REGULATING SEARCH ENGINES: TAKING STOCK AND LOOKING AHEAD

"To exist is to be indexed by a search engine"
(Introna & Nissenbaum)

URS GASSER*

TABLE OF CONTENTS

* Associate Professor of Law, S.J.D. (St. Gallen), J.D. (St. Gallen), LL.M. (Harvard), Attorney at Law, Director, Research Center for Information Law, Univ. of St. Gallen, Faculty Fellow, Berkman Center for Internet & Society, Harvard Law School. I owe special thanks to my colleague James Thurman and the research team at the Research Center for Information Law at the Univ. of St. Gallen. Particular thanks also to the organizers and participants of the Information Society Project's "Regulating Search?" conference at Yale Law School. Further thanks are due to Herbert Burkert, John Palfrey, and Sacha Wunsch-Vincent.

I. INTRODUCTION

Since the creation of the first pre-Web Internet search engines in the early 1990s, search engines have become almost as important as email as a primary online activity. Arguably, search engines are among the most important gatekeepers in today's digitally networked environment. Thus, it does not come as a surprise that the evolution of search technology and the diffusion of search engines have been accompanied by a series of conflicts among stakeholders such as search operators, content creators, consumers/users, activists, and governments. While few tussles existed in the initial phase of innovation where Internet search engines were mainly used by 'techies' and academics, substantial conflicts emerged once the technology got out of the universities and entered the commercial space. When search technology advanced and search services gained commercial significance, these conflicts became more severe and made their way into the legal arena. At the core of most of these disputes were controversies over intellectual property, particularly trademark and copyright issues.

Recently, the growing market power of a few search engine providers and their increased role in controlling access to information and agenda setting has triggered a new series of concerns and conflicts, permeating consumer protection, competition law, and free speech issues. Some of these issues have been subject to litigation; others have been dealt with in the context of industry self-regulation. However, certain issues are or will be considered by regulators and legislators. In contrast to the initial responses by the legal system to the new phenomena—responses that have been rather perfunctory and based on traditional doctrines—the emerging legal and regulatory issues are likely to concern the role and functionality of search engines in broader terms. At this inflection point, it becomes important to avoid premature legislative or other forms of governmental intervention. Rather, a thorough assessment of alternative regulatory approaches and strategies that might be applied in the future is required. Such an assessment, however, requires an open discussion and shared understanding of what fundamental policy objectives should underlie today's information society in the first place.

In this light, the paper has two objectives. First, it seeks to take stock and provide a brief summary of the current state of an emerging law of search engines, mainly from a U.S. perspective. Second, it aims to contribute to the development of an analytical framework that may provide guidance in assessing proposals aimed at regulating search engines in particular and search more generally. The paper is organized in three Parts. In Part I, I provide a brief history of search engines to set the stage for Part II, which will briefly discuss the initial responses by the legal system to the phenomenon "search engines," hereby focusing on the past and the present and looking at case law on the one hand and regulatory as well as legislative

interventions on the other hand. This discussion is not intended to be a detailed exposition, but rather will simply map out overall trends. Part III, in broader terms, identifies key policy themes of an evolving debate about the regulation of search engines that seems more comprehensive than previous discussions. Against this backdrop, I will briefly illustrate the need for a systematic evaluation of alternative (or competing) approaches to search regulation. The paper finally discusses core values of a democratic information ecosystem from which one might derive normative criteria for the assessment of search engine governance proposals.

II. A BRIEF (AND CASUAL) HISTORY OF SEARCH ENGINES

The history of Internet search tools starts in 1990,[1] when a group of McGill University of Montreal students created Archie, a script-based data gathering program that downloaded the directory listings of all the files located on FTP sites and created a searchable database of filenames.[2] Archie was a response to the primary method of storing and retrieving files in the pre-Web days, where files where scattered on public anonymous FTP servers and could only be located if someone announced the availability of the file via email to a message list, a discussion forum, or the like. A year later, a distributed document search and retrieval network protocol called Gopher was released by a group of researchers at the University of Minnesota,[3] followed by the appearance of the searching programs Veronica and Jughead, which searched the files sorted in the Gopher index systems and provided a keyword search of menu titles and listings on thousands of Gopher servers.[4]

Access to the Internet rapidly expanded outside its previous domain of academia and industrial research organizations once the World Wide Web (WWW), publicly available since August 1991,[5] gained critical mass in 1993 through the appearance of the web browser "Mosaic," the first program providing a graphical user interface.[6] Parallel to Mosaic's release, the first Web search engine emerged. Wandex was an index of captured

---

[1] *See, e.g.*, *Search Engine*, *in* WIKIPEDIA, http://en.wikipedia.org/wiki/Search_engine (last visited April 24, 2006) (providing a timeline of search engine development).
[2] See *Archie Search Engine*, *in* WIKIPEDIA, http://en.wikipedia.org/wiki/Archie_search_engine (last visited April 24, 2006).
[3] *See Gopher Protocol*, *in* WIKIPEDIA, http://en.wikipedia.org/wiki/Gopher_protocol (last visited April 24, 2006).
[4] *See Veronica (Computer)*, *in* WIKIPEDIA, http://en.wikipedia.org/wiki/Veronica_%28computer%29 (last visited April 24, 2006); *Jughead (Computer)*, *in* WIKIPEDIA, http://en.wikipedia.org/wiki/Jughead_%28computer%29 (last visited April 24, 2006).
[5] *See World Wide Web*, *in* WIKIPEDIA, *in* http://en.wikipedia.org/wiki/World_wide_web (last visited April 24, 2006).
[6] *See Mosaic Web Browser*, *in* WIKIPEDIA, http://en.wikipedia.org/wiki/Mosaic_web_browser (last visited April 24, 2006).

URLs and based on the first web crawler called World Wide Web Wanderer, originally designed at MIT to track the web's growth. At the same time, other search engines appeared, including Aliweb, where webmasters of participating sites posted their own index information for the pages they wanted to list, and which avoided the early web crawler's problem causing performance degradation. The first full-text crawler-based search engine, however, appeared in 1994. The search engine WebCrawler with its simple browser-based interface let users search for any word in any web page and became very popular within months.[7] Also in 1994, the search engine Lycos was created, born from a research project at Pittsburgh's Carnegie Mellon University. It was the first search engine to use (outbound) links to a web site to determine context and relevance, respectively.[8] Additionally, Lycos displayed not only the title and ranking of a page as its predecessor, but provided "snippets" of web pages,[9] and added features such as prefix matching and word proximity. Arguably, however, Lycos' main difference was the size of its catalog, which had reached 1.5 million documents by January 1995 and 60 million documents by November 1996, more than any other search engine back in the early days of the WWW.[10]

By 1995, several other search tools—providing different degrees of innovation—had emerged, including Infoseek, AltaVista, and Excite. Infoseek was based on existing technology; it introduced a complex system of search modifiers[11] and became popular due to a strategic partnership with web browser Mosaic Netscape.[12] AltaVista, developed and marketed by Digital Equipment Corporation (DEC), went online in late 1995 and soon became the "king of search."[13] It is considered to be the first high-speed search engine that enabled natural language search. AltaVista was also the first multi-lingual search engine, and included features such as advanced searching techniques (e.g. searching for phrases using quotes),[14] and the ability to search for sites that link to a particular URL.[15] Excite, created by a group of Stanford students, also launched in 1995 with a web directory

[7] *See Webcrawler*, *in* WIKIPEDIA, http://en.wikipedia.org/wiki/WebCrawler (last visited April 24, 2006).
[8] JOHN BATTELLE, THE SEARCH: HOW GOOGLE AND ITS RIVALS REWROTE THE RULES OF BUSINESS AND TRANSFORMED OUR CULTURE 53 (2005).
[9] *Id.* at 54.
[10] *See* Michael Maudlin, *Lycos: Design Choices in an Internet Search Service*, IEEE EXPERT, Jan.-Feb., 1997, at 8, *available at* http://www.lazytd.com/lti/pub/ieee97.html.
[11] *See Infoseek*, *in* WIKIPEDIA, http://en.wikipedia.org/wiki/Infoseek (last visited April 24, 2006).
[12] Wes Sonnenreich, A History of Search Engines (1997), http://www.wiley.com/legacy/compbooks/sonnenreich/history.
[13] BATTELLE, *supra* note 8, 51.
[14] *Alta Vista*, *in* WIKIPEDIA, http://en.wikipedia.org/wiki/Alta_Vista (last visited April 24, 2006).
[15] *See* Sonnenreich, *supra* note 12.

and a search engine. Reportedly, it was the first search engine "to transcend classic keyword-based searching with technology that grouped Web pages by their underlying concepts" to fine-tune search results to its users.[16] These full-text indexing search engines were in strong competition with Yahoo!, which made its debut in late 1994 and followed a different search paradigm by providing hierarchical, subject-classified directories of web content.[17]

Since competing search engines used different techniques, they produced different search results—a phenomenon that led in the mid 1990s to the development of meta-search engines such as MetaCrawler or Savvy Search. This generation of search engines forwarded search queries to all of the major web engines at once and compiled search results, although they were not able to synchronize the search syntaxes offered by the various search engines.[18] Another innovation was the introduction of personalized search, where search results were custom tailored to personal profiles or the like. HotBot, for instance, a search engine released in 1996 with a capacity to index over 10 million pages per day, made use of cookies to store personal search preferences. In a later version of the program, however, the functionality disappeared. In 2000, finally, major search engine providers including AltaVista introduced customized search.[19]

Several other search engines were released between 1995 and 2000, while others were acquired, integrated, or otherwise disappeared from the market.[20] By 2001, Google (launched in 1998 by Larry Page and Sergey Brin) had become one of the most prominent search engines.[21] Arguably, its success was based on its simple user-interface on the one hand, and the concept of link popularity and PageRank, "a method for rating Web pages objectively and mechanically, effectively measuring the human interest and attention devoted to them," on the other hand.[22] Since 2000, several other search engines have appeared, among them Yahoo! Search, MSN Search, and (Google-based) A9, to name just a few. The underlying technologies of

---

[16] BATTELLE, *supra* note 8, 55.
[17] *See, e.g.*, The History of Yahoo – How it all Started (2005), http://docs.yahoo.com/info/misc/history.html.
[18] Sonnenreich, *supra* note 12.
[19] *See, e.g.*, Greg Notess, *Customization Options for Web Searching*, ONLINE, Jan. 2001, *available at* http://www.onlinemag.net/OL2001/net1_01.html.
[20] For an overview, see *Search Engine*, *in* WIKIPEDIA, http://en.wikipedia.org/wiki/Search_engine#History (last visited April 24, 2006).
[21] See, e.g., Corporate Information, http://www.google.com/corporate/history.html (last visited April 4, 2006).
[22] Lawrence Page, Sergey Brin, Rajeev Motwani, Terry Winograd, *The PageRank Citation Ranking: Bringing Order to the Web* (Jan. 28, 1998), http://dbpubs.stanford.edu:8090/pub/showDoc.Fulltext?lang=en&doc=1999-66&format=pdf&compression=&name=1999-66.pdf. For a detailed account of the Google success story, see BATTELLE, supra note 8. For an overview, *see, e.g.*, *Google (Search Engine)*, *in* WIKIPEDIA, http://en.wikipedia.org/wiki/Google_%28search_engine%29 (last visited April 24, 2006).

search engines—web crawling, indexing, and searching—have become even more advanced and efficient. Recently, search engines are using new protocols such as XML or RSS that are increasingly provided automatically by websites such as weblogs and news sites and that allow for more efficient data indexing without requiring extensive crawling. Another recent innovation in search engine technology is the inclusion of geocoding, a process that matches search results to geographic locations such as street address, neighborhood, and the like.[23] Other trends in search are, among others, vertical search (e.g. image or product search), local, personal, and contextual search.[24]

The technological advancement has been accompanied by an enormous increase in the index size of search engines. Despite difficulties in measuring and comparing index sizes over time, the following numbers might illustrate the scale of growth in the size of search engines. By the end of 1999, for instance, major search engines indexed up to 200 million documents. In June 2000, Google set a new benchmark of 500 million indexed pages. In 2002, the largest search engines reportedly indexed already 3 billion pages, by the end of 2003 4 billion indexed pages (and other file formats.) By 2004, MSN indexed 5 billion documents, and in November 2004 Google increased its database index to a record of 8 billion documents.[25] By mid 2005, the Yahoo! Search index provided access to 20 billion items, including 19.2 billion web documents, 1.6 billion images, and over 50 million audio and video files.[26] It is expected that the trend will continue as new content is indexed, both in the form of existing online content (such as home videos)[27] and in offline materials (such as books)[28] that are digitized for the purpose of online search and accessibility.

---

[23] *See*, *Search Engine*, *in* WIKIPEDIA, http://en.wikipedia.org/wiki/Search_engine#Geospatially_enabled_search_engines (last visited April 24, 2006).

[24] "Vertical" search refers to specialized search engines. For instance, Indeed.com, LinkedIn.com, and SimplyHired.com are all vertical search engines designed for searching for jobs. Examples of "local" search are local.google.com, local.yahoo.com, and local.ask.com/local. Yahoo provides a "contextual" search tool which allows users to conduct searches relating to the content of a webpage while viewing that very webpage. *See*, Margaret Kane, *Yahoo Launches 'Contextual' Search*, NEWS.COM, Feb. 3, 2005, http://news.com.com/Yahoo+launches+contextual+search/2100-1038_3-5561712.html.

[25] These numbers have been taken from Danny Sullivan, *Search Engine Sizes*, SEARCH ENGINE WATCH, Jan. 28, 2005, http://searchenginewatch.com/reports/article.php/2156481.

[26] *See* Tim Mayer, *Our Blog is Growing Up – And So Has Our Index* (Aug. 8, 2005), http://www.ysearchblog.com/archives/000172.html. Google, however, questioned the accuracy of this number. *See, e.g.*, Elinor Mills, *Google to Yahoo: Ours Is Bigger*, NEWS.COM, Sept. 26, 2005, http://news.com.com/Google+touts+size+of+its+search+index/2100-1038_3-5883345.html.

[27] Google has begun a project in which they permit users to upload their personal videos to Google's servers. *See* Juan Carlos Perez, *Google Lets You Upload Your Own Videos*,

Since the early days of web search, search engine providers are not only in the search business, but to varying degree also in the *advertising business*.[29] In fact, advertisement is the main revenue source of many search engines—including players such as Google, Yahoo!, AskJeeves, and LookSmart.[30] Advertising in the search engine context can take different forms. On the one hand, traditional types of advertisements such as display ads, sponsorships, and listings or classified ads have been replicated by search engine providers.[31] On the other hand, search-specific advertising products have emerged.[32] The two most prominent types of search-specific advertisements are paid placement, where an advertisement is linked to a search term, and paid inclusion, where the advertiser pays a fee to the search engine provider in order to get a site included in the search index.[33] As will be discussed below, paid inclusion in particular has caused much controversy among users and even intervention on the part of regulators.[34] Current trends in advertising, as far as search engines are concerned, include portal advertising, such as that found on yahoo.com, "query-based paid placement," where favorable link positioning is sold or advertising is tied to particular search terms, and "content-targeted advertising," where a search service sends advertising to a web page upon determining relevant topics covered in the web page.[35] Google's AdSense program is the prime example of this last form of advertising. The revenue derived from advertising can be substantial. Google, which derives the majority of its

---

PCWORLD.COM, April 14, 2005, http://www.pcworld.com/news/article/0,aid,120434,00.asp.

[28] Google's library project involves the scanning of books in the collections of the Harvard, Stanford, Oxford and University of Michigan libraries as well as that of the New York Public Library. *See e.g.*, Jefferson Graham, *Google's Library Plan 'a Huge Help*', USA TODAY.COM, Dec. 15, 2004, http://www.usatoday.com/money/industries/technology/2004-12-14-google-usat_x.htm.

[29] *See*, e.g., Elizabeth Van Couvering, *New Media? The Political Economy of Internet Search Engines*, Sept. 2, 2004, at 6, *available at* http://personal.lse.ac.uk/vancouve/IAMCR-CTP_SearchEnginePoliticalEconomy_EVC_2004-07-14.pdf.

[30] According to Van Couvering's study, 95% of Google's, 82% of Yahoo!'s , 96% of AskJeeves, and 90% of LookSmart's total revenues in 2003 came from advertisement. *Id*. at 7. Some commentators, however, have questioned the wisdom of Google's (continued) dependence on advertising as well as the viability of advertising in web *applications* as opposed to web content. *See, e.g.*, the discussion on ZDNet from December, 2005, http://blogs.zdnet.com/SAAS/?cat=24 (last visited April 24, 2006).

[31] *See* Van Couvering, *supra* note 29, at 11-13.

[32] *Id.* at 13-17.

[33] *See, e.g.*, Rita Vine, *The Business of Search Engines*, at 26, *available at* http://www.workingfaster.com/2004_business_of_search_engines_final.pdf (last visited April 24, 2006).

[34] *Infra* Part B.

[35] *See, e.g.*, Michael Rappa, *Business Models on the Web*, http://digitalenterprise.org/models/models.html (last visited April 24, 2006).

revenue from advertising, posted income of $6,065,003,000 in 2005.[36] In the third quarter of 2005, Yahoo reported revenue of $442 million from search advertisements, compared with Google's $1.6 billion in that quarter.[37]

III. SEARCH ENGINE REGULATION: PAST AND PRESENT

A. OVERVIEW OF SEARCH ENGINE-RELATED CASES

1. Period before 2000

In the years before 2000, the number of cases concerning search engines and/or web search had been limited, although the importance of search engines was widely recognized only a few years after the web started off and the first full-text crawler-based search engine emerged. Courts, too, acknowledged the role of search engines in cyberspace. In mid 1996, the District Court for the Eastern District of Pennsylvania, for instance, described the situation based on a stipulation filed by the parties as follows:

> "… A variety of systems have developed that allow users of the Web to search particular information among all of the public sites that are part of the Web. Services such as Yahoo, Magellan, Altavista, Webcrawler, and Lycos are all services known as "search engines" which allow users to search for Web sites that contain certain categories of information, or to search for key words. For example, a Web user looking for the text of Supreme Court opinions would type the words "Supreme Court" into a search engine, and then be presented with a list of World Wide Web sites that contain Supreme Court information. This list would actually be a series of links to those sites. Having searched out a number of sites that might contain the desired information, the user would then follow individual links, browsing through the information on each site, until the desired material is found. For many content providers on the Web, the ability to be found by these search engines is very important."[38]

ACLU v. Janet Reno was among the first rulings where the functionality and importance of web search engines were explicitly discussed. The role of search engines was also mentioned in *Lockheed*

[36] Google Income Statement, http://investor.google.com/fin_data.html (last visited April 24, 2006).
[37] Saul Hansell, *Yahoo Reports Revenue Gains Bolstered by Online Ads*, NYTIMES.COM (Oct. 19, 2005), http://www.nytimes.com/2005/10/19/technology/19yahoo.html?ex=1287374400&en=bdafd1ae5ed986ac&ei=5090&partner=rssuserland&emc=rss and Google Income Statement, http://investor.google.com/fin_data.html. The New York Times' figure of $1.16 *million* for Yahoo's *total* advertising revenue for the third quarter of 2005 must certainly be a typographical error.
[38] ACLU v. Reno, 929 F. Supp. 824, 837 (E.D. Pa. 1996).

Martin Corp. v. Network Solutions, Inc., a trademark case brought by a company against the domain name registrar.[39] The ruling highlighted the importance of corporate names, trademarks or servicemarks as domain names, arguing that keyword searches on the web (as opposed to cases in which users know the exact address) "often yield thousands of possible Web sites," and that "[s]uch a cumbersome process is rarely satisfactory to businesses seeking to use the Web as a marketing tool."[40]

At the same time, the first search engine-specific cases were brought before courts. One might roughly distinguish between two categories of cases. First, there were disputes between web site providers (beneficiaries of search engines) who sought to use certain features of search engines in order to get more attention. Second, there emerged a few conflicts between web site providers on the one hand and search engine operators on the other hand.

– The first category, of course, refers to the use of meta tags by web page providers. Meta tags are HTML elements used to provide metadata about a web page. In the early days of web search, search engines had used meta tag data to classify a given web page and, based on this system, to generate and display a list of search results matching a given query.[41] However, webmasters quickly learned the commercial significance of having the 'right' meta tag, as it frequently led to a high ranking in the search engines and, consequently, to more 'hits.' One practice that soon became subject to litigation was "pagejacking," where the traffic to a web page was increased by "falsifying the information in metatags to emulate the appearance of another Web site in search engine results."[42] Among the first cases concerning meta tagging,[43] starting in mid 1997, were *Oppedahl & Larson v. Advanced Concepts* (no opinion issued),[44] *Insituform Technologies, Inc v. National Envirotech Group, LLC*,[45] *Playboy Enterprises, Inc. v. Calvin Designer Label*,[46] *Patmont Motor Werks, Inc. v. Gateway Marine, Inc.*,[47] *Playboy Enterprises,*

[39] 985 F. Supp. 949 (D. Cal. 1997).
[40] *Id.* at 952.
[41] Since early 2000, search engines have not relied on meta tags due to the inappropriate use of meta keywords or other practices aimed at increasing a web page's search engine ranking. Some search engines still take meta tags into consideration. In addition, techniques are applied to down-rank web sites that "game the system." *See, e.g.*, *Metatags*, WIKIPEDIA, http://en.wikipedia.org/wiki/Metatags (last visited April 24, 2006).
[42] DAVID W. QUINTO, LAW OF INTERNET DISPUTES, §10.01[A], 10-5 (2001 & Supp. 2003).
[43] *See, e.g.*, QUINTO, *supra* note 42, at § 10.01; Danny Sullivan, *Search Engine Lawsuits O'Plenty*, Dec. 16, 1999, http://searchenginewatch.com/sereport/article.php/2167671.
[44] No. 97-1592 (D. Colo. 1998).
[45] No. 97-2064 (E.D. La. 1997).
[46] 985 F. Supp. 1220 (N.D. Cal. 1997).
[47] 1997 WL 811770 (N.D. Cal. 1997).

Inc. v. AsiaFocus International, Inc.,[48] *Playboy Enterprises, Inc. v. Welles*,[49] *Niton Corp. v. Radiation Monitoring Devices, Inc.*,[50] and *Brookfield Communications, Inc. v. West Coast Entertainment Corp*.[51]

– The second group of early cases is more interesting from the perspective of search engine regulation, because here the lawsuits were directly targeted against search engine operators. Prior to 2000, at least three cases deserve particular attention. In *Ken Roberts Co. v. GoTo.com*,[52] the Ken Roberts Company brought suit for the unauthorized use of Roberts' name (in both web content and meta tags) and likeness on the part of several financial trading related websites. Although GoTo.com was dismissed from the suit on February 9, 2000, Hi-Tech Futures Trading, Inc. and Softrade, Inc. were found liable of Lanham Act trademark-related violations as well as violations of state-based laws, such as unfair business acts.[53] The suit in *Playboy Enterprises, Inc. v. Netscape Communications Corp*.[54] concerned the search engine's business practice of "keying" search terms (plaintiff's marks) to advertising banners for adult products. The plaintiff claimed, in essence, "'that Excite [and Netscape] has hijacked and usurped PEI's good will and reputation by exploiting a search based on a PEI mark as an opportunity to run banner advertisements and display directories specifically keyed to the PEI marks'"[55] and therefore sought a preliminary injunction against Netscape's and Excite's further use of the marks. The District Court held that Playboy had failed to show that Netscape had used Playboy's marks in interstate commerce—as opposed to generic terms of the English language, failed to show that there was likelihood for consumer confusion, failed to show sufficient evidence of trademark dilution, and additionally held that Netscape's use of search terms was protected by the First Amendment and constituted fair use as well.[56] The third case, *Kelly v. Arriba Soft Corp*.,[57] is neither linked to meta tagging nor keying. Rather, it involved copyright issues triggered by a technological

[48] 1998 WL 724000 (E.D. Va. 1998).
[49] 7 F. Supp. 2d 1098 (S.D. Cal. 1998).
[50] 27 F. Supp. 2d 102 (D. Mass. 1998).
[51] 174 F.3d 1036 (9th Cir. 1999).
[52] 2000 WL 33680439 (N.D. Cal. 2000).
[53] *Id.*
[54] 55 F. Supp. 2d 1070 (C.D. Cal. 1999).
[55] *Id*. at 1081.
[56] *See generally, Playboy Enterprises, Inc.*, 55 F. Supp. 2d 1070 (C.D. Cal. 1999). Part II.A.2 of this paper discusses later decisions involving this case.
[57] 77 F. Supp. 2d 1116 (C.D. Cal. 1999).

> innovation. The defendant operated a visual search engine on the Internet, which allowed users to search the web for pictures and produced a list of reduced, "thumbnail" pictures related to the user's query. The plaintiff, a photographer, claimed that some of his online images were indexed by the search engine's crawler and put in the defendant's image database, thus becoming available in thumbnail form to the search engine's users. He argued, among other things, that his copyrights in the images were infringed by the defendant's actions and claimed a violation of the DMCA. The court, on first impression, held the use of copyrighted images by the visual search engine as a prima facie copyright violation, but one that was justified under the fair use doctrine. It further held that the DMCA was not violated.

In sum, a rough overview of the case law prior to 2000 suggests that the growing importance of search engines was widely acknowledged and undisputed as early as 1996. Further, this brief analysis has made clear that initial conflicts surrounding search engine and search practices that made their way into courtrooms dominantly concerned intellectual property rights—a set of claims and issues that can be seen as typical for the transition from the phase of innovation to the phase of commercial exploitation. Interestingly, though, the majority of the early rulings concerned beneficiaries of search engines, i.e., web site providers who used legitimate and illegitimate practices to increase their visibility in cyberspace. Only in a few cases (that made it to the courts) claims were brought against search engine operators directly. In this context, it might be interesting to note that our survey has not proven the possible hypothesis that the subject of litigation would be closely related (although time-delayed) to the steps of evolution in search technology or the underlying business models as they have been outlined in Part I of this paper. Rather, the claims prior to 2000 involved rather basic and stable features of contemporary search engines. Only *Kelly v. Arriba Soft Corp.* concerning image search could be interpreted as a reaction to a more specific innovation in search technology.

2. Period after 2000

According to an extensive Westlaw search, the year 2000 marks the crossroad in search engine-specific case law, primarily from a quantitative, but to some extent also qualitative perspective. First, some of the cases decided by the courts of first instance got appealed and were decided in the new millennium by appellate courts. Among them were the above-mentioned *Playboy Enterprises, Inc. v. Netscape Communications Corp.*[58]

[58] 354 F.3d 1020 (9th Cir. 2004).

and *Kelly v. Arriba Soft Corp.*[59] In the former case, the Ninth Circuit reversed the lower court's grant of summary judgment in favor of Netscape and Excite, holding that there was a genuine issue of fact as to whether the keying practices constituted trademark infringement and dilution. The *Playboy* court heavily relied on the initial interest confusion analysis as set forth in an earlier case,[60] ruling that a banner ad that clearly identified its source with the sponsor's name might eliminate the existing likelihood of initial interest confusion. A week after the appeals court ruling, the companies reached a settlement under undisclosed terms.[61] *Kelly* was also appealed. The Ninth Circuit remanded the case in part, ruling that the use of the images as thumbnails was fair use, but declined to extend that holding to the use of full size images.[62]

Second, many more lawsuits against search engines concerning the sales of third party trademarks for use in sponsored links and banner ads were filed after 2000, since keyword advertising had become the key driver of the search engine business.[63] Some of them were settled or dismissed before judgment, others decided by courts. Among the cases that gained a lot of attention was *Geico v. Google*.[64] The plaintiff claimed, *inter alia*, that Google and Overture's sale of the marks GEICO and GEICO DIRECT as keywords constituted trademark infringement, contributory infringement, vicarious trademark infringement, unfair competition, and trademark dilution under the Lanham Act. A district court denied the defendants' motion to dismiss and held that the plaintiff had alleged facts sufficient to support its liability claims. While Geico and Overture reached a settlement, the trial court later held that Geico had not presented sufficient evidence that Google's sale of trademarks to others as keywords constituted trademark infringement since the ads themselves did not include the trademarks and there was no evidence that the relevant activity standing alone caused confusion. Other cases concerning similar trademark issues include *Google v. American Blind and Wallpaper Factory, Inc.*,[65] *Novak v.*

[59] 280 F.3d 934 (9th Cir. 2002).
[60] Brookfield Communications Inc. v. West Coast Entertainment Corp., 174 F.3d 1036 (9th Cir. 1999).
[61] *Update 14*, LINKS AND LAW, Feb. 14, 2004, http://www.linksandlaw.com/news-update14.htm.
[62] 336 F.3d 811, 2003 (9th Cir. 2003).
[63] For a comprehensive overview, *see, e.g.*, Heidi S. Padawer, *Google This: Search Engine Results Weave a Web for Trademark Infringement Actions on the Internet*, 81 WASH. U. L.Q. 1099 (Winter 2003); Lauren Troxclair, *Search Engines and Internet Advertisers: Just one Click Away from Trademark Infringement?*, 62 WASH. & LEE L.REV. 1365 (Summer 2005); Perry Viscounty & Jordan Kushner, *Order to Confusion: Trademark Infringement Liability for Search Engine Keying Ads*, 1 HASTINGS BUS. L.J. 151 (May 2005); *see also* Eric Goldman, *Deregulating Relevancy in Internet Trademark Law*, 54 EMORY L.J. 507 (2005).
[64] Government Employees Ins. Co. v. Google, Inc., 330 F. Supp. 2d 700 (D. Va. 2004).
[65] 74 U.S.P.Q.2d 1385, 2005 WL 832398, No. 03-05340 (N.D. Cal. 2005).

Overture Services Inc.,[66] and *800-JR-Cigar v. Overture*,[67] and (more recently) *Newborn v. Yahoo!, Inc*.[68]

Similarly, the number of copyright-related claims against search engine operators has increased, especially recently. The plaintiff in *Perfect 10 v. Google, Inc*. claimed, among other things, that Google directly infringed Perfect 10's copyrights in images by making those images available as thumbnails and was vicariously and contributorily liable for linking to third party sites which featured unauthorized full-size images belonging to Perfect 10.[69] In ruling on Perfect 10's motion for a preliminary injunction, the District Court for the Central District of California held with regard to Google that Perfect 10 was likely to succeed on its claim for direct infringement but not on the claims for vicarious and contributory infringement.[70] Another series of recent cases deals with the cache function as provided, for instance, by Google. In *Field v. Google, Inc.*,[71] the plaintiff claimed that Google directly infringed copyright when Google users clicked on a cached link to the web pages containing copyrighted materials and downloaded a copy of these works. The court, in contrast, held that it was the search engine user rather than the search engine operator that created and distributed copies of the copyrighted work in this process. Since Google remained passive in this process and only responded automatically to users' requests, Google's conduct did not constitute a direct copyright infringement. Further, the court held, *inter alia*, that Google held an implied license since the plaintiff took several steps to get his works included in the engine's search results, where he knew they would be archived. Further, the plaintiff deliberately ignored options that would have instructed Google not to present cached links. The court also ruled that the relevant use of the copyrighted materials constituted a fair use. A similar claim underlay *Parker v. Google*,[72] where the plaintiff alleged direct infringement from Google's automatic archiving of a USENET site that contained a posting of the plaintiff's ebook. The court found no direct infringement because of the automated and non-volitional nature of archiving.

Third, other types of conflicts emerged post-2000 and were brought to courts. A series of cases was triggered by the increased use of so-called "spiders" for the purpose of content aggregation. *EBay, Inc. v. Bidder's Edge, Inc*.[73] is among the landmark cases in this context.[74] EBay, as the

[66] 309 F. Supp. 2d 446 (E.D. N.Y. 2004).
[67] No. 2:00-03179 (D. N.J. 2000).
[68] 391 F. Supp. 2d 181 (D.D.C. 2005).
[69] Perfect 10 v. Google, Inc., 78 U.S.P.Q.2d 1072 (C.D. Cal. 2006).
[70] *Id*.
[71] F. Supp. 2d, 77 U.S.P.Q.2d 1738 (D. Nev. 2006).
[72] Parker v. Google, Inc., No. 04-CV-3918, 2006 WL 680916 (E.D. Pa. 2006).
[73] 100 F.Supp.2d 1058 (N.D. Cal. 2000).

provider of the famous Internet auction site, sued its competitor, Bidder's Edge, which, by using spiders, compiled listings for specific items from several online auction sites, including eBay and displayed them in aggregated form on its own website. After technological measures aimed at blocking the entry of the competitor's spiders failed, eBay filed suit and claimed that the defendant was committing a trespass to chattels. The district court granted preliminary injunction in favor of eBay. The court held that the use of spiders was likely to qualify as "trespassing" in eBay's servers, thereby consuming at least a portion of eBay's bandwidth and server capacity and therefore depriving eBay of the ability to use that portion of its personal property for its own purposes.[75]

Another problem involved the alleged manipulation of PageRanks by Google. SearchKing, a company selling ad space on sites ranked highly by the PageRank system, claimed that the search engine purposefully and maliciously manually decreased the PageRank of SearchKing and certain other web sites once it learned that SearchKing profited from the search engine's system. The plaintiff alleged that the down-ranking caused immensurable harm to its goodwill and business relations. Google, by contrast, considered PageRank to be a protected opinion under the First Amendment. The court in *Search King, Inc. v. Google Technology, Inc*.[76] agreed and held that Google's actions were privileged, although it could be argued that the search engine had acted maliciously and wrongfully as to SearchKing. The court ruled that the defendant (absent any business relationship with the plaintiff) had no duty to rank, or refrain from ranking, the plaintiff's or any other website. The court concluded that the plaintiff took the risk to build a business model that largely depended on a factor over which it had no control, and concluded that a unilateral change of the factor under such circumstances cannot give rise to a claim for tortious interference with contractual relations. The controversy over downgrading PageRanks, however, is not yet over. A more recent class action lawsuit has been filed in the Northern District of California.[77] Time will tell if the California District Court will reach a similar conclusion regarding the manipulation of rankings on the part of search engine providers.

Other issues up for discussion that recently emerged in the search engine context are privacy and defamation, respectively. In *Parker v. Google*, the plaintiff alleged that Google is liable, *inter alia*, for the tort of defamation, because the defendant archived defamatory messages posted by

---

[74] *But see* Ticketmaster Corp. v. Tickets.com, Inc., No. CV 99-7654 HCH (BQRx) (C.D. Cal. 2000); eBay, Inc. v. ReverseAuction.com, Inc., No. C-00 20023 RMW (N.D. Cal. 2000); Register.com, Inc. v. Verio, Inc., 126 F. Supp. 2d 238 (S.D.N.Y. 2000).

[75] Bidder's Edge, Inc., 100 F.Supp.2d at 1070-71.

[76] SearchKing, Inc. v. Google Technology, Inc., No. CIV-02-1457-M, 2003 WL 21464568 (W.D. Okla. 2003).

[77] Kinderstart.com, LLC v. Google, Inc., No. C-06 2057 (N.D. Cal. 2006), *available at* http://blog.searchenginewatch.com/blog/googlesuit_031806.pdf.

USENET users and because of defamatory statements that were located on a website that was in Google's cache. Further, the plaintiff claimed liability for invasion of privacy because "the act of Google users putting in a search query of his name led Google to produce a list of websites in which his name appeared, thus creating what he called 'an unauthorized biography of Plaintiff that is an invasion of his right to privacy.'"[78] The court held that the defendant is immune from such state tort claims under the Communication Decency Act.[79]

3. Conclusion

A high-level overview of cases against search engine operators since the mid 1990s leads to three tentative conclusions. First, the overview suggests that different types of concerns, tussles, and conflicts have evolved over time and made their way into the legal system. In the early days of web search and roughly up to 2000, meta tagging was apparently the most frequent subject of litigation involving search engine operators. The second generation of lawsuits against search engine operators, however, has become more diverse, although intellectual property issues—probably with a shift from trademark issues towards copyright issues—continue to play an important if not predominant role. An increased number of claims based on trespass to chattels, defamation, privacy, and other grounds might indeed signal that the conflicts surrounding search engines are broadening.

Connecting the evolution of case law with the history of search engines as outlined in Part I, it is interesting to observe that the different waves of litigation are in fact related to particular technological advancements (e.g. keyword search) and the evolution of business models (e.g. paid placement), but are less tightly connected to them as one might expect. On the one hand, important and potentially controversial innovations such as the introduction of web page summaries ("snippets") in search results, for instance, does not seem to have triggered waves of (copyright) litigation. On the other hand, conflicts that are clearly connected with an innovation in search technology—conflicts surrounding spiders, for example—found entry into the legal system only several years after mass-adaptation by users. Similarly, the timing of the legal system's response to certain business practices (like keying) is likely to depend on various factors besides the first appearance of the respective conduct, making both causal explanations and predictions difficult.

Third, the case law overview demonstrates that search engines, and search more generally, have been regulated to one degree or another since the early days of web search. Evidently, the emerging case law has a direct impact on the behavior of the involved parties. In *Bidder's Edge*,

[78] Parker v. Google, Inc., 2006 WL 680916, at *6.
[79] 47 U.S.C. § 230 (2000).

Register.com, and *Perfect 10*, for example, the plaintiffs succeeded in obtaining preliminary injunctions with respect to at least part of their claims. In some instances, the regulatory effects of litigation have been even broader. One example in this context is the strategic response of search engines to intense litigation regarding keyword advertisement. Vis-à-vis remaining uncertainty as to the applicable legal standards, some players have crafted and/or revised their respective keyword policies. Google, for instance, revised its keyword policy in 2004 in the light of recent case law, allowing U.S. advertisers to bid on trademarked keywords, but prohibiting the use of third party trademarks in the text of an advertisement.[80] Microsoft's current U.S. policy for its MSN keywords program allows informal uses of third party trademarks, but enforces its well-balanced policy by filters and other technologies, complaint procedures, and the like.[81] Yahoo! Search Marketing went a step further and recently announced that U.S. advertisers will no longer be allowed to bid on keywords trademarked by competitors.[82]

B. LEGISLATION AND REGULATION

Not only courts have been dealing with legal issues accompanying the emergence and further development of search engine technology and business. Legislators and regulators have addressed aspects related to online search in general and search engines in particular. As is not uncommon in other contexts as well, legislative proposals concerning the online environment have sometimes emerged in reaction to controversial cases. Based on the result analysis of an extensive search with terms such as "search engine," "internet directory" and "internet resources guide" on Westlaw and on THOMAS, one might discern areas of legislation where Congress clearly had implications for search engines in mind. On the other hand, amendments to Title 47 of the U.S. Code introduced new legislative terminology in response to the emerging digital revolution. Although terms such as "interactive computer service," "access software provider" and "information location tool" have become fairly common parlance in bill drafting, these terms do not always refer to the definitions contained within Title 47, nor are they always defined in the same manner. In some instances, the use of a particular term clearly implicates search engines

[80] Pamela Parker, *Google Shifts Trademark Policy*, CLICKZNEWS, April 13, 2004, http://www.clickz.com/news/article.php/3339581.

[81] *See, e.g.*, Jon M. Zieger, *Search Engine Liability for Trademark Infringement: Seeking a Balanced Policy Amidst Legal Uncertainty*, Position Paper presented at the "Regulating Search" conference at Yale Law School, December 3, 2005, *available at* http://islandia.law.yale.edu/isp/search_papers/zieger.doc.

[82] Kevin Newcomb, *Yahoo Modifies Trademark Keyword Policy*, CLICKZNEWS, Feb. 24, 2006, http://www.clickz.com/news/article.php/3587316.

while in some instances, search engines are clearly not implicated, and in yet other instances, the implications are not clear.[83] Irrespective of these problem areas, overall, one may roughly distinguish among three areas of law and regulation in which search engines have specifically gained policy-makers' and regulators' attention.[84]

The first area relates to *content regulation* and its limitations. Given the ubiquitous availability of online content and the absence of customary consumer controls that exist in brick-and-mortar stores of adult products, a number of these legislative proposals have concerned the protection of minors. The 1998 Senate Report on Commercial Distribution of Material Harmful to Minors on World Wide Web,[85] for instance, emphasized the role of search engines in cyberspace,[86] and described the problem of spoofing, where pornographers trick search engines by including innocent search terms on their web sites.[87] Similarly, the 1998 House Report on the Child Online Protection Act discussed the problem where children enter seemingly unrelated terms such as "toy" or "dollhouse" into a search engine and would be led to material harmful to minors.[88] On the other hand, search-related techniques such as meta tagging were considered as possible means of identifying harmful content and restricting its availability.[89] These issues had also been repeated, for instance, in the 1999 Senate Report on the Children's Internet Protection Act.[90] There, the Committee on Commerce, Science, and Transportation discussed the ease with which minors could come upon adult-oriented materials through the use of search engines, since search services contained no artificial intelligence to omit the content.[91]

---

[83] The term "information location tool," for instance, appears to always include search engines within the ambit of its meaning, whereas "access software provider," as defined in § 230 of the Communications Decency Act, clearly includes search engines, but as defined in the Internet Election Information Act of 1997 (H.R. 653.IH) would not likely include search engines.

[84] More obscure regulatory issues would include, for example, the SEC's statement issued March 27, 1998, in which the application of U.S. securities regulation to websites that promulgate "offering and solicitation materials" for offshore sales of investment services and securities was discussed. In a footnote, the SEC addressed the issue of meta-tagging and targeted communications, stating that it will generally not view the use of tags relating to securities or investments as transforming web sites into a targeted communication that would require additional measures to assure against sales to U.S. persons. *See* 63 Fed. Reg. 14806, 14807 (Mar. 27, 1998).

[85] S. REP. No. 105-225 (1998). The Report states that the bill was "in response to the Supreme Court ruling on the 'indecency' and 'patently offensive' provisions of the Communications Decency Act, and addresses the concerns of the Court in the case, Reno v. ACLU, 117 S. Ct. 2329 (1997)." *Id.* at 2.

[86] *Id*. at 2.

[87] *Id.* at 4.

[88] H.R. REP. No. 105-775, at 10 (1998).

[89] *Id.* at 17.

[90] S. REP. No. 106-141, at 3 (1999).

[91] *Id*.

Also with regard to the promotion of freedom of expression, the role of engines has recently been considered in the Global Online Freedom Act of 2006.[92] If the bill becomes law, it would prohibit search engines from locating any hardware associated with their services within a country designated by the act as Internet restricting,[93] and would prohibit operators from altering their search services within such a country.[94] Further, it would oblige search engine operators to provide a special committee with a list of terms intended for the filtering policy of an Internet restricting country.[95] Thus, there has been a desire on the part of Congress to limit access by certain classes to content on the one hand, and preserve the free expression of content on the other.

The second area where search engines attracted legislators' attention relates to *liability* of search operators. Search engines have been explicitly mentioned in the context of limitations on liability for copyright infringement. A bill aimed at providing limitations on copyright liability relating to material online (Digital Copyright Clarification and Technology Education Act of 1997), for instance, provided in section 102 (a proposed additional section 512 to chapter 5 of title 17 of the United States Code) a safe harbor from copyright infringement liability for search engines.[96] Similarly, the Senate Report on the Digital Millennium Copyright Act of 1998 mentioned search engines in discussion of the limitation on the liability for copyright infringement included in the bill.[97] The corresponding House Report, too, mentioned search engines in the context of the safe harbor provisions.[98] Opposition to the imposition of criminal liability on search engines, among other ISPs, for content supplied or controlled by a third party was expressed in a 2001 House of Representatives Resolution.[99] More generally, but without explicit reference to search engines, section 230 of the Communications Decency Act shields access software providers from liability derived from the "publication" of content. The term "access software provider" means a provider of software or enabling tools that, *inter alia*, cache, search, or organize content.[100] The Child Online Protection Act also contained a provision exempting persons in the business of providing an "Internet information location tool" as well as anyone engaged in the

[92] H.R. 4780, 109th Cong. (2006).
[93] *Id.* § 201.
[94] *Id.* § 202.
[95] *Id.* § 203.
[96] S. 1146, 105th Cong. § 102 (1997).
[97] S. REP. No. 105-190, at 48 (1998).
[98] H.R. REP. No. 105-551, at 56 (1998)
[99] H.R. Res. 12, 107th Cong., (2001).
[100] Pub. L. No. 104-104, tit. V, § 230(f)(4)(C), 110 Stat. 113, *invalidated by* Reno v. ACLU, 117 S. Ct. 2329 (1997). For a discussion of the Safe Harbor provisions under the DMCA and Communications Decency Act, *see generally* Jonathan Band & Matthew Schruers, *Safe Harbors Against the Liability Hurricane: The Communications Decency Act and the Digital Millennium Copyright Act*, 20 CARDOZO ARTS & ENT. L.J. 295 (2002).

"storage, retrieval, hosting, formatting, or translation" of internet communications.[101] Similarly, several other bills contemplated liability exemptions for information location tools or exempted them from the bill's purview altogether.[102]

The third area of intervention has been (general) *consumer protection*. The most prominent example belonging to this category are actions taken by the Federal Trade Commission, which issued a letter with recommendations to search engine operators in response to a complaint filed by Commercial Alert requesting the agency to investigate whether certain search engines were violating Section 5 of the Federal Trade Commission Act by failing to disclose that advertisements are inserted into search engine results lists.[103] In response, the FTC drafted a letter to search operators recommending that they review their web sites to ensure that (1) any paid ranking search results are distinguished from non-paid results with clear and conspicuous disclosures; (2) the use of paid inclusion is clearly and conspicuously explained and disclosed; and (3) no affirmative statement is made that might mislead consumers as to the basis on which a search result is generated.[104] Additionally, the Anti-Phishing Act of 2004[105] and the Internet False Identification Prevention Act of 2000[106] were both proposed as measures to combat online fraud. The Anti-Phishing Act would create criminal liability for search engines wherever they point to a fraudulent site with knowledge or intent to commit fraud or identity theft.[107] The False Identification Prevention Act, on the other hand, exempts search engines as "access software providers" or "interactive computer services" from liabilty that would be imposed by the bill with certain exceptions.[108]

C. SUMMARY

---

[101] H.R. REP. No. 105-775, at 30 (1998). Also consider the provisions of the Online Parental Control Act of 1996, H.R. 3089.IH relating to "access software providers."

[102] *See, e.g.*, Internet False Identification Prevention Act of 2000, Pub. L. No. 106-578, 114 Stat. 3075 (2000); Prisoner Web Site Disclosure Act of 1999, H.R. 1930, 106th Cong. (1999); Securely Protect Yourself Against Cyber Trespass Act , H.R. 29, 109th Cong. (2005); Ryan Haight Internet Pharmacy Consumer Protection Act of 2005, H.R. 840, 109th Cong. (2005); Medicare Drugs for Seniors (MED) Act of 2006, H.R. 4697, 109th Cong. (2006).

[103] Letter from Commercial Alert to Federal Trade Commission (July 16, 2001), http://www.commercialalert.org/PDFs/SearchEngines.pdf.

[104] Draft Letter from the Federal Trade Commission (June 27, 2002), http://www.ftc.gov/os/closings/staff/commercialalertattatch.htm.

[105] S. 2636, 108th Cong. (2004).

[106] S. 2924, 106th Cong. (2000).

[107] S. 2636, 108th Cong. § 3 (2004).

[108] S. 2924, 106th Cong. § 3(6) (Engrossed as Agreed to or Passed by Senate, October 31, 2000). The exceptions include, *inter alia*, where the service has knowingly permitted its service to be used to perpetrate an act prohibited under the bill's provisions and an officer, director, partner, or controlling shareholder has the specific intent that the service be used to that purpose. *Id.*

Part II of this paper has provided an overview of what one might call the emerging law of search engines. The previous sections have illustrated that certain search practices in general and certain forms of behavior of search engine operators in particular have been the subject of legal regulation—using the term regulation in its broad sense—since the early days when web search became a mass-phenomenon. The responses by the legal system have either been triggered by technological innovation in search or new business models, or by a combination of these factors.

In a first phase, trademark disputes were predominant issues to be resolved in courts. In a second phase, additional issues have entered the legal arena, including privacy concerns and free speech issues—although IPR disputes (including trademark and copyright) still play a very important role. At the legislative and regulatory level, content regulation and its limits, immunity from liability for copyright infringement as well as liability derived from publication of content, and consumer protection have been the key topics where the specific role of search engines has been taken into account.

The high-level analysis has shown that interventions by courts, legislators, and regulators alike have generally been issue-specific, ranging from specialties such as keying, meta tagging, spiders, to caching and paid inclusion. At the same time, however, more and more issues have become relevant from the legal and regulatory perspective, thus broadening over time the scope and reach of the law governing search and search engines. A brief overview of emerging legal and regulatory issues up for discussion in various fora, finally, has confirmed this trend.

III. POSSIBLE FUTURE: HETEROGENEOUS POLICY DEBATES AND THE NEED FOR A NORMATIVE FRAMEWORK

A. THEMES OF FUTURE POLICY DEBATES

The current state of search engine regulation as sketched in Part II has suggested that the emerging body of law is characterized by thematic diversity. In that regard, it mirrors the state of cyberlaw more generally.[109] Based on the analysis of past and present discourses in courts, parliaments, agencies, academic fora, etc., the following threads of discussion

[109] *See* Herbert Burkert, *Von künftigen Aufgaben des Informationsrechts*, *in* RECHT UND INTERNATIONALISIERUNG, 157-158 (Christian J. Meier-Schatz and Rainer J. Schweizer eds., 2000).

concerning the law and policy of search engines are likely to be the key topics of intensified regulatory debates in the future:[110]

– The *infrastructure* debate concerns the ordering of the physical and logical infrastructure necessary to provide search functionalities on the web. Issues such as the informational equivalent of common carrier rules for search engines, the obligation of providing even-handed listings, or the disclosure of a search engines' algorithm are topics belonging to this thread of discussion.[111] In some jurisdictions (particularly in Europe), this debate also includes the question of the state's role in information processes (*service public*) vis-à-vis privately owned and controlled search infrastructure. This debate intensifies in the current digital environment where the search engine market is rather concentrated and centralized.[112]

– The *content* debate covers at least three related, but analytically distinct issues. First, the discussion of search engines' role in promoting freedom of expression in general and political speech in particular.[113] Second, the controversies concerning the limitations on free speech and the search engines' responsibility in enforcing these limits, for example with regard to materials harmful to minors (should search engines remove objectionable content?). Third, the debate about the cultural bias of search engines and cultural diversity, respectively.[114]

– The *ownership* debate is directed at the future of intellectual property rights and similar claims in light of existing and evolving search technology and corresponding business models. At least three issues relate to this category. First, the discussion about the adequate scope of IP rights for search engine operators that enable them to protect their

[110] Inspired by Burkert's discussion of legal issues in cyberlaw, *supra* note 109, at 157. *See also* Urs Gasser, *What is Information Law – and what could it be?, in* INFORMATION LAW IN EENVIRONMENTS 11-12 (Urs Gasser ed., 2002).

[111] *See, e.g.*, Lucas D. Introna & Helen Nissenbaum, *Shaping the Web: Why the Politics of Search Engines Matters, available at* http://www.nyu.edu/projects/nissenbaum/papers/searchengines.pdf (last visited Apr. 24, 2006).

[112] A recent global user survey, for instance, suggests that Google's global usage share has reached 57.2%. Google User Share Rising (Feb. 7, 2005), http://www.webrankinfo.com/english/seo-news/topic-503.htm. In addition, not all search engines use their own technology. Instead, they rely on other search providers for their listings. *E.g.* Van Couvering, *supra* note 29, at 9.

[113] *See, e.g.*, the discussions surrounding the Global Online Freedom Act of 2006, *supra* note 92.

[114] See the rationale for building the above-mentioned Franco-German Search Engine "Quaero". *German Partners for European Search Engine 'Quaero'*, HEISE ONLINE, Mar. 11, 2006, http://www.heise.de/english/newsticker/news/70717.

algorithms and databases. Second, IPR issues that arise between competitors; such claims often involve patent disputes,[115] but might also include other copyright or trademark issues. Third, the obligations of search engine operators vis-à-vis the copyright and trademark claims of the providers of content that is indexed, categorized, linked, cached, etc.[116] Recent controversies regarding digitization projects suggest that these conflicts will even intensify in the months and years to come.[117]

– The *security* debate takes as central themes, among others, the security of the search infrastructure as well as security in search-related transactions. Recent disputes about click fraud attacks against search engines' advertising programs are illustrations of infrastructure security-related issues.[118]

– The *identity* and *privacy* debate comprises a broad spectrum of questions about identity management in search engine-mediated information processes, and issues about data protection and informational self-determination vis-à-vis large databases controlled by search engine operators.[119] Examples include the recent controversy surrounding the disclosure of a search engine's data requested by the Department of Justice for the purpose of monitoring sexually explicit materials on the Web,[120] the use of search history for marketing and

[115] For examples of suits brought by Digital Envoy, NetJumper, and Overture against Google for patent infringement, see Danny Sullivan, *Search Engines and Legal Issues*, Search Engine Watch, http://searchenginewatch.com/resources/article.php/2156541#Patents (last visited Apr. 24, 2006).

[116] *See supra* Part II.A. for illustrations of such conflicts.

[117] *See, e.g.*, the Google Print controversy: Author's Guild v. Google, Inc., No. 05CV8136 (S.D.N.Y. Sept. 20, 2005), and McGraw-Hill Co. v. Google, Inc., No. 05Civ8881 (S.D.N.Y. Oct. 19, 2005). See also the tussle over Google's News Services: Agence France Press v. Google, Inc., No. 05-00546 (D.D.C. filed Mar. 3, 2005). *See* Daniel Farey-Jones, *News Producers Single Out Google News in Battle Over Free Content*, BRAND REPUBLIC, Feb. 1, 2006, *available at* http://www.brandrepublic.com/bulletins/media/article/538934/news-producers-single-google-news-battle-free-content/ (last visited April 24, 2006).

[118] *See, e.g.*, Brian Quinton, *Will $90 Million Make Google Click Fraud Go Away?*, MULTICHANNEL MERCHANT, Mar. 21, 2006, http://multichannelmerchant.com/searchline/3-15-06-Google-settlement/, (discussing Lane's Gifts & Collectibles LLC v. Yahoo! Inc., No. CV-2005-52-1 (Ark. Cir. Ct. filed Feb. 17, 2005) and Advanced Internet Techs. v. Google, 2006 WL 889477 (N.D. Cal. Apr. 5, 2006)).

[119] *See, e.g.*, Herman T. Tavani, *Search Engines, Personal Information and the Problem of Privacy in Public*, 3 IRIE 39 (2005), *available at* http://www.i-r-i-e.net/inhalt/003/003_tavani.pdf.

[120] *See, e.g.*, Gonzales v. Google, Inc., 2006 WL 778720 (N.D.Cal. 2006); *Judge: Google Must Give Info to Feds*, CBS NEWS, Mar. 14, 2006, http://www.cbsnews.com/stories/2006/03/14/tech/main1401585.shtml.

other purposes, or practices such as "Google hacking," where search engines are used to gather sensitive information on the Internet.[121]

– The debate about *participation* focuses on the role of search engines in political and cultural processes and spaces. In the age of power-law distribution, what are the implications of technologies and techniques of search such as PageRank for information participation, individual dissent, and personal liberty?[122] The debate also includes questions concerning a potential "right to access search technology," and the possible need for a "right to get indexed."[123]

– The *ethics debate* concerns the reevaluation of basic concepts of right and wrong behavior in a dynamic and globalized information environment. The question is not only about the moral values shared in a given society, but also about the relationship between ethics and the law. The latter topic has gained relevance in the context of global business activities carried out by search engines, leading to conflicts between local laws and ethical commitments of U.S.-based Internet intermediaries.[124] Currently, non-legal rules for search engine providers such as code of ethics or best practices models, and the like are under consideration.[125]

In sum, this rough overview suggests that the law and policy discourse on search engines is still fairly fragmented. [126] However, given the search engines' important role in the digital society and the interdependencies between the policy areas outlined above, this discourse is likely to result in a broader governance discussion where the interactions among legal and regulatory measures, search engines, and other constituencies of the digitally networked environment need to be explored

---

[121] *See, e.g.*, Tom Sanders, *Worms turn on Google to hunt for victims*, VNUnet UK, Feb 15, 2006, http://www.vnunet.com/vnunet/news/2150292/worms-google-hunt-victims

[122] *See, e.g.*, Symposium, "Regulating Search?" Panel 4, held by the Yale Law School, *available at* http://islandia.law.yale.edu/isp/regulatingsearch.html#paneldescriptions (last visited April 24, 2006).

[123] Conversely, and linked to the privacy debate, is the issue of withholding or intentionally "down-ranking" undesirable materials with regard to search results. *See, e.g.*, Frank A. Pasquale, *Rankings, Reductionism, and Responsibility*, Seton Hall Public Law Research Paper No. 888327, Feb. 25, 2006, *available a*t http://ssrn.com/abstract=888327.

[124] *See, e.g.*, Andrew McLaughlin, *Congressional Human Rights Caucus Members' Briefing "Human Rights and the Internet – The People's Republic of China*," Feb. 1, 2006, http://googleblog.blogspot.com/2006/02/human-rights-caucus-briefing.html.

[125] *See, e.g.*, John G. Palfrey, Jr., Testimony to the U.S. House of Representatives Committee on International Relations, Feb. 15, 2006, *available at* http://blogs.law.harvard.edu/palfrey/stories/storyReader$1063.

[126] Among the most comprehensive studies is that of Rolf H. Weber & Dirk Spacek, RECHTSFRAGEN RUND UM SUCHMASCHINEN (2003).

carefully. Taking the recent Internet governance debate as a background and looking ahead, the following section seeks to sketch some of the emerging cross-sectional challenges for future policy-making concerning search engines.

B. CHALLENGES AHEAD

Policy-makers face a series of challenges when crafting governance frameworks aimed at regulating search engines in particular and online search in general. Some of the challenges are problems generally associated with law and policy-making, both in offline environments and cyberspace, and others are more search engine-specific. With regard to search engine regulation, one might identify, inter alia, the following key challenges:

- *Justification*: At least in Western societies, the burden of proof regarding the need for regulation is on the regulator. In the case of search engines, especially the existence of information asymmetries—e.g. regarding search algorithms[127]—and market power[128] may be considered justifications for future regulation.[129] However, cyberspace creates a "quicksilver technological environment"[130] that might make yesterday's regulation superfluous tomorrow. In fact, the brief history of search engines sketched in Part I of this paper not only illustrated how fast-paced innovation in search technology has been, but also demonstrated the power of new technologies to reallocate the market power of search engine operators.[131]
- *Prioritization*: Legislation and regulation, respectively, are costly processes, requiring that the many items on the broad policy agenda[132] are prioritized. As discussed in Part II, IPR issues have traditionally gained a significant amount of attention both by courts and legislators, while debates about content regulation, consumer protection, and privacy have intensified more recently. Vis-à-vis the complex interactions among powerful interest groups involved in

[127] *See, e.g.*, NIVA ELKIN-KOREN & ELI M. SALZBERGER, LAW, ECONOMICS AND CYBERSPACE: THE EFFECTS OF CYBERSPACE ON THE ECONOMIC ANALYSIS OF LAW 73 (2004).
[128] *Id.* at 77.
[129] For a general discussion, see, for example, STEPHEN BREYER: REGULATION AND ITS REFORM 15-35 (1982), and ROBERT BALDWIN & MARTIN CAVE, UNDERSTANDING REGULATION: THEORY, STRATEGY, AND PRACTICE 9-17 (1999).
[130] *See* MGM Studios, Inc. v. Grokster Ltd., 380 F.3d 1154, 1167 (9th Cir. 2004) (citing AT&T Corp. v. City of Portland, 216 F.3d 871, 876 (9th Cir.1999)).
[131] *Supra* Part I. *See also* Neil Gandal, *The Dynamics of Competition in the Internet Search Engine Market*, Univ. of Cal., Berkeley Competition Policy Ctr., Working Paper No. CPC01-17 (Jan. 2001), *available at* http://repositories.cdlib.org/iber/cpc/CPC01-017/.
[132] *See supra* Part III.A.

legislative processes, however, it remains an open question which policy area will be in the focus of a next wave of regulation.

- *Reconciliation*: Arguably, proposals of legal and/or regulatory interventions aimed at governing search engines in the policy areas outlined above pursue a wide range of policy goals, some of which will not be perfectly aligned. Such regulatory trade-offs—or at least tensions—may exist, for instance, between open access to search infrastructure and infrastructure security, or between privacy and content control.[133] The challenge to reconcile different policy objectives might thereby increase in the case of staggered legislation and regulation due to effects such as path-dependency or the like.
- *Timing and Change*: The history of technology-regulation is rich with examples of outdated laws.[134] As noted above, search technology has been evolving rapidly, too. Thus, policy-makers face the challenge of synchronizing technological innovation with legal evolution if they choose to regulate search engines.[135] Techniques such as "sunset-clauses" and fixed periods of evaluation will become particularly important in the search governance context.
- *Design*: In the case of search engine regulation, as in others, policy-makers have to make a series of design choices,[136] including decisions about the appropriate regulatory strategy (e.g., command and control regulation, incentive-based regimes, liability laws), and choices about institutions and structures. Most recently, the promises and limits of self-regulation of search engine operators

[133] The latter tension is illustrated by the law enforcement agencies' interest in search data. *See, e.g.*, Fred von Lohmann, *Could Future Subpoenas Tie You to 'Britney Spears Nude'?*, Special to Law.com, Feb. 6, 2006, http://www.eff.org/deeplinks/archives/004385.php.

[134] See, for example, the Audio Home Recording Act of 1992, 17 U.S.C. § 10 (2000). The Act was primarily aimed at DAT technology and sought to establish a system of royalty levies. But DATs were quickly supplanted by compact discs before DAT technology had a chance to take hold in the U.S. market, due probably in large part to threatened legal action. By the time recordable CD media became available which may have fallen within the Act's provisions, other digital recording technology—the MP3—had emerged and was held by the Ninth Circuit Court to escape the purview of the Act. *See* WILLIAM W. FISHER III, PROMISES TO KEEP: TECHNOLOGY, LAW, AND THE FUTURE OF ENTERTAINMENT 83-87 (2004).

[135] On the myth of technological neutrality in information regulation, see Herbert Burkert, *Four Myths About Regulating the Information Society – A Comment*, *in* STARTING POINTS FOR ICT REGULATION. DECONSTRUCTING PREVALENT POLICY ONE-LINERS 240-42 (Bert-Jaap Koops, Miriam Lips, Corien Prins et al. eds., 2006).

[136] *See, e.g.*, BALDWIN & CAVE, *supra* note 129, at 34-75.

have come up for discussion, especially in the context of speech regulation.[137]

– *Internationalization and transcultural issues*: Online search engines are operating in a globalized and networked environment. It is well established that this particular environment is characterized by a tension between the global scope of business activities and local laws that seek to regulate such activities,[138] a situation that poses manifold challenges for policy-making, both at the legislative[139] and judicial[140] level. Search engine operators themselves, in turn, are currently particularly concerned about the significant differences among national laws, regulations, and ethics that govern content and informational privacy, as a recent congressional hearing illustrated.[141]

---

[137] The German example of the Subcode of Conduct for Search Engine Providers of the Association of Voluntary Self-Regulating Multimedia Service Providers, *available at* http://www.fsm.de/en/SubCoC_Search_Engines (last visited April 10, 2006), aimed at improving consumer protection as well as protection of children and young persons with their use of search engines in Germany, illustrates in this context how blended governance models of state-based regulation and self-regulation can emerge.

[138] For a general overview of the cyber-internationalist discourse, see Viktor Mayer-Schönberger, *The Shape of Governance: Analyzing the World of Internet Regulation*, 43 VA. J. INT'L L. 605, 626-30 (2003).

[139] *See, e.g.*, Global Online Freedom Act of 2006, H. R. RES. 4780, 109th Cong. (2006). For another interesting example, see also H.R. RES. 12, 107th Cong., at 3 (2001), opposing the imposition of criminal liability on Internet service providers based on the actions of their users ("Whereas a number of European and Asian countries have held Internet service providers in the United States liable for content that is illegal under the laws of those countries, but protected by the first amendment to our Constitution . . . . ").

[140] Consider, for example, the long-running dispute between Yahoo!, U.S. courts, and French courts. See Ordonnance de référé rendue le 20 novembre 2000, *available at* http://www.juriscom.net/txt/jurisfr/cti/tgiparis20001120.pdf; Yahoo!, Inc. v. La Ligue Contre Le Racisme et L'Antisémitisme, 169 F. Supp. 2d 1181 (N.D.Cal. 2001); Yahoo!, Inc. v. La Ligue Contre Le Racisme et L'Antisémitisme, 379 F.3d 1120 (9th Cir. 2004) (holding that the French associations were not subject to personal jurisdiction in ISPs action.); Yahoo! Inc. v. La Ligue Contre Le Racisme**,** 433 F.3d 1199 (9th Cir. 2006) (where, before a panel of 11 judges, a majority of the bench concluded that the suit should be dismissed, but no majority agreed on the grounds for dismissal). For a legal analysis, see, for example, Joel R.Reidenberg, *The Yahoo Case and the International Democratization of the Internet*, Fordham Law & Economics Research Paper No. 11 (Apr. 2001), *available at* http://ssrn.com/abstract=267148. From a business ethics perspective, see Mark Hunter, Marc Le Menestrel, & Henri-Claude de Bettignies, *Ethical Crisis on the Internet*: *The Case of* Licra vs. Yahoo!, *in* BUSINESS ETHICS AND THE ELECTRONIC ECONOMY 177-208 (Peter Koslowski, Christoph Hubig & Peter Fischer eds., 2004).

[141] *The Internet in China: A Tool for Freedom or Suppression?: Joint Hearing of the House Subcommittee on Africa, Global Human Rights and International Operations and the Subcommittee on Asia and the Pacific*, 109th Cong. (2006), http://wwwc.house.gov/international_relations/109/af021506.htm; witness testimony available at http://wwwc.house.gov/international_relations/afhear.htm (last visited Apr. 24, 2006).

In sum, policy-makers—both at the national and international level—have to make a complex set of choices about sometimes complementary, sometimes competing policy goals, regulatory strategies and techniques, institutional designs, and timing, to name just a few, if they seek to establish a governance framework for search engines. In the discursive processes of policy-making, these choices—as the history of cyberlaw teaches us[142]–require an open discussion and shared understanding of what fundamental values should underlie today's information society in the first place. The next section seeks to contribute to this discourse.

C. NORMATIVE FOUNDATIONS

1. Democratic values

The heated global Internet Governance debate over the past few years has illustrated the extent to which information-related values, like others, are mostly culture-specific. However, despite all differences, overlapping consensus exists with regard to certain ethical convictions on the one hand and certain universal values—i.e., human rights—on the other hand.[143] It remains the challenge of future discourses in various fora to identify such clusters of basic norms, values, and rules. In the context of this paper, I would like to suggest three core values of a democratic ecosystem that are hopefully widely acceptable at least in the Western part of the world.[144] These core values are: (a) informational autonomy; (b) diversity; and (c) information quality.

The first value suggested here is *informational autonomy*. Viewed from an information law perspective,[145] autonomy in this sense includes at least three elements. First, an individual must have the freedom to make choices among alternative sets of information, ideas, and opinions. This

[142] See Burkert, *supra* note 109, at 171.

[143] *See*, e.g., Thomas Hausmanninger, *Controlling the Net: Pragmatic Actions or Ethics Needed?* IJIE Vo. 1 (June, 2004), *available at* http://www.i-r-i-e.net/inhalt/001/ijie_001_04_hausmanninger.pdf.

[144] Note that some of the values mentioned below, in fact, are fundamental rights, including human rights. I use the term value in this context as a generic term for various categories of policy goals. The following sections are based upon Urs Gasser, *The Good, The Bad, and The Ugly: Information Quality on the Internet* (unpublished manuscript, on file with author.)

[145] The relation between autonomy and information has been analyzed in great detail by Yochai Benkler, *Siren Songs and Amish Children: Autonomy, Information, and Law*, 76 N.Y.U. L. REV. 23 (2001) (discussing the potential effects of law on autonomy by structuring the information environment), and most recently in YOCHAI BENKLER, THE WEALTH OF NETWORKS: HOW SOCIAL PRODUCTION TRANSFORMS MARKETS AND FREEDOM, ch. 5 (133 et seq.) (2006).

includes the freedom to decide what information someone wants to receive and process.[146] Second, informational autonomy as an aspect of individual liberty necessitates that everyone has the right to express her own beliefs and opinions.[147] Third, informational autonomy in the digitally networked environment arguably requires that every user can participate in the creation of information, knowledge, and entertainment.[148] It is the shift from passive receivers of information to active users[149] that fosters individual participation and enables new forms of creative expression, thereby expanding the possibilities for the realization of a semiotic democracy.[150] The development of an individual's own personality and self-fulfillment[151] intersects with a second core value of a democratic information society: its diversity.

Diversity in the sense of a wide distribution of information from a great variety of competing sources as a societal value has traditionally been emphasized in First Amendment jurisprudence and scholarship, where it has long been considered to be essential to public welfare.[152] Diversity, in

---

[146] In the U.S., this right is an inherent corollary of the rights of free speech and free press. Thomas v. Collins, 323 U.S. 516, 534 (1944). "The dissemination of ideas can accomplish nothing if otherwise willing addressees are not free to receive and consider them." Jamie Kennedy, Comment, *The Right to Receive Information: The Current State of the Doctrine and the Best Application for the Future*, 35 SETON HALL L. REV. 789, 792 (2005) (quoting Lamont v. Postmaster General, 381 U.S. 301, 308 (1965) (Brennan, J., concurring)).

[147] The freedom to speak has long been recognized as an aspect of individual liberty and, consequently, as an end in itself. *See* Bose Corp. v. Consumers Union of United States, Inc., 466 U.S. 485 (1984). *See, e.g.*, Edwin Baker, *First Amendment Limits on Copyright*, 55 VAND.L. REV. 891 (2002) (conceptualizing "expressive liberty" as part of a person's autonomy that must be respected by the state).

[148] *See* Jack Balkin, *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, 79 N.Y.U. L. REV. 1 (2004) (arguing that digital technologies have altered the social conditions of speech and, thus, that free speech theory should focus on protecting and promoting a democratic culture; Balkin frames democratic culture both in terms of individual liberty as well as collective self-governance).

[149] *See, e.g.*, Yochai Benkler, *Viacom-CBS Merger: From Consumers to Users: Shifting the Deeper Structures of Regulation Towards Sustainable Commons and User Access*, 52 FED. COMM. L.J. 561, 562 (2000).

[150] *See, e.g.*, William W. Fisher III, PROMISES TO KEEP: TECHNOLOGY, LAW, AND THE FUTURE OF ENTERTAINMENT, 28-31 (Stanford University Press , 2004). *See also* Rosemary J. Coombe, *Author/izing Celebrity: Publicity Rights, Postmodern Politics, and Unauthorized Genders*, 10 CARDOZO ARTS & ENTERTAINMENT L.J. 365 (1992); Michael Madow, *Private Ownership of Public Image: Popular Culture and Publicity Rights*, 81 CAL. L. REV. 125 (1993); Sudakshina Sen, *Fluency of the Flesh: Perils of an Expanding Right of Publicity*, 59 ALBANY L. REV 739, 752-3 (1995). The phrase "semiotic democracy" goes back to cultural theorist John Fiske. JOHN FISKE, TELEVISION CULTURE 236-39 (1987).

[151] *See, e.g.*, Melville Nimmer, *The Right to Speak from Times to Time: First Amendment Theory Applied to Libel and Misapplied to Privacy*, 56 CAL. L. REV. 935 (1968).

[152] *See* Associated Press v. U.S., 326 U.S. 1, 20 (1945) ("[The First] Amendment rests on the assumption that the widest possible dissemination of information from diverse and antagonistic sources is essential to the welfare of the public…").

essence, can either be seen as a valuable mechanism to attain truth,[153] or as a crucial instrument for protecting democratic process and democratic deliberation.[154] However, a diverse information environment in its current incarnation not only improves deliberation and decision-making processes. Rather, the diversity of information, knowledge, and entertainment is an important aspect of the broader concept of cultural diversity which has been recognized as a fundamental value of our societies.[155] A diverse informational and cultural environment, in turn, has important feedback effects on individuals. The greater the variety in information, knowledge, and entertainment opportunities available to the members of a society, the more they are asked to decide for themselves what to think and how to act. In this process, users further develop their own informational skills and routines and, in turn, contribute to a richer and more diverse information environment.[156]

As individuals, groups, and societies, we heavily depend in our decision-making processes on information, which is increasingly acquired over the Internet. According to an April 2006 survey by the Pew Research Center, for instance, 45% of Internet users indicated that the Internet helped them make big decisions or negotiate their way through major episodes in

[153] The theory that free speech is an instrument of the search for truth on a "marketplace of ideas" underlies Holmes' famous dissent in *Abrams v. United States*: "the best test for truth is the power of the thought to get itself accepted in the competition of the market." 250 U.S. 616, 630 (1919). The truth and social utility approach to the legitimation of free speech has been contested. *See, e.g*., Derek Bambauer, *Shopping Badly: Cognitive Biases, Communications, and the Fallacy of the Marketplace of Ideas*, 77 UNIV. OF COLO. L. REV. (forthcoming, spring 2006).

[154] One school of thought sees freedom of speech as a mean to assure the effectiveness of democratic processes. *See, e.g.*, ALEXANDER MEIKLEJOHN, FREE SPEECH AND ITS RELATION TO SELF-GOVERNMENT (1948), *reprinted in* POLITICAL FREEDOM: THE CONSTITUTIONAL POWERS OF THE PEOPLE (1979). The consideration of as many facts and arguments as possible which can be put forth in support of or against a proposition, so the argument goes, is the best way to make sound and rational judgments. *See, e.g.*, Thomas I. Emerson, *Toward a General Theory of the First Amendment*, 72 YALE L.J. 877 (1963). Another approach focuses on democratic participation in the sense of collective self-determination. *See, e.g.*, OWEN M. FISS, THE IRONY OF FREE SPEECH (1996).

[155] Cultural diversity has been recognized in the international arena. *See* UNESCO, *Universal Declaration on Cultural Diversity* (November 2, 2001), *available at* http://unesdoc.unesco.org/images/0012/001271/127160m.pdf; *Convention on the Protection and Promotion of the Diversity of Cultural Expressions* (October 20, 2005), *available at* http://unesdoc.unesco.org/images/0014/001429/142919e.pdf; *see, e.g.*, Ivan Bernier, *A UNESCO International Convention on Cultural Diversity*, *in* FREE TRADE VERSUS CULTURAL DIVERSITY: WTO NEGOTIATIONS IN THE FIELD OF AUDIOVISUAL SERVICES 65-76 (Christoph Beat Graber, Michael Girsberger, Mira Nenova eds., 2004).

[156] *See also* FISHER, *supra* note 134, at 26-28 (discussing the social benefits of cultural diversity).

their lives in the previous two years.[157] Another earlier Pew study suggests that 67% of Americans expect that they can find reliable information about health or medical conditions online,[158] while 63% expect that businesses have a web site that provides information about a product they are considering to buy, and 65% of all Americans expect the Web to have information from a government agency.[159] A recent Pew Report suggests that online news takes center stage as a news source for 40% of broadband users,[160] while an earlier study indicates that 85% of American Internet users expect to be able to find reliable, up-to-date news online.[161] In order to make sound decisions in the above-mentioned and other areas of life, we depend on high-*quality information*. However, functional and cognitive aspects are only two dimensions of the information quality concept.[162] It also includes aesthetic and ethical requirements of different stakeholders such as users, creators, experts, and administrators. In order to increase an individual's opportunity to live her life according to her own informational preferences, legal and regulatory regimes should contribute to the creation and further development of a high-quality information ecosystem.

It is important to note that these core values are not necessarily always aligned. Unleashed diversity in the digitally networked environment, for instance, might have negative feedback effects on user autonomy because it increases an individual's risk to be exposed to undesired information. A regulatory approach aimed at ensuring high-quality information, by contrast, might be in tension with informational autonomy, because it may impose a quality requirement leading to a level of quality

---

[157] John Horrigan & Lee Rainie, *The Internet's Growing Role in Life's Major Moments*, Pew Internet & American Life Project (April 19, 2006), http://www.pewinternet.org/pdfs/PIP_Major%20Moments_2006.pdf.

[158] A recent study suggests that 79%of American Internet users have searched for health information online. *See* Susanna Fox, *Reports: Health Information Online*, Pew Internet & American Life Project (May, 2005), http://www.pewinternet.org/pdfs/PIP_Healthtopics_May05.pdf.

[159] John Horrigan & Lee Rainie, *Counting on the Internet*, Pew Internet & American Life Project (December 29, 2002), http://www.pewinternet.org/pdfs/PIP_Expectations.pdf. Compare with more recent studies conducted by UCLA and the USC Annenberg School, Center for the Digital Future, which find that user perception of the reliability and accuracy of information on the internet has been falling; 48.8% of users in 2005 indicated that they believed most or all information on the internet was reliable and accurate, whereas 81.3% of users indicated that they believed most or all information on sites they visit regularly was reliable and accurate. Center for the Digital Future, USC Annenberg School, *Fifth Study of the Internet by the Digital Future Project Finds Major New Trends in Online Use for Political Campaigns* (Dec. 7, 2005), at 4-5, http://www.digitalcenter.org/pdf/Center-for-the-Digital-Future-2005-Highlights.pdf.

[160] John Horrigan, *Online News: For many home broadband users, the internet is a primary news source*, Pew Internet & American Life Project (March 22, 2006), http://www.pewinternet.org/pdfs/PIP_News.and.Broadband.pdf.

[161] Horrigan, *Counting on the Internet*, *supra* note 159.

[162] *See, e.g.*, MARTIN EPPLER, MANAGING INFORMATION QUALITY (2003), 58 *et seq*.

that does not meet an individual's informational needs.[163] Thus, policy-makers seeking to regulate the digitally networked environment face the challenge of dynamically balancing among autonomy, diversity, and quality.

2. Quest for policy principles

However, the three fundamental information-related values outlined in the preceding paragraph set the stage for crafting *guiding principles* for policy-making. With regard to search engine regulation, specifically, one might derive, *inter alia*, the following policy principles, which may provide guidance for policy-makers in the public and private sector, respectively:

1. *Access*: Search engine governance frameworks should aim to maximize access to search engines both for users and content providers on non-discriminatory terms. The role of search engines as the new gatekeepers has been discussed elsewhere and does not have to be repeated here.[164] In any event, "access" has at least two important meanings from a normative perspective. Access in the sense of access to search infrastructure is crucial for users,[165] because it is the prerequisite for the

[163] In the case of search engine regulation, this problem is accentuated by the fact that search engines simultaneously affect all three aspects. For example, since search engine users often do not know in advance what specific piece of information they are looking for, the quality of the information that users get depends to a great extent on search engines. Consequently, the quality of information is intertwined with the quality of the search engine that defines which information becomes available based on any given query. Similarly, search engines have effects on autonomy and diversity in the digitally networked environment.

[164] This role has been particularly emphasized by German scholars. *E.g.* Marcel Machill, *Wegweiser im Netz: Qualität und Nutzung von Suchmaschinen*, *in* WEGWEISER IM NETZ (Marcel Machill and Welp Carsten. eds, 2003); WOLFGANG SCHULZ, THORSTEN HELD, AND ARNE LAUDIEN, SUCHMASCHINEN ALS GATEKEEPER IN DER ÖFFENTLICHEN KOMMUNIKATION (2005). *See generally* Introna & Nissenbaum, *supra* note 111; Nico van Eijk, *Search Engines: Seek and Ye Shall Find? The Position of Search Engines in Law*, IRIS PLUS 2006-02 (Jan. 2006), *available at* http://www.obs.coe.int/oea_publ/iris/iris_plus/iplus2_2006.pdf.en; Eszter Hargittai, *Online Gatekeepers: Myth or Reality*, http://tprc.org/papers/2002/82/hargittai-tprc2002paper.pdf (last visited Apr. 24, 2006); Niva Elkin-Koren, *Let the Crawlers Crawl: On Virtual Gatekeepers and the Right To Exclude Indexing*, 26 DAYTON L. REV. 179 (2001); Karine Barzilai-Nahon & Seev Neumann, *Gatekeeping in Networks: A Metatheoretical Framework for Exploring Information Control* (Nov. 2005), http://www.ischool.washington.edu/karineb/html/pub/GatekeepingMetatheory.pdf (providing a more theoretical discussion of gatekeepers in networked environments).

[165] Competing search engines, too, can have an interest in accessing the search infrastructure—or parts of it such as the index—of their competitors. For a German view on the competition law issues involved, see Wolfgang Schulz, Thorsten Held and Arne Laudien, *Search Engines as Gatekeepers of Public Communication: Analysis of the*

above-mentioned freedom to efficiently and effectively make choices among alternative sets of ideas, information, and opinions in the digital age. Consequently, policies that pursue the goal of fostering informational autonomy in the digitally networked environment would aim to create an ecosystem that tends to increase access to search infrastructures.[166] However, in an environment where consumers are no longer passive receivers of information, but increasingly active contributors to the information ecosystem, access also concerns the (controversial) debate about the entitlement of users (as creators) to be integrated into search indexes and ranking lists, or at least the possible remedies against discrimination in the indexing or ranking processes. Viewed from the autonomy and diversity perspective and as a matter of policy,[167] technologies and politics that are aimed at inclusion are therefore *prima vista* favorable over alternative approaches that would result in significant decrease in content inclusion.

2. *Informational self-determination*: A second principle that derives from the values outlined above and is closely related to informational autonomy is the users' right to make choices about the collection and use of personal search data collected by search engine operators. Thus, the respective policy principle asks for the creation of governance regimes where the collection and storage of personal search data—taking the different interests into account—is optimized or, preferably, minimized.[168] The problems associated with information collection practices by search engines have been illustrated both in the domestic and international contexts.[169]
3. *Transparency*: Another policy principle that might be derived from the values discussed above is transparency of search engines. Transparency requirements in the context of search engines are often considered as the

---

German framework applicable to internet search engines including media law and antitrust law, 5 GERMAN L.J. No. 10 – 1, 1424-27 (October 2005).

[166] The means to achieve this goal, of course, do not need to follow a command-and-control approach. Rather, the regulatory strategy might be a completely incentive-based, market-driven approach. However, interventionist proposals such as the above-mentioned idea of the creation of a *service public* search engine might be evaluated in the light of their impact on equal and universal access to search.

[167] For the current state of and developments in U.S. case law, see Part II. Access rights of this sort, in contrast, are considered in some European jurisdictions. *See, e.g.*, SCHULZ ET AL., *supra* note 164, at 1424 (differentiating between "normal" inclusion and "paid inclusion", id. 1425).

[168] A potential "right to search anonymously" was also on the agenda at the Regulating Search? Conference at Yale Law School in December 2005.

[169] *See, e.g.*, A Code of Conduct for Internet Companies in Authoritarian Regimes (Feb. 15, 2006), http://www.eff.org/deeplinks/archives/004410.php ("With the stakes so high in countries like China, no Internet company should gather more information than they absolutely need about their costumers …"); von Lohmann, *supra* note 133.

potential response to a problem of asymmetric information,[170] i.e., the fact that the algorithms of search engines are generally trade secrets[171] and might therefore result in undetected, inherent biases[172] that ultimately shape the construction of meaning in cyberspace.[173] A policy principle—applicable at the corporate level—might suggest that operators inform the users about the way in which the search engine works and explain the basic criteria of ranking.[174] Additionally, transparency as a policy principle can also relate to yet another controversial subject: the separation of advertisement from the list of unpaid results and the question of appropriate labeling of commercial communications. As a model for a policy principle one might consider § 2 of the German Subcode of Conduct for Search Engine Providers.[175] In a third interpretation, transparency as a mechanism can be applied to alleviate the impact of content filtering requirements imposed on search engines by legislation or regulations. Google, for instance, uses this mechanism in several jurisdictions if search results are removed for legal reasons. In response to a search on Google.de for the keyword "stormfront," for example, Google informs at the bottom of the result page how many results had to be removed due to legal requirements.[176] This notice links to the ChillingEffects.org project, where the user can

[170] *See, e.g.*, Introna & Nissenbaum, *supra* note 111, at 32; SCHULZ ET AL., *supra* note 164, at 1431.

[171] *See, e.g.*, SearchKing, Inc. v. Google Technology, Inc., No. CIV-02-1457-M, 2003 WL 21464568, at *3 n.2 (W.D. Okla. May 27, 2003).

[172] For a detailed discussion, see Eric Goldman, *Search Engine Bias and the Demise of Search Engine Utopianism* in this volume.

[173] On search engines' role in construction meaning, *see, e.g.*, ELKIN-KOREN & SALZBERGER, *supra* note 127.

[174] *See* Subcode of Conduct for Search Engine Providers of the Association of Voluntary Self-Regulating Multimedia Service Provider, *supra* note 141, at § 2 Rules of Conduct, clause 1 ("The Code signatories agree to clarify to the user the functioning method of the search engine. In the same way, the signatories shall describe the circumstances that will cause an exclusion from the search results. This information should be easily accessible to the user."). *See also* Carsten Welp and Marcel Machill, *Code of Conduct. Transparency in the Net: Search Engines*, 3 IRIE (June 2005), *available at* http://www.i-r-i-e.net/inhalt/003/003_code.pdf. For a critical view on regulatory interventions, see Goldman, *supra* note 179.

[175] Subcode of Conduct for Search Engine Providers of the Association of Voluntary Self-Regulating Multimedia Service Provider, *supra* note 137, at § 2 Rules of Conduct, clause 2 ("Within the framework of its possibilities, the Code signatories agree to transparently structure its search results pages. Search engine results which owe their position on the search results page to a commercial agreement with the respective search engine provider shall be reasonably designated. This can occur, in particular, by use of the terms 'Advertisement', 'Sponsor Link', 'Sponsored Link' or 'Sponsored Web Site'.").

[176] "Aus Rechtsgründen hat Google 3 Ergebnis(se) von dieser Seite entfernt. Weitere Informationen über diese Rechtsgründe finden Sie unter ChillingEffects.org." Stormfront – Google-Suche, http://www.google.de/search?hl=de&q=stormfront&btnG=Google-Suche (last visited Apr. 24, 2006).

> learn more about the reasons that led to the filtering of the results, and can compare search results across national domains.[177] This practice is well suited to contribute to all of the three values outlined above and should be considered as a minimum transparency principle for search engines in particular and Internet intermediaries more generally.[178]

The rough sketch of three basic principles might illustrate how concrete guidance for policy-makers both in the public and private sector can be derived from core values that underlie today's information society. The proposed policy principles may also serve as an initial basis for a systematic comparison and thorough normative evaluation of future governance regimes aimed at regulating search engines in particular and searches in general.

IV. CONCLUSION

Building upon a brief history of the technological innovations that underlie web search and corresponding business models, this paper has traced the emerging law of search engines in broad strokes. This analysis illustrates how and in what respect the legal system has responded to search engine-related legal issues. Past and present issues considered by courts, regulators, and legislators reveal seven core themes of future policy debates: infrastructure, content, ownership, security, identity and privacy, participation, and the ethics debate. For these policy areas, policy-makers have to deal with the manifold challenges touched upon in this paper, including the task of prioritizing items on the regulatory agenda, reconciling competing policy goals, ensuring the legal system's ability to learn in response to technological change, and managing transcultural issues, among others. Three basic values—informational autonomy, diversity, and, information quality—intersect the policy debates surrounding the role and function of search engines within the digital environment. Taken together, these considerations may chart out a more comprehensive governance framework which effectively addresses total policy concerns, yet retains the flexibility to respond to technological change and innovation.

---

[177] Chilling Effects Google Search Comparator, http://www.chillingeffects.org/images/search-comparator/ (last visited Apr. 24, 2006).

[178] A best practice-oriented approach could go further by obliging search engine operators, if not prohibited by law, to report data on search terms and web sites that are considered to be sensitive under the applicable law and by the respective authorities, respectively.