (stating that she lost hundreds of pounds on the Subway diet but has cheated by sneaking junk food snacks), partly for amusement but partly to throw searchers off the track of true information about her.

[174] *See* Lior Strahilevitz, *Friendster and Symmetrical Privacy*, University of Chicago Law School Faculty Blog (Oct. 6, 2005), http://uchicagolaw.typepad.com/faculty/2005/10/friendster_and_.html.

[175] One exception is Toback v. Google, Inc., No. 06-007246 (NY Sup. Ct. complaint filed May 4, 2006), a private suit in nuisance and intentional infliction of emotional distress that accused Google of allowing traffic child pornography. The causes of action were obviously preempted by Section 230, and the plaintiff, a New York state legislator, dropped it within two months, after Google "offered to sit down and discuss the issues."

[176] *But see Hacking the Great Firewall of China*.

[177] For now, some endpoint-abetted controls are surviving judicial scrutiny in the United states. Public libraries may be required to install filtering software on their computers; adult website providers are governed by strict age-certification requirements for the performers they show.

The extension of such requirements to search engines is also not new. In 2000, French groups discovered that Yahoo's chat rooms and auctions included Neo-Nazi material, and ultimately won a court order requiring Yahoo to block French access not merely to such auctions, but also "to any other site or service that may be construed as an apology for Nazism or contesting the reality of Nazi crimes." Yahoo's attempt to block United States enforcement of the order was ultimately inconclusive; long before the Ninth Circuit dismissed Yahoo's declaratory judgment action for lack of personal jurisdiction, Yahoo! had decided that it preferred to do business in France on French terms and came into compliance with the order. This has been the pattern for many United States technologies companies doing business abroad ever since—initial protestations, followed by compliance with local law.[178]

China is today the site of the most aggressive search-blocking regime.[179] Search engines—like most other Internet intermediaries—are expected to self-monitor to block a slightly nebulous set of forbidden content (the vagueness of the standards leads to inconsistent and unpredictable overblocking).[180] Falun Gong and political dissidence are the principal, but by no means the only, targets. These policies—particularly when adopted by "do no evil" Google—have led to a domestic political backlash, including the introduction of the Global Online Freedom Act, which would have required search engines not to alter results at the request of governments of "internet restricting countries" and required that search engines notify the United States government of their blocklists. The complexity of varying international standards and the concern that compliance with local law may lead to human rights violations have led some companies to ask the United States government for help, possibly through using trade negotiations to push for more rights-protective laws abroad.[181]

Governments have also learned that intermediaries can be asked or forced to identify Internet users who are exchanging forbidden content. One Chinese journalist convicted and imprisoned with information supplied by Yahoo! is apparently about to sue Yahoo! in the United States.[182] Search companies have been inconsistent in their willingness to contest demands for private information from local authorities—contrast Yahoo's response in the Shi Tao case with Google's strong resistance to the Department of Justice subpoena in *Gonzales v. Google* or its

---

[178] *See* JACK GOLDSMITH & TIM WU, WHO CONTROLS THE INTERNET? (2006). Yahoo!, having learned this lesson first, has been the most complaisant of the United States search engines in adhering to local repression. Wikipedia has refused to implement self-censorship to satisfy Chinese authorities; blocking of Wikipedia in China has been inconsistent.

[179] *See* Thompson, *supra* note 60.

[180] This self-censorship system is not unique to China; search providers in Germany use a similar system to suppress hate speech.

[181] *See* Anne Broache, Web Giants Ask for Feds' Help on Censorship, CNet News (Jan. 31, 2007), http://news.com.com/2102-1028_3-6154930.html. This is one, perhaps charitable, interpretation of these companies' requests. Their opposition to legislation on the subject makes it less than clear what practical assistance they would like, other than a magic wand to make moral quandaries go away.

[182] *See* Dan Nystedt, Jailed Chinese journalist to file US suit versus Yahoo, Computer World (Sept. 25, 2006), at http://www.computerworld.com.au/pp.php?id=1316513149&fp=4&fpid=1398720840.

ambivalent response to a Brazilian court order requiring it to identify users of its Orkut social-networking site spreading child pornography or engaged in hate speech.[183]

The tensions between government interests in censoring or finding the authors of unwanted speech is in direct tension with many user and provider interests in search. Blocking is a deliberate and heavy-handed form of bias. Government recruitment of search also directly threatens both user and provider privacy. On the user side, search queries—whether obtained from the search engine or from other sources—can provide highly incriminating evidence.[184] And on the provider side, search provides another tool in the government's surveillance toolkit. Unlike with most private surveillance, the government may be interested in searching entire populations for what it considers suspicions activity. Think of the NSA surveillance program, which amounts to a gargantuan search project across the huge daily volume of phone calls, searching for, it appears, patterns of calls fitting specified profiles.

D. *Search Engines' Interests*

1. Preventing SEO and Click Fraud

Search fraud can be described as a deliberate attempt by providers to introduce bias and distortion in search results. As such, it causes harms to the users whose searches are less useful, and to legitimate providers whose content is made less visible. Search engines, dependent on users to search, already have strong market incentives to stamp out search fraud. Search fraud therefore creates a technical arms race between engines and manipulators. Search engines jealously guard their ranking algorithms as a way of maintaining an edge in this race.[185] Because search results are generally public, however, manipulators are typically able to engage in fairly extensive reverse engineering. Even if they do not know precisely how the engine ranks, they can create reasonable functional models of the sorts of content it favors and disfavors.[186]

Pure search fraud has not generated much in the way of litigation. Although suits by providers and users upset at fraud-heavy results are not inconceivable, these parties are in the same boat as engines, engines almost never make specific promises to stamp out search fraud, and there is little more most engines could do. Direct litigation against manipulators is potentially more significant. The trademark theories mentioned above were originally developed in direct suits by trademark holders against alleged manipulators.[187] Courts have recognized that some techniques of content design are deceptively manipulative and cause harm to legitimate providers, and it is possible that innovative pleading could properly state other business torts against manipulators. Similarly, luring users to one's content through search fraud raises significant false advertising concerns. Where the search engine is a passive conduit, it seems

---

[183] *See* Anne Broache, Google to hand over Brazilian user data, Google Blog, C|Net News.com (Sept. 5, 2006), at http://news.com.com/2061-10812_3-6112176.html.

[184] *See United States v. Schuster*, No. 05-4244 (7th Cir. Oct. 27, 2006); K.C. Jones, *Murder Suspect's Google Searches Spotlighted in Trial*, CRN (Nov. 11, 2005), available at http://www.crn.com/components/weblogs/article.jhtml?articleId=173602157.

[185] *See, e.g.*, Matt Cutts, *Notifying Webmasters of Penalties*, Matt Cutts: Gadgets, Google, and SEO (Apr. 26, 2006), http://www.mattcutts.com/blog/notifying-webmasters-of-penalties/ (discussing tradeoffs involved in notifying providers that their content uses forbidden techniques).

[186] *See, e.g.*, Dave Tiberio, *Reverse Engineering Search Engine Ranking Algorithms*, ArticleCentral.org, http://www.articlecentral.org/SEO_23141.html (explaining basics of reverse engineering of search engine algorithms).

[187] *See, e.g.*, Brookfield, *supra* note 128.

appropriate to examine the actual effect of SEO techniques to determine whether they are misleading users.

The hard part, both practically and legally, is telling whether a given technique actually constitutes search fraud. Because being caught engaging in SEO is the most common way to be stricken from an engine's index, the question of whether a provider is engaging in SEO has sometimes figured in placement suits.[188] Search optimization techniques are neither necessarily commercial nor necessarily centralized. Consider Googlebombing: the process of creating hyperlinks using a particular phrase and pointing to a particular page with the goal of convincing Web search engines to return that page on a query for that phrase.[189] Search engines collect and focus distributed intelligence, but they are also vulnerable to attempts to misrepresent what the collective general will actually thinks.[190] The most famous Googlebomb of all time may be the linking of the phrase "miserable failure" to the White House home page.[191] Similarly, companies have learned how to engage in "sock puppetry," creating fake content and personae to express their point of view and suppress negative opinions of them in search results.[192] The misrepresentations, if any, involved in these techniques are not easily characterized. Opinions differ as to whether search engines should try to suppress such techniques.[193]

Click fraud, because it involves money taken directly from advertisers' pockets and the contractual relationship between advertisers and engines, has led to more litigation.[194] The Coase Theorem suggests that systematic click fraud will simply be reflected in lower prices for

---

[188] *See, e.g.*, Search King, *supra* note 135.

[189] See Adam Mathes, *Filler Friday: Google Bombing*, Über.nu (Apr. 6, 2001), at http://uber.nu/2001/04/06/, available at http://web.archive.org/web/20050401091851/http://uber.nu/2001/04/06/.

[190] *See generally* Jean-Jacques Rousseau, The Social Contract (1762) (defining the "general will").

[191] *See Political Google Bombs*, Wikipedia, *at* http://en.wikipedia.org/wiki/Political_Google_bombs. Whether the critics of George W. Bush who created this Google Bomb will feel as proud of it when the White House has its next occupant is another matter entirely.

[192] *See, e.g.* Mark Glaser, *Companies subvert search results to squelch criticism*, Online Journalism Review (June 1, 2005), at http://www.ojr.org/ojr/stories/050601glaser/. *See also* Wikipedia, *Sock Puppetry*, at http://en.wikipedia.org/wiki/Wikipedia:Sock_puppet ("Use of sock puppets is discouraged in most cases").

[193] *See generally* Clifford Tatum, *Deconstructing Google bombs: A breach of symbolic power or just a goofy prank?*, First Monday (Oct. 2005), at http://www.firstmonday.org/issues/issue10_10/tatum/index.html. *Compare* Saul Hansell, *Foes of Bush Enlist Google to Make Point*, New York Times (Dec. 8, 2003) ("We just reflect the opinion on the Web . . . ."), *with* Danny Sullivan, *Googlebombing Now a "Prank" And Not Web's Opinion, Says Google*, Search Engine Watch (Sept. 19, 2005), at http://blog.searchenginewatch.com/blog/050919-095321.

[194] *See* Crafts by Veronica v. Yahoo, Inc., No.2:06-cv-01985 (D.N.J. complaint filed May 1, 2006); Draucker Development v. Yahoo, Inc., No. CV06-2737 (C.D. Cal. complaint filed May 4, 2006); Lane's Gifts v. Google, No. CV-2005-52-1 (Ark. Circuit Ct. 8th Dist. settlement approved July 27, 2006); Checkmate Strategic Group, Inc. v. Yahoo, Inc., No. 2:05-CV-04588-CAS-FMO (C.D. Cal. preliminary settlement approved June 28, 2006); Advanced Internet Technologies v. Google, 2006 WL 889477 (N.D. Cal. stay entered Apr. 5, 2006); CLRB Hanson Industries LLC v. Google, Inc., Case No. 1-05-CV-046409 (Cal. Sup. Ct. complaint filed August 3, 2005). *See also* David A. Vise, *Clicking to Steal*, Washington Post (Apr. 17, 2005) (describing lawsuit by Google against Auctions Expert alleging click fraud).

advertising, since it increases the number of clicks charged per sale generated.[195] Particular advertisers targeted by competitors for fraud (or singled our for some reason) raise a more difficult problem. One might expect that advertisers would demand clauses relieving them of the responsibility of paying for fraudulent clicks if they are targeted.[196] But enforcing such clauses requires that advertisers be able to monitor the clicks for which they are charged.[197] Search engines have been reluctant to share the necessary data, out of fear that it would permit reverse engineering of their fraud-detection algorithms.[198] On the other hand, because the penalties unilaterally imposed by search engines against sources of click fraud can include being banned from search advertising, advertisers are also upset by overly zealous enforcement of anti-fraud policies. [199] Independent auditing of click counts and of anti-fraud programs may be the wave of the future.[200]

Click and search fraud, because of their unsavory links to the parasite economy,[201] have also embroiled some search engines in litigation involving those links.[202] Consumers and state attorneys general have begun to sue spyware makers and the advertising networks linked to them.[203] Following the chain of business relationships leads inevitably back to the search engines whose affiliate networks provide the ad delivery and billing at the heart of many of these schemes.[204] Those who distribute spyware also do so through advertisements placed on search engine sites (and on sites highly ranked due to search fraud), suggesting that search engines (which often forbid these tactics in their advertiser guidelines) may bear some moral and legal

---

[195] Some have argued that search engine have no incentive to police click fraud because they can charge for illegitimate clicks. *See, e.g.* Martin Fleischmann's remarks in Brian Grow & Ben Elgin, *Click Fraud: The dark side of online advertising*, Business Week (Oct. 2, 2006) ("'I told Yahoo years ago,' he says, '"If this was costing you money instead of making you money, you would have stopped this."'"). Not so. The Coasean exchange returns to search engines the necessary incentive, as they will be able to increase their cost-per-click if they reduce fraud rates.

[196] *But see, e.g.*, Yahoo! Search Marketing, *Click Fraud FAQs*, at http://searchmarketing.yahoo.com/legal/clickfraud.php ("*[E]ven though we are not obligated to*, we voluntarily designed the Click Protection System") (emphasis added).

[197] *See* Google, Inc. Click Quality Team, How Fictitious Clicks Occur in Third-Party Fraud Audit Reports (Aug. 8, 2006), at http://www.google.com/adwords/ReportonThird-PartyClickFraudAuditing.pdf.

[198] *See* Alexander Tuzhilin, The Lane's Gifts v. Google Report, available at http://googleblog.blogspot.com/pdf/Tuzhilin_Report.pdf.

[199] *See* Battelle. *supra* note 1.

[200] *See* Interactive Advertising Bureau, *The Interactive Industry Commits to the Development of Click Measurement Guidelines* (Aug. 2, 2006), available at http://www.iab.net/news/pr_2006_08_02.asp.

[201] *See* Damien Cave, *The parasite economy*, Salon (Aug. 2, 2001), http://archive.salon.com/tech/feature/2001/08/02/parasite_capital/index.html (coining the term); Stefan Gorling, *An Introduction to the Parasite Economy*, EICAR 2004 Conference Proceedings (U.E. Gattiker, ed. 2004), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=683081 (generalizing it).

[202] *See, e.g.*, *Crafts by Veronica*, *supra* note 194.

[203] *See, e.g.*, Washington v. Secure Computer, LLC, No. C-06-0126 (W.D. Wash. complaint filed Jan. 24, 2006); New York v. DirectRevenue, LLC (N.Y. Sup. Ct. complaint filed Apr. 4, 2006); Sotelo v. DirectRevenue, LLC No. 05-C-2562 (N.D. Cal. settlement preliminarily approved Mar. 10, 2006).

[204] *See* Ben Edelman, *The Spyware - Click-Fraud Connection -- and Yahoo's Role Revisited* (Apr. 4, 2006), at http://www.benedelman.org/news/040406-1.html (describing role of Yahoo's affiliate network in placing advertisements in spyware). (Edelman is co-counsel for plaintiffs in the *Crafts by Veronica* suit.)

responsibility for the problems created by spyware.[205] Their central placement athwart many information flows may make search engines least-cost avoiders for some of these ecological problems.

2. Innovation

Search engines compete with other search engines for users.[206] They compete by offering more complete indexes of the Internet, by providing more responsive results, and by integrating their search with other features valued by users. (Their actual revenue source, advertising, is dependent on their ability to attract users, so market share by searches is a good indicator of competitive success.) All three of these techniques involve a sometimes-frenetic pace of innovation. Search engines have been zealous about using intellectual property rights to protect their innovations from being appropriated by each other.[207]

First off, search engines use trade secret techniques extensively. Their ranking and indexing algorithms are closely guarded, and the need to protect that secrecy is routinely invoked in litigation that might expose operational details.[208] As noted above, in practice, this secrecy is incomplete, since the public disclosure of results permits some degree of reverse engineering. Search engines are often quite reluctant to take technical steps that would discourage large-scale reverse-engineering of the algorithms; they have discovered instead that there is value is allowing nearly unlimited usage of their search facilities, including through automated APIs.[209] Further, the pressures of PR encourage search engines to trumpet various advances and tweaks to their algorithms, if only in general terms.[210]

Search engines are also careful to guard the protected status of their trade secrets with appropriate measures. They routinely require employees to sign nondisclosure and noncompetition agreements, to prevent the departure of search secrets to rivals with key

---

[205] *See* Ben Edelman, *Pushing Spyware through Search* (Jan. 26, 2006), at http://www.benedelman.org/news/012606-1.html (analyzing prevalence of spyware-laden sites in Google results and advertisements for some common searches).

[206] Currently, according to ComScore and Nielsen NetRatings, Google is the market leader with a roughly 45% share, Yahoo! has about 25%, and MSN another 10%. AOL and Ask also have significant shares of perhaps 5% each, and the remainder is split among a great many minor search engines.

[207] Jonathan Thaw & Susan Decker, Google to Subpoena Yahoo, Microsoft on Book Scanning, Bloomberg News (Oct. 5, 2006), at http://www.bloomberg.com/apps/news?pid=20601103&sid=amfuMLMq__H8.

[208] *See* Gonzales v. Google (alleging that revealing random URLs and user queries would expose trade secret).

[209] *See* Robert D. Hof, *Mix, Match, and Mutate*, BusinessWeek (July 25, 2005), available at http://www.businessweek.com/@@76IH*ocQ34AvyQMA/magazine/content/05_30/b3944108_mz063.htm (describing "mashup" applicataions made by combining multiple programmatic APIs). Such mashups would be nearly impossible if search engines and other web companies treated their results as resources to be closely guarded.

[210] *See, e.g.*, Matt Cutts, A quick word about Googlebombs, Official Google Webmaster Central Blog (Jan. 25, 2007), http://googlewebmastercentral.blogspot.com/2007/01/quick-word-about-googlebombs.html (describing Google's algorithmic changes to end the effect of Googlebombs on its rankings).

employees.[211] They also typically follow the tech-company way of life of requiring visitors to agree to quite comprehensive NDAs.[212] Secrecy at Google, in particular, is almost a way of life.

Search engines also use the federal IP regimes extensively. Google's patent on PageRank is well-known,[213] but all of the major search engines have patent portfolios. They are also regularly patent defendants.[214] In the run-up to its IPO, Google settled, for stock worth around $300 million, an outstanding patent infringement lawsuit brought by Yahoo relating to a patent on the technique of displaying keyword-triggered search ads based on the bids submitted by advertisers—the technique at the heart of most search advertising today.[215]

Search engines also use copyright and trademark to protect their business models, although these matters are less litigated. Search engines possess valid copyrights in their software and interfaces; they have valuable trademark and trade dress rights in their brands.[216] Whether search results are copyrightable as such is a debatable proposition.

3. Competition

Competition *qua* competition also creates another concern for search engines: that other search engines not monopolize the market for search or use improper competitive techniques. Search engines have therefore occasionally been the subject of antitrust speculation.

As is common in many high-tech industries, the scope of the search market are not obvious. The business model of web search makes it a two-sided market: selling search to users and users' attention to advertisers. Moreover, because the boundary of search technology is fuzzy, the choice to classify some technologies as "search" or not has significant implications. Currently, Google is the big fish among major search engines, but its share would drop if traffic to other sites with specialized search functionality—such as eBay—were included. Given these complexities, and the relative lack of litigation so far, the shape of search engine antitrust law remains largely undefined.

The law of unfair competition—unconnected with monopolization as such—provides a baseline of legitimate and illegitimate business practices in the search market. Thus, some of the

---

[211] *See, e.g.*, Google Inc. v. Microsoft Corp., 415 F. Supp. 2d 1018 (N.D. Cal. 2005) (discussing attempt to prevent Kai-Fu Lee, former vice president at Microsoft, from becoming vice president at Google).

[212] *See* GOOGLE: This NDA Never Existed, Valleywag (Jan. 22, 2007), http://valleywag.com/tech/google/this-nda-never-existed-230407.php (reprinting Google's standard NDA).

[213] U.S. Pat No. 7,058,628.

[214] Hyperphrase Techs., LLC v. Google Inc., 2006 U.S. Dist. LEXIS 64918 (W.D. Wisc. 2006); Skyline Software Sys. v. Keyhole, Inc., 421 F. Supp. 2d 371 (D. Mass. 2006); Netjumper Software, L.L.C. v. Google Inc., 2005 U.S. Dist. LEXIS 27813 (S.D.N.Y. 2005); FindWhat.com v. Overture Servs., 2003 U.S. Dist. LEXIS 2450 (S.D.N.Y. 2003).

[215] *See* George Mannes, *Yahoo! Gets Bigger Stake in Google*, TheStreet.com (Aug. 9, 2004), http://www.thestreet.com/pf/tech/georgemannes/10177217.html.

[216] *See* Penelope Patsuris, *The Making of a $2 Billion Brand*, Forbes.com (Feb. 21, 2003), at http://www.forbes.com/2003/02/21/cx_pp_0221google.html; Beth Lipton Kriegel, *Yahoo not amused by pot parody site*, C|Net News.com (Jan. 11, 1999), at http://news.com.com/2100-1023-219986.html; Frank Ahrens, *So Google Is No Brand X, but What Is 'Genericide'?,* Washington Post D01 (Aug. 5, 2006), at http://www.washingtonpost.com/wp-dyn/content/article/2006/08/04/AR2006080401536.html (describing baseless attempt by Google trademark attorneys to object to newspaper's statement that 'Google' was in common use as generic verb).

trespass-to-chattels and copyright theories raised by providers are also arguably unfair competition claims among search engines themselves. In China, Yahoo and the local Qihoo have sued each other for unfair competitive practices.[217]

In terms of antitrust suits themselves, some search ranking suits by disappointed providers have raised antitrust theories, although such theories generally have failed to explain why the poor ranking of a provider constitutes an antitrust injury.[218] Where market power in a related market—e.g. desktop operating systems or provision of broadband telecommunications—is more clearly established, search engines have appealed to the law and policy of antitrust to prevent the leveraging of that power into the search market. Thus, competing search engines have objected to the integration of a search engine into other applications—e.g. through a search box in a browser—but have not so far convinced antitrust regulators that the practice is worthy of concern.[219] Similarly, search engines have appealed to telecommunications regulators, asking that network operators not be given too much market power to discriminate among different traffic flows across their networks.[220] Search engines fear that network operators would use this power to extract rents from them.[221]

For now, while Google is the big fish in the search pond, other search engines have been maintaining market share. The greater competition concern may come from the rising costs of entry; the Internet continues to grow rapidly and the SEO arms races have meant that sophisticated and computationally expensive algorithms appear to be part of the price of offering useful search. Switching costs for users have historically been quite low, but the rise of personalized search may be raising them, as users are locked into greater personalization and customization with the search engine they have been using.

III. INTERCONNECTIONS IN SEARCH ENGINE LAW

Already some of the connections and conflicts among these problems in search engine law should be evident. This Part will examine five in more depth, to show how the twelve doctrinal areas described in the previous Part are deeply intertwined.

First, many forms of relief against search engines are functional substitutes for each other. Those denied one cause of action may turn to another; those objecting to search may even attempt to borrow each others' legal claims. Some of the hardest-fought issues in search policy are all but moot in light of doctrines in nearby areas. The debates over trespass to chattels and

---

[217] *Cf.* Doug Young, Yahoo and former China head in brewing legal battle, Reuters (Aug. 17, 2006), at http://today.reuters.com/news/articleinvesting.aspx?view=CN&storyID=2006-08-17T080011Z_01_SHA226749_RTRIDST_0_MEDIA-CHINA-ALIBABA.XML&rpc=66&type=qcna (Qihoo sues Yahoo for defamation), *with* Sophie Taylor, Yahoo! China sues local firm for unfair competition, Reuters (Sept. 29, 2006), at http://today.reuters.com/news/articleinvesting.aspx?view=CN&storyID=2006-09-29T094620Z_01_SHA357681_RTRIDST_0_TECH-CHINA-ALIBABA-UPDATE-1.XML&rpc=66&type=qcna (Yahoo sues Qihoo for unfair competition).

[218] *See* KinderStart; Person v. Google Inc., No. 06 Civ. 4683, 2006 WL 2884444 (S.D.N.Y. Oct. 11, 2006).

[219] *See* Ina Fried, *Vista search seems fair, regulators say*, C|Net News.com (May 12, 2006), at http://news.com.com/2102-1014_3-6071851.html.

[220] *See* Eric Schmidt, *A Note to Google Users on Net Neutrality*, Google Help Center, at http://www.google.com/help/netneutrality_letter.html.

[221] *See* Arshad Mohammed, *Verizon Executive Calls for End to Google's 'Free Lunch'*, Washington Post (Feb. 7, 2006).

initial copying, in particular, seem irrelevant in light of browsewrap restrictions on indexing. Taking a broad view of search forces us to ask whether these differentiations are sustainable, and to assess the quality of the fit between the interests we wish to protect and the legal theories that protect them.

Second, search engines raise a number of problems of unaccountable discretion. Two natural tools to investigate and remedy abuses of discretion are greater disclosure of the basis for ranking decisions and mandated corrections to misleading rankings. But these tools run squarely up against search engines' operational interests in fighting search fraud, against their competitive interest in trade secrecy, and perhaps even against user interests in privacy. Taking a broad view of search forces us to weigh the costs and benefits of these remedies carefully.

Third, user privacy concerns must be understood in the context of search engine operations and third-party concerns. Search engines use query and clickthrough data to target advertisements, to refine search quality, and to personalize search. Prohibiting these uses outright could have significant negative effects on users, including exacerbating search engine bias. At the same time, third parties may have quite legitimate interests in learning user identities, so that query privacy policies reflect a struggle between users and third parties for relative advantage. Given the profoundly private information users entrust to search engines, balancing out these concerns will not be easy. A broader view of search makes clear the competing values at stake.

Fourth, many legal theories raised by and against search engines turn on the speech content of search recommendations. Search engines have encouraged a view that such recommendations are subjective statements of opinions about page quality, and as such are entitled to substantial First Amendment protection. But there is a sense in which it is precisely the subjectivity of search rankings that make them problematic—the more individually tweaked and the less automatic that results are, the greater the concern that the search engine is using its privileged position to engage in unfair discrimination against particular targets. The question of search engine speech needs to be regrounded on a more stable foundation than the subjective-statement-of-opinion analysis alone can provide. A broader view of search can provide a start.

Fifth, thinking about trademark disputes without placing them in context is a recipe for trouble. Search engines provide enormous value to consumers in quite literally reducing search costs, and they are fueled in this ability by advertising sales. Trademark holders' demands for veto power over keyword sales must be understood as a tactic in the ranking wars; giving them that power would frustrate the policies of trademark law and hamper search innovation. At the same time, however, simply excluding keyword sales and search engine manipulation from trademark scrutiny altogether, as some courts have done, is also dangerous. SEO tactics by providers increase consumer confusion and are socially wasteful; it is easy to envision search engines and near relatives that flout the goals of unfair competition law. A broader view of search can point to a healthier balance.

A. *Claims Against Search Engines as Functional Substitutes*

The multiplicity of relationships between search engine and other parties means that there are multiple legal lines of communication open. Those concerned with one particular form of harm—or simply interested in extracting rents—are not limited to legal theories directly addressing that harm. If they can gain relief against a search engine[222] on another theory, it may be just as good. The relief from the two theories may be targeted at the exact same action, or the

[222] Or, in the engine's case, gain a blanket immunity from suit.

second may simply give enough negotiating leverage over the search engine that it can be traded for the relief the plaintiff actually desires. In such cases, one might be concerned that providers not be able to subvert the policies of a doctrine by using it to obtain relief from the 'wrong' form of harm.

Consider first providers' desire to prevent searching. Suits based in unwanted access (whether in indexing or by users) can often be interchanged with unfair competition suits. Many lawsuits by providers run together both forms of harm in their complaints—the search engine both imposes technical burdens and weakens providers' connections with users.[223] Indeed, many unwanted access suits are substantially motivated by competitive concerns.[224] Conversely, providers often bring unfair competition suits when, for some other reason, they cannot turn to the law or to self-help to prevent access. Thus, for example, the Authors Guild has no ability to prevent Google from obtaining physical access to books. [225] Its lawsuit against Google is therefore exclusively based in copyright infringement.[226]

Considering the three theories providers have used to argue against access for indexing—trespass to chattels, anti-intrusion statutes, and contract—reveals another concern. Here, pro- and anti-access camps have fought intensely over the proper scope of trespass to chattels. *Intel v. Hamidi* seems to have settled the issue, for the time being, in favor of the pro-access side by requiring proof of actual harm to the chattel to support a trespass to chattels action. This victory, however, is largely symbolic, given the broad scope of some interpretations of the scope of the Computer Fraud and Abuse Act and related statutes—and particularly in light of the broad deference courts have been willing to give to browsewrap contracts once they have found that the browser had sufficient notice of the contract. Whether or not all terms in such contracts would be upheld, a simple statement that browsing for some purposes or by some users is forbidden is not a particularly onerous one, and contract law gives little reason to invalidate it. This collision is a clear case of legal convergence; three distinct theories apply to exactly the same acts, and ultimately only one rule can prevail.

It is not only providers who can substitute one claim against a search engine for another. Private third parties have recognized that copyright claims provide much stronger relief against intermediaries than other private law claims do.[227] Two ISP examples show the phenomenon at work. In the pre-CDA case of *Religious Technology Center v. Netcom On-Line Communications*

---

[223] *See, e.g.* Tickets.com, *supra* note 88 (raising copyright and trespass to chattels theories); Kelly v. Arriba Soft, 336 F.3d.811 (9th Cir. 2003) (raising copyright theories predicated both on initial copying and on framing).

[224] *See* eBay, Register.com. Many of the these cases are brought by plaintiffs who are providing factual information not protectable in copyright, so that the most promising unfair competition theories are unavailable. *See also* Tickets.com, *supra* note 88; eBay, *supra* note 86; Farechase, *supra* note 92..

[225] *See* 17 U.S.C. § 109(a) (first sale). *But see* Sharon Billington, *Relief from Online Used Book Sales During New Book Launches*, 29 Colum. J.L. & Arts 497, 497-98 (2006) (arguing, against empirical evidence and without considering effects on libraries and critics, that publishers should be allowed to prevent resales of used books within six months of first publication).

[226] *See* Authors' Guild, *supra* note 106. *Cf* Ty, Inc., v. Publications Int'l Ltd., 292 F.3d 512 (7th Cir. 2002) (copyright infringement suit involving collectors' guide to Beanie Babies, with allegedly infringing photographs thereof). *Cf.* Kevin Emerson Collins, *Cybertrespass and Trespass to Documents*, 54 Clev. St. L. Rev. 41-68 (2006) (discussing cases involving subsequent use of information obtained through improper access to physical documents).

[227] *See, e.g.*, 47 U.S.C. § 230(e)(2) ("Nothing in [the CDA intermediary immunity] shall be construed to limit or expand any law pertaining to intellectual property.")

Services, Inc., the Church of Scientology attempted to prevent the distribution of unpublished L. Ron Hubbard manuscripts by asserting copyright claims against an ISP; the Church has been a regular user of DMCA notices under Section 512 against search engines. In *Diebold v. Online Policy Group*, an electronic voting machine company attempted to use the Section 512 subpoena process to suppress distribution of the source code to its voting machines, relying on Section 512(d), the "information location tools" provision applicable to search engines. Both were acting principally to avoid embarrassment, but claims in which the embarrassment constituted the actionable harm were not likely to succeed against the ISPs involved. Instead, copyright theories promised more legal leverage.

Lawyers in search engine suits will not respect boundaries between legal fields when framing their cases. Those who make law and policy for search engines must be alert to these overlaps and end-runs. Considering the various strands of search engine law together will help make such possibilities clear.

B. *The Pros and Cons of Disclosure and Mandated Results*

If we were convinced to a moral certainty that we could identify particular instances in which search engines were returning incorrect results, the natural response would be to mandate that they return instead the results we knew to be correct. Even if we do not believe that we know which results are correct, we might know that some particular results are incorrect and require their deletion. Frank Pasquale claims that subjects of search should not have search engines return misleading information about them, and identifies two cases in which the correct results can be specified precisely enough that legal intervention is warranted: searches on proper names and searches on trademarked terms. Although he stops short of claiming that such searches should return instead that content provided by the person named or the holder of the trademark, he does propose that they be allowed to annotate misleading results with an asterisk.[228]

A related claim—one supportable even without any objective conception of correct and incorrect results—starts from the observation that users depend on search engines to find information for them. Users in general can comparison-shop between search engines,[229] and so can pick the engine that most consistently returns them the best results, or use multiple search engines for the same search. Their ability to know whether to trust an engine, however, depends on answering precisely the question that the search engine seeks to solve: *What content is available?* Especially in personalized search, users need information about a search engine's inputs and its reasoning to make informed choices. This claim leads, therefore, to an argument that search engines should disclose to users information about their algorithms. (In general terms, search engines do, but these arguments would demand more specific disclosures.)

Providers and search subjects also can articulate legitimate reasons for greater disclosure of search engine operational details. Google has been accused of manipulating results during litigation to make the judge at a crucial hearing unable to replicate the behavior of which its adversary complained.[230] Providers who feel that they have been unfairly ranked—particularly

---

[228] Frank Pasquale, *Rankings, Reductionism, and Responsibility*. *But see* Eric Goldman, *Search Engine Bias and the Demise of Search Engine Utopianism*. Pasquale's precise specification of these cases responds to Goldman's critique that regulators cannot in general identify correct results as well as search engines can, by restricting intervention to clearly-identified mistakes.

[229] The possibility of such comparisons depends, of course, on there being a genuine diversity of search engine providers and search algorithms. *See* Goldman.

[230] *See* BATTELLE, *supra* note 1.

in retaliation for not purchasing advertising, complaining about the search engine, or competing with it—may need access to operational details to evaluate whether they really have been targeted. Third parties may need explanations as to why unflattering content appears prominently. And many disputes about click fraud cannot be evaluated without examining details of search engine billing. Again, disclosure seems the natural remedy.

That said, excessive mandating and disclosure can have dangerous consequences not only for search engines but for the entire search ecosystem. Search engines' innovative interests mean that too much disclosure can suppress their incentives to create new algorithms and reduce the diversity of options, thereby actually reducing users' genuine choice among engines. Mandated results are even more restrictive for competition and diversity, in that they enforce uniform policies about results across engines. Moreover, too much transparency in relationship to personalized search will require at least some disclosure of user queries, raising privacy concerns.

These concerns, however, pale next to the consequences of mandated results and disclosure for the SEO arms race. Search engine manipulators make their living by reverse engineering search algorithms (something that is always plausible, given that search engine results are easy to observe). Search engines, on the other hand, are able to preserve a layer of genuine, useful results largely through a combination of keeping precise algorithmic details secret and changing their algorithms to foil detected SEO techniques. Mandated disclosure undermines the former; mandated results undermine the latter. Legal interventions here threaten to hand search engine algorithms to manipulators on a platter. Even Pasquale's limited proposals are partially vulnerable to manipulation. What proof would a search engine require of one's real name before awarding an asterisk? And what would stop manipulators from registering trademarks on popular search terms on unlikely categories of goods? Consider a registration of REAL ESTATE as used to sell lip gloss (a product category for which it is fanciful, and thus registrable)—perhaps a pointless trademark but excellent for search engine placement.

This is not the place to evaluate when disclosure and results mandates are appropriate policy. For present purposes, it should suffice to note that these remedies raise concerns that cut across many areas of search engine law. There are reasons why they may be useful interventions, and reasons why they may be dangerous. Considering the one without considering the other would be reckless policy-making.

C. *User Privacy Concerns Implicate Others' Interests*

The privacy problems posed by private stockpiles of user data are well-understood and do not require extensive rehashing here.[231] What does require attention is how the solutions to these problems may have different inflections in the search engine context than in other domains. Users' privacy interests must be understood in relation to other interests in query data.

First, it is actually inherent in the technology of search—as most commonly practiced today—that users' own computers disclose to those providers they visit the users' use of a search engine and the query terms they used. The "referrer" information that a browser by default gives to any site from which it requests a Web page includes the URL of the Web page that referred the user to the provider's page. Most search engines include the query terms that a user entered in the URL of the page displaying the results. This automatic leak of query information—which can be blocked by technical measures either at the search engine or the

[231] *See* SIMSON GARFINKEL, DATABASE NATION (2000); DANIEL SOLOVE, THE DIGITAL PERSON (2004).

user's computer—mean that search engines are not the only institutions that can easily accumulate query data.[232] Put another way, the interaction of results flows from engines to users with content flows from providers to users creates an additional query data flow from users to providers.

Second, third parties harmed by search may have legitimate interests in learning some private information about users. This tension has been most clearly noted in the case of flows of copyrighted information in the ISP context; the RIAA has been particularly active in attempting to breach user privacy.[233] Those whose privacy is breached by a search also have an interest in learning about it: they may need to know that they need to take precautions against stalkers, they may need to take action against the provider releasing this information, and they may need to deter the searchers from searching.[234] There is something uncomfortable about a rule that assigns different weights to the privacy interests of search users and search subjects.

Third, the uses to which search engines put their query data warehouses are relevant to users' other interests in search. Massive stockpiles of queries are useful for improving search in general. Indeed, the AOL data release was neither the product of poor security nor a concession to corporate pressure for valuable data—it was an ill-advised attempt to further academic research into better search technologies.[235] Extensive collection of query data is also a prerequisite for personalizing search.[236] Personalization of information reception and its concomitant promotion of diversity, in turn, can be an important technique for countering media bias.[237] Thus, a privacy-mediated concern with preventing individual manipulation by search engines[238] is in tension with a concern with preventing manipulation by monolithic one-size-fits-all information sources.

To repeat, problems of online privacy protection are subtle and tangled. Considering the various threads of search engine law all at once reveals just how many of them are connected to privacy in one way or another. Once again, any rational attempt to make sensible policy (here, privacy policy) in the search context demands careful engagement with these many interests and pressures.

---

[232] Simply cutting and pasting a search result URL into one's address bar will hide from the provider the search query that led one to their site. One can also install software to much the same end. *See, e.g.* RefControl, https://addons.mozilla.org/firefox/953/. Although blocking of such query leaks by search users is possible, it also seems to be quite rare.

[233] *See* Sonia Katyal, *The New Surveillance*, 54 Case W. L. Rev. 297 (2004).

[234] David Brin has written about this tension, and argues that the solution is greater transparency in general. We cannot stop people surveilling others, but can at least let those being surveilled know about it. *See* DAVID BRIN, THE TRANSPARENT SOCIETY.

[235] *See* Katie Hafner, *Researchers Yearn to Use AOL Logs, but They Hesitate*, New York Times (Aug 23, 2006) (discussing academic ethics of research using AOL query data).

[236] Consider in this light Amazon's practice of recommending books by linking books that users searched for or bought in the same browsing session. On such "collaborative filtering" systems in general, see JOHN RIEDL ET AL, WORD OF MOUSE: THE MARKETING POWER OF COLLABORATIVE FILTERING (2002).

[237] *See* Yochai Benkler, *Siren Songs and Amish Children: Autonomy, Information, and Law*, 76 N.Y.U. L. Rev. 23 (2001).

[238] *See* Tal Zarsky, *Online Privacy, Tailoring, and Persuasion*.

D. *Search Engine Results as Speech*

Many legal questions involving search will require articulating a theory of search engine speech. The First Amendment rights of search engines, users, and providers may provide defenses to third parties' attempts to impose liability for harmful content flows. Search engines' speech-facilitating roles may provide them with a thumb on the scales in debates over access. And the conflicting speech claims of search engines and providers will have a significant effect on how we think about search engine rankings. Ultimately, a theory of search engine speech will need to integrate all of these concerns.[239]

Such a theory is beyond the scope of this Article. Instead, this Section will consider the cross-cutting problems raised by one attempt at framing the question. Google has asserted a theory of search rankings as subjective statements of opinion. Under a claim of tortious interference with contract, such as that raised in the *Search King* suit, it is a complete defense if the allegedly harmful act consisted of protected speech. Statements of opinion on matters of public concern are protected unless provably false. Thus, for example, a negative bond rating is not a statement that could be proven "true" or "false," and thus cannot support defamation liability.[240] Since a search engine's rankings are merely a claim about the engine's subjective assessment of pages' relevance to particular users' queries, goes the reasoning, the search engine is not making a claim that could be shown false, and is therefore protected.[241]

The relationship of subjective opinion to objective fact, however, is not simple. Thus, for example, *Milkovich v. Lorain Journal Co.*,[242] while stating the rule that the Constitution shields opinions, leaves in place two significant exceptions. A statement of opinion may imply an underlying fact (the Court's example was "In my opinion John Jones is a liar."); and even a statement of opinion may be false if not honestly held (the Court's example was "I think Jones lied," where the speaker thought nothing of the sort.).

In this light, the *Search King* rule suggests several counterarguments. First, as Search King alleged, the purported "opinion" is in fact the output of a computer algorithm, and the computer is perfectly predictable and objective. The court rightly dismissed this argument, distinguishing process from result. The subjectivity entered the algorithm when it was programmed, and although the intervening process is mechanical, what emerges at the end are the subjective judgments made by Google programmers about Web page relevance and quality.[243]

Second, as Search King and KinderStart have alleged, search engines themselves emphasize the objective quality of their results and should be held to those statements.[244] This argument has slightly more bite, given that search engines have not been careful in their public

---

[239] Substantial guidance may come from telecommunications law, which has long struggled with intermediaries' dual roles as speakers and as conduits. The greater interactivity of search engines—whose rankings are driven by providers' attempts to seek placement, by users' queries, and by search engine algorithms themselves—may require even more subtle analyses.

[240] Jefferson County Sch. Dist. No. R-1 v. Moody's Investor's Services, Inc., 175 F.3d 848 (10th Cir. 1999).

[241] *Search King*.

[242] 497 U.S. 1 (1990).

[243] This view of the subjectivity of search results is consistent with an argument that they involve sufficient selection and arrangement to satisfy copyright's originality requirement.

[244] *See also* Maughan v. Google Technology, Inc., No. B183969, 2006 Cal. App. LEXIS 1574 (Oct. 11, 2006) (noting Google's emphasis that its search processes are completely automated).

statements.[245] By alleging claims of objectivity, provider plaintiffs also nudge their pleadings closer to consumer fraud causes of action—in which it is the initial claim of evenhanded objectivity, and not the later ranking decision, that provides the necessary falsehood. Ultimately, however, search engines have at most implied objective correctness rather than explicitly claimed it, and certainly have the power to amend their public relations materials to escape from this trap prospectively. Moreover, this counter misses a basic reality of search engine business. Search engines are trying hard to maximize the subjective satisfaction of users with their search results, and on a query-by-query basis, it is extremely hard to find an objective "fact" in the degree of user satisfaction.

A more troubling counterargument, however, combines the first two. On this theory, it is *hand manipulation* of results that is troubling.[246] This theory abandons any claim to object to broad algorithmic decisions, but argues that specific deviations from algorithmic choices—or specific algorithmic tweaks to hurt particular providers—should be actionable. There are, it should be noted, good reasons to be skeptical of this counterargument. The same programmers both write the algorithms and tweak the results, so the distinction is not entirely coherent.[247] And, indeed, they are acting at their most subjective when they make changes by hand.

But it is this last point that is key to the counterargument against hand tweaks. Here, it is the algorithmic baseline that provides the "objective" claim about fact. We do not necessarily expect that the search engine's overall rankings reflect a falsifiable claim of fact, but the output of the generic ranking algorithm provides a stable reference against which claims of later manipulation can be measured. In terms of the *Milkovich* analysis, the search engine is lying not about the poorly-ranked page's quality, but about its own belief that the page is of low quality. *The ranking algorithm believes that the page is important, so returning a worse rank is a lie about what the ranking algorithm believes*. This way of phrasing the argument captures its strong intuitive appeal. Even while defending their rankings as subjective assessments, search engines have been highly reluctant to make hand adjustments. Google's annotation of the "Jew" search results required substantial public handwringing over whether to take such an extraordinary step.[248] There is something unsettling about hand tweaks, even if that something does not rise to the level that would permit a suit in tort.

Even this brief survey of one issue's implications has stinted on other significant connections. A concern with hand tweaks suggests that greater transparency is a necessary tonic. It is also intimately related to the ongoing struggles against SEO, since the most frequent algorithmic changes are counters deliberately targeted at new SEO techniques. It connects up with fears of government censorship, with the individual deletions of DMCA takedowns, and

---

[245] Eric Goldman and Rebecca Tushnet have both noted this tension. Raymond Nimmer observes that *Search King* did not involve a claim by a user "who detrimentally relied on the rankings themselves" and contrasts cases in which safety ratings services "voluntarily assumed [a] role [that] invited reliance by the public." RAYMOND T. NIMMER, INFORMATION LAW § 10:77 (2006).

[246] *See* James Grimmelmann, *Google Replies to Search King Lawsuit*, LawMeme, at http://research.yale.edu/lawmeme/modules.php?name=News&file=print&sid=807.

[247] *See, e.g.*, William Slawsky, 20 Ways Search Engines May Rerank Search Results, SEO by the Sea (Oct. 14, 2006) (describing systematic changes search engines might make to results after the basic relevancy algorithm but before showing results to users).

[248] *See* Google, *Offensive Search Results*, http://www.google.com/explanation.html ("Our search results are generated completely objectively and are independent of the beliefs and preferences of those who work at Google.")

with broader questions of individual (and personal) actions versus massive parallelism that arise in privacy and access to content. A proper theory of search engine speech should consider these issues all together, rather than singly and in isolation.

E. *Trademarks and Search Engines in Context*

The problem of searches on trademarked terms is, as noted above, one of the most litigated issues in search. It also provides a case study in the respective perils of too much and too little deference to search engines' decisions.

Courts dealing with the trademark implications of provider and search engine behavior have been much influenced by the *Brookfield* analogy of a highway billboard. A user who is misdirected by an information location tool during on a search for a trademarked term, the analogy asserts, is like a driver who has been misdirected by a billboard to taking the wrong exit. She may not ultimately be confused about the source of the goods she acquires, but the additional search costs of going back to find the true source outweigh her desire for the real McCoy.

As applied to search engines, this analogy misses two important features of the online context. First, it misapprehends some of the technical realities of how various advertising techniques actually translate into user visits. In particular, if keywords in hidden metatags are billboards, they are invisible ones; most search engines now ignore them.[249] Similarly, contextual ads that are clearly disclosed as such and do not use the trademarks in their text never appear to users as billboards would; one might analogize them instead to billboards placed near the plaintiff's store but not using the plaintiff's trademarks. Second, the search costs involved in going back to find the originally desired source are much lower online than offline—a few seconds of clicking rather than a few minutes of driving—so far fewer users, even if diverted, will actually be locked into the wrong source.

Indeed, these advertisements reduce search costs, first by offering users information about alternatives possibly responsive to their queries and second by funding search itself. Contextual ads are actually a substantial improvement, from a consumer confusion point of view, over earlier search business models, such as direct results placement purchases and generic non-contextual banner ads.

Commentators and courts, however, have articulated a slightly questionable basis for finding no liability. Instead of finding no consumer confusion, they have instead found no use in commerce, cutting off the trademark inquiry at the threshold. The *interactivity* of the search engine's dealings with the user has created analytical confusion; any possible "use" of the trademark is in the user's search query, rather than in the results. It has therefore seemed

---

[249] *Compare* Pop Warner Little Scholars, Inc. v. New Hampshire Youth Football & Spirit Conference, 2006 WL 2591480 (D. N.H. Sept. 11, 2006) ("Google and other search tools continue to associate defendants' web site with plaintiffs' marks [due to metatags]"), *with* Eric Goldman, *Outdated Metatags Don't Infringe--Pop Warner v. NH Youth Football & Spirit Conference,* Technology & Marketing Law Blog (Sept. 25, 2006), at http://blog.ericgoldman.org/archives/2006/09/outdated_metata.htm (criticizing *Pop Warner* and stating,"[S]earch engines aren't that inefficient or inaccurate given that THEY ARE SMART ENOUGH NOT TO RECOGNIZE KEYWORD METATAGS IN THE FIRST PLACE.).

plausible to say that triggering ads or results based on a trademarked query is not a "use in commerce" by the search engine.[250]

To see why this blanket rule may be inappropriate, consider a line of interactive offline cases: trademark suits against restaurants that serve one cola when a customer has requested another.[251] The customer who receives a Pepsi after ordering a Coke is a victim of passing-off; whether the deception falls within the Lanham Act should not depend on whether the restaurant has use the COCA-COLA trademark on its menu or whether the server repeated "Coke" to confirm the customer's order.

The search engine's proper defense is that it is not misleading users, not that it is not using the trademark. It is easy to imagine search engines that deliberately cause serious confusion. Think of what would happen if Froogle or Amazon—search engines specifically oriented towards finding particular goods for purchase—were to adopt a policy of steering all searches for COCA-COLA to purchase pages for substitute brands. The law should not wholly ignore this possibility. Similarly, blatant SEO tactics cause enormous consumer confusion—particularly when they push genuine results entirely out of view—and a rule that such tactics are categorically immune from trademark scrutiny because search engine spamming is not trademark use seems perverse.

Finally, as in so many other contexts, the degree of concern increases with the opacity of the search engine's processes and the paucity of useful alternative search engines available to users. Decisions that affect these other matters affect the trademark inquiry. Thus, at the risk of sounding like a broken record, I will reiterate the theme of this Part. *Looking at various strands of search engine law together makes important connections clear*. Trademark law itself tries to incorporate many of these concerns, so being aware of how they play out in the search engine context improves the clarity of the doctrinal trademark inquiry itself.

V. CONCLUSION

This Article has argued that search engine law is important and that it is complicated. It is important because it exists at the point of convergence of many strands of Internet law. It is complicated for the same reason. The bulk of this Article has been an examination of the many doctrines from which we must assemble a coherent law of search engines, and of their many interrelationships. Search engine law is a system with many moving parts but few degrees of freedom, and the challenge will be to satisfy as many competing policy demands at once as possible.

Legal scholars have much work to do in the search space. Isolated patches—the trademark law of keyword advertising sales, access to computer systems, intermediary liability for the flow of copyrighted materials, and a few others—have received careful and sustained scholarly attention. But these efforts must be linked up, and supplemented with equally

---

[250] Another way of analyzing this distinction is to argue that "use in commerce" incorporates a requirement that the "use" be "as a mark," and that search engine keyword sales do not use the trademark *qua* trademark. *See, e.g.*, Stacey L. Dogan & Mark A. Lemley, *Trademark and Consumer Search Costs on the Internet*, 41 Hous. L. Rev. 777 (2004). *But see* Graeme B. Dinwoodie & Mark B. Janis, *Confusion Over Use: Contextualism in Trademark Law*, University of Iowa Legal Studies Research Paper No. 06-06 (2006), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=927996 (arguing that "use in commerce" incorporates no such requirement and that such a requirement would be unwise, particularly for search engines).

[251] *See, e.g.* Coca-Cola Co. v. Overland, Inc., 692 F.2d 1250 (9th Cir. 1982).

thoughtful analyses of the many other specific conflicts created by search. And they must be connected with more overarching studies of larger themes in search engine law.

This Article has repeatedly referred to some of these themes: the tension between transparency and secrecy in search engine operations; the relationship of competition among providers and competition among search engines; the power of search engines to promote and to infringe upon privacy of users, providers, and third parties; the role of search engines in enhancing and inhibiting free speech; and the political economy of innovative freedom and others' claims upon search engines. A fuller discussion of these themes will need to await other days and other articles. The need for such further study should by now be apparent.

As of this writing, Google lists 583 results for "search engine law." That number will not remain so small for long.