

INTERNET INFOGLUT AND INVISIBLE INK:  
SPAMDEXING SEARCH ENGINES WITH META TAGS

Ira S. Nathenson\*

*Unseen treasure and hidden wisdom, of what use is either?*  
— 14th century librarian<sup>1</sup>

*The more we know, the less we know.*  
— 20th century author<sup>2</sup>

TABLE OF CONTENTS

I. INTRODUCTION .....	44
II. WHERE DO YOU WANT TO GO <i>TODAY?</i> : INTERNET INFOGLUT AND SEARCH ENGINE ABUSE .....	51
III. DEFINING THE INTERESTS .....	72
A. <i>Relevance and Noise</i> .....	72
B. <i>Reasonable Expectations</i> .....	77
1. The Public .....	78
2. Webmasters .....	82
3. Proprietary-Rights Holders .....	84
C. <i>Domain Names, Linking, and Framing Compared</i> .....	86
D. <i>Externalization Costs</i> .....	89
IV. STRIKING A BALANCE .....	92
A. <i>Deceptiveness: "Protecting" the Public</i> .....	93
1. Section 5 of the Federal Trade Commission Act .....	93

---

\* J.D. 1998, University of Pittsburgh School of Law; B.A. 1985, University of Pittsburgh College of Arts and Sciences. I would like to thank Professor Julie E. Cohen for her encouragement and suggestions, and Jeffrey R. Waxman for other helpful comments. None of this would have been possible without the love, support, and patience of my wife, Helen, and my children, Katie and Jenny. I may be reached at <ira.nathenson@worldnet.att.net>.

1. Roger K. Summit, *Knowledge Online: Current Implications and Future Trends*, in INFORMATION SCIENCE: STILL AN EMERGING DISCIPLINE 101, 102 (James G. Williams & Toni Carbo eds., 1997) [hereinafter INFORMATION SCIENCE].

2. David Shenk, *Data Smog: Surviving the Info Glut*, TECH. REV., May 15, 1997, at 18.

2.	How Much Deception Is Sufficient? . . . . .	98
3.	FTC Action Is Needed . . . . .	101
B.	<i>Unfair Competition: Balancing Competitive Interests</i> . .	106
1.	Rationale of Unfair Competition Law . . . . .	106
2.	Trademark Infringement: Conceptually Inapposite .	111
3.	Trademark Dilution: Tailor-Made . . . . .	118
4.	Protection for Less-Than-Famous Marks? . . . . .	122
5.	Defenses to Infringement and Dilution . . . . .	124
V.	WHERE DO YOU WANT TO GO <i>TOMORROW?</i> : TAMING THE INTERNET INFOGLUT . . . . .	131
VI.	CONCLUSION . . . . .	146

## I. INTRODUCTION

Imagine a never-ending traffic jam on a ten-lane highway. Road signs can't be trusted: the sign for Exit 7 leads to Exit 12, the sign for Cleveland leads to Erie. If you ask the guy at the Kwik-E-Mart how to get to I-79, he gives you directions to Route 30. To top it off, when you ask for a Coke, he gives you a Pepsi. Enough already. You stop at a pay phone to call directory assistance for the number to the local auto club, and instead get connected to "Dial-a-porn."

Road trip from hell? Not exactly. Welcome to the Internet, the "information super-highway" that lacks reliable road signs. Suppose, for example, you use an Internet search engine to look for Web pages on the late Princess Diana. Instead, you may find get-rich schemes and pornography.<sup>3</sup> Or you search for an attorney — by name — and instead get Internet service providers ("ISPs") and software companies.<sup>4</sup> If you search for pages on "Monica Lewinsky," you might be shocked to find that the top listing from one search engine is "CityAuction," an Internet classified advertising site.<sup>5</sup>

3. See Elizabeth Weise, *Some Web Pages Take Search Engines for a Ride*, USA TODAY, Sept. 29, 1997, at 4D.

4. See James Evans, *Fraud: Hidden Tags Leading to Court Battles*, INTERNET WORLD, Dec. 1997, at 17.

5. A search for "Monica Lewinsky" using the HotBot search engine on Feb. 27, 1998 resulted in 1,126 matches (on file with *Harvard Journal of Law & Technology*). Of these, CityAuction was the very first site listed. See CityAuction, *Monica Lewinsky, Please Use CityAuction Classifieds* (visited Feb. 27, 1998) <<http://www.cityauction.com>>.



Microsoft's advertising, waxing poetic, asks, "Where do you want to go today?"<sup>6</sup> Instead, maybe it should ask, "How in the hell do you get there?"<sup>7</sup> After all, even those who know where they want to go may need help finding their way. Such help exists in the form of search engines, but some Internet content providers abuse these tools to the frustration of many. This Article is about those who search, those who want to be found, and those who would stand in between.

The primary means of finding information on the Internet is via search engines.<sup>8</sup> In compiling their databases, search engines rely on those who publish Web pages — webmasters — to supply indexing information.<sup>9</sup> Because an increase in visits, or "hits," means more exposure or revenue, webmasters are strongly motivated to do anything to increase their chances of getting hits.<sup>10</sup> Many therefore "spam" search

---

com/lewinsky.asp>. This page is not the main CityAuction page, but rather a "bridge" page: one designed to draw in the public and then direct it to another place. *See infra* note 111 and accompanying text. In light of the discussion of relevance in Part III, it is ironic to note the attempt made by this page to appear relevant to Ms. Lewinsky by stating an "offer": "Monica Lewinsky, we hereby invite you to use CityAuction." CityAuction, *supra*. A look at the code for this page indicates that "Monica Lewinsky" is also used as a meta tag and title tag. *See id.*; *see also infra* notes 79–83 and accompanying text (discussing various ways Web pages are indexed by search engines).

6. Michael Newman, *Sticky Wickets: Yahoo! and Its Rivals Now Want You to Stay Around and Not Go Searching*, PITT. POST-GAZETTE, Apr. 19, 1998, at E1.

7. *See Connected: Click Here to Go to Hell*, LONDON DAILY TELEGRAPH, June 24, 1997, at 2.

The background music for Microsoft's current television campaign is the Confutatis from Mozart's Requiem. The words of the final blast of music that accompanies the slogan "Where do you want to go today?" are "confutatis maledictis, flammis acribus addictis" which means "the damned and accused are convicted to flames of hell." All of which is grist to the mill of those who believe that Microsoft boss Bill Gates is the devil in disguise.

*Id.*

8. *See infra* notes 68–83 and accompanying text.

9. Strictly speaking, a webmaster is one who designs or maintains a Web page. In the spamdexing context, *see infra* note 16 and accompanying text, the implementation of irrelevant or deceptive indexing terms may be done by an actual webmaster, or by a webmaster acting as an agent on behalf of his principal. For simplicity, "webmaster" will refer to any party who provides content on the World Wide Web, either personally or via a webmaster-agent.

10. "[A] high ranking can make all the difference. Consider that a search for the words 'tennis racket' on AltaVista yields more than 850 pages. The user is much more likely to glance through the first 20 sites listed than the last 20." *Net Interest: Web-Search-2: It's Up to You to Stand Out*, DOW JONES NEWS SERV., Oct. 9, 1997.

engines in hopes of appearing in as many searches, and as highly ranked in those searches, as possible.<sup>11</sup>

Some webmasters use indexing terms that are completely irrelevant to actual Web page content in order to gain hits from searches on those topics.<sup>12</sup> Others “stuff” multiple instances of a term, relevant or not, in hopes of appearing more relevant to the search engine, and therefore higher in its output.<sup>13</sup> Yet others use the trademarks and trade names of third parties, hoping to take advantage of the goodwill of those asserting proprietary rights<sup>14</sup> in an indexing term.<sup>15</sup> *USA Today* calls these practices “spamdexing.”<sup>16</sup>

---

11. See DON SELLERS, GETTING HITS: THE DEFINITIVE GUIDE TO PROMOTING YOUR WEB SITE 21–22 (1997); Weise, *supra* note 3; see also *Internet and Unix Dictionary* (last modified Nov. 17, 1998) <<http://www.msg.net/kadow/answers/s.html#spam>> (defining “spam” as “bulk, mass, or repeated posting or mailing of substantially identical messages. The emphasis is on the multiple sending, either many copies to one destination, or one copy to many destinations. This is a reference to the famous Monty Python Spam sketch.”).

12. See *infra* note 97 (noting site that used buried text referring to “Seinfeld”).

13. See SELLERS, *supra* note 11, at 32–33; Weise, *supra* note 3 (“A scuba-diving site might hide the word ‘scuba’ a hundred times, so search engines would rank it above others.”). A view of the source code for an Internet site dedicated to the World Wrestling Federation indicates that the term “WWF” is repeated five times as a meta tag. See *WWF No Holds Barred* (visited Jan. 31, 1998) <<http://www.wwf2000.com>> (source code on file with *Harvard Journal of Law & Technology*).

14. The term “property” is somewhat loaded, so I will avoid its use in this Article. After all, merely calling something property ascribes to it the conclusion to the question presented: should the right of exclusion apply? See Felix S. Cohen, *Transcendental Nonsense and the Functional Approach*, 35 COLUM. L. REV. 809, 815 (1935) [hereinafter Cohen, *Transcendental Nonsense*]; see also Zechariah Chaffee, Jr., *The Disorderly Conduct of Words*, 41 COLUM. L. REV. 381, 390–91 (1941). Therefore, references to “proprietary” or “trademark” rights provide a somewhat less volatile label. The proprietary rights to which I refer are more in the nature of quasi-property rights between competitors than rights between the holder and the public. See *International News Serv. v. Associated Press*, 248 U.S. 215, 236 (1918).

15. As one commentator wryly notes, “[this] practice has been compared to calling [directory assistance] to get the listing of Apple Computer, only to receive the number for Microsoft.” Evans, *supra* note 4; see also *Cyberscam*, DATA COMM., Mar. 1, 1998, at 15 (quoting one attorney who calls this practice “[g]arden-variety commercial fraud”); Martin J. Elgison & James M. Jordan III, *Trademark Cases Arise from Meta-Tags, Frames; Disputes Involve Search-Engine Indexes, Web Sites Within Web Sites, as Well as Hyperlinking*, NAT’L L.J., Oct. 20, 1997, at C6 (noting that it has become popular for competitors to “bury the trademarks of a better known competitor”). Ironically, several Internet directories (which are similar to search engines) became embroiled in a meta tag dispute. See Elizabeth Gardner, *Trademark Battles Simmer Behind Sites*, WEB WEEK, Aug. 25, 1997, available at <<http://www.webweek.com/1997/08/25/news/19970825-battles.html>>.

16. See Weise, *supra* note 3 (“Spamdexing, named for its attempt to spam, or flood,



Spamdexing techniques are invisible to the typical person browsing the World Wide Web. These techniques include the use of stuffing, buried text, meta tags, and other items that search engines use to index Web pages.<sup>17</sup> “Stuffing” is the practice of repeating an indexing term multiple times in an effort to boost the Web page’s relevance to that term. This is often accomplished through “buried text,” which is the use of characters that are invisibly hidden in a background color.<sup>18</sup> More commonly abused toward similar purposes are “meta tags,” which are codes that are invisibly embedded into a Web page to help search engines index the page.<sup>19</sup> By themselves, meta tags represent a modern variant on long-utilized systems of keyword indexing, such as those used by libraries in their card catalogs.<sup>20</sup> Today, however, meta tags are commonly used to spam rather than to inform:

- Adult-oriented sites have used the registered trademarks “Playboy” and “Playmate;”<sup>21</sup>
- Businesses have used the names of direct competitors;<sup>22</sup> and

---

search engine indexes, is an ongoing and escalating fight.”); *see also* Elgison & Jordan, *supra* note 15; Brian Livingston, *More on Finding, or Not Finding, Your Special Web Site*, INFOWORLD, Nov. 10, 1997, at 48 (discussing what he calls “spam-dexing” or “search-engine spamming”); Christopher Stamper, *The War on Spamdexing*, ABCNEWS.COM (Oct. 31, 1997) <<http://archive.abcnews.com/sections/scitech/spamdex1030/index.html>>.

17. *See infra* notes 84–111 and accompanying text (discussing types of deceptive behavior).

18. *See* Rebecca Quick, *Web-Search Services Can Set Your Site’s Visibility Higher — But It’s Up to You to Make the Most of Page’s Listing*, WALL ST. J., Oct. 9, 1997, at B8; *infra* note 86 and accompanying text; *see also* SELLERS, *supra* note 11, at 21–22.

19. *See* SELLERS, *supra* note 11, at 22; *infra* notes 91–107 and accompanying text.

20. *See infra* notes 433–44 and accompanying text.

21. *See* Playboy Enters., Inc. v. Welles, 7 F. Supp. 2d 1098 (S.D. Cal.), *aff’d*, No. 98-55911, 1998 U.S. App. LEXIS 27739 (9th Cir. Oct. 27, 1998) (finding no trademark infringement); Playboy Enters., Inc. v. AsiaFocus Int’l, Inc., No. 97-734-A, 1998 U.S. Dist. LEXIS 10359 (E.D. Va. Feb. 2, 1998) (Mag. J.), *adopted by* 1998 U.S. Dist. LEXIS 10459 (E.D. Va. Apr. 10, 1998); Playboy Enters., Inc. v. Calvin Designer Label, 985 F. Supp. 1220 (N.D. Cal. 1997); *see also* Jeffrey R. Kuester & Peter A. Nieves, *Hyperlinks, Frames and Meta-Tags: An Intellectual Property Analysis*, 38 IDEA 243, 274–75 (1998); Maureen A. O’Rourke, *Defining the Limits of Free-Riding in Cyberspace: Trademark Liability for Metatagging*, 33 GONZ. L. REV. 277, 289–91 (1998); Steven M. Weinberg, *Cyberjinks: Trademark Hijinks in Cyberspace Through Hyperlinking and Metatags*, 87 TRADEMARK REP. 576, 588–89 (1997); Ann Davis, *Web Weaves a Tangled Trademark Issue*, WALL ST. J., Sept. 15, 1997, at B10; Nikki Tait, *Playboy Wins Internet Case*, FIN. TIMES, Apr. 23, 1998, at 7; *infra* notes 113–18, 132–45, 383–403 and accompanying text.

22. *See* Insituform Technologies, Inc. v. National Envirotech Group, L.L.C., Civ. Act No. 97-2064 (E.D. La. Aug. 26, 1997), *available at* <<http://www.cll.com/>>

- ISPs have used the names of attorneys Oppedahl & Larson, presumably hoping to take advantage of their reputation as Internet experts.<sup>23</sup>

Some spamdexers use their competitors' trademarks and trade names as meta tags, oftentimes with an arguably relevant connection to actual content at the spamdexer's site.<sup>24</sup> In contrast, other spamdexers incorporate terms that are completely irrelevant to actual content.<sup>25</sup> Both forms of spamdexing frustrate the public's ability to use search engines without the undue introduction of "noise" — i.e., information pollution.<sup>26</sup>

A fundamental tension exists among the public, webmasters, and those with proprietary rights, because each seeks to externalize its costs of information dissemination and retrieval.<sup>27</sup> Overly broad liability for spamdexing could make search queries underinclusive because Web page publishers would hesitate to use questionable but relevant indexing terms, such as the names of competitors and related products.<sup>28</sup> On the

case1.htm>; see also *Niton Corp. v. Radiation Monitoring Devices, Inc.*, No. Civ. A. 98-11629-REK, 1998 WL 812685 (D. Mass. Nov. 18, 1998); Kuester & Nieves, *supra* note 21, at 275; Weinberg, *supra* note 21, at 589; Davis, *supra* note 21.

23. See Davis, *supra* note 21; see also Kuester & Nieves, *supra* note 21, at 276; Weinberg, *supra* note 21, at 589; *infra* notes 122–30 and accompanying text.

24. See *infra* Part IV.B. See generally Marcelo Halpern, *Meta-Tags: Effective Marketing or Unfair Competition?*, CYBERSPACE LAWYER, Oct. 1997, at 2.

25. See *infra* notes 96–100 and accompanying text.

26. See *infra* Part III.A.

27. Other commentators, particularly M. Ethan Katsh and James Boyle, have noted the tensions that arise from new media technologies. See JAMES BOYLE, *SHAMANS, SOFTWARE, AND SPLEENS: LAW AND THE CONSTRUCTION OF THE INFORMATION SOCIETY* (1996); M. ETHAN KATSH, *THE ELECTRONIC MEDIA AND THE TRANSFORMATION OF LAW* (1989) [hereinafter KATSH, *TRANSFORMATION OF LAW*]; M. ETHAN KATSH, *LAW IN A DIGITAL WORLD* (1995) [hereinafter KATSH, *DIGITAL WORLD*]; James Boyle, *A Politics of Intellectual Property: Environmentalism for the Net?*, 47 DUKE L.J. 87 (1997) [hereinafter Boyle, *Politics*]; James Boyle, *A Theory of Law and Information: Copyright, Spleens, Blackmail, and Insider Trading*, 80 CAL. L. REV. 1415 (1992); M. Ethan Katsh, *The First Amendment and Technological Change: The New Media Have a Message*, 57 GEO. WASH. L. REV. 1459 (1989) [hereinafter Katsh, *New Media*]; M. Ethan Katsh, *Rights, Camera, Action: Cyberspatial Settings and the First Amendment*, 104 YALE L.J. 1681, 1689 (1995) ("[I]nteracting with machines will require reading and writing in a new context.") [hereinafter Katsh, *Cyberspatial Settings*]; see also Mark A. Lemley, *Romantic Authorship and the Rhetoric of Property*, 75 TEX. L. REV. 873 (1997) (reviewing Boyle's book); Pamela Samuelson, *The Quest for Enabling Metaphors for Law and Lawyering in the Information Age*, 94 MICH. L. REV. 2029 (1996) (comparing and critiquing Boyle and Katsh).

28. See KATSH, *TRANSFORMATION OF LAW*, *supra* note 27, at 170–71 ("If the new technologies are significantly expanding the production and movement of information, then the ultimate effect of restrictive laws will be, at best, to slow down the increase in



other hand, overly narrow liability could have an equally undesirable effect: webmasters would have little incentive to exercise self-restraint and would spamdex freely, making searches overinclusive.<sup>29</sup> Both scenarios render the Internet a less effective medium of information dissemination and retrieval.<sup>30</sup> Therefore, finding a balance is essential.

These issues must be addressed now because spamdexing lawsuits are starting to proliferate.<sup>31</sup> Part II describes how the Internet infoglut has created a need for search engines, which in turn have made spamdexing possible. Although spamdexing is a form of abuse brought on by new technology, placing the infoglut in its historical context shows that the problem runs far deeper than any technical solution. Even if a solution to meta tag abuse could be found, the broader problem of indexing abuse will not disappear.

Part III shows that the real harm of spamdexing is the introduction of “noise” into search engine databases that render searches “imprecise” — i.e., overinclusive. In addition, Part III provides an analytic framework that looks to relevance and reasonable expectations. Specifically, in spamdexing disputes, it is necessary to consider the relevance between the search term and the search goal, and between the index term and actual content, as informed by both information science

---

the flow of information rather than to reduce it.”).

29. See Terry Brock, *Internet Searches Can Yield Dividends*, JACKSONVILLE BUS. J., Dec. 26, 1997, at 18 (“[Search engines] can often give you thousands of ‘hits’ on general searches.”).

30. Such issues may be addressed in terms of the advantages and problems posed by the growth of networks:

The death of Usenet as an effective means of communication resulted from an overabundance of participation coupled with the lack of limits on the relevance of that participation. . . . [P]eople may well want to access only a subset of information available on the Web, if that subset is filtered or tailored in such a way as to make it more useful to them.

Mark A. Lemley & David McGowan, *Legal Implications of Network Economic Effects*, 86 CAL. L. REV. 479, 560 (1998) (footnote omitted); see also Kevin Kelly, *New Rules for the New Economy*, WIRED, Sept. 1997, at 140, 192 (suggesting that the primary focus of businesses in a networked world should be to maximize the value of the infrastructure as a whole); Peter Schwartz & Peter Leyden, *The Long Boom: A History of the Future: 1980–2020*, WIRED, July 1997, at 115, 121 (suggesting that open, rather than closed, models will help ensure an economic boom that will carry over into the next century). Some commentators, however, are more hesitant to adopt rose-colored glasses: “Dazzled by computers and communications theory, we’ve been misled into thinking that experience can be broken down into bits and bytes. . . . The Internet, that great digital dumpster, confers not power, not prosperity, not perspicacity.” CLIFFORD STOLL, *SILICON SNAKE OIL: SECOND THOUGHTS ON THE INFORMATION HIGHWAY* 194 (1995).

31. See *infra* notes 112–45 and accompanying text.

and the Federal Rules of Evidence. It is also necessary to consider the reasonableness of each party's desire to use, or exclude the use of, such a term for indexing and searching purposes. There is a tension between these interests, as each party will try to externalize its costs of information dissemination and retrieval on the others. Part III further fleshes out the nature of the various interests via a brief comparison of spamdexing to the analogous areas of domain name, hyperlinking, and framing liability.

Part IV applies the framework of relevance and reasonable expectations to the laws of consumer protection and unfair competition. Although consumer protection laws are applicable to spamdexing, the diffuse nature of the consumer spamdexing injury makes it unlikely that the Internet-surfing public will bring consumer protection actions. Regulators, better equipped to track and to publicize the dangers of spamdexing, should therefore enforce these laws on the public's behalf. On the unfair competition front, trademark infringement is not a proper conceptual basis for spamdexing actions; instead, trademark dilution provides a more sound conceptual basis for liability. Although courts may be tempted to expand unfair competition liability beyond the scope of current standards, they should be hesitant to do so, because overly broad liability could be anti-competitive.

Due to the specter of chilled competition, Part V provides a broader view of the concerns implicated by new forms of information dissemination and retrieval. Harkening back to Part II, it examines the roles that libraries have traditionally played, and those that the Internet-as-library may yet play. Not all dissemination and retrieval needs are best served by search engines. The existence of a number of reasonable alternatives serves to undercut the desire of webmasters, the public, and proprietary-rights holders to assert their interests at the expense of the others. Each must adapt its expectations in an effort to make efficient the dissemination and retrieval of information. The best way to maximize all parties' interests is to treat the public interest as paramount.



## II. WHERE DO YOU WANT TO GO *TODAY*?: INTERNET INFOGLUT AND SEARCH ENGINE ABUSE

*There is no traffic control system for today's information technology.*

— Paul Virilio<sup>32</sup>

*What's important — increasingly important [in the Information Economy] — is the process by which you figure out what to look at. This is the beginning of the real and true economics of information — not who owns the books, who prints the books, who has the holdings. The crux today is access, not holdings. And not even access itself but the signposts that tell you what to access — what to pay attention to. In the Information Economy everything is plentiful — except attention.*

— Bruce Sterling<sup>33</sup>

It has become a truism of the Information Age that the Internet is growing quickly.<sup>34</sup> Phenomenal growth in the amount of information available online has resulted in a corresponding difficulty in sorting

32. Paul Virilio, *Red Alert in Cyberspace!*, RADICAL PHIL., Nov.–Dec. 1995, at 2–3 (Malcolm Imrie trans.), available at <<http://www.nytimes.com/library/cyber/surf/020597surf-redalert.html>>; see also Ashley Dunn, *As Bookends to the 20th Century, 2 Views of Technology's Promise*, CYBERTIMES (N.Y. Times) (Feb. 5, 1997) <<http://www.nytimes.com/library/cyber/surf/020597surf.html>> (discussing Virilio); Bruno Giussani, *For a Philosopher, the Net Is a Whole New Perspective*, CYBERTIMES (N.Y. Times) (Dec. 9, 1997) <<http://www.nytimes.com/library/cyber/euro/120997euro.html>> (same).

33. Bruce Sterling, *Free as Air, Free as Water, Free as Knowledge*, in THINKING ROBOTS, AN AWARE INTERNET, AND CYBERPUNK LIBRARIANS: THE 1992 LITA PRESIDENT'S PROGRAM 25, 28 (R. Bruce Miller & Milton T. Wolf eds., 1992) (emphasis removed) [hereinafter THINKING ROBOTS]. As Sterling puts it:

What's information *really* about? It seems to me there's something direly wrong with the *Information Economy*. It's not about data, it's about attention. In a few years you may be able to carry the Library of Congress in your hip pocket. So? You're never gonna read the Library of Congress. You'll die long before you access one tenth of one percent of it.

*Id.*

34. See, e.g., *Reno v. ACLU*, 117 S. Ct. 2329, 2334 (1997) (“About 40 million people used the Internet at the time of [the trial in the *Reno* case], a number that is expected to mushroom to 200 million by 1999.”); cf. Samuelson, *supra* note 27, at 2029 (stating the analogous sentiment that “[i]t has become a truism, if not a cliché, that developments in information technologies are causing a fundamental transformation in society, taking us out of the industrial era and into an information age”). The amount of traffic on the Internet is currently doubling every 100 days. See U.S. DEP'T OF COMMERCE, THE EMERGING DIGITAL ECONOMY 8 (1998), available at <<http://www.doc.gov/ecommerce/emerging.htm>> [hereinafter THE EMERGING DIGITAL ECONOMY].

through it — infoglut.<sup>35</sup> The amount of information has expanded far beyond our ability to process or comprehend it.<sup>36</sup> It also frustrates those

35. See, e.g., Shenk, *supra* note 2.

If the concept of too much information seems odd and vaguely inhuman, that's because, in evolutionary-historical terms, it is. For 100,000 years people have been able to examine and consider information about as quickly as they have been able to create and circulate it. A range of communication technologies from the drum and smoke signal to the telegraph and telephone enabled us to develop and sustain culture and overcome our fear of others, diminishing the likelihood of conflict. But in the middle of this century the introduction of computers, microwave transmissions, television, and satellites abruptly knocked this graceful synchrony off track. These hyper-production and hyper-distribution mechanisms have surged ahead and left us with a permanent processing deficit — what Finnish sociologist Jaako Lehtonen calls an “information discrepancy.”

*Id.*; see also DOUGLAS S. ROBERTSON, *THE NEW RENAISSANCE: COMPUTERS AND THE NEXT LEVEL OF CIVILIZATION* 8 (1998); THEODORE ROSZAK, *THE CULT OF INFORMATION* 163 (1986) (quoting John Naisbitt as stating that “[r]unning out of [data] is not a problem, but drowning in it is. Data is now doubling every twenty months.”); DAVID SHENK, *DATA SMOG: SURVIVING THE INFORMATION GLUT* 28 (rev. ed. 1998) [hereinafter SHENK, *DATA SMOG*]; Kathy Fladland, *Earth and the Internet*, in *THINKING ROBOTS*, *supra* note 33, at 113, 118 (“[T]he problem isn’t gaining access to information. The problem is screening out the irrelevant bits of information.”); Reid Goldsborough, *Staying Afloat in a Rising Tide of Data*, *OFFICE SYSTEMS*, July 1, 1998, at 44; Matt Lake, *Desperately Seeking Something: Using Search Engines Can Be a Frustrating Experience, But You Can Learn to Make Them Find What You Seek*, *HOUS. CHRON.*, Sept. 18, 1998, at 1.

36. See Vannevar Bush, *As We May Think*, *ATLANTIC MONTHLY*, July 1945, at 101, 105 (“[I]f the scholar can get at only one [piece of knowledge] a week by diligent search, his syntheses are not likely to keep up with the current scene.”).

The problem of the infoglut is more subtle than one of quantity. While much of the problem is due to the sheer *amount* of information, it is also due in large part to a lack of effective navigation tools. To the extent that these tools exist, they are *part* of the information they help to circumnavigate. In other words, information often *includes* the tools used to interact with it. This combination of information-as-technology and technology-as-information is not historically unique. For instance, directory assistance, as a form of indexing and searching, presupposes the infrastructure of the entire telephone system, as do phone books and classified advertising. See *infra* notes 449–63 and accompanying text. It is therefore somewhat disingenuous to assert that only in the modern era does the amount of information increase faster than our ability to keep up with it, and complain that previously, “over our long history, people have been able to examine and consider information about as quickly as it could be created and circulated.” SHENK, *DATA SMOG*, *supra* note 35, at 27–28. In reality, as technology advances to increase the amount of information, it also acts to develop organizing and filtering tools to allow people to manage information efficiently. See ROBERTSON, *supra* note 35, at 18.



who hope to profit from the Internet's growth.<sup>37</sup> James Boyle muses that this scenario may someday be viewed as "information pollution."<sup>38</sup> French thinker Paul Virilio believes that the "suffocation of meaning" arising from the infoglut is itself a kind of "disinformation" that represents "another major danger for humanity."<sup>39</sup> Rather than being a mere inconvenience, the confusion and speed of modern communications may even cause a "disorientation of the individual that can be manipulated by those who control the flow of information — typically governments and businesses."<sup>40</sup> The infoglut poses challenges

37. One source estimates online sales over a five-year period as follows:

1994	\$ 8 million
1995	\$ 436 million
1996	\$ 2,908 million
1997	\$13,086 million
1998	\$45,808 million

ActivMedia, Inc., *Trends in the World Wide Web Marketplace, Estimated On-Line Sales, 1993–1998*, reprinted in IntellectualCapital.com, *Ideas: Charts & Graphs* (Nov. 21, 1996) <<http://www.intellectualcapital.com/issues/96/1121/icchart.html>>. The CEO of Cisco Systems stated that annual Internet commerce could grow to \$1 trillion "shortly after the turn of the century." Mitchell Martin, *Big Recomputation of Sales on Net*, INT'L HERALD TRIB., Nov. 19, 1997, at 16, available at <<http://www.iht.com/IHT/MM/97/mm111997.html>>. Some estimate that just the online classified advertising market will reach \$1.9 billion in domestic sales in 2002. See *Nearly \$2 Billion in Online Classified Ad Sales Likely by 2001, Say Analysts at Ground-Breaking Symposium*, PR NEWSWIRE, Feb. 2, 1998.

From 1993 to 1998, the IT [Information Technology] share of the economy will have risen from 6.4 percent to an estimated 8.2 percent . . . . With such rapid expansion, IT's share of total nominal GDP growth has been running almost double its share of the economy, at close to 15 percent. . . .

. . . .

In recent years, IT industries have been responsible for more than one-quarter of real economic growth . . . .

THE EMERGING DIGITAL ECONOMY, *supra* note 34, at 4–6.

38. BOYLE, *supra* note 27, at 180.

39. Virilio, *supra* note 32, at 3.

40. Dunn, *supra* note 32; see also ROSZAK, *supra* note 35, at 161 ("[I]t is not facts that determine policy, but more often policy that determines the facts — by selection, adjustment, distortion."). This is hardly a novel sentiment. In 1838, James Fenimore Cooper wrote:

If newspapers are useful in overthrowing tyrants, it is only to establish a tyranny of their own. . . . Under the pretence [sic] of protecting publick [sic] morals, it is corrupting them to the core, and under the semblance of maintaining liberty, it is gradually establishing a despotism as ruthless, as grasping, and one that is quite as vulgar as that of any christian [sic] state known.

JAMES FENIMORE COOPER, THE AMERICAN DEMOCRAT 123 (Alfred A. Knopf 1931)

to even the Internet-literate (“neterate”) public, as well as the potential for abuse.<sup>41</sup> Thus, it becomes imperative to develop effective ways to filter information pollution.<sup>42</sup> Indeed, when properly handled, infoglut can be transformed into opportunity.<sup>43</sup>

The problem of information pollution is neither recent nor historically unique. The explosion of printing that followed the invention of the printing press led to the development of searching and indexing tools, such as page numbering, indices, and bibliographies.<sup>44</sup> James Russell Lowell said over one hundred years ago: “[W]e are getting buried alive under this avalanche of earthly impertinences . . . . [W]e . . . are willing to become mere sponges saturated from the stagnant goosepond of village gossip.”<sup>45</sup> Fifty years ago, at the dawn of the digital age, Vannevar Bush recognized the existence of a post-war infoglut and the lack of useful tools with which to navigate and create.<sup>46</sup> Therefore, it should not surprise us that new media will cause us to reexamine paradigmatic ways of disseminating and retrieving information.<sup>47</sup>

---

(1838), *quoted in* STEVEN LUBAR, INFOCULTURE: THE SMITHSONIAN BOOK OF INFORMATION AGE INVENTIONS 23–24 (1993); *see also* ROSZAK, *supra* note 35, at 164 (stating that “data glut” is “a strategy of social control, deliberately and often expertly wielded” by “modern governments and interest groups [to] obfuscate issues to their own advantage”).

41. *See* ROSZAK, *supra* note 35, at 163 (noting that with the infoglut, “the strategy of government is not to censor but to counter fact with fact, number with number, research with research. It even becomes advantageous to have lots of contention about facts and figures, a statistical blizzard that numbs the attention.”).

42. *See id.*

43. *See* Gary McWilliams & Marcia Stepanek, *Knowledge Management: Taming the Info Monster*, BUS. WK., June 22, 1998, at 170.

44. *See* GREGORY J.E. RAWLINS, MOTHS TO THE FLAME: THE SEDUCTIONS OF COMPUTER TECHNOLOGY 47–48 (1996). Prior to the invention of the printing press, only a few thousand hand-scribed manuscripts existed. *See id.* at 47. “But by the early sixteenth century, twenty million books had been printed — almost as many books as there are in the U.S. Library of Congress today.” *Id.* Another commentator ascribes the development of the scientific method to sixteenth-century infoglut. *See* ROBERTSON, *supra* note 35, at 18.

45. Sterling, *supra* note 33, at 30 (some alterations in original).

46. *See* Bush, *supra* note 36, at 101–02. “The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of square-rigged ships.” *Id.* at 102.

47. *See* ROBERTSON, *supra* note 35, at 18 (“[F]ollowing the invention of printing we find, first, a massive increase in the production of information, and then an equally massive development and application of old techniques for refining that information.”); *see also* Katsh, *Cyberspatial Settings*, *supra* note 27, at 1689 (“[I]nteracting with



Many see the Internet as a huge library that will someday contain the sum of human knowledge.<sup>48</sup> Library- and Internet-navigation tools share much in common,<sup>49</sup> including the use of filtering and indexing.<sup>50</sup> Although today's Internet contains a treasure trove of information, like a super-library, its lack of efficient ways to navigate, organize, and search render it very different from traditional conceptions of "library."<sup>51</sup> In fact, the open-ended nature of the Internet makes the comparatively rigid indexing systems used by libraries a dubious metaphor.<sup>52</sup>

Indeed, the Internet represents a major shift in what we traditionally think of as an information repository.<sup>53</sup> It provides an opportunity for

machines will require reading and writing in a new context."); Katsh, *New Media*, *supra* note 27, at 1464 ("The new qualities of printed works focused more attention on basic organizing principles and concepts . . .").

48. See PAUL GILSTER, *DIGITAL LITERACY* 164 (1997) ("The ultimate goal of network visionaries is the construction of an on-line reference work that contains the sum of all human knowledge from the days of the first cave paintings to the latest scientific breakthroughs . . .").

49. See Fladland, *supra* note 35, at 118 ("Information overload has been the monster under the bed of librarianship since at least the fifties.").

50. See Sterling, *supra* note 33, at 28–29 (stating that librarians and others, such as stores and distributors, act as information and attention filters).

51. See Brian Trench, *Assessing the Internet: Is All Information Equal?*, *IRISH TIMES*, Dec. 3, 1997, at 20.

The Internet is often compared with a huge library, holding more information than could ever be held in a library full of paper. But no known library would want to hold the Internet's stores — a library's information is organised and classified; the lack of organisation and classification define the Internet.

*Id.*; see also GILSTER, *supra* note 48, at 164 ("If the Internet is to achieve [the goal of containing the sum of all human knowledge], it will take all the skills of present-day library science and all the power of tomorrow's technology . . . to make it happen.").

52. Vannevar Bush recognized this fifty years ago:

The real heart of the matter of selection [i.e., searching], however, goes deeper than a lag in the adoption of mechanisms by libraries . . . . *Our ineptitude in getting at the record is largely caused by the artificiality of systems of indexing.* When data of any sort are placed in storage, they are filed alphabetically or numerically, and information is found (when it is) by tracing it down from subclass to subclass.

Bush, *supra* note 36, at 106 (emphasis added). Bush foresaw a new form of information creation and retrieval that would use associative links: by the old system, information "can be in only one place, unless duplicates are used . . . . [But, t]he human mind does not work that way. It operates by *association*." *Id.* (emphasis added). His description of the "new" way of creating information is eerily evocative of the system we now use: the World Wide Web. See *id.* at 106–08.

53. See KATSH, *DIGITAL WORLD*, *supra* note 27, at 56 ("The new media represent the equivalent of an earthquake hitting the library, not because the electronic library may

those eager to devise new ways to navigate the infoglut.<sup>54</sup> In particular, it presents the possibility of finding, or even creating, useful information that was not available before.<sup>55</sup> As Bush suggested in his description of his hypothetical “memex”<sup>56</sup> — an eerie foreshadowing of hyperlinks and the World Wide Web — “[t]he inheritance from the master becomes, not only his additions to the world’s record, but for his disciples the entire scaffolding by which they were erected.”<sup>57</sup>

Whether the Internet can act as a democratizing and value-adding institution depends in large part on the extent to which society develops tools that allow effective dissemination and filtering of relevant information.<sup>58</sup> The advent of the printing press took the control of information out of the hands of yesterday’s information elite — priests,

replace the physical library but because the role of the library, whether it is electronic or print, is shifting as information becomes accessible from new sources.”).

54. See Bush, *supra* note 36, at 108 (“Wholly new forms of encyclopedias will appear, ready-made with a mesh of associative trails running through them . . . . There [will be] a new profession of trail blazers, those who find delight in the task of establishing useful trails through the enormous mass of the common record.”).

55. The “reworking of information is a key ingredient of the information society or information economy in which new information, or old information processed into a new form, acquires substantial economic value.” Katsh, *New Media*, *supra* note 27, at 1478; see also STEVEN JOHNSON, *INTERFACE CULTURE: HOW NEW TECHNOLOGY TRANSFORMS THE WAY WE CREATE AND COMMUNICATE* 155–59 (1997) (describing how high-speed computer textual analysis showed which roles Shakespeare likely performed as an actor in his own plays); ROSZAK, *supra* note 35, at 163 (quoting John Naisbitt as stating that “[i]nformation technology brings order to the chaos of information pollution and therefore gives value to data that would otherwise be useless”).

56. See Bush, *supra* note 36, at 108; see also LUBAR, *supra* note 40, at 46 (discussing Bush as well as Douglas Englebart, who, inspired by Bush, created the first hypertext system at Stanford Research Institute in the mid-1960s). Appropriately, Bush’s article is available online at <<http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm>>.

57. Bush, *supra* note 36, at 108. Katsh would appear to agree:

Are these new technologies merely more efficient versions of the old? . . . Do they simply move information faster? Or does the use of information in a new form, particularly by an institution for whom information is a highly valued commodity, change the institution, the user, and those who come in contact with the user? Does it create a new type of institution where it is possible to do new things with information and relate to and interact with information differently?

KATSH, *DIGITAL WORLD*, *supra* note 27, at 8.

58. See ROSZAK, *supra* note 35, at 165 (noting that “in a vital democracy, it is not the quantity but the *quality* of information that matters,” and that “[r]elevance, coherence, and insight” are important criteria).



aristocrats, and guildsmen — and made it available to the masses.<sup>59</sup> This devalued the old forms of information dissemination and created possibilities for new kinds of value: engineering, science, technology transfer, and the Renaissance.<sup>60</sup> Thus, the advent of the Internet is little more than the most recent stepping stone in six thousand years of advances that arguably “drop the price of thinking.”<sup>61</sup> Yet as the *cost* of thinking has reduced, the *amount* of thinking seems to have expanded.

Indeed, the cause of the infoglut may be, in part, the infoglut itself. David Shenk suggests that the “widening pool” of information on “every side of every question . . . paradoxically [leads] to less clarity.”<sup>62</sup> As a result, people focus on the process of argumentation rather than reaching conclusions.<sup>63</sup> In turn, “communicators of all types resort to barrier-piercing countermeasures, feeding a vicious spiral in which the data smog gets thicker and thicker and the efforts to cut through the smog ever more desperate.”<sup>64</sup> This leads to our present difficulty of finding relevant, reliable information.

The interfaces that we use to navigate information are therefore necessary and corrective prerequisites for making sense of information overload.<sup>65</sup> A wide variety of these “interfaces” — tools, programs, and techniques<sup>66</sup> — are used to make sense by filtering out that which we

---

59. See RAWLINS, *supra* note 44, at 48.

60. See *id.*

61. *Id.* at 49 (noting the advent of clay tablets, papyrus, parchment, and movable type, and stating that “[w]henver we improve the production, handling, and distribution of information we drop the price of thinking”).

62. Shenk, *supra* note 2.

63. See *id.* (“Because there is always an opportunity to crunch some more numbers, spin them a bit, and prove the opposite, the winner has become argumentation itself.”).

64. *Id.*

65. See JOHNSON, *supra* note 55, at 236–37.

66. This includes hyperlinks, framing, domain names, “hotlist” pages, “cool” sites, “Top Ten” award certification sites, WebRings, “push” content, e-zines, advertising, magazines, and books with lists of websites. See *infra* note 445 and accompanying text. These other methods of organizing and retrieving information are relevant to this Article in two respects. First, abuse of these other forms is a potential source of liability that may be fruitfully compared to spamdexing. Cases arising under these forms of information organization may be helpful in determining how to balance rights of the public and webmasters against those who seek a legal remedy for deception or free riding. See *infra* Part III.C. Second, these other methods provide alternative ways to navigate the infoglut. The existence of alternatives also helps to determine a proper balance of interests by identifying alternatives for those information-retrieval scenarios to which search engines are not well-suited. See *infra* Part V.

neither want nor need.<sup>67</sup> Perhaps the most important tool, and the focus of this Article, is the search engine.<sup>68</sup> Five of the ten most-visited websites in August 1997 were search engines.<sup>69</sup> Because of search engines' importance, popularity, and potential for profit, those who operate them seek any competitive advantage.<sup>70</sup>

---

67. See JOHNSON, *supra* note 55, at 238–39.

Cathedrals, remember, were “infinity imagined,” the heavens brought down to earthly scale. The medieval mind couldn’t take in the full infinity of godliness, but it could subjugate itself before the majestic spires of Chartres or Saint-Sulpice. The interface offers a comparable sidelong view onto the infosphere, half unveiling and half disappearing act. It makes information sensible to you by keeping most of it from view — for the simple reason that “most of it” is far too multitudinous to imagine in a single thought.

*Id.* at 239.

68. See Shenk, *supra* note 2 (“Many journalists haven’t yet come to terms with the implications of our society’s fundamental shift from scarcity to glut, which is why Ya[h]oo, Alta Vista, and other World Wide Web search engines are on their way to becoming our primary information sources.”). Although search engines are the most thorough tools currently available to index the Internet’s content, they are by no means even close to complete. One study suggests that even the best search engines index no more than 40% of the approximately 320 million Web pages in existence. See Steve Lawrence & C. Lee Giles, *Searching the World Wide Web*, SCIENCE, Apr. 3, 1998, at 98. Although indexing even 100 million pages is impressive, the incompleteness of indexing is itself a problem independent of spamdexing, especially in light of the fact that the number of Web pages should increase 1,000 percent over the next several years. See *id.*

69. See *Yahoo! Wins on New Ratings*, INTERNET WORLD, Dec. 1997, at 24, available at <<http://www.internetworld.com/print/monthly/1997/12/news.html>>. These search engines were Yahoo!, Infoseek, Excite, Lycos, and AltaVista. See *id.*; see also *Lycos: It’s Official! The Only Profitable Search Engine on the Net*, M2 PRESSWIRE, Nov. 19, 1997, available in 1997 WL 15144355. Although this Article does not consider Yahoo! — a human-created directory — to be a search engine, it provides an excellent example of the potential profit for those who offer a better way to get around on the Internet.

70. See Jean Rowan, *Infoseek Gets Patent for Novel Search Engine Technique*, SAN ANTONIO BUS. J., Oct. 31, 1997, at 16. The amount of money at stake is huge. As of May 1998, Wall Street valued Yahoo!, Excite, Infoseek, and Lycos at \$9 billion. See Saul Hansell, *Turning Search Engines into Money Machines*, CYBERTIMES (N.Y. Times) (May 11, 1998) <<http://www.nytimes.com/library/tech/98/05/biztech/articles/11web-eyes.html>>. However, these stocks may be overvalued. See Saul Hansell, *There Is a Why? in Yahoo Stocks: It’s the Hottest Investment in Cyberspace, But It Might Be Wise to Question the Cost*, ORANGE CO. REG. (Cal.), Feb. 15, 1998, at K1; *Lycos: It’s Official! The Only Profitable Search Engine on the Net*, *supra* note 69. The competition among the search engines to offer extra value, or to seek competitive alliances, has heightened. See Michael Newman, *A World of Competition: Lycos Buys WiseWire in Its Continuing Drive to Unseat Yahoo! as the Internet’s No. 1 Search Engine*, PITT. POST-GAZETTE, May 1, 1998, at E1; see also *infra* note 479 and accompanying text.



Search engines allow users to query for specific information. Novices, information providers, and information experts alike hope that these tools will make the Internet attractive and profitable.<sup>71</sup> A directory such as Yahoo! — not technically a search engine — is like an enhanced yellow pages: it contains searchable, hierarchical categories of interest.<sup>72</sup> Yahoo! accepts requests from webmasters to be added to its directory, but whether and in which category a Web listing appears is at the discretion of human Yahoo! staffers.<sup>73</sup> A search engine, in contrast, is a database that relies upon software robots, as described below.<sup>74</sup> Search

---

71. See *AltaVista for Intranets of All Shapes and Sizes*, IMAGING WORLD, Apr. 7, 1997, at 59 (quoting Pat Condo, CEO of Excalibur Technologies, as stating that “[b]usinesses are faced with limited resources at a time of maximum information input — the ‘info-glut’ — and there is a critical requirement to manage that situation. . . . [T]here is just too much information, and it is increasingly important to be able to extract what’s most useful from that flood.”); *The Epidemic of Infoglut*, PROVIDENCE J.-BULL., May 12, 1997, at B5 (quoting Robert L. Olson, research director of the Institute for Alternative Futures, as stating “[b]etter search engines” will help us filter out the excess electronic junk; educators “will focus more on skills of knowledge navigating and learning to learn” in navigating the ocean of facts and factoids).

72. See SELLERS, *supra* note 11, at 3–4.

73. See Dawn C. Chmielewski, *Defining Parameters Allows You to Limit Your Research Sites*, BUFFALO NEWS, July 7, 1998, at D7 (“Yahoo! uses staff searchers to evaluate, review and categorize a site. It’s a relatively small directory (about 500,000 sites), with each site hand-picked and assigned a subject classification.”).

74. See *infra* notes 79–80 and accompanying text.

It is unclear whether search engines themselves could be liable under trade regulation or unfair competition law for how they classify and present their listings. Cf. Daniel Ovanezian, Comment, *Internet Search Engine Copying: Fair Use Defense to Copyright Infringement*, 14 SANTA CLARA COMPUTER & HIGH TECH. L.J. 267, 301–02 (1998) (concluding that although search engines arguably engage in copyright infringement in constructing their databases, they are nevertheless eligible for fair use and estoppel defenses). Apparently nobody has ever brought suit against a search engine for trademark infringement and dilution, presumably because current technology makes it extremely difficult for search engines to choose how to arrange their query output. Such suits are possible, however, because trademark infringement and dilution do not require a showing of intent. See 3 J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 23:107, at 23-207 (4th ed. 1996) (“The courts have always held that proof of defendant’s intent to deceive or confuse is not required for infringement of a federally registered mark.”); *id.* § 24:89 (listing elements of prima facie case of dilution).

To the extent that search engines and other such tools rely upon human agents, it is more likely that traditional principles of vicarious liability will apply. To the extent that they are automated, it is correspondingly less likely. For example, with old versions of Netscape Navigator, when a user entered a single word into the browser’s Uniform Resource Locator (“URL”) field, the browser automatically defaulted to that term plus “.com” as the URL. For example, if the user entered “patents” into the URL field, the user would go to <http://www.patents.com>. The current version of Navigator/

engines and directories are suited to different types of searches: whereas search engines are good for finding specific information by using information-rich terms, directories are more suited to categorical searches.<sup>75</sup> All-at-once search engines, such as MetaCrawler, take a search request, submit it to multiple search engines and directories simultaneously, and then return collated output.<sup>76</sup>

Although many search engines include directories,<sup>77</sup> and most directories are searchable,<sup>78</sup> this Article focuses exclusively on search engines because human-created directories like Yahoo! are by definition immune to spamdexing. Nevertheless, the existence of alternative forms of information dissemination and retrieval is highly relevant to determining the reasonable expectations of webmasters, the public, and proprietary-rights holders.

Search engines such as Lycos, Excite, Infoseek, and AltaVista create their databases by using automated programs to index Web pages.<sup>79</sup> These programs have been aptly dubbed "spiders," "crawlers," "robots," and "bots":

---

Communicator, however, incorporates a "keyword" feature that diverts the user to the so-called "proper" URL. Thus, if the user types "patents" into the URL field, the user is transported to the United States Patent and Trademark Office, which is at <<http://www.uspto.gov>>. This feature appears to be intended to meet the user's "reasonable" expectations as to where she expects to go, and to side-step those who register an Internet domain name in hopes of benefitting from those expectations. That *human* agents presumably implement Netscape's keyword feature greatly increases the possibility of lawsuits from disgruntled webmasters and trademark owners.

75. See Lake, *supra* note 35 (stating that "Web directories like Yahoo and Web search engines may look the same on the surface, but each type of site is good for finding different types of information"); see also *infra* Part V. Note, however, that hybrid search engines, such as Lycos, also have directories. To the extent that portions of Lycos's content is human-created, spamdexing has no effect.

76. Because all-at-once search engines are derivative of "normal" search engines and directories, they need not be addressed separately. To the extent that spamdexing deceives a normal search engine, it will also deceive an all-at-once search engine.

77. For example, Lycos has its A-Z directory, which lists websites by subject, and its Top 5% list, which is a selection of the "Best of the Web." See Lycos <<http://www.lycos.com>>; see also Ted Witulski, *Using Search Engines Effectively*, in PC NOVICE GUIDE TO GOING ONLINE 89 (1997). Excite automatically compiles its directory from its search engine's database via artificial intelligence routines. See SELLERS, *supra* note 11, at 4.

78. Yahoo! can also give information supplied by Inktomi, a company that fuels search engines such as HotBot.

79. See Search Engine Watch, *How Search Engines Work* (visited Nov. 17, 1998) <<http://searchenginewatch.internet.com/webmasters/work.html>>.



A spider (also known as a robot, crawler or indexer) is a program that scans the Web, crawling from link to link, visiting Web pages, recording URLs (uniform resource locators) and building an index for the search engine. . . . [It] will request documents from web sites, access all links from each web page and deposit its findings in the [search engine's] index.<sup>80</sup>

Indexing information is garnered from actual content and other items supplied by webmasters, including a Web page's text, title, URL,<sup>81</sup> files, and even sound, video, graphics, Java programs,<sup>82</sup> and last — but by no means least — meta tags, which are invisible terms provided by webmasters for the sole purpose of indexing. With the exception of meta tags, most everything else that is indexed will reflect the site's content because these other items generally are the visible content.<sup>83</sup>

Reliance on webmasters to supply indexing information encourages schemes to spam the search engines with deceptive index terms — hence “spamdexing.” After all, webmasters want their pages to appear in as many query results as possible. They also want to appear high in these query results.<sup>84</sup> This reliance is tantamount to a bank counting on its depositors to know how much money remains in each account. Not surprisingly, an entire cottage industry has developed in the wake of the World Wide Web to give tips on how to “beat” or “spam” the search engines.<sup>85</sup>

---

80. John M. Mrsich & Meeka Jun, *Terms You Need to Know: Search Engines, MULTIMEDIA & WEB STRATEGIST*, May 1997, at 3; *see also* GILSTER, *supra* note 48, at 165 (stating that “the general principle is to send a so-called software robot out onto the Internet to follow links from one Web page to another, cataloging the content at each”); Search Engine Watch, *supra* note 79.

81. Uniform Resource Locator, the Internet “address” of the site.

82. Java, an enhancement and addition to “normal” Web pages, is a computer language that allows websites to give executable instructions to a computer displaying the page. It is essentially a way for websites to make displaying computers run programs. Ideally, Java runs the same way across all platforms — Unix, Windows, and Macintosh. As another form of information, it offers yet another way to index. For instance, HotBot allows searching via Java files. *See* HotBot <<http://www.hotbot.com>>.

83. One exception is irrelevant buried or stuffed text. *See infra* notes 86–90 and accompanying text.

84. For example, a search on InfoSeek conducted on January 10, 1998 for “Princess Diana” gave 16,961 results (on file with *Harvard Journal of Law & Technology*). *See* InfoSeek <<http://www.infoseek.com>>.

85. *See, e.g.*, V.A. SHIVA AYYADURAI, *THE INTERNET PUBLICITY GUIDE* 83–87 (1997); SELLERS, *supra* note 11, at 27–44; Bruce Clay, *Search Engine Ranking Tools* (visited Nov. 17, 1998) <[http://www.bruceclay.com/web\\_rank.htm](http://www.bruceclay.com/web_rank.htm)>; Bruce Grossan,

Spamdexing includes a variety of conduct. Early on, webmasters realized that they could use “buried text” — tiny words camouflaged in the background color — to incorporate extra text, relevant or not, that would be visible to the search engines but not the public.<sup>86</sup> Some webmasters engage in “stuffing” a term, which is the use of multiple repetitions in buried text or meta tags in an attempt to boost the apparent relevance of the page to the stuffed term.<sup>87</sup> Stuffing is analogous to pumping up the volume to cut through the noise; as David Shenk points out, such activities, while temporarily effective, ultimately serve to “feed[] a vicious spiral in which the data smog gets thicker.”<sup>88</sup> Search engines quickly caught on to buried text and stuffing; now many penalize sites engaging in these practices.<sup>89</sup> These forms of abuse have therefore become less commonplace and will not be addressed in detail.<sup>90</sup>

More commonplace today is the misuse of meta tags to accomplish the same abusive purposes. “Tags” are lines of code in HTML,<sup>91</sup> invisible to the viewer, that provide various information, including indexing information for search engines. For example, the “title” tag allows the title of the Web page to appear in the title bar of a browsing

---

*Search Engines: What They Are, How They Work, and Practical Suggestions for Getting the Most Out of Them* (visited Nov. 17, 1998) <<http://webreference.com/content/search/index.html>>; Jim Rhodes, *The Art of Business Web Site Promotion* (visited Nov. 17, 1998) <<http://deadlock.com/promote/>>; Search Engine Watch, *A Webmaster's Guide to Search Engines* (visited Nov. 17, 1998) <<http://searchenginewatch.internet.com/webmasters/>>.

86. A Web page with buried text usually is easy to spot: excessive blank space at the bottom of the document should raise a suspicion that can be confirmed by highlighting the “blank” space with a click-and-drag of the mouse, revealing any “invisible” text.

87. See Davis, *supra* note 21; Weise, *supra* note 3.

88. SHENK, DATA SMOG, *supra* note 35, at 102.

89. See Paul C. Judge, *Revenge of the Search Engine*, BUS. WK., Nov. 17, 1997, at 146B; Weise, *supra* note 3; Search Engine Watch, *Search Engine Features Comparison Chart* (visited Nov. 17, 1998) <<http://searchenginewatch.internet.com/webmasters/features.html>> [hereinafter *Comparison Chart*].

90. This gives yet another example of how the technologies may change (just as meta tags may eventually not be used), but human deception and unfairness always continue. This Article's analysis of meta tags applies equally well to buried text, which is essentially a more primitive version of meta tag abuse. This Article's analysis also generally applies to stuffing, which is essentially an attempt to artificially boost relevance. Where necessary, I will point out any differences that apply to stuffing.

91. HTML stands for “Hypertext Mark-Up Language,” the language in which Web pages are written. Most major browsing programs allow users to “look under the kitchen sink” and view the underlying HTML code.



program, as well as to provide an additional source of indexing.<sup>92</sup> There are two kinds of “meta” tags: keyword and description tags. Keyword meta tags are invisible to the Web-surfing viewer, but not to search engine robots, which use them to index the page.<sup>93</sup> Description meta tags, also invisible to the “naked” viewing eye, are made visible by search engines, which use them to supply the description of a Web page for search query output.<sup>94</sup> Because spamdexing generally involves keyword rather than description meta tag abuse, this Article will focus on the former.<sup>95</sup>

Because keyword meta tags are invisible to those browsing the Internet, many webmasters are tempted to use keywords that are irrelevant (or remotely related) to actual Web page content. Some “overload their pages with frequently requested key words (‘sex,’ ‘money,’ or ‘Pamela Anderson,’ for instance) or with thousands of repetitions of the same key word.”<sup>96</sup> Some use the names of

---

92. See HotWired, *Webmonkey: Beginners: Make Sure Your Site Gets Indexed: The Meta Tag Attributes* (visited Nov. 17, 1998) <[http://www.hotwired.com/webmonkey/geektalk/97/18/index4a\\_page3.html](http://www.hotwired.com/webmonkey/geektalk/97/18/index4a_page3.html)>.

93. An example of a keyword meta tag in HTML would be <META name=“keywords” content=“article, meta tags, internet, unfair competition”>.

94. An example of a description meta tag in HTML would be <META name=“description” content=“An article about whether the use of meta tags constitutes unfair competition”>.

95. Because search engines generally use the description meta tag for the text in query outputs, *see* SELLERS, *supra* note 11, at 22, spamdexers do not benefit from spamdexing *both* the keyword and description meta tags. By spamdexing the keyword, webmasters hope to appear in query output. Once the public sees this output, it is to the benefit of the webmaster to be more forthcoming in the description that will be read. Although some webmasters have spamdexed the description tag, the majority of spamdexing is with keyword meta tags. *But see* Niton Corp. v. Radiation Monitoring Devices, Inc., No. Civ.A. 98-11629-REK, 1998 WL 812685, at \*3 (D. Mass. Nov. 18, 1998) (granting “preliminary injunction subject to modification” where defendant copied verbatim the description meta tags of plaintiff competitor and used them in its own Web page); *see also* note 333 and accompanying text. Therefore, references to meta tags throughout this Article should be understood as references to keyword meta tags.

96. Randy McClain, *Snared by the Web? Even the Cyber-Savvy Will Need Patience to Snag Internet Success*, BATON ROUGE ADVOC., Sept. 20, 1997, at 1E.

celebrities,<sup>97</sup> trademarks and trade names of others,<sup>98</sup> popular generic terms,<sup>99</sup> analogous terms,<sup>100</sup> and categorical generic terms abstracted from actual content.<sup>101</sup> Honest webmasters, frustrated by spamdexers, may themselves be forced to spamdex so that they can be found by those who actually seek them.<sup>102</sup>

As noted above, not all search engines use the same set of indexing information or sorting criteria.<sup>103</sup> Most search engines use meta tags, but

---

97. One page, in invisible buried text, mentions "Gillian Anderson," "Seinfeld," "Jeff Daniels," "Cameron Diaz," and "Jim Carrey" on a page offering "pictures of nude girls." *Dirty Lobster Warning* (visited Jan. 10, 1997) (Web page on file with *Harvard Journal of Law & Technology*). The offending page was found by doing a search for "Gillian Anderson" on AltaVista. See AltaVista (visited Jan. 10, 1997) <<http://www.altavista.digital.com>> (search on file with *Harvard Journal of Law & Technology*). The search turned up 64,997 documents; the offending page was the third page listed. See *id.*

98. See *Dirty Lobster Warning*, *supra* note 97 (using "Coca Cola," "Playboy," and "Zippo" in buried text).

99. See *id.* This site lists generic terms, with variations of spellings and cases, presumably to ensure it shows up in a variety of searches. For example, the site uses "Free," "free," "Porno," "porno," "Lesbians," "LESBIANS," and "lesbians." *Id.* It even uses all the letters of the alphabet as separate terms. See *id.*

100. The *Advanced Concepts* suit, see *infra* notes 122–30 and accompanying text, is a good example of this.

101. For instance, a site created by a law student that contained legal links, course outlines, and lawyer jokes might use keyword meta tags that would include "law," "jokes," "funny," "study," and "education."

102. David Landgren stuffs his own name multiple times into buried text at the bottom of his home page. See *Landgren.net* (visited Nov. 17, 1998) <<http://www.landgren.net>>. Landgren, in the same buried text, wryly apologizes for this act: "I hate to spamdex like this but otherwise there are too many hits for me over at [another website], and people would otherwise never hit this page when they searched for me. Please accept my apologies." *Id.*

On a more personal note, when researching this Article, I performed electronic searches using all possible variations of "meta tags" — "meta tags," "metatags," and "meta-tags" — to get complete results. However, a casual researcher might only search for "metatags" and consequently overlook scholarship on "meta tags." Thus, there is something to be said for webmasters that come up with every imaginable variant of a legitimate keyword to help ensure that their sites will be found.

103. Search engines supply varying levels of detail on what types of information they use to index, and how they use these terms to index. See AltaVista, *AltaVista Help — FAQ* (visited Dec. 19, 1998) <[http://www.altavista.com/av/content/ques\\_webmaster.htm](http://www.altavista.com/av/content/ques_webmaster.htm)>; AltaVista, *AltaVista Help — Meta Tag* (visited Nov. 17, 1998) <[http://www.altavista.com/av/content/addurl\\_meta.htm](http://www.altavista.com/av/content/addurl_meta.htm)>; Excite, *Getting Listed Help* (visited Nov. 17, 1998) <<http://www.excite.com/Info/listing.html>>; HotBot, *HotBot Help: Common Questions: How Do I Improve My Site's Ranking?* (visited Nov. 17, 1998) <<http://www.hotbot.com/help/questions/question2.asp>>; Infoseek, *Infoseek: Help: How Do I Submit My Web Site?* (visited Nov. 17, 1998) <[http://www.infoseek.com/Help?pg=meta\\_tag](http://www.infoseek.com/Help?pg=meta_tag).





search engines.<sup>110</sup> Yet another is the use of “bridge” pages to get around search engines that penalize for keyword repetitions.<sup>111</sup> Thus, even if meta tags eventually diminish in importance, similar kinds of abuse will continue to emerge, to which the framework of relevance discussed in this Article should apply equally well. The real issue — resolving the tension between webmasters, the public, and proprietary-rights holders — will not go away merely because of a change in the form of indexing.

---

110. WWF Impact uses frames for its visible content. See *WWF Impact* (visited Jan. 31, 1998) <<http://wwfimpact.simplenet.com/main.html>>. Framing is the practice of “set[ting] up a Web site so that the viewer will click on a hyperlink and find a second Web site displayed within a ‘frame’ which is on the first Web site.” Elgison & Jordan, *supra* note 15. A peek at WWF Impact’s source code, however, shows additional text that would appear for those few of us still using non-frame compatible browsers, such as an early version of Netscape Navigator. This text includes *seventeen* repetitions of the phrase “WWF, World Wrestling Federation, Monday Night Raw, Raw Is War, Steve Austin, Sunny, The Undertaker, Shawn Michaels, Degeneration-X, nWo, WCW, Pro-Wrestling, Prowrestling.” *WWF Impact, supra* (source code on file with *Harvard Journal of Law & Technology*). This page appeared as the second out of 68 results of a query for “wwf wwf wwf wwf wwf” on HotBot (visited Jan. 31, 1998) <<http://www.hotbot.com>>. Here, the indexing term was used neither as a meta tag, nor as invisible text in the traditional sense (where the text appears *on* the page, but in a background color). As such, it is yet another excellent example of how smart webmasters will come up with new ways to spamdex, even where search engines try to screen out known forms of abuse.

111. A “bridge” page is one that exists only to draw attention to another page. It will generally contain meta tags keyed to a particular interest. See SELLERS, *supra* note 11, at 32–33. One might even create separate bridge pages designed with the indexing priorities of different search engines in mind. See *id.* The use of “Monica Lewinsky” by CityAuction is one such example. See *supra* note 5 and accompanying text.



Lawsuits are beginning to proliferate.<sup>112</sup> The two earliest cases involved indexing terms that were at least indirectly relevant to the offending site,<sup>113</sup> but involved competitive injuries. In *Playboy Enterprises, Inc. v. Calvin Designer Label*,<sup>114</sup> both the defendant and plaintiff Playboy Enterprises, Inc. ("PEI"), ran adult Internet sites. The defendant's Internet site used the domain names "playmatelive.com" and "playboyxxx.com," and the "Playboy" and "Playmate" trademarks on the site itself in buried text.<sup>115</sup> PEI asserted claims for federal trademark infringement, false designation of origin and unfair competition, federal trademark dilution, and assorted state law claims.<sup>116</sup> The Northern District of California granted PEI a preliminary injunction on all claims, ordering the defendants to cease using PEI's trademarks in buried text

---

112. Prohibitions on key word, meta tag, and other abuse of trademarks have started to appear in court injunctions. See *McGraw v. Salmon*, No. CV 98-2495R, 1998 U.S. Dist. LEXIS 10987, at \*2-\*4 (C.D. Cal. June 30, 1998) (permanently enjoining defendant from using the names of plaintiff country music performers as meta tags, domain names, or otherwise); *Cardservice Int'l v. McGee*, 950 F. Supp. 737, 743 (E.D. Va.) (prohibiting any use of plaintiff's trademarks in any part of an Internet site), *aff'd*, 129 F.3d 1258 (4th Cir. 1997); *Toys "R" Us, Inc. v. Abir*, No. 97 Civ. 8673, 1997 U.S. Dist. LEXIS 22430, at \*3 (S.D.N.Y. Dec. 31, 1997) (order for preliminary injunction requiring defendants to "immediately notify in writing all publishers of directories in which [Defendants' infringing] names appear, including internet search engines, to delete all references to these names from their public databases"); *Toys "R" Us, Inc. v. Akkaoui*, No. C 96-3381, 1996 U.S. Dist. LEXIS 17090 (N.D. Cal. Oct. 29, 1996) (ordering defendant to contact all search engines and databases to remove references to defendant's site); see also *Patmont Motor Werks, Inc. v. Gateway Marine, Inc.*, No. C 96-2703, 1997 U.S. Dist. LEXIS 20877, at \*17 n.7 (N.D. Cal. Dec. 17, 1997) (noting that although plaintiff appeared to have argued for trademark infringement because defendant "'inserted' [plaintiff's] mark into internet search engines," the court granted summary judgment for defendant because plaintiff failed to describe how "websurfers were lured away from [plaintiff's] site," and therefore did not present a genuine issue for trial).

113. The defendants offered pictures of naked women, as does Playboy. Whether the use of "Playmate" and "Playboy" — as terms connoting pictures of naked women — is trademark infringement is a separate issue from that of whether such use is relevant to actual content at defendants' site. This point will be developed further, *infra* Part IV.B.

114. 985 F. Supp. 1220 (N.D. Cal. 1997); see also *Kuester & Nieves*, *supra* note 21, at 274-75; *Weinberg*, *supra* note 21, at 588-89.

115. See *Calvin Designer Label*, 985 F. Supp. at 1221-22; see also First Amended Complaint, *Calvin Designer Label*, No. C-97-3204, ¶¶ 27, 31 (N.D. Cal. 1997), available at <<http://www.patents.com/ac/playcpt.sht>> (visited Nov. 15, 1998); *Abate*, *supra* note 106.

116. See *Calvin Designer Label*, 985 F. Supp. at 1221; see also First Amended Complaint, *Calvin Designer Label*, No. C-97-3204, ¶¶ 37-57.

or meta tags.<sup>117</sup> The decision, however, may have rested on domain name usage, and therefore provides little guidance on meta tag abuse.<sup>118</sup>

The second early case also does not provide much guidance. In *Insituform Technologies, Inc. v. National Envirotech Group, L.L.C.*,<sup>119</sup> the Eastern District of Louisiana entered a consent judgment against the use of meta tags by a direct competitor.<sup>120</sup> The defendant was ordered to remove plaintiff's federally registered trademarks and service marks from the meta tags of its Web page, and to resubmit its corrected Internet sites to the major search engines to purge the allegedly infringing references.<sup>121</sup> Because it was a consent judgment, the court did not reach the meta tag issue on its merits.

A third case, *Oppedahl & Larson v. Advanced Concepts*,<sup>122</sup> had the potential for more far-reaching consequences. Unlike *Insituform* and *Calvin Designer Label*, this was a pure meta tag case that did not involve directly competitive or similar services; further, the meta tags were used in an indirectly relevant way.<sup>123</sup> Plaintiffs, themselves attorneys, asserted that the use of their names as meta tags violated federal unfair competition, federal dilution, and state law.<sup>124</sup> The defendants were neither lawyers nor providers of legal services; however, they did offer domain name registration and other Internet services.<sup>125</sup> Oppedahl is a

---

117. See *Calvin Designer Label*, 985 F. Supp. at 1221. The court also ordered defendants to cease using the disputed domain names. See *id.*

118. See Kuester & Nieves, *supra* note 21, at 274-75.

119. Civ. Act No. 97-2064 (E.D. La. Aug. 26, 1997), available at <<http://www.cll.com/case1.htm>>.

120. See *Insituform*, Civ. Act. No. 97-2064, slip op. at 2; see also Kuester & Nieves, *supra* note 21, at 275; Weinberg, *supra* note 21, at 589.

121. See *Insituform*, Civ. Act. No. 97-2064, slip op. at 2-3. Doing this would have the effect of wiping clean the indexing to plaintiff's site. The defendant was also ordered to follow up with the search engines in writing to ensure that the allegedly infringing meta tags were purged. See *id.* at 3-4, Exhibit 1.

122. Civ. Act. No. 97-Z-1592, 1998 U.S. Dist. LEXIS 18359 (D. Colo. Feb. 6, 1998), 1997 U.S. Dist. LEXIS 23105 (D. Colo. Dec. 19, 1997), 1997 U.S. Dist. LEXIS 23108 (D. Colo. Dec. 19, 1997).

123. Trademark infringement requires that consumers are likely to confuse defendants' use of the trademark with that of the plaintiff. See 15 U.S.C. §§ 1114(1)(a), 1125(a) (1994). Dilution, on the other hand, precludes uses that reduce the distinctive quality of plaintiff's trademark, regardless of whether the defendants' use leads to a likelihood of confusion. See 15 U.S.C. §§ 1125(c), 1127 (Supp. I 1995). Both doctrines are discussed in detail, *infra* Part IV.B.

124. See Complaint, *Advanced Concepts*, ¶ 1 (D. Colo. filed July 23, 1997), available at <<http://www.patents.com/ac/complain.sht>>; see also Kuester & Nieves, *supra* note 21, at 276.

125. The online version of the Complaint contains hyperlinks to copies of the allegedly infringing pages. See Complaint, *Advanced Concepts*, *supra* note 124.



well-known authority on Internet law and domain name disputes.<sup>126</sup> It appears that the defendants used plaintiffs' names in hopes of gaining exposure to persons who use "Oppedahl" as a search term because of their interest in domain names or Internet legal issues.<sup>127</sup> Permanent injunctions have been granted, two on joint motion and another by default.<sup>128</sup> Unfortunately, the court did not have the opportunity to reach a decision on the merits. Such a decision might have helped to determine just how relevant a meta tag must be to actual content to avoid liability, especially considering that the litigants were not competitors.

These suits raise troubling questions that were not answered by their respective courts. In the first two suits, the offending indexing terms were directly or indirectly relevant. "Playboy" is evocative of adult entertainment; further, the use of one competitor's name is evocative of another.<sup>129</sup> In the third case, *Advanced Concepts*, there was no real likelihood of confusion, and although the use of the firm's name was at best indirectly relevant, "Oppedahl & Larson" is arguably a reasonable search term for "Internet-savvy lawyers."<sup>130</sup> Whether its status as a trade

---

126. A search in LEXIS, News library, Curnws file on Nov. 23, 1997 turned up 48 hits referring to Oppedahl, an attorney who maintains a detailed site on developments on the topic of Internet domain names. See Oppedahl & Larson, *NSI Flawed Domain Name Policy Page* (visited Nov. 30, 1998) <<http://www.patents.com/nsi.shi>>.

127. Defendants, by using Oppedahl's name, likely hoped that (1) people interested in Oppedahl might also be interested in defendants; or (2) people would use Oppedahl's name as a conceptual search term, i.e., as a proxy for "Internet law" or "domain name dispute information." See *infra* Part III.B.1 (discussing types of searches).

128. See Oppedahl & Larson v. Advanced Concepts, Civ. Act. No. 97-Z-1592, 1998 U.S. Dist. LEXIS 18359 (D. Colo. Feb. 6, 1998) (granting motion for permanent injunction against defendant George Williams, and treating it as one for default judgment); Oppedahl & Larson v. Advanced Concepts, Civ. Act. No. 97-Z-1592, 1997 U.S. Dist. LEXIS 23105 (D. Colo. Dec. 19, 1997) (granting joint motion for judgment and permanent injunction against defendants Welch and Advanced Concepts); Oppedahl & Larson v. Advanced Concepts, Civ. Act. No. 97-Z-1592, 1997 U.S. Dist. LEXIS 23108 (D. Colo. Dec. 19, 1997) (granting joint motion for judgment and permanent injunction against defendants MSI Marketing, Inc., Professional Website Development, and Internet Business Services).

129. Such uses directly implicate unfair competition law. See *infra* Part IV.B.

130. One commentator suggests that a celebrity may lose control over his persona when it becomes a generic mark. See Paul J. Heald, *Filling Two Gaps in the Restatement (Third) of Unfair Competition: Mixed-Use Trademarks and the Problem with Vanna*, 47 S.C.L. REV. 783, 804 (1996).

Not all commercial references to celebrity personae are, however, pure exploitations of the unique or intrinsic personal qualities of the individual celebrity. Because the meaning of celebrity is often determined by the public, frequently certain aspects of a celebrity persona are employed generically. For

name renders its use (however relevant) as a meta tag unfair is another matter. Because the Internet community is highly responsive to even the mere threat of a lawsuit,<sup>131</sup> it is vital to determine how *much* relevance is necessary, and how *unfair* spamdexing must be to impose liability.

*Playboy Enterprises, Inc. v. Welles*<sup>132</sup> addresses some of these concerns. Defendant Terri Welles was Playboy "Playmate of the Year" in 1981 and alleged that she had referred to herself as such since then.<sup>133</sup> In 1997, she established a website at "www.terriwelles.com," which included photos, a fan club posting board, and other information.<sup>134</sup> The site included visible references to "Playboy," "Playmate of the Year," and "PMOY," along with disclaimers.<sup>135</sup> The site also used the terms "Playboy" and "Playmate" as meta tags, along with a number of relevant generic terms.<sup>136</sup> PEI moved for a preliminary injunction on the basis of trademark infringement, false designation of origin, and trademark dilution.<sup>137</sup> The court denied the motion, because Welles's uses of the terms raised a fair use defense under 15 U.S.C. §§ 1115(b)(4) and

example, . . . . John Wayne = rugged frontier spirit and patriotism . . . .

*Id.*

131. See Cindy Collins, *Web Site Coding Practices Turn Up Tricky Trademark Issues*, INSIDE LITIG., Jan. 1998, at 15 (quoting attorney William Cook, "who characterizes the Internet community as 'heavily responsive' to litigation or threatened litigation").

132. 7 F. Supp. 2d 1098 (S.D. Cal.), *aff'd*, No. 98-55911, 1998 U.S. App. LEXIS 27739 (9th Cir. Oct. 27, 1998); see also *Descriptive Use of Playboy Marks is Fair Use*, Pat. Trademark & Copyright J. (BNA), May 7, 1998, at 12; John Naughton, *Internet: Stripped for Action — Playboy Versus the Pin-Up Queen*, LONDON OBSERVER, Mar. 29, 1998, at 9.

133. See *Welles*, 7 F. Supp. 2d at 1100.

134. See *id.* The site "garners about \$2,500 per month." Carl S. Kaplan, *Case Against Playboy Model Could Set Precedent*, CYBERTIMES (N.Y. Times) (Apr. 17, 1998) <<http://www.nytimes.com/library/tech/9804/cyber/cyberlaw/17law.html>>.

135. See *Welles*, 7 F. Supp. 2d at 1100–01. The title bar stated "Terri Welles — Playboy Playmate of the Year 1981," and the page itself was entitled "Terri Welles — Playmate of the Year 1981." *Id.* at 1100. Each page at her site also used "PMOY '81" as a repeating watermark. See *id.* The disclaimers noted that Ms. Welles was not affiliated with PEI, and also noted that PEI's marks were federally registered. See *id.* at 1100–01.

136. See *id.* at 1101. A viewing of the source code shows that the site also uses generic terms including "model," "naked," and other such terms. See *Terri Welles* (visited Apr. 19, 1998) <<http://www.terriwelles.com>> (on file with *Harvard Journal of Law & Technology*).

137. See *Welles*, 7 F. Supp. 2d at 1101–02. The claims were asserted under the Lanham Act, §§ 32(1), 43(a), and 43(c). See 15 U.S.C. §§ 1141(1), 1125(a), (c) (1994 & Supp. I 1995).



1125(c)(4)(A).<sup>138</sup> Unlike the cases mentioned above, the defendant's use of PEI's trademarks were in good faith and for a descriptive (i.e., non-trademark) use:

[T]rademarks such as *Playmate* are not only trademarks related to *Playboy* magazine, but they are titles bestowed upon particular models who appear in that magazine. . . ., who then use the title to describe themselves. Much like Academy Award winners, crowned Miss Americas, and Heisman Trophy winners, *Playboy* Playmates are given a title which becomes part of their identity and adds value to their name. Indisputably, these winners represent the awarding organization or sponsor, but the title becomes part of who they are to the public.<sup>139</sup>

*Welles* may provide guidance as to when the use of a term is sufficiently relevant, and in sufficient good faith, to constitute fair use.<sup>140</sup> A helpful contrast is provided by *Playboy Enterprises, Inc. v. AsiaFocus International, Inc.*<sup>141</sup> In *AsiaFocus*, PEI asserted claims for trademark infringement, unfair competition, and dilution.<sup>142</sup> As in *Welles*, defendants used PEI's trademarks in their websites' source code.<sup>143</sup> However, defendants also registered domain names that contained PEI's trademarks, used PEI's trademarks in its Web pages, and sold playing cards, calendars, wrist watches, and key chains with PEI's trademarks.<sup>144</sup> Unlike *Welles*, where the plaintiff had a good faith reason to use PEI's trademarks descriptively, the defendants in *AsiaFocus* used "Playboy" and "Playmate" without authorization or reason. The district judge, in adopting the report and recommendation of the magistrate judge, assessed statutory damages of \$3,000,000 plus costs and reasonable attorney's fees<sup>145</sup> — a very different result, and one that should caution

---

138. See *Welles*, 7 F. Supp. 2d at 1104–05.

139. *Id.* at 1102.

140. See *infra* notes 383–96 and accompanying text.

141. No. 97-734-A, 1998 U.S. Dist. LEXIS 10359 (E.D. Va. Feb. 2, 1998) (Mag. J.), adopted by 1998 U.S. Dist. LEXIS 10459 (E.D. Va. Apr. 10, 1998); see also *infra* notes 396–403 and accompanying text.

142. See *AsiaFocus*, 1998 U.S. Dist. LEXIS 10359, at \*1.

143. See *id.* at \*8.

144. See *id.*

145. See *AsiaFocus*, 1998 U.S. Dist. LEXIS 10459, at \*1. Note, however, that the defendants did not appear and were subject to a default judgment. See *id.*

those who do not know whether they sit on the *Welles* or *AsiaFocus* side of the fence.

I do not suggest that these courts reached incorrect conclusions. Rather, these cases show the importance of discussing in greater detail the nature and the importance of the interests involved. For instance, might Pepsi incur the wrath of Coke by using “coke” as a meta tag on its home page? If yes, might it be able to do so on a sub-page that discusses a Pepsi-Coke taste test? Would I, referring to the analysis in this Article, incur liability by using “Oppedahl & Larson” as a meta tag? Two issues emerge, both of which will remain central to the analysis. First, how much relevance is sufficient? Part III seeks to answer this question. Second, assuming that liability is proper where meta tags are used irrelevantly, when is liability also proper where meta tags are used relevantly? Part IV explores these questions.

### III. DEFINING THE INTERESTS

*Often [searching] involves tinkering with [a] combination of words, so as to weed out extraneous information — or “noise” — from the search results. It’s the digital equivalent of jostling the TV antenna to get better reception.*

— Stephen Johnson<sup>146</sup>

Prior to attempting to balance the interests of the public, webmasters, and those with proprietary rights, it is necessary to define these interests. This Part examines the nature and weight of these interests in light of well-established principles of relevance and unfair prejudice, using as guidance information science and the Federal Rules of Evidence. To help determine the reasonableness of each party’s expectations, this Part briefly considers analogous disputes involving domain names, hyperlinking, and framing. It concludes with a brief exploration of spamdexing in light of economic theory.

#### *A. Relevance and Noise*

The public wants to find information, and webmasters want to be found. The Internet infoglut has encouraged the development of search engines that seek to fill these needs.<sup>147</sup> Those asserting proprietary rights, however, typically want to control their “property” regardless of

---

146. JOHNSON, *supra* note 55, at 163–64.

147. *See supra* notes 32–83 and accompanying text.



others' desire to find or be found. The perfect Internet, from the public's perspective, should not suffer undue interference from trademark owners, whose rights can make searches underinclusive, or from webmasters, whose spamdexing makes searches overinclusive. To reach a proper balance, it is necessary to define and weigh the interests of the public, webmasters, and those asserting proprietary rights.

The weight of the interests of the public and webmasters varies with the keywords at issue and their relevance to either the goal of the search (for the public) or the actual content at issue (for webmasters).<sup>148</sup> The interest of those asserting proprietary rights is measured, not just through relevance analysis, but also through the laws of unfair competition, which is addressed in further detail in Part IV.B.

The key concept in the spamdexing context is relatedness, or relevance.<sup>149</sup> The harm in spamdexing is the intentional introduction of

---

148. The weight of a particular party's interest depends in large part on the reasonableness of its expectations of success in obtaining what it wants. The label "reasonable" by itself, however, is not helpful, because it invites courts to engage in post-hoc adjudication.

149. The issue of relevance is a separate — yet intertwined — issue from that of reliability. On one hand, even where information is dead-on relevant, it may not be reliable. On the other hand, some authorities suggest that the Internet infoglut has a leveling effect that makes people treat unreliable sources as credibly as reliable sources of information. See Scott Canon, *Look Out: It's Media Overload; Proliferation of TV Shows, Periodicals and Web Sites Can Be Hard to Cope With*, KANSAS CITY STAR, July 6, 1997, at A1 (noting that a University of Missouri study has shown that "people gave no more credibility to established newspapers than they afforded to individuals they had never heard of before"). Thus, it is essential that we develop ways of ensuring both the relevance and reliability of information.

Analyzing the differences and interplay between relevance and reliability is beyond the scope of this Article. In the context of "traditional" (i.e., presumptively reliable) news media, one either assumes a threshold of forgivable irrelevance, or one develops habits to skip the irrelevant. For example, we have all sat still through twenty minutes of irrelevant news when all we are interested in is the nightly weather forecast. In fact, advertisers depend on our willingness to sit through several minutes of commercials that are irrelevant to our favorite television show. On the other hand, with media that does not force us to endure irrelevance, such as a newspaper, we simply flip past that which does not interest us. The newspaper itself is organized in a fashion to expedite this process.

The Internet, of course, has much more information than the morning newspaper, with exponentially less structure. It is hard to separate relevance from reliability; we like our newspaper because it is both relevant and reliable. With institutions such as the Drudge Report, one sees the emergence of new forms of relevant information, but with no corresponding assurances of reliability. See *The Media Are Split Over Appearance of Drudge on 'Meet The Press'*, ST. LOUIS POST-DISPATCH, Feb. 10, 1998, at D3. I suspect that just as institutions developed to ensure reliability in the context of print and television, we will see the same occur eventually in the context of the Internet.

*irrelevant* documents into a search query. As defined by information science, relevance “is the possibility — deemed possibility or probability — of the document’s helping to solve a problem.”<sup>150</sup> This possibility — whether there is probative value in what appears in one’s query output — is a concept familiar to anyone acquainted with the well-established standard of relevance embraced in Federal Rules of Evidence 401 and 402.<sup>151</sup>

Whether the webmaster interest is significant thus requires looking at the relationship between meta tag and actual content. Paraphrasing the Advisory Committee’s Note, “[r]elevancy is not an inherent characteristic of any item of [data] but exists only as a *relation between* an item of [data] and a matter [sought from a search].”<sup>152</sup> Thus, in the context of online searching, Rule 401, which defines relevance, could be rewritten as follows: “‘Relevant data’ means query output having *any tendency* to be of use to one who is seeking such data.”<sup>153</sup> As will be developed below, the standard of relevance should not be overly strict,

---

150. Roger R. Flynn, *Introduction to Information Retrieval*, in INFORMATION SCIENCE, *supra* note 1, at 63, 66.

151. Rule 401 states: “‘Relevant evidence’ means evidence having *any tendency* to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.” FED. R. EVID. 401 (emphasis added). Rule 402 states: “All relevant evidence is admissible, except as otherwise provided . . . . Evidence which is not relevant is not admissible.” FED. R. EVID. 402.

152. FED. R. EVID. 401 (Advisory Comm. note). The notion of relevance in the Federal Rules is one of “logical” relevance. See Norman M. Garland, *An Overview of Relevance and Hearsay: A Nine Step Analytical Guide*, 22 SW. U. L. REV. 1039, 1042–43 (1993) (“Evidence will be logically relevant if it has any tendency (even the slightest) to make the fact of consequence more or less likely. Thus, the evidence does not have to conclusively prove the fact.”) (footnote omitted); Edward J. Imwinkelried, *A Brief Defense of the Supreme Court’s Approach to the Interpretation of The Federal Rules of Evidence*, 27 IND. L. REV. 267, 288 (1993). Logical relevance, at the common law, embraces two concepts: (1) materiality, or the logical relation between the fact asserted and the fact sought to be proved; and (2) underlying logical relevance and authenticity. See Edward J. Imwinkelried, *Federal Rule of Evidence 402: The Second Revolution*, 6 REV. LITIG. 129, 139 (1987); see also Keith Burgess-Jackson, *An Epistemic Approach to Legal Relevance*, 18 ST. MARY’S L.J. 463, 465 (1986) (suggesting that “[t]he central concepts in the law of relevance are relevance, admissibility, materiality, and sufficiency”).

153. See CHARLES T. MCCORMICK, MCCORMICK’S HANDBOOK OF THE LAW OF EVIDENCE § 185 (4th ed. 1992) (stating that “the evidence must be more (or less) probable when the disputed fact is true rather than false. Such evidence often is said to have ‘logical relevance’ . . .”).



as long as the data has *any tendency* to be of use.<sup>154</sup> Thus, as long as output is even remotely related to the query goal, it is relevant.<sup>155</sup>

Using relevance as a lodestone sharpens the analysis of spamdexing in at least two ways. First, it provides a standard against which liability may be assessed. Second, and more significantly for present purposes, it provides a definition of the spamdexing harm in terms used by information science. Table One shows how relevance is expressed by information science. "Relevant hits" are the total number of relevant documents that show up in query output.<sup>156</sup> "Noise" refers to the number of irrelevant documents, or false drops, that nevertheless show up in query output.<sup>157</sup> "Precision" is a measure of the percentage of relevant hits to noise in query output.<sup>158</sup> This ratio measures "the human effort required in perusing and evaluating the output."<sup>159</sup> In other words, it measures a "search's success in not retrieving irrelevant" information.<sup>160</sup> "Recall" measures the ratio between relevant hits retrieved against the total relevant documents, retrieved or not.<sup>161</sup> Recall reflects a "search's success in retrieving" relevant information.<sup>162</sup>

---

154. Cf. FED. R. EVID. 401 (Advisory Comm. note).

The fact to be proved may be ultimate, intermediate, or evidentiary; it matters not, so long as it is of consequence in the determination . . . .

. . . Evidence which is essentially background in nature can scarcely be said to involve disputed matter, yet it is universally offered and admitted as an aid to understanding.

*Id.*; see also MCCORMICK, *supra* note 153, § 185 (stating that materiality, a component of relevance, "in its more precise meaning looks to the relation between the propositions for which the evidence is offered and the issues of the case"). In court, relevance is determined by the relation between the evidence and the legal issues before the court, *see id.*, whereas in the spamdexing context, relevance is determined by the relation between data obtained in a query (query output), data sought in such a query (query goal), and the index term.

155. I will use the terms "relatedness" and "relevance" interchangeably.

156. See Flynn, *supra* note 150, at 66–67.

157. See *id.*; see also SHENK, DATA SMOG, *supra* note 35, at 30.

158. See Flynn, *supra* note 150, at 66–67.

159. *Id.* at 66; see also NAHUM GOLDMAN, ONLINE INFORMATION HUNTING 116 (1992).

160. GOLDMAN, *supra* note 159, at 116.

161. See Flynn, *supra* note 150, at 66–67.

162. GOLDMAN, *supra* note 159, at 116. There is, in general, an inverse relationship between recall and precision. See *id.* at 117.

Table One: Spamdexing as Described by Information Science<sup>163</sup>

	Relevant	Irrelevant	
Query output	Relevant Hits (a)	Noise (b)	Total retrieved (total hits) (a + b)
Not in query output	Relevant Misses (c)	Rightly rejected (d)	Total not retrieved (c + d)
	Total Relevant (a + c)	Total irrelevant (b + d)	Total documents in search engine database (a + b + c + d)
	Precision: relevant hits divided by relevant hits + noise. a ÷ (a + b)  Recall: relevant hits divided by relevant hits + relevant misses. a ÷ (a + c)		

The precision of a search rests in large part on the depth of information contained in the search terms used.<sup>164</sup> More useful are words that are distinctive, infrequent, or capable of a limited number of meanings; such words or phrases are *information-rich*.<sup>165</sup> It is unreasonable to rely solely upon *information-poor*, or frequently used words, for a search.<sup>166</sup> In all cases, but especially in the case of words

163. This table is adapted from Flynn, *supra* note 150, at 67.

164. See JOHNSON, *supra* note 55, at 153 (“Any text can be reduced to an inventory of words, arranged not by syntactical order but by frequency . . . . Some of these numbers are more relevant than others . . . . Any English speaker will immediately comprehend that the latter number reveals more than the former.”).

165. As Steven Johnson puts it, once you “eliminate the high-frequency words . . . [w]hat remains are the more distinctive words: *interface, bitmap, mouse, Overlook, blood.*” *Id.* at 154 (demonstrating this concept by running a word-frequency analysis using a chapter of Johnson’s book, a segment of an Apple computer manual, and a selection from Stephen King’s *The Shining*). The rarity of a word or phrase “makes [it] meaningful, in the same way that the description ‘he has a nose’ has less information than ‘he has a Roman nose.’” *Id.*

166. See *id.* at 155 (“[T]he results tend to overemphasize the coincidences, placing undue importance on the terms that happen to overlap.”). A look at the most popular search terms show that people tend to overuse information-poor terms. According to one source, the top ten search terms are (1) free; (2) sex; (3) nude; (4) pictures; (5) warez; (6) xxx; (7) diana; (8) pics; (9) new; and (10) university. See David Segal, *The XXX-Files*, WASH. POST, Oct. 26, 1997, at W7.

It is illustrative to sit back and watch people as they search the Internet. See EVAN I. SCHWARTZ, WEBONOMICS 21 (1997). “Voyeur” sites allow one to spy on searches



that are information-poor, one must use terms together to reduce the noise that results in imprecise searches.<sup>167</sup> In any case, it is always better to “rig up” a list of information-rich words.<sup>168</sup>

### B. Reasonable Expectations

By spamdexing the search engines with irrelevant keywords, webmasters vastly increase the amount of noise in query output, making the precision of searches low — rendering searches vastly overinclusive. This increase in noise is the harm of spamdexing. Of course, a searcher relying on only one keyword may also miss many relevant documents, resulting in a low recall ratio. However, it is within the power of searchers to use a variety of searching techniques that can increase *both* precision and recall,<sup>169</sup> such as using multiple keywords, Boolean operators, phrase searching, and using superior alternatives when applicable.<sup>170</sup> Thus, it is necessary to look to more than just spamdexing’s effect on the precision ratio. First, even if an indexing term is irrelevant from the consumer’s point of view, it may be relevant from that of the alleged spamdexer.<sup>171</sup> Although the supposedly offending term may be sufficiently ambiguous such that its use in searches invites imprecision, a webmaster may still have a legitimate interest in using it. Second, because searchers may employ techniques

---

being conducted by other parties on the search engines. *See, e.g.,* Magellan Voyeur (visited Feb. 7, 1998) <<http://voyeur.mckinley.com/cgi-bin/voyeur.cgi>>; *see also* Bill Husted, *Personal Technology; Cyberscene: What a Fun Web You Can Weave Snooping in Cyberspace*, ATLANTA J.-CONST., Nov. 30, 1997, at P5.

167. *See* JOHNSON, *supra* note 55, at 163–64 (“Often this process involves tinkering with [a] combination of words, so as to weed out extraneous information — or ‘noise’ — from the search results. It’s the digital equivalent of jostling the TV antenna to get better reception.”).

168. *See id.* at 162.

169. “[R]elevance and recall[] are the cornerstone on which the evaluation of systems rests. They indicate the effectiveness of the system (recall, in performing the search, i.e., obtaining the search results) and efficiency (eliminating nonwanted articles, i.e., without wasted effort in the output stage).” Flynn, *supra* note 150, at 67. Although there may be an inverse relationship between precision and recall, *see* GOLDMAN, *supra* note 159, at 117, where this point lies may well depend on whether the search is targeted or conceptual, as discussed in this Section. A well-tailored search should boost both precision and recall, at least up to some point of diminishing returns. After that point, an increase in precision would indeed reduce recall.

170. *See infra* notes 446–63 and accompanying text.

171. Suppose one were searching for information on Coca-Cola and used “coke” as a search term. Some pages in the query output might not be about soft drinks, but cocaine. Here, the indexing term “coke” is irrelevant from the searching party’s point of view, but highly relevant from that of the webmaster.

to narrow their searches, it is reasonable to expect them to use such techniques to increase precision and recall. While new media offers new resources, it also requires that users adopt new skills.<sup>172</sup>

Table Two compares relevance from the public’s and webmasters’ points of view. The nature of the query (the ways that the searching public uses a search term) is expressed by the y-axis (vertical), and that of the output (the relevance of the search term to actual content) by the x-axis (horizontal). As one moves to the right and downward in Table Two, the interests for both webmasters and the public are less significant.

Table Two: Defining and Comparing the Interests \*

	Directly related content	Indirectly related content	Remotely related content	Unrelated content
Narrow Target	P = high W = high A1	P = high W = medium A2	P = high W = low A3	P = high W = none A4
Broad Target	P = high W = high B1	P = high W = medium B2	P = high W = low B3	P = high W = none B4
Categorical	P = medium W = high C1	P = medium W = medium C2	P = medium W = low C3	P = medium W = none C4
Relational	P = low W = high D1	P = low W = medium D2	P = low W = low D3	P = low W = none D4

\* P = Members of the public searching on the Internet  
W = Webmasters

1. The Public

When using a search engine, an individual generally knows her goal.<sup>173</sup> The nature of this goal, however, may dictate whether a

---

172. See Katsh, *New Media*, *supra* note 27, at 1465.

173. Of course, this will not always be the case. One might surf the Internet blindly, using search terms with no expectation of a particular output. Alternatively, one might have a specific goal and be diverted by what one sees. Indeed, much of spamdexing is premised on that possibility. Nevertheless, what may be an attraction to some is a nuisance to many. For purposes of this analysis, I will assume that an Internet searcher



particular search technique or tool is reasonable.<sup>174</sup> Thus, analysis of the public's interest in finding relevant information focuses on the nature of the query itself, as expressed by the y-axis of Table Two.

Searches fall along a spectrum that may be roughly broken into two categories, each with two subcategories. The first type, *targeted* searches, are more specific and often more effective. Targeted searches include "narrow" targets and "broad" targets. The second type, *conceptual* searches, are less well-suited for search engines. Conceptual searches include the use of "categorical" and "relational" terms. As searches become less targeted and more conceptual, the expectations of the public to find relevant output become less reasonable, meriting a lower interest.<sup>175</sup>

With a targeted search, one is looking for specific information. For instance, one might want to find the home page of an institution or individual.<sup>176</sup> Another might want to find a particular document, image, or the source of a quotation.<sup>177</sup> These may be described as narrow target searches. In other cases, a searcher may want to find general information on a particular topic. For instance, one might want to locate all fan sites for the "X-Files" or "Princess Diana."<sup>178</sup> Another might want to find references to a particular word, phrase, or title, such as "Copyright Act of 1976," or "Law in a Digital World." Yet another

has some idea of that which she seeks.

174. What is relevant depends not only on *what* one is searching for, but also the depth of detail or sophistication appropriate to the user's needs. See Toni Carbo & David A. Wallace, *National and Global Information Infrastructure: Status, Issues, and Challenges*, in *INFORMATION SCIENCE*, *supra* note 1, at 153, 156 ("How will individuals perform reliable information discovery and retrieval where they will be presented with relevant information that not only matches their query but is appropriate to them based upon their language, literacy level, and level of subject expertise?").

175. As Part IV.A suggests, the public's interest in finding relevant information must be balanced against that of webmasters to determine whether the latter's deception rises to the level where liability should attach. Because the public may use alternate forms of searching in cases where search engines are ineffective, *see infra* notes 442–60 and accompanying text, liability should attach only where there is a substantial imbalance of interests.

176. One might be looking for a particular company or institution, such as Sony or Harvard. These searches are basically attempts to find the "address" of an entity. As such, they are similar in nature to the guessing games one plays with Internet domain names to find a company or institution. For instance, it is reasonable to assume that NBC may be found at nbc.com. See *infra* notes 211–13 and accompanying text.

177. Many search engines, such as HotBot, allow searching by phrase. Some, such as Lycos and HotBot, allow searching by image type, e.g., ".jpg." HotBot also allows searching by the type of program or executable file, such as Java, DirectX controls, etc. See HotBot <<http://www.hotbot.com>>; Lycos <<http://www.lycos.com>>.

178. See Weise, *supra* note 3.

might want to find out what third parties have to say about him.<sup>179</sup> These may be described as broad target searches. Broad target searches are similar to narrow target searches in that the user looks for references to a specific thing, but differ in that they seek more than one particular result.

Conceptual searches also fall into two subcategories: categorical and relational. Conceptual *categorical* searches look within a hierarchy, downward from the general to the specific. Conceptual *relational* searches, rather than looking downward, require an inferential leap to something else, whether general or specific.<sup>180</sup> Like targeted searches, however, conceptual searches also exist along somewhat of a continuum, and it will not always be easy to determine whether a search is relational or categorical.<sup>181</sup>

In a categorical search, one searches for information within a term's category or hierarchy. One might search for "cars" to find information on different brands, or "sitcoms" to find listings of popular programs and stars. Although categorical search terms that are generic do not generally conflict with proprietary rights,<sup>182</sup> some terms are generic in one context but distinctive and propertied in another. Thus, one might search for lawyers with the term "esquire," which is also the name of a magazine.<sup>183</sup>

Relational searches seek output that is inferred from the query term. Here, no discrete category exists to deal with the searches' inferences,

---

179. "Ego surfing" (searching for references to one's own name) has its benefits: Carl Oppedahl initially learned that his firm was being spandexed when searching for "his firm's name to see where it's turning up and who's linking to [his] Web site." Gardner, *supra* note 15.

180. The search may be reverse-hierarchical, such as using "Kordell Stewart" as a search term for the Pittsburgh Steelers; in other words, one would use a specific term to find a broader category in which the search term lies. The terms may also be common members of a group, such as using "whiskey" to find information on vodka. Other terms, such as "book," are capable of multiple meanings: something one reads, arresting somebody, etc. See Chaffee, *supra* note 14, at 387.

181. Indeed, a "category" is little more than a set of items that are "related" in some organized or commonly used way. The extent to which the category itself is pre-existing will determine whether something is categorical rather than relational.

182. See 15 U.S.C. § 1052(f) (1994).

183. Cf. *Hearst Corp. v. Goldberger*, No. 96 CIV. 3620, 1997 WL 97097 (S.D.N.Y. Feb. 26, 1997) (dispute between attorney with domain name "esqwire.com" and publishers of *Esquire* magazine).



analogies, or metaphors.<sup>184</sup> For instance, one might use “Princess Diana” to search for information about royalty in general.<sup>185</sup>

As one moves along the continuum from targeted to conceptual, one’s expectations of high precision or recall become less reasonable. One might assume that the targeted/conceptual continuum reflects the differences between information-rich and information-poor terms — after all, information-rich terms are better-suited to targeted searches. However, this assumption would be incorrect. The same search may fall anywhere along the continuum,<sup>186</sup> depending on the goal of each individual.<sup>187</sup> Suppose one wants to find out the name of the actress on the NBC sitcom *Friends* who plays the character “Monica” (i.e., Courtney Cox). Here, a search with the term “Friends” would be categorical, because Monica is one of the “Friends.” One may also use “Friends” as a narrow target term to search for the official “Friends” website, or as a broad target term to find references to that sitcom by others. Finally, “friends” could be used in a relational manner, to search for information on love and companionship.

Two important points emerge from this analysis. First, whether a term is information-rich is a separate concern from its use for searching. Second, a term may be used in a multitude of ways. Both conclusions will weigh heavily in balancing the interests.

---

184. See BOYLE, *supra* note 27, at 111 (“Analogies and metaphors are standard fare in legal doctrine; the analogy is used to connect one set of facts, as yet unclassified, to another set of facts which has already been subsumed under some principle.”).

185. One writer notes that “[Gerry] Adams’s regular protestation that ‘Sinn Fein is not the IRA’ sits oddly with the fact that the keywords embedded in the meta-tags at the top of his home page contain the words: ‘Irish Republican Army, IRA, I.R.A., PIRA, P.I.R.A.’” John Naughton, *Internet: Funny How the Keywords for the Sinn Fein Site Include ‘IRA’*, LONDON OBSERVER, May 17, 1998, at 9.

186. See Flynn, *supra* note 150, at 67 (“[Indexing] is a difficult problem because of the vagaries of English (ambiguity, multiple meanings) and the variance in the terms used over the years in a field and the variation in indexers.”); see also Brock, *supra* note 29 (“[S]uppose you are doing some research on ‘Star Trek: The Next Generation’ actress Gates McFadden. By typing in ‘Gates,’ you’ll turn up sites on Bill Gates, fence gates, Heaven’s gates, Gates Mills, Ohio, and more.”); Chaffee, *supra* note 14, at 387.

187. See generally Eric Lease Morgan, *A Day in the Life of Mr. D*, in THINKING ROBOTS, *supra* note 33, at 151, 153 (stating that librarians “know different answers will satisfy the needs of different people even if the questions are the same. This is because people’s knowledge bases are different. The question, ‘What is a financial planner?’ asked by Thomas would be answered quite differently for Mrs. Reid.”).

## 2. Webmasters

Lured by the promise of cash<sup>188</sup> or prestige, webmasters may be tempted to use any indexing term that presents a remote possibility that somebody will stumble across the page. Thus, a spamdexing webmaster potentially injures not only the public, but also other webmasters whose ability to be found is impaired.<sup>189</sup> In either case, however, the issue is the degree of relevance<sup>190</sup> an indexing term has to actual content at the attached Internet site or page. The relevance of an indexing term in this context must be measured objectively from a reasonable webmaster's point of view, in light of actual page content.

Webmasters have a high interest in an indexing term when it is directly relevant to actual content at the page or site. Terms that have only indirect or remote relevance merit lower expectations. The fact that the webmaster's use is irrelevant from the point of view of a particular consumer is not germane to whether it is relevant from the point of view of the webmaster.<sup>191</sup> Due to the broad standard of relevance embraced in this Article, many uses of a term will be either relevant (in some sense) or not.<sup>192</sup> To make such a determination, one must consider how

---

188. See *ActivMedia*, *supra* note 37.

189. In many cases, the webmaster-as-victim should be able to assert proprietary rights. See *infra* Part III.B.3. In others, the webmaster-as-victim may have standing similar to members of the searching public. See *supra* Part III.B.1. Webmaster injuries may exist under both categories of liability discussed in this Article: trade regulation statutes and unfair competition law. To the extent that their proprietary rights are not implicated under unfair competition law, see *infra* Part IV.B, webmasters may still have recourse to trade regulation statutes. See *infra* Part IV.A.

190. The word choice is intentional. This standard should be similar to the broad standard of relevance in the Federal Rules of Evidence. See FED. R. EVID. 401. Thus, the issue of whether a term is itself general or broad in scope, while relevant to determining whether the term is an *effective* index term, is not itself germane to whether the term is *relevant* from the perspective of the webmaster.

191. See Lars-Erik Nelson, *Internet Is Virtually Awash in Porn*, LAS VEGAS REV.-J., Dec. 10, 1997, at 15B (noting that a search for "toys" brought up information on sex toys); cf. Roger E. Schechter, *Additional Pieces of the Deception Puzzle: Some Reactions to Professor BeVier*, 78 VA. L. REV. 57, 57 (1992) ("Designing sensible false advertising rules has proven to be surprisingly difficult. In part, this is because different consumers interpret the same advertisements differently.").

192. It is tempting to use the targeted/conceptual framework for webmasters, but that may be unhelpful. The public interest, as measured by the y-axis of Table Two, measures the effectiveness of particular search techniques. The webmaster interest, as measured by the x-axis, does not consider the effectiveness of a particular search term, but rather its degree of relevance. Consider, for example, a website for a repair shop that specializes in Subaru repairs. Terms such as "automobile" and "repair" are generic. By themselves, such terms are not particularly useful, but webmasters must be able to use



large an inferential leap must be made from index term to actual content.<sup>193</sup> As noted, when directly relevant, both information-rich<sup>194</sup> and information-poor<sup>195</sup> terms merit a high interest.

Relevance must be measured not just by the logical relation of the term to actual page content, but also to two other factors: proximity and degree. First, meta tags should only be attached to the pages to which they are actually relevant.<sup>196</sup> Second, stuffing should almost never be

---

them. First, these terms are directly relevant to actual Web page content. Second, a consumer may string together several terms, such as "automobile" plus "repair" to find the site.

193. See MORGAN, BASIC PROBLEMS OF EVIDENCE 185–88 (1961); see also MCCORMICK, *supra* note 153, § 185 (noting that while direct evidence is offered toward the fact at issue, circumstantial evidence is offered to prove a proposition that itself requires an inferential leap toward ultimate proof of the fact at issue).

194. Take, for instance, this Article. Assuming that it was converted into a Web document, terms such as my name and the document title would certainly be directly relevant. So too would terms that are the specific subject of the Article: "spamdexing" and "meta tags." Names of persons, places, and things discussed here would also be relevant, though to a lesser degree. For example, I might have a higher interest in using "Oppedahl & Larson" as an indexing phrase than would the defendants in the *Advanced Concepts* lawsuit because that suit is directly addressed in this Article. See *supra* notes 122–30 and accompanying text.

Along these lines, a parodist put up a Web page using multiple instances of Oppedahl & Larson's name just for the purpose of stating that the page had nothing to do with them. See *This Page Has Nothing to Do with Carl Oppedahl or Oppedahl & Larson* (visited Nov. 30, 1998) <<http://www.geocities.com/CapitolHill/Lobby/6620/index.htm>>. Oppedahl and Larson, rather than joining the "infringing" webmaster as a defendant, mentioned this page on their Internet page detailing the progress of the *Advanced Concepts* suit. See Oppedahl & Larson, *Advanced Concepts Lawsuit* (visited Nov. 30, 1998) <<http://www.patents.com/ac/index.sht>>.

195. For instance, "law," "Internet," "deceptiveness," and "unfair competition" are directly relevant to the content of this Article. By themselves, these terms are not particularly descriptive. Still, the fact that they are not particularly useful indexing terms does not render them deceptive.

196. In some cases, the issue may concern whether the indexing term must be relevant to the particular *page* it is attached to, or to the *site* in general. A good rule of thumb is that if the indexing term is relevant only to a sub-page within the site, rather than to the home page itself, then the indexing term should only appear on the sub-page. For example, suppose Pepsi did a market study which compared Pepsi and Coke, and the results appeared on a sub-page, rather than the home page at pepsi.com. See Search Engine Watch, *More About Meta Tag Lawsuits* (subscriber-only page) (on file with *Harvard Journal of Law & Technology*). Supposing that it was otherwise appropriate for Pepsi to include "coke" as an indexing term on the sub-page, it would be less appropriate for it to do so on its home page. Certainly, the market study, existing on one page, would not give Pepsi the right to use "coke" as an indexing term on every page.

considered a valid use, because it is used to overstate a term's degree of relevance.<sup>197</sup>

One must be careful not to confuse the issue of relevance with effectiveness. The fact that a generic term (such as "law") may not be an *effective* form of indexing from the public's point of view does not determine whether it is relevant from the webmaster's point of view.<sup>198</sup> Also, just as the public may search in a relational fashion, so too may webmasters index with relational terms. However, such terms, by requiring a larger inferential leap, are more likely to be indirectly or remotely related to actual site content.<sup>199</sup> At some point, however, the connection between website and indexing term will eventually become so tenuous that the term is simply not relevant.<sup>200</sup>

### 3. Proprietary-Rights Holders

The interests of those holding proprietary rights in a term — generally trademark rights — may sometimes conflict with members of the public who want to use that term to discover "unauthorized" information relating to the trademark holder, or in ways unrelated to the

---

197. A term may appear *within* actual text many times. For instance, the term "meta tag" appears in this document well over one hundred times. In that light, the repetitions are indeed a valid way of measuring how relevant "meta tag" is to this Article. The same reasoning does not apply to any repetitions of the phrase "meta tag" *within meta tags themselves*. It may be argued that multiple repetitions are an appropriate proxy where the term is not used many times, but is nonetheless highly relevant to the page. This rule, however, would be unmanageable; the end result would pump up the volume of Internet noise, increasing the infoglut. See, e.g., Landgren.net, *supra* note 102 (webmaster who felt the need to stuff keywords relevant to himself so that people could find him amidst the spam).

198. Even though "law" by itself is not a particularly effective indexing term, using "law," "trademarks," "Internet," and "meta tags" together would be.

199. For instance, a *Star Trek: Deep Space Nine* ("ST:DS9") fan site might use the names of other *Star Trek* shows, such as "Star Trek: Voyager," to appeal to *Star Trek* fans. What if the site also used "Babylon 5" as an index term? After all, one who is interested in *Babylon 5* might also be interested in ST:DS9. Both shows are about space stations. Both include long-developing story lines involving a large cast of alien characters. Such a use may be indirectly relevant. Taking this one step further, suppose that the ST:DS9 site used "World Wrestling Federation" or "WWF" as an index term, assuming that people who like science fiction also like wrestling? This use is most probably remotely relevant, if relevant at all. The use of Oppedahl's name by Internet domain name registrars is an example of a remotely related indexing term. See *supra* notes 122–30 and accompanying text.

200. Certainly were the hypothetical ST:DS9 site, *supra* note 199, to use "Microsoft" or "Princess Diana," it would have long crossed the line demarcating relevance from irrelevance.



trademark. More commonly, proprietary interests will also conflict with those of webmasters who seek to use the trademark as an indexing term. Although trademark holders are not necessarily content providers on the Internet,<sup>201</sup> a trademark holder is often also a webmaster competing with other spamdexing webmasters.<sup>202</sup>

In either webmaster-conflict situation, the trademark owner often wants to prevent others from using its name. This is the competitive or exclusionary interest: the desire to exclude those who may divert one's customers.<sup>203</sup> In other cases, the trademark owner wants to control, but not completely exclude, others from using its name.<sup>204</sup> A sports league, for instance, might encourage the development of fanzines, fan sites, and the like, as long as the league is portrayed in a favorable light. This is the associational interest: the desire to determine who may use another's trademark.<sup>205</sup>

Note that unlike assessments of public and webmaster interests, determinations of a trademark holder's interests do not *necessarily* involve a relevance test.<sup>206</sup> A trademark holder wants to maximize its exclusive use as against others even when the "unauthorized" use is otherwise relevant — unless the third-party use is favorable.<sup>207</sup> Of course, trademark holders also want to maximize their chances of

---

201. Cf. *Stern v. Delphi Internet Servs. Corp.*, 626 N.Y.S.2d 694 (N.Y. Sup. Ct. 1995).

202. Paramount, as owner of the valuable *Star Trek* franchise, is such a party. See <<http://www.startrek.com>>. The World Wrestling Federation is another. See <<http://www.wwf.com>>; see also *supra* notes 112–45 and accompanying text (discussing spamdexing cases).

203. This interest is protected by trademark infringement and dilution, as described *infra* Part IV.B.

204. Paramount has conducted a not-altogether successful campaign to prevent fan sites from using its trademarks and copyrighted works on the Internet. Fan reaction has been swift and negative. See Erika S. Koster & Jim Shatz-Akin, *Set Phasers on Stun: Handling Internet Fan Sites*, COMPUTER LAWYER, Jan. 1998, at 18; see also Eric Blom, *Mainers Join Online War to Protect Star Trek Sites; Fans Protest Paramount's Campaign Against Internet Sites Based on the Popular Sci-Fi TV Show*, PORTLAND PRESS HERALD, Dec. 19, 1997, at 1A; *Cybertrekkies Get Warnings; 'Star Trek' Creators Say Web Sites Violate Copyrights*, BANGOR DAILY NEWS, Dec. 20, 1997, available in 1997 WL 16994037.

205. This broader interest overlaps with the exclusionary interest and is protected by both trademark infringement and dilution, discussed *infra* Part IV.B.

206. But see *infra* notes 360–63 and accompanying text (using relevance analysis to contrast trademark infringement and dilution).

207. Here, both dilution and infringement might apply, depending on whether the relevant use rises to a "likelihood of confusion." See 15 U.S.C. §§ 1114(1)(a), 1125(a), (c), 1127 (1994 & Supp. I 1995).

appearing in query output without undue interference from irrelevant spamdexing.<sup>208</sup>

### C. Domain Names, Linking, and Framing Compared

The challenge of balancing the competing interests of the public, webmasters, and holders of proprietary rights is underscored through a comparison of these interests to those in other contexts, namely disputes involving domain names, hyperlinking, and framing. All three have led to a number of suits by those asserting trademark or copyright protection,<sup>209</sup> and have engendered a prodigious amount of scholarly commentary.<sup>210</sup> Table Three displays the types of interests under each category of party and dispute.

Domain names, such as "yahoo.com," allow a trademark to serve as a unique Internet address.<sup>211</sup> This uniqueness allows people to "guesstimate" that an organization may be found by appending ".com" onto a famous trademark, such as "sony.com."<sup>212</sup> Correspondingly, domain name "squatters" have registered famous names of others in hopes of free riding or otherwise gaining financially.<sup>213</sup>

208. Thus, irrelevant spamdexing would seem to implicate only dilution situations, because situations that involve the "likelihood of confusion" standard of trademark infringement are also situations where the indexing term was used in a relevant, albeit infringing, fashion. *See id.*; *see also infra* notes 360–63 and accompanying text.

209. *See, e.g.,* Data Concepts, Inc. v. Digital Consulting, Inc., 150 F.3d 620 (6th Cir. 1998) (domain name conflict); Panavision Int'l, L.P. v. Toeppen, 141 F.3d 1316 (9th Cir. 1998) (same); Futuredontics Inc. v. Applied Anagramics Inc., 45 U.S.P.Q.2d 2005 (C.D. Cal.), *aff'd*, 152 F.3d 925 (9th Cir. 1998) (framing dispute); Ticketmaster Corp. v. Microsoft Corp., No. 97-3055 (C.D. Cal. filed Apr. 28, 1997) (hyperlinking dispute); Washington Post Co. v. TotalNews, Inc., No. 97 Civ. 1190 (S.D.N.Y. filed Feb. 20, 1997) (framing dispute).

210. *See, e.g.,* Dan L. Burk, *Trademarks Along the Infobahn: A First Look at the Emerging Law of Cybermarks*, 1 RICH J.L. & TECH. 1 (April 10, 1995) <<http://www.urich.edu/~jolt/v1i1/burk.html>>; Walter A. Effross, *Withdrawal of the Reference: Rights, Rules, and Remedies for Unwelcomed Web-Linking*, 49 S.C. L. REV. 651 (1998) (discussing hyperlinking and framing); Kuester & Nieves, *supra* note 21 (discussing hyperlinking and framing); Carl Oppedahl, *Remedies in Domain Name Lawsuits: How Is a Domain Name Like a Cow?*, 15 J. MARSHALL J. COMPUTER & INFO. L. 437 (1997); Ira S. Nathenson, Comment, *Showdown at the Domain Name Corral: Property Rights and Personal Jurisdiction Over Squatters, Poachers, and Other Parasites*, 58 U. PITT. L. REV. 911, 925–29 (1997).

211. *See* Burk, *supra* note 210, ¶¶ 29–39.

212. *See* Nathenson, *supra* note 210, at 920–21.

213. *See* Panavision, 141 F.3d at 1319; Nathenson, *supra* note 210, at 925–29 (discussing categories of disputes, including squatters).



Table Three: Interests Implicated by Other Forms of Internet Information Dissemination and Retrieval

	Public	Webmaster <sup>214</sup>	Proprietary-Rights Holder
Search engines	Relevant output	Being found	Exclusiveness, association
Domain names	Guesstimation, easy to remember	Guesstimation, easy to remember	Exclusiveness
Hyperlinks	Ease of navigation	Ease of navigation	Control over who links to you and where
Framing	Multiple screens	Maintain visibility when public leaves site	Exclusiveness, association

The conceptual difficulty underlying these suits is the fact that each domain name is unique, allowing one party exclusive use. In contrast, trademark law allows concurrent use of the same name by multiple parties where there is no likelihood of confusion.<sup>215</sup> On one hand, domain name squatting may significantly frustrate a trademark holder’s ability to conduct business over the Internet. On the other, when a domain name registrant has a legitimate reason — such as good faith concurrent use — to use the trademark in a domain name, it is much more reasonable to expect one concurrent user to assume the risk of loss of another party having the domain name.

Domain name disputes are unlike spamdexing in several ways. First, spamdexing allows concurrent use; indeed, the abuse of concurrent use is the root of the spamdexing injury. Second, the expectation of finding things, and being found, via an indexing term is often much lower with search engine indexing than with domain names. Third, a searcher can often “fix” a faulty search by making the query more specific. This cannot be done with a domain name. Sony either is or is not at “sony.com.” These differences suggest that trademark holders have a lower interest in spamdexing than in domain name disputes;

---

214. Webmasters include both spamdexers and those asserting proprietary rights. Both share similar interests as webmasters. Additional interests of the webmaster as holder of proprietary rights belong in the “proprietary rights” category.

215. See 15 U.S.C. §§ 1114(1)(a), 1125(a)(1)(A) (1994).

courts should therefore be more hesitant to grant broad proprietary rights in this context.

Hyperlinking presents another form of potential liability.<sup>216</sup> With hyperlinking, the public expects to be able to go quickly from one document or website to another. Webmasters provide links and hope for reciprocal links from others, so that the public may easily navigate within a site and beyond. Those asserting proprietary rights, however, may seek enjoinder of associational or competitive injuries. A party may assert an associational injury when one is linked-to from a site whose reputation may disparage them.<sup>217</sup> Competitive injuries may occur when webmasters, rather than linking to a third party's home page, link deep within the third party's site,<sup>218</sup> or to a subsection of a page or to a particular image file.<sup>219</sup>

Hyperlinking cases are problematic. The World Wide Web was designed for linking.<sup>220</sup> The interest of both the public and webmaster in using hyperlinks is quite high. With spamdexing, however, the public's expectation of finding things is much lower: people expect multiple hits in a query. The interest in using hyperlinks, on the other hand, is quite high for both the public and webmasters.

The injury alleged in hyperlinking cases is either disparagement, or the loss of traffic that might have been gained by making the public go

---

216. See Kuester & Nieves, *supra* note 21, at 261–70; Martin H. Samson, *Hyperlink at Your Own Risk*, N.Y.L.J., June 24, 1997, at 1. For resources on hyperlinking and framing, see Stefan Bechtold, *The Link Controversy Page* (updated Nov. 18, 1997) <<http://www.jura.uni-tuebingen.de/~s-bes1/lcp.html>>; *Internet Copyright: Hyperlinks*, L.J. EXTRA (visited Nov. 20, 1998) <[http://www.ljx.com/internet/ir\\_copy.html#links](http://www.ljx.com/internet/ir_copy.html#links)> (linking controversies).

217. Cf. *Hasbro, Inc. v. Internet Entertainment Group, Ltd.*, 40 U.S.P.Q.2d (BNA) 1479 (W.D. Wash 1996) (enjoining adult website from continued use of the domain name “candyland.com,” where “Candyland” was also name of famous children’s board game).

218. In *Shetland Times Ltd. v. Wills*, Lord Hamilton disapproved of deep linking, believing that access to material at a site should be obtained only through a link to the site’s home (main) page. See *Shetland Times Ltd. v. Wills*, 24 F.S.R. 614 (Sess. 1996); see also Kuester & Nieves, *supra* note 21, at 262–63.

219. See Kara Beal, Comment, *The Potential Liability of Linking on The Internet: An Examination of Possible Legal Solutions*, 1998 B.Y.U.L. REV. 703, 716 (discussing dispute arising when a third party created a Web page that used hyperlinks to import the comic strip *Dilbert* into his home page); Oppedahl & Larson, *Web Law FAQ* (last modified June 18, 1998) <<http://www.patents.com/weblaw.sht>>.

220. One court has arguably held that there is a First Amendment right to hyperlink. See *ACLU v. Miller*, 977 F. Supp. 1228, 1232–33 & n.5 (N.D. Ga. 1997); see also Jeffrey R. Kuester & Peter A. Nieves, *Hyperlinks: A Form of Protected Expression?*, L.J. EXTRA (Jan. 26, 1998) <<http://www.ljx.com/internet/0126hyperlink.html>>.



through the plaintiff's home page. Providing a remedy for this "injury," however, could drastically rearrange the Internet landscape, forcing content providers to adopt centralized licensing arrangements, and vastly increase the costs of disseminating information. Although those asserting proprietary rights may claim some of these practices are unfair, the unfair competition in hyperlinking will probably have to be quite egregious to merit liability. Irrelevant spamdexing, on the other hand, is always unfair, and may often lead to liability. Arguably relevant spamdexing, of course, presents a more difficult case.

Framing — the bordering of one website around another — is more clearly a practice smacking of unfair competition, because of the increased possibility of false affiliation.<sup>221</sup> Further, there is little public interest in preferring framed sites over non-framed sites, because the frame unnecessarily takes up valuable real estate on the user's computer screen. Indeed, the only party to find frames attractive is the framing webmaster, who wants to sell advertising and maintain visibility even when the public goes to another site. In contrast, in spamdexing disputes, the public has a high interest in using targeted search terms, as do webmasters in using terms directly relevant to actual content. In both framing and spamdexing cases, however, those asserting proprietary rights can argue credibly that such practices are unfair.

#### *D. Externalization Costs*

As the discussion above suggests, courts faced with close cases must broach the difficult question of when a practice is sufficiently "unfair" or "deceptive" to merit liability. The more an activity appears to be geared toward profit rather than information dissemination, the more likely that courts will find for the plaintiff. Economic analysis is instructive. Absent incentive or punishment, parties will often try to externalize their costs.<sup>222</sup> It is more economically efficient to place the risk of loss on the "least cost avoider," who has superior knowledge, and can be encouraged — by carrot or stick — not to externalize its costs.<sup>223</sup> For example, advertisers are normally expected to internalize the costs of building goodwill and public awareness. Merely shifting costs back to businesses, however, does not necessarily reduce overall costs,

---

221. See Kuester & Nieves, *supra* note 21, at 271–73; Beal, *supra* note 219, at 717.

222. See RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* § 8.4, at 285 (5th ed. 1998) (suggesting that the common law represents an effort to "attach costs to the violation of moral principles . . . to enhance the efficiency of a market . . . economy").

223. See Kenneth Sutherland Dueker, Note, *Trademark Law Lost in Cyberspace: Trademark Protection for Internet Addresses*, 9 HARV. J.L. & TECH. 483, 509 (1996).

because businesses will try to shift costs back to consumers via higher prices or lower quality.<sup>224</sup> However, truthful advertising, as an internalized cost of the advertiser, can be efficient, because “[a]dvertising contributes to consumer welfare by providing information, which lowers consumer search costs, which in turn facilitates competition.”<sup>225</sup>

Spamdexing hurts everyone. Spamdexers try to reduce their costs of advertising by taking what are, from their perspective, inexpensive steps to increase hits. In other words, they seek to internalize little and externalize much. The public, however, is forced to internalize the higher costs of more cumbersome searching. Although spamdexing might not lead to direct pecuniary losses, its introduction of information noise frustrates searching,<sup>226</sup> which inflicts indirect pecuniary losses. By making individuals spend more time to find information, spamdexing also increases searcher’s opportunity costs, and harms holders of proprietary rights by diminishing the goodwill that they have accumulated. In some cases, spamdexing may force owners of trademarks and trade names to internalize the additional costs of defending their good names; in others, the aggrieved will suffer irretrievable losses of goodwill. Even worse, spamdexing is economically short-sighted. By increasing the amount of info-noise, spamdexers ultimately contribute to making it more difficult for anybody — including themselves — to be found, thus indirectly re-internalizing their own and others’ costs of being found.

In conventional advertising situations, consumers who are deceived can retaliate by withholding future purchases, or by sharing their distrust of the advertiser with others.<sup>227</sup> Even assuming that the consumer has some ability to self-protect, the deterrent effect of retaliation is premised on the possibility of repeat visits by a consumer.<sup>228</sup> With spamdexing, though, a similar deterrent effect is absent. First, if the public is not interested in a spamdexed site, it might not even visit the site that first time. Second, if members of the public do like the site, they will return based on the actual, *truthful* website content, disregarding the *deceptive*

---

224. See Gregory E. Maggs, *Internet Solutions to Consumer Protection Problems*, 49 S.C. L. REV. 887, 889 (1998).

225. Lillian R. BeVier, *Competitor Suits for False Advertising Under Section 43(a) of the Lanham Act: A Puzzle in the Law of Deception*, 78 VA. L. REV. 1, 8 (1992).

226. See *supra* Part III.A.

227. See BeVier, *supra* note 225, at 11.

228. See Schechter, *supra* note 191, at 66–67 (arguing that Professor BeVier overstates the consumer’s ability to self-protect through “trial and error” and “consumer skepticism”).



spamdexing that brought them there in the first place. Because consumer action will have little deterrent effect, it is necessary for the state, through liability, to impose disincentives external to the consumer-spamdexer transaction.

By creating a real risk of liability, trade regulators and proprietary-rights holders can force webmasters to internalize the social costs of spamdexing by using alternative forms of information dissemination.<sup>229</sup> This may encourage those webmasters who have a good faith commitment to providing content to reach their audience by legitimate alternative means.<sup>230</sup>

The specter of liability, when appropriate, will force webmasters to be circumspect about how they use meta tags. Finding a balance, however, will be difficult when the index term has some good faith relevance to actual content. Even though many people will not find the meta tag relevant in the way intended by the webmaster, others may find such indexing to be socially beneficial.<sup>231</sup> Too much liability could also encourage webmasters to take too much care, chilling valuable commercial discourse (the underinclusiveness problem).<sup>232</sup> Too little liability, though, will have the opposite effect (the overinclusiveness problem).<sup>233</sup>

Economic analysis also informs the weighing of the interests of the public and proprietary-rights holders. The public must internalize some of the costs of searching by engaging in reasonable searching habits: using information-rich terms, Boolean operators, and alternative search

---

229. See Roger D. Blair & Thomas F. Cotter, *An Economic Analysis of Damages Rules in Intellectual Property Law*, 39 WM. & MARY L. REV. 1585, 1644 (1998).

The infringer will internalize these costs only if, in addition to being required to disgorge any profits attributable to the infringement, he also is liable for any actual damages resulting from injury to the owner's reputation and from the deception of consumers. In theory, this latter interest could be vindicated either by a direct action on the part of those consumers or by allowing the trademark owner to recover enhanced damages.

*Id.*; see also Paul Heald, Comment, *Money Damages and Corrective Advertising: An Economic Analysis*, 55 U. CHI. L. REV. 629, 643 (1988).

230. See *infra* notes 472–78 and accompanying text.

231. See Fred S. McChesney, *Deception, Trademark Infringement, and the Lanham Act: A Property-Rights Reconciliation*, 78 VA. L. REV. 49, 53 (1992). “[A]dvertising imparts benefits (truthful information) to some, although possibly imposing costs (false information) on others. The correct analysis of false advertising is therefore the standard evaluation of externalities — activities that generate benefits for some while imposing costs on others, but on net are socially beneficial.” *Id.*

232. See Heald, *supra* note 229, at 643.

233. See *id.*

tools when appropriate.<sup>234</sup> These steps are simple to take, inexpensive, and well within the control of the public. Similarly, proprietary-rights holders cannot expect to control every use by webmasters of their trademarks and trade names. If the mere use of a trademark were enough for liability, then webmasters would have to go to ridiculous lengths to describe themselves or others. Imagine trying to write about Coca-Cola without using the term. As the discussion of fair use doctrine will show,<sup>235</sup> merely descriptive or nominative uses of trademarks cannot be enjoined by proprietary-rights holders.

This discussion of law and economics, however, provides only a starting point. It begs the question to state that an “efficient” level of liability would be the best; it is first necessary to address where the costs and liabilities should lie in determining that efficiency. Part IV, which addresses liability under consumer protection and federal unfair competition law, seeks to do just that.

#### IV. STRIKING A BALANCE

*Overprotecting intellectual property is as harmful as underprotecting it. Creativity is impossible without a rich public domain. Nothing today, likely nothing since we tamed fire, is genuinely new: Culture, like science and technology, grows by accretion, each new creator building on the works of those who came before. Overprotection stifles the very creative forces it's supposed to nurture.*

— Judge Alex Kozinski<sup>236</sup>

Spamdexing may lead to liability on two broad fronts: public deception and unfair competition.<sup>237</sup> Public deception implicates consumer protection statutes, which may be invoked by a variety of parties.<sup>238</sup> The diffuse nature of consumer harm makes it unlikely that specific consumers will be motivated to cry for help; it is therefore appropriate for regulators to act on the public's behalf.

---

234. See *infra* notes 446–63 and accompanying text.

235. See *infra* Part IV.B.5.

236. *White v. Samsung Electronics America, Inc.*, 989 F.2d 1512, 1513 (9th Cir. 1993) (en banc) (Kozinski, J., dissenting from denial of rehearing en banc).

237. As one commentator notes, although spamdexing is sometimes referred to as “invisible trademark infringement,” it is more accurate to refer to it as “a deceptive trade practice or unfair competition.” Elgison & Jordan, *supra* note 15, at C7. Here, I will address spamdexing from the points of view of both doctrines.

238. See *infra* Part IV.A.



Unfair competition doctrine — specifically trademark law — seeks to prevent one party from unfairly free riding off the goodwill of another. Under unfair competition law, it is necessary to balance the interests of plaintiff, defendant, and the public.<sup>239</sup> Whereas trademark infringement is conceptually inapposite to spamdexing, trademark dilution provides a better-suited remedy, at least for famous names. However, overly broad liability must be avoided to keep from chilling the fair use, where applicable, of relevant indexing terms.

*A. Deceptiveness: “Protecting” the Public*

*The tension between avoiding consumer deception while maximizing the flow of truthful speech is apparent. [One] must decide whether the total information environment would be improved by the elimination of particular misleading claims, or whether consumers would be better off with some slightly misleading information about an attribute rather than no information at all.*

— *Federal Trade Commission*<sup>240</sup>

1. Section 5 of the Federal Trade Commission Act

Government regulators and others may challenge spamdexers under consumer protection statutes. Although the Federal Trade Commission (“FTC”), originally established in 1914, was not initially created to deal with false advertising claims, this omission was corrected in 1938 with the addition of section 5 of the Federal Trade Commission Act (“FTC Act”), which prohibits “[u]nfair methods of competition . . . and unfair or deceptive acts or practices in or affecting commerce.”<sup>241</sup> Private parties do not have standing to assert an action under the FTC Act.<sup>242</sup> Many states, however, have enacted variations of the Unfair Trade Practices and Consumer Protection Act, a model act developed by the FTC in 1967.<sup>243</sup> In the vast majority of states, a private cause of action

---

239. *See infra* Part IV.B.

240. J. HOWARD BEALES & TIMOTHY J. MURIS, STATE AND FEDERAL REGULATION OF NATIONAL ADVERTISING 34 (1993) (quoting FEDERAL TRADE COMMISSION, CONSUMER INFORMATION REMEDIES 283–84 (1979)).

241. 15 U.S.C. § 45(a)(1) (1994); *see also* EARL W. KINTNER, A PRIMER OF THE LAW OF DECEPTIVE PRACTICES: A GUIDE FOR THE BUSINESSMAN 15–16 (1971).

242. *See* 15 U.S.C. § 45(b) (1994); *Holloway v. Bristol-Myers Corp.*, 485 F.2d 986, 1002 (D.C. Cir. 1973) (no implied private cause of action under FTC Act); PETER C. WARD, FEDERAL TRADE COMMISSION: LAW, PRACTICE AND PROCEDURE § 3.04 (1997).

243. COUNCIL OF STATE GOV'TS, 1970 SUGGESTED STATE LEGISLATION 141 (1969); *see also* RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 1, stat. note (1995). Other

may be maintained.<sup>244</sup> Many of these states use the language of section 5 of the FTC Act in these statutes; it is therefore important to consider the FTC Act even in construing the "little FTC Acts."<sup>245</sup>

Acts or practices are "unfair" under section 5 if they "cause[] or [are] likely to cause substantial injury to consumers which [are] not reasonably avoidable by consumers themselves and not outweighed by countervailing benefits to consumers or to competition."<sup>246</sup> Such an act or practice is violative if it "(1) has a tendency or capacity to mislead (2) a substantial number of consumers (3) in a material way."<sup>247</sup> If these criteria are met, the FTC has three options: sending cease and desist orders; filing suit in federal district court for injunctive relief, redress for consumers, or restitution of unfairly received gains; or seeking assistance from the Department of Justice to file contempt or criminal proceedings against those violating court orders.<sup>248</sup>

Persons using search engines should be considered "consumers" for purposes of the Act, because they will make purchases based on their actions on the Internet, or will face the frustration of spamdexing in their attempts to purchase. Further, consumers using a search engine will be exposed to advertising via irrelevant query output. Courts and the FTC have often stated that "deceptiveness under section 5 is to be determined by focusing on the unsophisticated and gullible."<sup>249</sup> Thus, the fact that many Internet surfers may be experienced, even savvy, in dealing with

uniform acts include the Uniform Deceptive Trade Practice Act and the Uniform Consumer Sales Practices Act. See Robert E. Reyna, *State Little FTC Acts and Unfair Methods of Competition*, ALI-ABA COURSE OF STUDY, Mar. 6, 1997, at 47, 51-54.

244. See RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 1, stat. note (1995); see also MICHAEL M. GREENFIELD, CONSUMER LAW: A GUIDE FOR THOSE WHO REPRESENT SELLERS, LENDERS, AND CONSUMERS § 4.4.1 (1995) ("A private damages remedy exists in almost every state.").

245. RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 1, stat. note (1995); A. Michael Ferrill, *Federal Law of Unfair Competition*, in BUSINESS TORTS & UNFAIR COMPETITION: A PRACTITIONER'S HANDBOOK 41, 70-71 (A. Michael Ferrill ed., 1996). For brevity's sake, this Section considers spamdexing solely from the perspective of the FTC Act.

246. 15 U.S.C. § 45(n) (1994).

247. WARD, *supra* note 242, § 6.03[1].

248. See *Federal Trade Commission Responses to Questions Regarding Electronic Commerce*, The Honorable Tom Bliley, Chairman, U.S. House Committee on Commerce (visited Oct. 6, 1998) <<http://www.ftc.gov/os/1998/9804/blileyt.htm>> [hereinafter *FTC Responses*]; see also 15 U.S.C. § 45(b) (1994) (authorizing cease and desist orders); *id.* § 53 (authorizing injunctive relief); *id.* § 54(a) (authorizing civil penalties); *id.* § 56 (authorizing the Commissioner to request the assistance of the Attorney General); *id.* § 57b(b) (authorizing consumer redress).

249. GREENFIELD, *supra* note 244, § 3.3.1.



spamdexing, does not preclude liability. Indeed, the “fact that a false statement may be obviously false to those who are trained and experienced does not change its character, nor take away its power to deceive others less experienced.”<sup>250</sup>

The FTC Act also requires that the deception be material, i.e., likely to affect a consumer’s decision.<sup>251</sup> Materiality is measured by a reasonable person standard.<sup>252</sup> There need not be injury to a specific consumer; instead, it is sufficient that the act or practice “*interfere[s] with the consumer’s exercise of choice* in determining whether to enter a transaction and, if so, with whom and on what terms.”<sup>253</sup> Spamdexing, by giving consumers irrelevant output, or output that is so diluted that it becomes useless, will often interfere with a consumer’s ability to choose with whom to transact. The value or legitimacy of the spamdexer’s product should be immaterial: “[I]t is no defense that the seller’s product is worth the price paid for it.”<sup>254</sup>

What is “deceptive” for purposes of the FTC Act? Traditionally, there need not be proof of actual confusion; the mere “capacity to deceive” is sufficient for a section 5 violation.<sup>255</sup> This “capacity to deceive” standard will be met if “the act or practice [tends] to deceive the average consumer or [has] the capacity to deceive a substantial number of consumers.”<sup>256</sup> In 1983, however, the FTC reformulated the “deception” and “consumer” elements, requiring that the act be likely (rather than have a mere capacity) to mislead consumers who act reasonably under the circumstances (rather than presuming gullibility of consumers).<sup>257</sup> As redefined, a deceptive communication is one “that is

---

250. *FTC v. Standard Educ. Soc.*, 302 U.S. 112, 116 (1937). It should also be noted that competitors and those with proprietary interests at stake in the indexing terms may be able to state a claim under consumer protection acts. *See* RESTATEMENT (THIRD) UNFAIR COMPETITION § 1 cmt. d. Thus, in many jurisdictions, whatever protections apply to consumers also apply to those with proprietary interests. For present purposes, this Article will focus on the consumer, addressing proprietary rights in the next Section.

251. *See* GREENFIELD, *supra* note 244, § 3.3.1.

252. *See id.*

253. *Id.* (emphasis added).

254. *Id.*

255. *Charles of the Ritz Distrib. Corp. v. FTC*, 143 F.2d 676, 680 (2d Cir. 1944); *see also* GREENFIELD, *supra* note 244, § 3.3.1.

256. GREENFIELD, *supra* note 244, § 3.3.1.

257. *See id.* (citing Policy Statement on Deception, in Letter to John D. Dingell, Chairman, Subcommittee on Oversight and Investigations, Committee on Energy and Commerce, Oct. 14, 1983, *reprinted in In re Cliffdale Assocs., Inc.*, 103 F.T.C. 110, app. at 174 (1984)).

likely to mislead consumers acting reasonably under the circumstances and is also material.”<sup>258</sup>

“It is unclear whether the new articulation changes the standard of deception under section 5 [of the FTC Act].”<sup>259</sup> Some believe so,<sup>260</sup> whereas others are doubtful.<sup>261</sup> Without concluding which approach better protects consumers in general, it is prudent to limit liability in the spamdexing context to acts that harm reasonable (rather than merely gullible) consumers. For one thing, it is important for society to develop a neterate sensibility. More importantly, the deceptiveness of spamdexing is fairly obvious to even a non-neterate consumer. Protecting even the gullible has its merits in encouraging people to interact on the Internet without fear (and thus encouraging the growth of Internet commerce), but in the spamdexing context, such a low threshold would be unnecessary. Further, due to the broad standard of relevance articulated in Part III, there must be substantial imbalance between the consumer and webmaster interest for liability to attach.

Deception may arise under the FTC Act in two forms: false advertising and bait and switch tactics. Although false advertising is a well-established type of deceptive conduct, it is not conceptually well-suited to spamdexing cases.<sup>262</sup> False advertising is “an advertisement,

---

258. Ivan L. Preston, *The Definition of Deceptiveness in Advertising and Other Commercial Speech*, 39 CATH. U.L. REV. 1035, 1037 (1990) (citing *Cliffdale*, 103 F.T.C. app. at 174–84); see also Richard Craswell, *Interpreting Deceptive Advertising*, 65 B.U. L. REV. 657 (1985) [hereinafter Craswell, *Interpreting Deceptive Advertising*]; Richard Craswell, *Regulating Deceptive Advertising: The Role of Cost-Benefit Analysis*, 64 S. CAL. L. REV. 550 (1991).

259. GREENFIELD, *supra* note 244, § 3.3.1.

260. See Jack E. Karns & Alan C. Roline, *The Federal Trade Commission's Deception Policy in the Next Millennium: Evaluating the Subjective Impact of Cliffdale Associates*, 74 N.D. L. REV. 441, 443 (1998); see also *Southwest Sunsites, Inc. v. FTC*, 785 F.2d 1431, 1435 (9th Cir. 1986).

261. See GREENFIELD, *supra* note 244, § 3.3.1; see also Jack E. Karns, *The Federal Trade Commission's Evolving Deception Policy*, 22 U. RICH. L. REV. 399, 409 (1988) (noting that FTC Chairman Miller saw the new articulation as “merely an accurate articulation of the standard” already established by FTC case law); Candace Lance Oxendale, Comment, *The FTC and Deceptive Trade Practices: A Reasonable Standard?*, 35 EMORY L.J. 683, 684 (1986).

262. See Arthur Best, *Monetary Damages for False Advertising*, 49 U. PITT. L. REV. 1, 38 (1987) (“There is little dispute that deceptive and false advertisements harm consumers and competitors.”) (footnote omitted); John T. Cross, *Language and the Law: The Special Role of Trademarks, Trade Names, and Other Trade Emblems*, 76 NEB. L. REV. 95, 132 (1997) (“False advertising laws arguably are more effective in dealing with [certain types] of deception, especially considering that false advertising laws focus on the message instead of the privileges of the messenger.”). Professor Craswell suggests that advertising should be considered deceptive when “the advertiser fail[s] to take some



other than labeling, which is misleading in a material respect.”<sup>263</sup> Yet spamdexing misleads search engines, not consumers. The placement of irrelevant indexing terms on a page is tantamount to intentionally placing ads in the wrong section of the yellow pages. Assuming the correct standard requires a likelihood (and not a mere capacity) of misleading reasonable (and not merely gullible) consumers, it is doubtful that reasonable people will not recognize the irrelevance of spamdexed query output, making it hard to say that liability should attach. Further, the reasonableness standard is appropriate to the spamdexing context. After all, anybody can tell that a pet shop ad does not belong next to ads for plumbers. Further, the public can and should be expected to search reasonably. When using search engines, people should combine keywords to filter out noise and maximize precision. Thus, false advertising under the FTC Act will generally be conceptually inapposite to spamdexing.

The bait and switch, “an attractive but insincere offer,”<sup>264</sup> provides a better conceptual fit. It requires an alluring “bait” that attracts consumers, and a “switch” to another product that may or may not be more expensive.<sup>265</sup> Spamdexing does not pressure the consumer into buying a different, more expensive product. However, the consumer may be diverted to view or to buy something different, or be frustrated, confused, or even prevented from finding wanted information. “The problem with bait and switch is not the substance of the *ultimate* transaction, it is the *deception used to get there*.”<sup>266</sup> Here, assuming a consumer is using a reasonable search method, bait and switch remedies should thus be available under either the traditional or heightened

precaution that would have reduced the net injury caused by the ad.” Craswell, *Interpreting Deceptive Advertising*, *supra* note 258, at 660.

263. 15 U.S.C. § 55(a)(1) (1994).

264. GREENFIELD, *supra* note 244, § 3.3.2.4; *see also* *In re Tashof*, 74 F.T.C. 1361, 1388 (1968), *aff’d*, 437 F.2d 707 (D.C. Cir. 1970).

265. *See* GREENFIELD, *supra* note 244, § 3.3.2.4.

266. *Id.* (emphasis added). Greenfield notes that bait and switch for the mere purpose of *diverting* consumers to a destination is enough:

Even milder forms of bait and switch may be deceptive. One function of advertising is to attract the consumer to the seller’s place of business. Many sellers hope that once the consumer is in the store, he or she will purchase items in addition to the advertised one. There is nothing deceptive about this, *so long as the advertised item really is available*. It is deceptive, however, if the seller does not have the advertised item at all.

*Id.* (emphasis added). Thus, if a consumer is diverted to a place where such items, or information, is not available at all, or not in the form he reasonably expects, then the spamdexer herself is a perpetrator of bait and switch.

standard. Interestingly, the present treatment of false advertising and bait and switch parallels this Article's rejection of trademark infringement and its acceptance of trademark dilution as appropriate causes of action.<sup>267</sup>

## 2. How Much Deception Is Sufficient?

Assuming that liability might attach under the FTC Act, how much deception is enough? The easy case is where the indexing terms are completely irrelevant. Cases involving indirectly or remotely related keywords are more difficult.<sup>268</sup> Recall that the same term may be relevant to consumers and webmasters in different ways.<sup>269</sup> Thus, it is generally not helpful to look at how a particular consumer uses an indexing term. One may use it in a targeted sense, and another may use it in a relational manner. It is difficult indeed to balance interests when each consumer may have a different meaning in mind. Therefore, liability should attach only where there is substantial imbalance in the interests of webmasters and the public. Just as the relevance standard described in this Article is evocative of Federal Rule of Evidence 401, courts should balance the competing interests in a standard patterned after Rule 403:

Although relevant, indexing terms may be deceptive if their relevance is *substantially outweighed* by the

---

267. See *infra* Part IV.B.

268. "Virtually everyone agrees that deceptive advertising is bad. Few, however, agree about how best to tell whether an advertisement is deceptive." Craswell, *Interpreting Deceptive Advertising*, *supra* note 258, at 658.

269. See *supra* notes 186–87, 191 and accompanying text. The spamdexer has a defensible interest in using directly or indirectly relevant, if nonetheless ineffective, indexing terms, as suggested by Table Two (columns one and two). Suppose, for example, that a webmaster has a site about first aid, and uses "poison" as an indexing term. Such use is directly relevant to actual site content. A consumer looking for information on the defunct 1980s heavy metal band "Poison" would be disappointed by a query that included such a first aid site, but could not state a claim. Here, *both* consumer and webmaster interests are high, as represented by cell A2 of Table Two.

Also notable are situations where the consumer may have a high interest, but the webmaster has a medium-to-low interest. For example, take the *Advanced Concepts* case. See *supra* notes 122–30 and accompanying text. A consumer wanting to find Oppedahl & Larson's Internet site would have a high interest in the narrow target search of using "Oppedahl & Larson" or "Carl Oppedahl" as query terms. Parties like defendants have a medium-to-low interest in using the same terms to stand for "Internet-savvy lawyer," with the goal of getting hits from parties that might use the same terms in a relational context.



danger of unfair searching prejudice, confusion of consumers, or misleading the public, or by considerations of undue delay, waste of time, or needless presentation of information noise.<sup>270</sup>

There should be substantial, rather than a mere, imbalance before liability attaches, for at least three reasons. First, consumers can often avoid injury by acting reasonably, as required by the plain language of the FTC Act, which requires “substantial injury” that is not “reasonably avoidable by consumers.”<sup>271</sup> When a webmaster uses a term in an irrelevant manner, it is easy to conclude that the injury is substantial. When it is used in a way even remotely relevant, though, it will be difficult to determine whether liability is appropriate. For example, when a term is information-poor, it is within the control of consumers to narrow their searches to reduce the number of irrelevant hits. Thus, consumers can often act to “reasonably avoid” a “substantial injury.” They could search by using a phrase<sup>272</sup> or by using Boolean connectors to increase the odds of receiving a more precise list of hits.<sup>273</sup>

Reasonable searchers should list the words and phrases they would expect to find in their target document.<sup>274</sup> A consumer who uses a sole conceptual term, or even a targeted term that has alternative conceptual constructions, cannot reasonably expect such a search to act as an effective filter of a search engine’s database.<sup>275</sup> Such searches will result in extremely low precision. Further, although some consumers may be

---

270. The actual text of Rule 403 states: “Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.” FED. R. EVID. 403.

271. 15 U.S.C. § 45(n) (1994).

272. HotBot allows phrase searching. See HotBot <<http://www.hotbot.com>>.

273. See CAROL H. FENICHEL & THOMAS H. HOGAN, *ONLINE SEARCHING: A PRIMER* 48–50 (1981); GILSTER, *supra* note 48, at 167; GOLDMAN, *supra* note 159, at 99–114 (discussing Boolean logic, truncation, and other methods). Consumers should become familiar with their preferred search engines, because some default to an “OR” connector and others to an “AND” connector. See GILSTER, *supra* note 48, at 167; see also *id.* at 172 (suggesting using adjacent terms, excluding terms, capitalizing proper names, and using quotation marks to specify phrases).

274. See GILSTER, *supra* note 48, at 168. For example, “[a] search for Web pages specializing in on-line education . . . might include words like *syllabus*, *class*, *faculty*, *lecture*, and, of course, *education*.” *Id.*

275. See *id.*

annoyed by uses of an index term irrelevant to what they seek, others will appreciate the term's "countervailing benefits."<sup>276</sup>

On the other hand, if a webmaster uses terms that are remotely related but not reasonably necessary, and uses so many that consumers constructing queries cannot "reasonably avoid" the site, then liability may be appropriate in some cases. Put differently, it may be necessary to examine a webmaster's good faith in cases involving remotely related terms, whereas in cases involving irrelevant terms, bad faith may be presumed.

Second, because alternative forms of information retrieval may be better suited to categorical and relational searches, more than a mere imbalance should be shown prior to attaching liability. As is developed more fully in Part V, search engines are better-suited to targeted searches, which can incorporate a number of information-rich terms in Boolean unison.<sup>277</sup> In search scenarios involving conceptual terms, it is often unreasonable for consumers to rely on search engines; instead, directories such as Yahoo! better serve searching needs, and consumers should not cry foul when they fail to use the better tool.

Third, liability for a mere imbalance of interests may have chilling effects on the free-speech rights of content providers.<sup>278</sup> Thus, as a practical matter, a balance substantial enough to merit trade-regulation liability will exist only in cases of irrelevant spamdexing. In cases where the relation between meta tag and actual content is remote rather than nonexistent, the FTC should not act in the absence of affirmative indicia of extreme bad faith.<sup>279</sup>

---

276. 15 U.S.C. § 45(n) (1994). Supposing that a site about pen pals uses the term "friends" as a meta tag, a consumer seeking information on the NBC sitcom *Friends* would be disappointed, but a person looking for information on pen pals would be pleased. See *supra* notes 186–87 and accompanying text.

277. See *infra* notes 442–58 and accompanying text.

278. "[T]he law of false advertising should discourage some deceptive advertising, but not deceptions which are the innocent by-products of conduct which is otherwise desirable." Jeff Sovern, *Private Actions Under the Deceptive Trade Practices Acts: Reconsidering the FTC Act as Rule Model*, 52 OHIO ST. L.J. 437, 457 (1991).

First Amendment concerns will be briefly noted in Part IV.B.5. See *infra* notes 406–10 and accompanying text. However, this issue as a whole is beyond the scope of this Article and will not be addressed in detail.

279. For instance, consider the CityAuction page, *supra* note 5, that "welcomes" Monica Lewinsky. That site is apparently designed with actual content geared pretextually toward "legitimately" using Ms. Lewinsky's name to gain hits from search engines. The same may be said of the uses made of "Oppedahl" and "Larson" in the *Advanced Concepts* suit. See *supra* notes 122–30 and accompanying text.



### 3. FTC Action Is Needed

As noted above, consumers do not have standing under the FTC Act, although they do under many of the mini-FTC Acts.<sup>280</sup> Private parties, though, will have little incentive to file suit under the state statutes that grant consumers standing. As noted above in the discussion of bait and switch, the nature of the harm is diversion, frustration, and confusion. Spamdexing, a diffuse injury, does not generally involve the kind of direct pecuniary loss that would prompt a lawsuit by an aggrieved consumer. Without a specific aggrieved victim, it is doubtful that there will be a cry for action from either the public or from someone claiming to act on its behalf.<sup>281</sup> If the FTC does not act, protection will be limited to proprietary-rights holders acting in their own interest.

We should not sit back idly and let the public quietly internalize the info-noise of spamdexing. The FTC should therefore send a strong message that spamdexing is wrong.<sup>282</sup> Initially, the FTC may be hesitant to act. Although it has actively fought false advertising and fraud on the Internet,<sup>283</sup> it might not be strongly motivated to assert actions in the

---

280. See *supra* notes 242–45 and accompanying text.

281. A good example of a context in which there are no intellectual property rights, and where consumers have a real, but diffuse, interest, is that of e-mail spamming. The FTC has shown an interest in acting against e-mail spammers. See *EduPage, Spammers Warned by FTC* (Feb. 8, 1998), available at <<http://xpression.net/currentevents/edupage8february1998.txt>> (“More than 1,000 ‘spammers,’ many of whom are suspected of being involved in fraudulent schemes, will receive letters from the Federal Trade Commission’s Bureau of Consumer Protection warning them to clean up their act.”); see also *CompuServe, Inc. v. Cyber Promotions, Inc.*, 962 F. Supp. 1015 (S.D. Ohio 1997); *Cyber Promotions, Inc. v. America Online, Inc.*, 948 F. Supp. 456 (E.D. Pa. 1996); *Cyber Promotions, Inc. v. American Online, Inc.*, 948 F. Supp. 436 (E.D. Pa. 1996). E-mail spamming presents the best example of the public interest reigning supreme, even in the absence of a so-called “property” interest. This interest, however, has been protected in large part because it is coterminous with the interests of ISPs. E-mail spamming floods ISPs and their customers with millions of unsolicited junk messages every day. Of course, in these cases, spam is not free: the costs are borne by ISPs, who would like to pass the cost on to users, but cannot due to current flat-pricing trends for Internet service. This scenario suggests that third parties will consider non-proprietary interests only when there is either: (1) a public outcry loud enough to lead to a political response; or (2) a property owner who claims to be speaking on behalf of the public interest.

282. Cf. SHENK, *DATA SMOG*, *supra* note 35, at 209–10 (arguing that consumers are distracted by the data smog because they are more vulnerable to deceptive advertising, and recommending that the FTC act in the public interest); Shenk, *supra* note 2 (“The [FTC] can . . . be an important player in limiting data smog. . . . We need a rejuvenated FTC that criticizes questionable marketing practices and imposes fines.”).

283. See *FTC, State Securities Regulators Target Business Opportunity Fraud on the*

realm of spamdexing. First, the FTC and its state counterparts have limited financial and human resources.<sup>284</sup> Second, there may be literally millions of spamdexing Web pages, making wide enforcement difficult.<sup>285</sup> Many spamdexers are essentially anonymous, and their pages come and go quickly.<sup>286</sup> Third, spamdexing injuries by themselves are often diffuse, particularly where no proprietary right is implicated.<sup>287</sup> Going after one party will not cure the problem; each spamdexing act by itself still causes a separate, marginal increase in noise. Therefore, unlike competitive injuries, only in the aggregate does spamdexing rise to a public concern.<sup>288</sup>

The objections noted above may be met by at least three responses. First, the FTC has shown a strong interest in preventing deceptive commercial activities in Cyberspace.<sup>289</sup> For example, from 1996 to

*Internet*, M2 PRESSWIRE, Apr. 25, 1997, available in 1997 WL 10366503; Thomas C. Morrison & Robert W. Lehrburger, *FTC Targets Deceptive Cyberspace Advertising*, NAT'L L.J., Aug. 12, 1996, at B7.

284. See Sovern, *supra* note 278, at 442 & n.28 (noting limited resources of FTC); *id.* at 448 & n.62 (noting limited resources of state agencies). But see William C. MacLeod, *Consumer Protection Developments*, 60 ANTITRUST L.J. 657, 659 (1991) ("Now that the FTC has begun to collaborate with state and local law enforcement, limited federal resources may no longer be an obstacle to local regulation.").

285. Further, Web pages may quickly come and go, making it difficult to track down offenders, let alone follow up with a lawsuit. See Dee Pridgen, *How Will Consumers Be Protected on the Information Superhighway?*, 32 LAND & WATER L. REV. 237, 249 (1997).

286. See Koster & Shatz-Akin, *supra* note 204, at 20-21.

287. The common abuse of generic terms does not give rise to an action under trademark law because generic terms are not registrable. See 15 U.S.C. § 1052(f) (1994). The fact that such a term is used by countless faceless spamdexers ultimately contributes toward that term's uselessness for indexing purposes.

288. Spamdexing of generic terms is different from spamdexing of a trademark, because the owner of a name has a direct interest in taking immediate action. In fact, if a trademark owner *does not* so act, it risks losing its rights via trademark abandonment. See 15 U.S.C. § 1127 (1994).

289. In July 1997, the FTC issued an opinion letter asserting its jurisdiction over deceptive practices conducted over the Internet. See John D. McClain, *Taking Kids' Data Online a No-No: FTC Says Internet Sites Don't Get Parents' OK*, DAILY REC. (Baltimore), Dec. 16, 1997, at 11. Its interest, however, seems currently limited to issues involving large-scale fraud and potential danger to children. In one scheme, con artists promised free pornography to parties who downloaded the defendants' software; this software, in reality, caused the victims' computers to run up large telephone bills calling Moldova, Russia. See Beylen Telecom, Ltd., 62 Fed. Reg. 59708 (FTC 1997). Another scheme, shut down by the FTC in November 1997, involved "an online pyramid scheme that allegedly bilked consumers with promises of skyrocketing investment returns." Maria Seminerio, *FTC Razes Net Pyramid Scheme*, ZDNET NEWS CHANNEL (Nov. 5, 1997) <<http://www.zdnet.com/zdnn/content/zdnn/1105/206422.html>>. The FTC is also



1998, the Bureau of Consumer Protection increased the percentage of its budget devoted to consumer protection in electronic commerce from four to sixteen percent.<sup>290</sup> The FTC has also issued a Staff Advisory Letter that frowns on the use of deceptive Internet domain names,<sup>291</sup> and has engaged in a number of “Surf Days,” which are synchronized searches of the Internet by law enforcement officials.<sup>292</sup> On one Surf Day, it sent warnings to over 500 sites with potential pyramid schemes; on another, it sent warnings to 1,000 junk e-mailers.<sup>293</sup> It has also recently requested public comment on its proposal to issue a policy statement regarding the applicability of FTC rules to electronic media, including the Internet.<sup>294</sup> In its proposal, the FTC notes the “special attributes of advertising on electronic media,” such as hyperlinks and scroll bars.<sup>295</sup> Although the proposal does not mention meta tags, the FTC does note that the

unique features [of the Internet] may require the Commission to give special consideration to certain factors in determining whether a disclosure is effectively communicated on electronic media. As is true for any medium, the specific elements necessary to effectively communicate a disclosure may vary depending on the nature of the advertisement and the nature of the claim.<sup>296</sup>

Thus, the FTC should consider invisible advertising involving the abuse of meta tags to be actionable within its statutory authority. This would be consistent with its self-described “role as one of continuing to

---

concerned about unauthorized collection of information from children who surf the Web. “In a ‘snapshot survey’ of 126 websites on Oct. 14, ‘Kids Privacy Surf Day,’ the FTC found that 86 percent were collecting names, e-mail and postal addresses and telephone numbers.” McClain, *supra*.

290. See *FTC Responses*, *supra* note 248.

291. See Staff Advisory Letter from David Medine, Assoc. Director, Federal Trade Comm’n, to David M. Graves, Internet Business Manager, Network Solutions, Inc. (Aug. 21, 1997), available at <<http://www.ftc.gov/os/1997/9708/internic.let.htm>>.

292. See *FTC Responses*, *supra* note 248.

293. See *id.*

294. See Interpretation of Rules and Guides for Electronic Media; Request for Comment, 63 Fed. Reg. 24,996 (1998) (request for public comment); see also *Consumer Protection — Advertising; FTC Proposes to Clarify Applicability of Its Regulations to Internet Advertising*, 66 U.S.L.W. 2693 (1998).

295. Interpretation of Rules and Guides for Electronic Media, 63 Fed. Reg. at 25,002.

296. *Id.* (footnotes omitted).

foster the development of electronic commerce by acting to prevent fraud and deception, which otherwise may deter consumers from using the Internet to transact business.”<sup>297</sup>

The second response is that even if effective enforcement would be difficult, the FTC should act in hopes of general deterrence. Parties may reasonably debate the general deterrent effect of FTC action.<sup>298</sup> Nevertheless, such action would undoubtedly attract a large amount of coverage in the Internet law and webmaster community.<sup>299</sup> Those who spamdex for financial gain will presumably consider cost-benefit concerns more readily than a seventeen-year old running a Web page from his bedroom. In the case of those motivated by profit, one can expect those who abandon spamdexing to seek alternative routes of publicity, legitimate or otherwise.<sup>300</sup> Also, although many spamdexers are anonymous, one commentator suggests getting subpoenas or suing ISPs to obtain contact information.<sup>301</sup> Encouraging ISPs to release private information is troubling;<sup>302</sup> perhaps a better approach would be to use the WHOIS database to find contact names of those parties who own domain names of sites that spamdex.<sup>303</sup>

The third response is that enforcement of consumer laws in this context furthers both corrective and distributive justice.<sup>304</sup> Although the

---

297. *FTC Responses*, *supra* note 248.

298. Given the defiant attitude of hackers, FTC action may simply further encourage spamdexers to break the rules.

299. Hopefully it would have an effect similar to that of random IRS audits: although it would not eliminate deceptive conduct, it could reasonably be expected to reduce it. *Cf. Collins*, *supra* note 131 (“Preventing [spamdexing] is simply a matter of bringing law suits, according to [attorney William] Cook, who characterizes the Internet community as ‘heavily responsive’ to litigation or threatened litigation.”).

300. *See infra* notes 472–78 and accompanying text.

301. *See Koster & Shatz-Akin*, *supra* note 204, at 20–21.

302. The scope of vicarious or contributory infringement is beyond this article, as is the scope of the bar against tort remedies under the Communications Decency Act. *See Zeran v. America Online, Inc.*, 129 F.3d 327, 330 (4th Cir. 1997) (tort suit held barred), *cert. denied*, 118 S. Ct. 2341 (1998). The concept of terrorizing ISPs into releasing personal information is quite troubling, but will not be addressed here.

303. WHOIS is a tool that may be used to determine contact information for Internet domain names. It is available on the World Wide Web at <<http://rs.internic.net/cgi-bin/itits/whois>>. It lists the addresses, telephone numbers, and e-mail addresses for administrative, technical, and billing contacts. This list will include spamdexers, or those who will know how to reach them. Relying on WHOIS will not be sufficient in cases where the spamdexer runs a website from a subdirectory off somebody else’s domain name, as is the case with GeoCities <<http://www.geocities.com>>.

304. Corrective justice is that which seeks to correct a wrong, to guarantee “correction of harms inflicted on the consumer.” Thomas Wilhelmsson, *Consumer Law and Social Justice*, in *CONSUMER LAW IN THE GLOBAL ECONOMY* 218 (Iain Ramsay ed.,



public interest is diffuse, it is nevertheless an important one, and should be addressed by the government<sup>305</sup> — especially in light of the apparent hesitance of search engines to remove spamdexers from their databases.<sup>306</sup> As discussed in the next Section, although enforcement of proprietary rights will help, in many cases the public's interest in using generic terms ("free beer") or distinctive but nonproprietary terms ("William Shakespeare") will substantially outweigh a spamdexer's interests in using such terms irrelevantly.

The FTC should engage in Surf Days targeting spamdexing. It should publicize this event through the major search engines and other news media both before and after the fact. Because of the difficulties involved in determining the degree of relevance in borderline cases, the FTC should focus solely on irrelevant spamdexing. It should have little trouble finding thousands of culprits who should be sent warnings. The FTC might even choose to follow up with enforcement actions against the more egregious spamdexers. The publicity may well make many spamdexers think twice.

---

1997). Corrective justice would be served by preventing spamdexers from frustrating the efforts of the public to reasonably locate information. Distributive justice, "on the other hand, is connected to the allocation of resources." *Id.* This too would be served by aggressive use of consumer laws, because the resource of database bandwidth, needed by the public and webmasters alike, should not be wasted by irrelevant indexing practices. Further, it will help economically disadvantaged consumers who cannot afford expensive search services. "By guaranteeing all consumers rights which, without the protective measures, would be enjoyed only by the more advantaged consumers, consumer law strengthens social justice." *Id.* at 225.

305. Unlike typical unfair trade practices or consumer fraud, consumers are not, by the mere act of spamdexing, likely to suffer *direct* economic harm. Still, they may suffer *indirect* economic harm, due to the diversion, delay, or frustration in seeking good information. Bloom states that:

A market is viewed as having consumer information problems if consumers are making poorer choices for themselves than they would make if they could gain access to the amount and type of information they would like to have to guide their choices. In markets with such problems, desired information is either unavailable or available in inaccurate, misleading, difficult-to-process, or unreasonably expensive forms.

Paul N. Bloom, *Identifying and Resolving Consumer Information Problems: A New Approach*, reprinted in *MARKETING AND ADVERTISING REGULATION: THE FEDERAL TRADE COMMISSION IN THE 1990S* 286–87 (Patrick E. Murphy & William L. Wilkie eds., 1990). Because there are so many offenders, it makes it hard to single out the more obvious transgressors, unlike the case of major e-mail spammers, who have been sued by ISPs. Here, there are no ISPs to act on the public's behalf.

306. See Judge, *supra* note 89 (noting that very few spamdexing sites are actually banned).

In locating spamdexers, the FTC should not limit itself to sites that spamdex trademarks and famous names, because holders of rights in such terms have standing to seek legal relief. Conversely, because generic terms fall outside of federal trademark law, nobody has the standing, and few would have the incentive, to come forward to seek relief against spamdexing of such terms.<sup>307</sup> Further, barring competitive injury, consumers generally do not have standing under the federal law of trademarks and unfair competition.<sup>308</sup> The FTC should therefore also target spamdexers who abuse popular generic terms such as “free,” “money,” etc. Although the FTC should act against the abuse of trademarks and trade names, private parties, by protecting their own interests, can also act as proxies for the public interest. The law of trademarks and unfair competition, under which most private claims will be brought, is addressed in the next Section.

### *B. Unfair Competition: Balancing Competitive Interests*

#### 1. Rationale of Unfair Competition Law

Disputes also arise between spamdexers and those claiming proprietary rights under the auspices of unfair competition. The Third Restatement of Unfair Competition acknowledges that while much of unfair competition law is codified, it remains at heart a doctrine of the common law.<sup>309</sup> Even if an existing category does not squarely apply, “harm result[ing] from . . . other acts or practices [may] be actionable as

---

307. See 15 U.S.C. § 1052(e) (1994).

308. See *Serbin v. Ziebart Int’l Corp.*, 11 F.3d 1163, 1179 & nn.25–26 (3d Cir. 1993) (acknowledging that the FTC has not effectively protected consumer interests, but refusing to allow consumers to bring section 43(a) claims unless they allege competitive injuries); Maury Tepper, Comment, *False Advertising Claims and the Revision of the Lanham Act: A Step in Which Direction?*, 59 U. CIN. L. REV. 957, 965 (1991) (recognizing lack of remedy for consumers in false advertising cases).

309. See RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 1 (1995). As Professor Green states:

New cases which are not easily identified by some specific tag are indiscriminately lumped together as “unfair competition.” . . . In the area of advertising and marketing of [other products and services], hurtful practices have developed which do not readily fall under the general doctrine or any one of the more specific doctrines. Difficult to pinpoint doctrinally, these practices, too, are termed “unfair competition.”

Leon Green, *Protection of Trade Relations Under Tort Law*, 47 VA. L. REV. 559, 566 (1961); see also ANSELM KAMPERMAN SANDERS, UNFAIR COMPETITION LAW: THE PROTECTION OF INTELLECTUAL AND INDUSTRIAL CREATIVITY 13 (1997).



an unfair method of competition, taking into account the nature of the conduct and its likely effect on both the person seeking relief and the public.”<sup>310</sup> Unfair competition encompasses a broad category of conduct, one far too extensive to catalog here in detail.<sup>311</sup> No laundry

310. RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 1 (1995); *see also* Monica Y. Youn, Case Note, *Neither Intellectual Nor Property*: National Basketball Ass’n v. Motorola, Inc., 107 YALE L.J. 267, 271 (1997) (noting that courts must consider all three parties and that sometimes the defendant must be viewed as a proxy for the public interest).

311. Some causes of action that may be applicable include:

*Trademark infringement* — *see* 15 U.S.C. § 1114(1)(a) (1994);

*Trademark dilution* — *see* Federal Trademark Dilution Act of 1995, Pub. L. No. 104-98, 109 Stat. 985 (codified at 15 U.S.C. §§ 1125(c), 1127 (Supp. I 1995));

*Federal unfair competition* — *see* 15 U.S.C. § 1125(a) (1994); *see also* CHARLES E. MCKENNEY & GEORGE F. LONG III, *FEDERAL UNFAIR COMPETITION: LANHAM ACT* § 43(A) (1996);

*Misappropriation* — *see* 15 U.S.C. § 1125(a)(1)(A) (1994); *International News Serv. v. Associated Press*, 248 U.S. 215, 240 (1918); *Toho Co., Ltd. v. Sears, Roebuck & Co.*, 645 F.2d 788, 794 (9th Cir. 1981). Section 43(a) of the Lanham Act (15 U.S.C. § 1125(a)) includes the federal common-law tort of misappropriation enunciated in *International News Service*. *See* MCKENNEY & LONG, *supra*, § 2.05. This doctrine includes cases “where no fraud is perpetrated but where defendant, for commercial advantage, has nonetheless misappropriated the benefit or property right of plaintiff and has exploited plaintiff’s business values.” *Id.*; *see also* RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 38 (1995);

*False advertising* — *see* 15 U.S.C. 1125(a)(1)(B) (1994); RESTATEMENT (THIRD) OF UNFAIR COMPETITION §§ 3–8 (1995);

*Right of publicity* — *see, e.g.*, *Stern v. Delphi Internet Servs. Corp.*, 626 N.Y.S.2d 694 (N.Y. Sup. Ct. 1995); Mitchell D. Kamarck, *Empowering Celebrities in Cyberspace: Stripping the Web of Nude Images*, ENT. & SPORTS LAW., Winter 1998, at 1, 14–15. The Restatement of Unfair Competition states, “[o]ne who appropriates the commercial value of a person’s identity by using without consent the person’s name, likeness, or other indicia of identity for purposes of trade is subject to liability for relief . . . .” RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 46 (1995); and

*State and Federal unfair trade practice statutes* — *see supra* Part IV.A. Although some state statutes against unfair trade practices limit enforcement to state officials, the vast majority of them also allow private individuals to have standing under these acts. *See* RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 1, stat. note (1995). Many of these states also allow competitors to seek redress for harm to commercial relations through these statutes. *See id.* Trade regulation statutes, discussed in Part IV.A, will not be discussed further in this Section.

Discussion of the intricacies of each and every one of these causes of action, however tempting, is beyond both the scope and the purpose of this Article. First, most unfair competition causes of action use the same “likelihood of confusion” standard that will be discussed here in detail. The obvious exceptions, trademark dilution and trade practice statutes, either have been discussed, or will be discussed. Second, whether or not rule application under current doctrine merits liability is not (excuse the term) *relevant* to the present discussion. These doctrines have historically adapted to new

list could ever be exhaustive, because this body of law retains the flexibility to address new situations.<sup>312</sup> It seeks to promote competition and redress anticompetitive business practices.<sup>313</sup> In other words, “[w]here exploitation of another’s achievement becomes inequitable, unfair competition law . . . provides a remedy.”<sup>314</sup>

Defining unfair competition in terms of inequitable conduct is somewhat circular. As Justice Holmes pointed out in *International News Service v. Associated Press*,<sup>315</sup> the mere existence of economic value does not by itself require legal protection as “property.”<sup>316</sup> Further,

---

technologies and new situations. Just because liability may lie under strict rule application does not mean that it is proper for courts to do so under the policies underlying these doctrines. Similarly, even if liability does not so lie, perhaps it should under a rereading or rewriting of the applicable statutes. My purpose here, then, is to provide an analytic framework that informs this process. After all, although meta tags may come and go, indexing abuse is here to stay. Therefore, a reasoned analysis of how relevance applies to unfair competition and trade regulation may be a valuable contribution.

312. See RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 1 cmt. g (1995).

A primary purpose of the law of unfair competition is the identification and redress of business practices that hinder rather than promote the efficient operation of the market. Certain recurring patterns of objectionable practices form the basis of the traditional categories of liability . . . . However, these specific forms of unfair competition do not fully exhaust the scope of statutory or common law liability for unfair methods of competition, and [the Restatement] therefore includes a residual category encompassing other business practices determined to be unfair.

*Id.*

313. See *id.*

314. SANDERS, *supra* note 309, at 8.

315. 248 U.S. 215 (1918).

316. As Justice Holmes stated:

*Property, a creation of law, does not arise from value. . . . Many exchangeable values may be destroyed intentionally without compensation. Property depends upon exclusion by law from interference, and a person is not excluded from using any combination of words merely because someone has used it before, even if it took labor and genius to make it.*

*Id.* at 246 (Holmes, J., dissenting) (emphasis added); see also Felix S. Cohen, *Dialogue on Private Property*, 9 RUTGERS L. REV. 357, 378 (1954) (noting the possibility of valueless property and propertyless value); Cohen, *Transcendental Nonsense*, *supra* note 14, at 815; Wendy J. Gordon, *On Owning Information: Intellectual Property and the Restitutionary Impulse*, 78 VA. L. REV. 149, 178–79 (1992) (quoting Holmes and noting that “[i]f the courts believe they must protect all existing value in order to ensure productivity, one can point out that only sometimes will granting legal rights increase the amount of value in the world”); Green, *supra* note 309, at 565 (“The term [‘unfair



overly broad protection can be anticompetitive.<sup>317</sup> Therefore, it is necessary to balance the interests involved.<sup>318</sup> To resolve this dilemma, unfair competition law traditionally looks at “the regulation of market behaviour rather than the protection of market interests.”<sup>319</sup>

It is difficult to predict the effects of too much or too little protection.<sup>320</sup> Trying to balance such disparate interests is also inherently problematic.<sup>321</sup> Overly broad protection of trademarks and

competition’] gives no hint of what is ‘unfair’ and is thus only a general label for the types of practices which have been held by the courts in specific cases to be ‘unfair.’”); Youn, *supra* note 310, at 272 & n.28; *see also supra* note 14 and accompanying text (noting the problematic nature of the term “property”). Although at the time of *International News Service v. Associated Press*, the case “was supposed to be a special case of limited importance,” the development of that day’s new media — radio and later television — caused the “doctrine underlying the case [to be] utilized in many subsequent cases.” Green, *supra* note 309, at 567. The *INS* rationale was just recently used by the Second Circuit. *See National Basketball Ass’n v. Motorola, Inc.*, 105 F.3d 841, 843 (2d Cir. 1997) (holding that a narrow *INS*-style “hot news” exception survived preemption under the Copyright Act).

317. *See* Green, *supra* note 309, at 565 (“[T]he mere fact that another’s achievement is being exploited does not call for any impediment on the basis of unfair competition provisions. On the contrary, appropriating and building on others’ achievements is the cornerstone of cultural and economic development.”); *see also* Michael H. Davis, *Death of a Salesman’s Doctrine: A Critical Look at Trademark Use*, 19 GA. L. REV. 233, 236–37 (1985); Robert N. Klieger, *Trademark Dilution: The Whittling Away of the Rational Basis for Trademark Protection*, 58 U. PITT. L. REV. 789, 861 (1997); J.H. Reichman, *Design Protection and the New Technologies: The United States Experience in a Transnational Perspective*, 19 U. BALT. L. REV. 7, 117–18 (1989).

318. *See* Anthony L. Clapes, *Proceed With Caution — Information Superhighway Under Construction: Selecting the Proper Intellectual Property Rights Paradigm to Apply to Passengers on the Interim-net*, 11 ST. JOHN’S J. LEGAL COMMENT. 621, 622 (1997); Lemley, *supra* note 27, at 888; Peter A. Wald et al., *Standards for Interoperability and the Copyright Protection of Computer Programs*, PLI INTELL. PROP./ANTITRUST, June–July 1993, at 891, 975.

319. SANDERS, *supra* note 309, at 8.

320. *See* BOYLE, *supra* note 27, at 115 (noting the “difficulty that economists have in theorizing about the amount of information that will be produced in an ‘unregulated’ market”).

321. *See id.* at 114.

[Intellectual property law] must explain how it is that we can motivate individuals, who are sometimes postulated to be essentially self-serving, and sometimes to be noble, idealistic souls, to produce information. If the answer is “by giving them property rights,” it must also explain why this will not diminish the common pool, or public domain, so greatly that a net decrease in the production of information will result.

*Id.*

trade names can be economically and socially inefficient.<sup>322</sup> Further, as a practical matter, the ease of using others' proprietary rights calls into question the concepts underlying these rights.<sup>323</sup> Courts should therefore resist the current tendency to "propertize" everything relating to information, because this tendency may reduce, not enhance, the number of works (i.e., information) available via search engines.<sup>324</sup> As Part V addresses in greater detail, the "reworking of information is a key ingredient of the information society or information economy in which new information or old information processed into a new form acquires substantial economic value."<sup>325</sup>

Historically, by signifying consistency in source or quality, trademarks "served to lower consumers' 'search costs,' . . . [becoming] an abbreviated informational proxy."<sup>326</sup> Put differently, trademarks are information-rich terms: the greater the goodwill and distinctiveness of a mark, the greater its informational value. This explains why trademarks, trade names, and the names of celebrities are attractive targets for abuse by spamdexers. The spamdexing harm occurs not necessarily by disparaging the quality of the owner's name, nor by suggesting affiliation, endorsement, or sponsorship, but by eliminating

---

322. Cf. *id.* at 119 (noting that in the copyright realm, "[a]n author-centered regime can actually *slow down* scientific progress, *diminish* the opportunities for creativity, and *curtail* the availability of new products"). However, many commentators, including Boyle, note the indeterminacy of economic arguments that seek to show that a particular regime will lead to either overproduction or underproduction of information. See *id.* at 41; Lemley, *supra* note 27, at 890. Professor Lemley believes that this indeterminacy is not limited to economic analysis, but is also part of "the unquestionable indeterminacy of intellectual property law." *Id.* at 893.

323. See KATSH, TRANSFORMATION OF LAW, *supra* note 27, at 170 ("Those who benefit from [intellectual property] law are, therefore, threatened as the inherent controls of print are being lifted. In addition, however, the concepts themselves are vulnerable as the public acquires experience with information that was previously suppressed."). It is erroneous to assume "that legal concepts that are widely accepted today will be unaffected by the new media." *Id.* at 171. It is necessary to evaluate intellectual property laws in light of these technologies. "Whether [previously suppressed] information was unavailable because of legal restrictions or the limitations of prior modes of communication, the ultimate effect of the new communications environment will be a new kind of accommodation, a new balance, and a new meaning." *Id.*

324. See Lemley, *supra* note 27, at 898–900 (criticizing this tendency and noting that "it is far from clear what valuable new works the *White v. Samsung Electronics America, Inc.* decision will encourage"); see also *White v. Samsung Electronics Am., Inc.*, 971 F.2d 1395 (9th Cir. 1992), *reh'g en banc denied*, 989 F.2d 1512 (1993); Heald, *supra* note 130, at 804–08 & n.114 (criticizing the *White* case and listing other critics).

325. KATSH, TRANSFORMATION OF LAW, *supra* note 27, at 172.

326. Keith Aoki, *Authors, Inventors and Trademark Owners: Private Intellectual Property and the Public Domain, Part II*, 18 COLUM.-VLA J.L. & ARTS 191, 236 (1994).



a reasonable probability of finding the true holder of the mark online.<sup>327</sup> In this context, the harm is dual: not only might the owner be harmed, but the spamdexer may be unjustly enriched.<sup>328</sup> I will examine this concept using the two main forms of legal protection for trademarks: infringement and dilution of trademarks under the Lanham Act.<sup>329</sup>

## 2. Trademark Infringement: Conceptually Inapposite

Trademark infringement is prohibited under sections 32(1)(a) and 43(a) of the Lanham Act.<sup>330</sup> Spamdexing may be actionable under one

---

327. See Green, *supra* note 309, at 567. For purposes of this Article, I will treat references to trademarks (or just “marks”) as also referring to other similar forms of goodwill, such as trade names, service marks, collective and certification marks, and rights of publicity.

328. See SANDERS, *supra* note 309, at 14.

329. 15 U.S.C. §§ 1051–1072, 1091–1096, 1111–1128 (1994 & Supp. I 1995); see also O’Rourke, *supra* note 21. Extended discussion of every cause of action is beyond the scope of this Article. However, it should be noted that most forms of unfair competition mentioned above use essentially the same standards, especially causes of action arising under section 43(a) of the Lanham Act, such as trade dress, false advertising, and right of publicity. Also note that consumers generally do not have standing to bring a suit under the Lanham Act. See *Colligan v. Activities Club of N. Y.*, 442 F.2d 686, 687 (2d Cir. 1971); see also *Barrus v. GTE*, 55 F.3d 468, 470 (9th Cir. 1995); *Serbin v. Ziebart Int’l Corp.*, 11 F.3d 1163, 1180 (3d Cir. 1993); MCKENNEY & LONG, *supra* note 311, § 9.03[1]. Although the language of section 43(a) of the Lanham Act seems to indicate that a private cause of action exists, even courts which have allowed private parties to state a claim have limited it to those with a commercial interest. See Ferrill, *supra* note 245, at 59–61.

330. Trademark infringement under section 32(1)(a) is a subset of the broader category of protection against competitive injuries embodied in section 43(a) of the Lanham Act. See 15 U.S.C. §§ 1114(1)(a), 1125(a) (1994); MCKENNEY & LONG, *supra* note 311, § 3.01 (“The protection afforded by [Section 43(a)] is broader than that provided under [Section 32], in that it was enacted to protect consumers and competitors alike against *all* forms of misdescription of products and services in commerce.”) (quoting *Vuitton et Fils, S.A. v. Crown Handbags*, 492 F. Supp. 1071, 1077 (S.D.N.Y. 1979)) (alterations in original). Unlike section 32(1), which applies only to federally registered trademarks, unfair competition under section 43(a) provides broader protection, both in scope of the protected subject matter as well as the prohibited activity. Section 43(a) protects any “word, term, name, symbol, or device, or any combination thereof.” 15 U.S.C. § 1125(a) (1994). Section 43(a) prohibits confusion by “passing off” or “reverse passing off” of “affiliation, connection, or association” of persons, or of the origin, sponsorship or approval of goods or services. 15 U.S.C. § 1125(a)(1)(A) (1994). Passing off occurs when A passes off his goods as B’s (such as fake Rolex watches). See, e.g., *Smith v. Montoro*, 648 F.2d 602, 604 (9th Cir. 1981). Reverse passing off is when A passes off B’s goods as his own. See *id.* at 606. Section 43(a) also prohibits misrepresentation of the nature or origin of goods via false advertising or disparagement. See 15 U.S.C. § 1125(a)(1)(B) (1994). False advertising

or both sections; the two will be analyzed together because they use the same multi-factor test.<sup>331</sup> Trademarks will be protected against infringement if they are valid, distinctive marks, and if the unauthorized use creates a likelihood of confusion, mistake, or deception.<sup>332</sup>

Conceptually, it is difficult to apply trademark infringement to spamdexing because of the difficulty of demonstrating a likelihood of consumer confusion stemming from it. For example, search engine output for a search using "Disney" might include spamdexed references to unrelated Web pages, but will also contain additional information that, in most cases, makes it clear that the supposedly infringing pages are not sponsored, endorsed by, or affiliated with Disney.<sup>333</sup> In most cases, this

exists if A misrepresents the qualities or origin of her own or another's goods or services in a deceptive, material way. *See Southland Sod Farms v. Stover Seed Co.*, 108 F.3d 1134, 1139 (9th Cir. 1997).

331. *See, e.g., Amstar Corp. v. Domino's Pizza, Inc.*, 615 F.2d 252, 258–59 (5th Cir. 1980).

332. *See* 15 U.S.C. §§ 1114(1)(a), 1125(a) (1994). All courts apply a multi-factor test to determine whether a likelihood of confusion exists. *See* 3 MCCARTHY, *supra* note 74, § 23:19; *see also* RESTATEMENT (THIRD) OF UNFAIR COMPETITION §§ 21–23 (1995); RESTATEMENT OF TORTS §§ 729, 731 (1938).

For present purposes, the term "trademark" will include service marks, certification marks, and collective marks, which may be federally registered, and trade names, which may not. *See* 15 U.S.C. § 1127 (1994). In practice, courts provide protection against infringement for both federally registered and unregistered trademarks. The former are covered under 15 U.S.C. § 1114(1)(a) (1994), and the latter under 15 U.S.C. § 1125(a) (1994). *See* 3 MCCARTHY, *supra* note 74, § 27:14. The same standard, "likelihood of confusion," applies in both cases. *See id.* § 23:1.

Only distinctive marks merit protection. *See* 15 U.S.C. § 1052(f) (1994). Marks that are inherently distinctive are immediately eligible for legal protection upon adoption and use. *See* 2 MCCARTHY, *supra* note 74, § 15:1. Such marks may be fanciful, arbitrary, or suggestive. *See id.* Fanciful marks are "coined" words that are created with the sole purpose of acting as a trademark, such as Exxon. *See* 1 *id.* §§ 11:5, 11:8. Arbitrary marks are when common words or images are used in an arbitrary way, such as using V-8 to name a mixture of eight kinds of vegetable juices. *See id.* § 11:11. Suggestive marks, while similar to arbitrary marks, suggest, however remotely, the product at issue, such as using "Greyhound" for a bus line. *See id.* § 11:12. However, marks that are merely descriptive do not qualify for protection unless they acquire secondary meaning; in other words, a term like "General Motors" did not qualify for protection until the term brought to consumers the mental image of a single source of the product or service. *See id.* § 11:15, 2 *id.* § 15:5.

333. Search engine output typically lists the page's title, URL (uniform resource locator, or Internet address), and a brief description or synopsis of the page's text. This text is typically culled from the first 250 words of the Web page, or recites the text of the description meta tag. This information is likely to prevent (before the fact) or quickly dispel (after the fact) any likelihood of confusion. Also, actual page content will, by definition, reflect actual content. Assuming that actual content does not independently give rise to a likelihood of confusion, it should work toward showing a lack of



information will be sufficient to preclude or dispel confusion. If the additional information itself creates a likelihood of confusion, then it would amount to an independent act of infringement, and should not bootstrap infringement onto the initial act of spamdexing. Any confusion arising from meta tags is that of the search engines themselves, because they lack the sophistication to cull the relevant from the irrelevant. Further, reasonable consumers should be expected to engage in practices to narrow their queries to increase precision.<sup>334</sup> As shown in Part III.B, even consumers who use targeted, information-rich terms or phrases (such as “Princess Diana” or “Battle of the Bulge”) must reasonably expect noisy output from queries that are not combined with other terms.

One might intuitively object to this argument due to the willful misbehavior of those who engage in blatant irrelevant spamdexing. However, this is not to say that spamdexing does not merit liability, but rather that the “likelihood of confusion” standard provides a poor conceptual basis for spamdexing liability.<sup>335</sup> After all, a major rationale for trademark rights is the protection of consumers’ expectations.<sup>336</sup> This rationale has no application, however, when consumers have no reasonable expectations to protect. To blindly apply trademark infringement may itself be anti-competitive, by overprotecting trademarks where no reasonable expectations exist.

---

confusion. Even though a webmaster may spamdex the keyword meta tag, it is not to most webmasters’ interest to spamdex the description meta tag. *See supra* note 95. *But see* Niton Corp. v. Radiation Monitoring Devices, Inc., No. Civ.A. 98-11629-REK, 1998 WL 812685, at \*2–\*3 (D. Mass. Nov. 18, 1998) (granting “preliminary injunction subject to modification” where defendant copied verbatim the description meta tags of plaintiff competitor and used them in its own Web page). Thus, even assuming the listing includes unrelated Web pages due to spamdexing, once a user views the information provided in the query output, a likelihood of confusion will rarely exist. Even if the query output is ambiguous, once the user goes to the “offending” page, the content (assuming that it is unrelated) would certainly dispel any possibility of confusion. If confusion exists at this stage, then the “offender” has the much bigger problem of a likelihood of confusion that stems from the page itself, and not just from the meta tag.

334. *See infra* notes 446–63 and accompanying text.

335. A recent article on meta tags addresses a similar point. *See* O’Rourke, *supra* note 21. I agree with Professor O’Rourke’s conclusion that the likelihood of confusion test is inapposite to spamdexing, as are claims for false advertising. *See id.* at 293–99. Professor O’Rourke’s analysis is based in large part on her correct observation that the reasonable consumer is savvy enough not to be confused. *See id.* at 294. In this Section, I attempt to detail the conceptual difficulties that explain why infringement is generally inapplicable to the mere use of a meta tag.

336. *See* Heald, *supra* note 130, at 788 (“Protecting consumer expectation is reasonable and, moreover, almost certainly economically efficient.”).

It might be argued that frustration of the public's ability to find unspamdexed information by itself satisfies the requisite likelihood of confusion. As noted above, the additional information included in search results might "preclude or dispel" confusion. "Dispel" suggests the possibility of *initial* infringement that is quickly dispelled, sometimes called "initial interest confusion"<sup>337</sup> by the courts. This might occur when the defendant's use of a deceptive term causes initial confusion to the consumer that attracts subsequent non-confused consumer attention.<sup>338</sup> This has been likened to a "bait and switch" that "effectively allow[s] the competitor to get its foot in the door by confusing consumers."<sup>339</sup>

Initial interest confusion, however, is a weak hook on which to sustain trademark infringement. First, if this confusion ever truly exists, it is fleeting. Second, by allowing liability to attach "even if confusion as to the source of the goods is dispelled by the time any sales are consummated,"<sup>340</sup> it cannot be said that reasonable consumer expectations are being protected. If consumer confusion is the loadstone for trademark infringement, then why impose liability when no realistic possibility of confusion exists? For these reasons, this expansive concept is not embraced by all courts.<sup>341</sup> The Federal Circuit, although not the final arbiter of trademark matters, has further called the concept into question by specifically refusing to adopt it in a case where it affirmed the Trademark Trial and Appeal Board on other grounds.<sup>342</sup>

Initial interest confusion does describe well the effect of spamdexing. However, its use is conceptually a misapplication of trademark infringement. After all, if infringement is possible on a bare

---

337. "[T]he Lanham Act forbids a competitor from luring potential customers away from a producer by initially passing off its goods [or services] as those of the producer's even if confusion as to the source of the goods is dispelled by the time any sales are consummated." *Rust Env't & Infrastructure, Inc. v. Teunissen*, 131 F.3d 1210, 1217 (7th Cir. 1997) (quoting *Dorr-Oliver, Inc. v. Fluid-Quip, Inc.*, 94 F.3d 376, 382 (7th Cir. 1996)) (alterations in original); see also Michael J. Allen, *The Scope of Confusion Actionable Under Federal Trademark Law: Who Must Be Confused and When?*, 26 WAKE FOREST L. REV. 321 (1991); Charles E. Bruzga, *Sophisticated Purchaser Defense Avoided Where Pre-Sale Confusion Is Harmful — A Brief Note*, 78 TRADEMARK REP. 659 (1988).

338. See 3 MCCARTHY, *supra* note 74, § 23:36.

339. *Dorr-Oliver*, 94 F.3d at 382.

340. *Id.*

341. See *Munsingswear Inc. v. Jockey Int'l Inc.*, 31 U.S.P.Q.2d (BNA) 1146, 1149 & n.6 (D. Minn.), *aff'd*, 39 F.3d 1184 (8th Cir. 1994) (noting split in authority); Allen, *supra* note 337, at 339.

342. See *Weiss Assoc., Inc. v. HRL Assoc., Inc.*, 902 F.2d 1546, 1549 (Fed. Cir. 1990); Allen, *supra* note 337, at 342–43 (discussing *Weiss*).



meta tag, then why would one inquire whether there were “use[s] in a permissible manner . . . . in the visible portion of the web page [that] might even render non-infringing the additional use of the mark” in a meta tag?<sup>343</sup> This may be illustrated by examining three possible scenarios for the imposition of infringement liability.

Table Four: When Infringement is Conceptually Possible

Scenario	Infringement conceptually possible?
1. The meta tag infringes by itself.	NO
2. The meta tag plus ambiguous actual content combine to make a single act of infringement.	YES
3. The actual content infringes by itself.	YES

Consider scenario #1. Suppose I used the term “Oppedahl & Larson” as a meta tag.<sup>344</sup> The initial interest confusion doctrine might suggest that I am liable for trademark infringement. I would assert, based on the content of this Article, non-trademark use, or alternatively, the First Amendment, as a defense to infringement.<sup>345</sup> Most likely, those defenses would be successful.<sup>346</sup> Where, then, was the infringement?<sup>347</sup> If looking at actual text showed that I had the right to use the term in the first place, how can it be said that the use of the meta tag — by itself — was infringing? This suggests that meta tags should not by themselves create trademark infringement, thus precluding scenario #1 as appropriate for infringement liability. After all, if there was initial

---

343. Halpern, *supra* note 24, at 2.  
344. *See supra* notes 122–30 and accompanying text.  
345. *See infra* notes 406–10 and accompanying text.  
346. *Cf. supra* note 194.  
347. The *Welles* case provides a subtle and wonderful example of the difficulties involved in trying to apply trademark infringement to spamdexing. The court seems to base its holding on a lengthy fair use analysis, which might suggest that the use of the meta tag might be separately infringing, subject to a fair use based on actual content. *See Playboy Enters., Inc. v. Welles*, 7 F. Supp. 2d 1098, 1103–04 (S.D. Cal.), *aff’d*, No. 98-55911, 1998 U.S. App. LEXIS 27739 (9th Cir. Oct. 27, 1998). But then the court goes ahead and, in dictum, says that there was no likelihood of confusion for the same reasons. *See id.* at 1104. So was there infringement, made into fair use by actual content, or did the court not know whether there was infringement until it looked at both meta tag and actual content? It seems that the court took this conceptual difficulty into account, but did not specifically address it.

confusion, why would we have to look at later confusion to demonstrate whether the initial confusion ever existed?

The argument that infringement is generally inapposite to spamdexing is better shown by a case in which plaintiff argues that both meta tag and actual content combine for a likelihood of confusion (scenario #2). Suppose I were an attorney using Oppedahl's name as a meta tag to draw his clientele. Here, the existence of actual content at my site may not even mention Oppedahl. Still, it is doubtful, though possible, that parties seeing "Ira S. Nathenson, Attorney at Law" at the main page would believe that I am associated with Oppedahl & Larson.

Suppose, however, that actual content is sufficiently ambiguous that sponsorship, endorsement, or affiliation might be inferred by an unsophisticated consumer. Consider the facts of the *Advanced Concepts* dispute. Suppose a consumer searches for Oppedahl for information on domain names. An unsophisticated consumer — drawn to a Web page belonging to a domain name broker due to that broker's spamdexing with Oppedahl's name — might reasonably infer that the broker is somehow affiliated with Oppedahl & Larson. Thus, it is possible that cases may arise where the meta tag, when *combined* with ambiguous actual content, could reasonably cause a likelihood of confusion.<sup>348</sup>

---

348. Suppose Pepsi used "Coke" as a meta tag, and that no actual content at its site mentioned Coke products. *See supra* note 196. By themselves, the meta tag and content do not infringe. The meta tag does not infringe by itself for the reasons mentioned above in the discussion of scenario #1. The actual content cannot infringe because it does not use any of Coca-Cola's trademarks. The two together, however, in theory could combine to create one act of infringement. Although it may turn out that reasonable consumers would not be confused, one can imagine scenarios where the same conduct would indeed cause a likelihood of confusion. For instance, suppose Pepsi used "Mountain Dew" or "Surge" as meta tags. Reasonable consumers could not be expected not to know who manufactures these products, suggesting a likelihood of confusion.

A recent case further illustrates this possibility. Two companies that sold products in the same field were embroiled in ongoing litigation. *See Niton Corp. v. Radiation Monitoring Devices, Inc.*, No. Civ.A. 98-11629-REK, 1998 WL 812685, at \*1 (D. Mass. Nov. 18, 1998). They were not direct competitors, but might become such as their product lines grew. *See id.* The plaintiff discovered that the defendant's Web pages had copied description meta tags from plaintiff's Web pages. *See id.* at \*2. Judge Keeton of the United States District Court for the District of Massachusetts therefore granted a "preliminary injunction subject to modification" to prevent defendant from using its website or other means of attracting visitors that was likely to lead the public to believe that, *inter alia*, defendant was known as, affiliated with, or made any products marketed by plaintiff. *See id.* at \*3–\*4.

Although the court did not reach the issue's merits or even cite the relevant law, it is apparent that the court, by using the "likely to lead users to believe" language, was considering trade name infringement under section 43(a) of the Lanham Act. The court reached the right result. Here, the description meta tag, when combined with actual



However, this would not be an initial interest case: initial ambiguity does not blossom into a real likelihood of confusion until one views actual content. Thus, only scenario #3 (actual content infringing by itself), or #2 (meta tag in combination with ambiguous actual content combine to produce one act of infringement) fall under trademark infringement.

Cases involving telephone mnemonics (such as 1-800-CALL-ATT) are instructive. The Sixth Circuit, in *Holiday Inns, Inc. v. 800 Reservation, Inc.*,<sup>349</sup> held that the act of obtaining a commonly misdialed variant of a mnemonic telephone number did not cause a likelihood of confusion.<sup>350</sup> The defendants did not advertise the fact that they had the number 1-800-H[zero]LIDAY, a commonly misdialed variant of plaintiff's mnemonic 1-800-HOLIDAY.<sup>351</sup> Had they done so, they would have taken an affirmative act to *create* confusion. However, here any confusion (misdialing) was in the mind of the consumers; the defendants, while taking advantage of this confusion, neither created nor furthered it through affirmative means. The Sixth Circuit drew the correct line in the murky soil that separates fair from unfair competition. Holiday Inns could have registered common misdialings of its telephone number, but did not. Defendants took no steps to create or further the

---

content, could indeed give rise to a likelihood of confusion.

Still, *Niton* is somewhat different from the typical meta tag abuse case because it involves description meta tags, which provide a short description that appears along with search query output. Thus, a searcher could actually see the defendant's description, which would state "The Home Page of Niton Corporation, makers of the finest lead, radon, and multi-element detectors." *Id.* at \*3. This, by itself, may be a separate act of trade name infringement. However, this possibility does not disprove this Article's contention that initial interest confusion is inapplicable to *keyword* meta tag abuse, because keyword meta tags — which are the focus of both this Article and almost all meta tag abuse — are not viewed by a search engine user. Conversely, a spamdexed description meta tag is viewed by the user, and can provide sufficient information to give rise to an independent instance of infringement.

349. 86 F.3d 619 (6th Cir. 1996), *cert. denied*, 117 S. Ct. 770 (1997).

350. *See id.* at 626. *See generally* Lisa D. Dame, Comment, *Confusingly Dissimilar Applications of Trademark Law to Vanity Telephone Numbers*, 46 CATH. U.L. REV. 1199 (1997); Elizabeth A. Horky, Note, *1-800-I-AM-VAIN: Should Telephone Mnemonics Be Protected as Trademarks?*, 3 J. INTELL. PROP. L. 213 (1995); Nathenson, *supra* note 210, at 966–74; Terry Ann Smith, Comment, *Telephone Numbers That Spell Generic Terms: A Protectable Trademark or an Invitation to Monopolize a Market?*, 28 U.S.F. L. REV. 1079 (1994).

351. *Compare Holiday Inns*, 86 F.3d at 621, 625–26, with *American Airlines, Inc. v. A 1-800-A-M-E-R-I-C-A-N Corp.*, 622 F. Supp. 673, 684–86 (N.D. Ill. 1985) (defendant published its deceptive mnemonic in yellow pages under "Airline Companies").

confusion, and, in fact, gave clear disclaimers to people once the phone was answered.<sup>352</sup>

In the case of spamdexing, the search engine is confused, while the public is generally savvy enough to realize what is going on. If information-poor terms are used, then the consumer does not have a reasonable expectation of finding much of anything without narrowing the search. This is analogous to misdialing, where the confusion pre-exists in the mind of the consumer. If information-rich terms are used, then the consumer will almost immediately realize the lack of sponsorship, affiliation, or endorsement with the trademark holder prior to or upon reaching the spamdexed page. In cases where the consumer is confused, it will stem from independent confusion created by actual content (scenario #3), or from a meta tag creating confusion in combination with ambiguous content (scenario #2). If the webmaster falls under scenarios #2 or #3, however, then she is taking affirmative steps to create confusion independently of or in tandem with the meta tag. Otherwise, the fact that the defendant merely waits for hits (scenario #1) should not be enough to merit infringement, absent adoption of the conceptually flawed initial interest confusion doctrine.

One might argue that spamdexing is less like *Holiday Inns* and more like other cases where the defendant advertised the misdialed mnemonic.<sup>353</sup> After all, a spamdexer takes the affirmative step of including an index term in a Web page. Telephone mnemonics, however, serve as a strong indication of source, while most indexing terms by themselves do not — suggesting that the mere use of a meta tag is not enough. Further, *Holiday Inns* may stand for the proposition that initial interest confusion requires some “plus” factor to ripen into infringement. To hold that the mere use of the meta tag is sufficient would bootstrap claims, imposing infringement liability for scenario # 1.

### 3. Trademark Dilution: Tailor-Made

Whereas trademark infringement is conceptually inapposite to spamdexing, trademark dilution is not. Dilution, while falling within the family of unfair competition actions, does not require proof of a likelihood of confusion.<sup>354</sup> In that respect, it provides very strong protection for plaintiffs, because they do not have to prove when and whether confusion exists. Dilution protection is available only for

---

352. See *Holiday Inns*, 86 F.3d at 625–26.

353. See *A 1-800-A-M-E-R-I-C-A-N*, 622 F. Supp. at 684–86.

354. See 15 U.S.C. §§ 1125(c), 1127 (Supp. I 1995).



“famous” marks, limiting liability to only those who spamdex well-known names.<sup>355</sup> Although the definition of dilution is rather vague, spamdexing arguably dilutes famous names by “blurring” their distinctive qualities.<sup>356</sup> Blurring occurs when a distinctive mark is used by others for a “plethora of different goods and services.”<sup>357</sup> Such use may “reduce the ability of the plaintiff’s mark to identify the appropriate goods.”<sup>358</sup> By blurring, a spamdexer may render a trademark useless for search engines by introducing so much noise that the mark becomes useless as a query term.<sup>359</sup> Rendering a famous name near-useless as an indexing term would therefore appear to be the textbook definition of dilution. Distinctive terms — those that deserve trademark protection — are information-rich and useful for searches high in recall and precision. Dilution, in information science terms, is an injury that renders an information-rich term information-poor. Attorneys call this “genericide.” Thus, by whittling away a term’s distinctiveness, a spamdexer introduces noise into searches, rendering the term near-useless for high-precision searching. Conceptually, dilution aptly describes the spamdexing injury.

The standard of relevance shown in Part III.B is instructive in contrasting infringement with dilution. Recall that in the context of

---

355. *See id.* § 1125(c)(1).

356. The definition of “dilution” is somewhat indefinite: “the lessening of the capacity of a famous mark to identify and distinguish goods or services, regardless of the presence or absence of — (1) competition between . . . the parties, or (2) likelihood of confusion, mistake, or deception.” 15 U.S.C. § 1127 (Supp. I 1995); *see also* 3 MCCARTHY, *supra* note 74, § 24:94. Another form of dilution is “tarnishment,” which is “when a famous mark is linked to products of poor quality or is portrayed in an unwholesome manner.” *Panavision Int’l, L.P. v. Toeppen*, 945 F. Supp. 1296, 1304 (C.D. Cal. 1996), *aff’d*, 141 F.3d 1316 (9th Cir. 1998); *see also* *Toys “R” Us v. Akkaoui*, No. C 96-3381, 1996 U.S. Dist. LEXIS 17090 (N.D. Cal. Oct. 29, 1996); 3 MCCARTHY, *supra* note 74, § 24:95.

357. 3 MCCARTHY, *supra* note 74, § 24:94; Klieger, *supra* note 317, at 823–27; Frank I. Schechter, *The Rational Basis of Trademark Protection*, 40 HARV. L. REV. 813, 831 (1927).

358. Terry R. Bowen, *The Federal Trademark Dilution Act of 1995 — Does It Address the Dilution Doctrine’s Most Serious Problems?*, 7 DEPAUL-LCAJ. ART & ENT. L. 75, 79 (1996); *see also* 3 MCCARTHY, *supra* note 74, § 24:94.

359. Indeed, the *AsiaFocus* court states that the defendants’ “purposeful tactic of embedding [plaintiff’s] trademarks . . . in the hidden computer source code . . . epitomizes the ‘blurring’ of [plaintiff’s] trademarks.” *Playboy Enters., Inc. v. AsiaFocus Int’l, Inc.*, No. 97-734-A, 1998 U.S. Dist. LEXIS 10359, at \*21 (E.D. Va. Feb. 2, 1998) (Mag. J.), *adopted by* 1998 U.S. Dist. LEXIS 10459 (E.D. Va. Apr. 10, 1998); *see also* O’Rourke, *supra* note 21, at 301; *supra* notes 141–45 and accompanying text; *infra* notes 396–403 and accompanying text.

consumer protection laws, relevance is the litmus test.<sup>360</sup> Relevance is also determinative in many infringement cases. Indeed, the “likelihood of confusion” standard itself is a kind of relevance test: in the normal (non-spamdexing) infringement case, a court must be able to reasonably connote a relationship of affiliation, sponsorship, or source, which is required to find a likelihood of confusion. A lack of relevance between meta tag and content suggests a lack of a likelihood of confusion. On the other hand, although a lack of relevance may be indicative of no infringement, the existence of relevance, by itself, does not conclusively prove infringement. The actual content itself must independently create or act together with the meta tag to create a likelihood of confusion, per scenarios #2 and #3 in Table Four.

Relevance analysis is also useful, although not determinative, in the dilution context. One must still look to actual content to determine whether dilution is applicable. Unlike infringement, however, a dilution claim may stem from use of a meta tag alone, because no likelihood of confusion need be shown. If the index term is a famous mark and relevant to actual content, there are a number of possibilities. First, the meta tag might dilute, but the actual content — silent as to the meta tag — might not. For example, one might use “Coca-Cola” as a meta tag for a beverage-oriented Web page that does not mention Coca-Cola. Second, both meta tag and actual content may dilute. For example, the meta tag might say “Exxon,” and actual content might say “Exxon shoes.” Third, actual content may suggest that the meta tag, while relevant, is not being used with the same meaning as the plaintiff’s famous mark, and does not dilute. For example, one might use “coke” to index a site about the dangers of cocaine; here, the relevance is not to “Coca-Cola.”<sup>361</sup> Fourth, even if the meta tag does not dilute, actual content might. This might happen if somebody privileged to use a trademark uses the term properly as a meta tag, but oversteps his bounds in actual content. For example, a Buick car dealer might use “Buick” as a meta tag, but dilute in actual content by referring to Buick aspirin and Buick pizza sauce. Finally, even if the meta tag dilutes, an affirmative defense may be applicable based on actual content.

---

360. See *supra* notes 268–79 and accompanying text.

361. It might also be argued in such a case that a mark that is capable of multiple meanings is perhaps not so famous and therefore undeserving of dilution protection. The existence of alternative uses for the same or similar marks by others is a factor that may be considered in the determination of whether a mark is sufficiently famous for dilution purposes. See 15 U.S.C. § 1125(c)(1)(G) (Supp. I 1995). Of course, “Coke” is about as famous as a trademark can get (although not as distinctive as “Coca-Cola”), which shows that no one particular factor is determinative.



When the famous name is used as a meta tag irrelevant to actual content, one faces the situation mentioned above regarding initial interest confusion.<sup>362</sup> Here, actual content must be examined to confirm the meta tag’s irrelevance. Unlike trademark infringement, which looks to consumer expectations, dilution looks at the effect of spamdexing on the trademark itself: does it cause dilution by “lessening the capacity of a famous mark to identify and distinguish goods or services”?<sup>363</sup> By rendering a famous name less distinctive as an indicator of source or origin, dilution is probable. Table Five illustrates how relevance analysis informs the determination of infringement and dilution.

Table Five: Interplay Between Relevance and Infringement/Dilution

	Meta tag relevant to actual content	Meta tag irrelevant to actual content
Infringement	Look to actual content to see if it creates likelihood of confusion (scenarios #2 and #3 in Table Four).	Infringement is highly unlikely (or even impossible)
Dilution	1) Meta tag dilutes, but actual content does not; 2) Both meta tag and actual content dilute; 3) Meta tag, based on actual content, is used in a different sense than trademark, and does not dilute; 4) Meta tag does not dilute, but actual content does; or 5) Actual content suggests an affirmative defense.	Dilution is highly likely (or even inevitable).

Dilution is not without problems. It risks conferring property rights in gross, granting trademark owners rights against everyone, whereas infringement, as a quasi-property right, is available only against competitors.<sup>364</sup> Thus, dilution rewards trademark holders for their efforts in establishing a famous mark, rather than merely regulating marketing

---

362. See *supra* notes 337–53 and accompanying text.  
363. 15 U.S.C. § 1127 (Supp. I 1995).  
364. See SANDERS, *supra* note 309, at 79; Lisa H. Johnston, *Drifting Toward Trademark Rights in Gross*, 85 TRADEMARK REP. 19 (1995); Klieger, *supra* note 317, at 851; Jonathan E. Moskin, *Dilution Law: At a Crossroads? Dilution or Delusion: The Rational Limits of Trademark Protection*, 83 TRADEMARK REP. 122, 132 (1993).

efforts.<sup>365</sup> Significantly, courts have not paid close attention to the fame requirement, finding almost any registered trademark to be sufficiently famous for protection.<sup>366</sup> Overreaching application of trademark dilution to less-than-famous marks might produce a “diminishing” effect, by chilling the use of allegedly proprietary indexing terms.<sup>367</sup>

#### 4. Protection for Less-Than-Famous Marks?

Thus, dilution protection should be limited to only truly famous marks. Still, those holding less-than-famous marks may be without a remedy in cases that do not involve a likelihood of confusion. It might be argued that cases involving irrelevant spamdexing of less-than-famous marks be considered a new category of unfair competition under section 43(a) of the Lanham Act. After all, unfair competition has historically adapted to situations where the competitive injury is novel or cumulative.<sup>368</sup> As mentioned before, the spamdexing injury from the consumer’s point of view is diffuse.<sup>369</sup> This may also be said of the injury to a trademark holder. Each separate act of spamdexing may have a small effect, but together, they add up to a large injury.

Should quasi-dilution protection under section 43(a) be extended to less-than-famous marks? The mere fact of trademark ownership cannot, by itself, be proof of the need for trademark protection: to conclude otherwise would be to engage in circular reasoning.<sup>370</sup> Felix S. Cohen

365. See David Sven Villwock, *The Federal Trademark Dilution Act of 1995*, 6 DEPAUL-LCA J. ART & ENT. L. 213, 223 (1996); see also SANDERS, *supra* note 309, at 103.

366. See Klieger, *supra* note 317, at 849–51. Unless courts limit the scope of “famous” to marks that have more national scope, the fame requirement will not be much of a limitation. Courts should pay greater attention to the listing of factors in the Dilution Act. See 15 U.S.C. § 1125(c)(1)(A)–(H) (Supp. I 1995). The House Report on the Dilution Act also suggests that “famous” marks are limited to those national in breadth, such as “Kodak” or “DuPont.” See H.R. REP. NO. 104-374, at 3 (1995), *reprinted in* 1995 U.S.C.C.A.N. 1029, 1030.

367. See BOYLE, *supra* note 27, at 114, 119.

368. See *supra* notes 310–14 and accompanying text.

369. See *supra* note 281 and accompanying text.

370. As Felix S. Cohen stated:

[Such reasoning] purports to base legal protection upon economic value, when, as a matter of actual fact, the economic value of a sales device depends upon the extent to which it will be legally protected. . . .

The circularity of legal reasoning in the whole field of unfair competition is veiled by the “thingification” of property.

Cohen, *Transcendental Nonsense*, *supra* note 14, at 815; see also SANDERS, *supra* note



noted over fifty years ago that terms such as “property,” “due process,” and “fair value,” rather than being defined empirically, were themselves used to answer empirical and ethical questions.<sup>371</sup> Therefore, by limiting dilution protection to famous marks only, it may well be that Congress has made a determination that some kinds of injuries do not merit a remedy.<sup>372</sup>

Considering the potential chilling effects on webmasters, we should be cautious to broaden dilution liability. In those cases where a webmaster is tempted to spamdex with irrelevant terms, he or she generally does so because the term is a famous, information-rich name or trademark. Spamdexers who use irrelevant meta tags tend to use either popular generic terms or extremely famous names, rather than less-than-famous marks.<sup>373</sup> After all, spamdexing is a numbers game — the rational spamdexer will include famous rather than less-than-famous marks.<sup>374</sup> In such cases, dilution is appropriate.

Conversely, it is doubtful that less-than-famous names would appeal to webmasters for purposes of irrelevant spamdexing. Put differently, those whose businesses or actual content is relevant — even indirectly — to a less-than-famous name might be tempted to spamdex, because they want to reach the same audience as the owner of the less-than-famous term. In such cases, trademark infringement is appropriate, probably under scenario #2 of Table Four (meta tag plus actual content combine for actual infringement).<sup>375</sup> But webmasters with actual content having nothing whatsoever to do with a less-than-famous term have little to gain by using such a term.

Nevertheless, one must assume that some spamdexers will use irrelevant, less-than-famous marks. These are the hard cases. Consider the *Advanced Concepts* dispute, in which the spamdexer used a remotely related, less-than-famous mark.<sup>376</sup> Assuming the defendant’s use is highly unrelated to its actual content, then a likelihood of confusion may not exist; because Oppedahl and Larson’s names may not be sufficiently

---

309, at 85.

371. See Cohen, *Transcendental Nonsense*, *supra* note 14, at 820.

372. See H.R. REP. No. 104-374, at 3 (1995), *reprinted in* 1995 U.S.C.C.A.N. 1029, 1030.

373. The most popular search terms on the Internet are generic. See *supra* note 166. While the popularity of a search term does not necessarily translate into popularity as an indexing term, it is not an unreasonable inference to make.

374. See O’Rourke, *supra* note 21, at 302. Actually, the rational spamdexer uses popular trademarks and trade names (“Princess Diana”) as well as popular generic terms (“free,” “love,” “money,” etc.).

375. See *supra* note 348 and accompanying text.

376. See *supra* notes 122–30 and accompanying text.

famous, dilution also might not apply. Although it may be tempting to broaden liability to punish such free riding, or to create a new category of liability, to do so could have chilling effects in cases with less plaintiff-friendly facts. It is therefore probably better that courts deal with these cases under traditional trademark infringement or other established concepts of liability, rather than under quasi-dilution.<sup>377</sup> First, dilution does not directly consider the concerns of the public; therefore, there is a risk that courts will not perform the balancing of public interests that is traditionally required when creating new categories of unfair competition.<sup>378</sup> Second, as a well-established category of liability, trademark infringement has firmly established principles that help to protect consumer interests.<sup>379</sup> Courts must consider the public interest in infringement analysis; similarly, they must also take care to consider the public interest by considering whether defenses to infringement and dilution apply, as discussed below.

## 5. Defenses to Infringement and Dilution

There are several defenses applicable to spamdexing. If defendant's use is irrelevant to actual content, it is hard to imagine fair use ever being applicable.<sup>380</sup> Fair uses, such as descriptive use, nominative use, or use

---

377. One commentator suggests that possible causes of action include interference with prospective business relations, interference with contractual relationships, misappropriation, deceit, and defamation or disparagement. *See Halpern, supra* note 24. In the case of misappropriation of names, another commentator suggests common law or statutory right of publicity, false light, and false endorsement. *See Kamarck, supra* note 311, at 14–15. As noted before, it is beyond the scope of this Article to provide analysis of the multitudes of causes of action. *See supra* note 311. Still, the policy concerns of this Article should apply equally well to other torts. For example, although fraud is quite different from trademark infringement, determining scienter should look to the good faith use of the meta tag, which in turn requires looking at the relevance of the term and the reasonable expectations of the webmaster. In the case of misappropriation, a key consideration would be whether the user had a privilege to use the name, which would require a similar analysis.

378. *See* RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 1 (1995); *see also supra* notes 315–25 and accompanying text.

379. *See* Karen S. Frank, *The Federal Trademark Dilution Act of 1995 Protecting Fame and Fortune*, 454 PLI/PAT. 523, 526 (1996) (“[W]here traditional trademark principles focus on protecting the consumer from confusion, the anti-dilution statute focuses on the trademark owner’s investment in its mark and the goodwill the mark has earned.”); *see also* O’Rourke, *supra* note 21, at 306.

380. Note, however, that a defendant might prevail on an equitable defense, such as laches or acquiescence, or on a showing that plaintiff had abandoned his mark, or that the mark was invalid. *See, e.g.,* 15 U.S.C. § 1127 (1994) (defining abandonment); *Lincoln Logs Ltd. v. Lincoln Pre-Cut Log Homes, Inc.*, 971 F.2d 732, 734 (Fed. Cir.



as other than as a trademark will therefore be applicable only when the defendant has used the plaintiff's trademark in some way relevant to actual content.<sup>381</sup> Assuming relevance exists, the next inquiry would be the nature of the relevance. If used as a form of sponsorship or affiliation, then no defense would apply. However, if, at the Web page itself, the term was used to describe the defendant, plaintiff, or their products/services, then the use might be fair.<sup>382</sup> The *Playboy*

---

1992) (defining laches and equitable estoppel). These defenses rest upon acts or omissions of the plaintiff, or defects in the trademark or its registration. As defenses that are independent to the conduct of the defendant, they are irrelevant to this Article's analysis of spamdexing.

381. The "classic fair use case [is] where the defendant has used the plaintiff's mark to describe the defendant's own product." *New Kids on the Block v. News Am. Publ'g, Inc.*, 971 F.2d 302, 308 (9th Cir. 1992); see also *Volkswagenwerk Aktiengesellschaft v. Church*, 411 F.2d 350, 352 (9th Cir. 1969) (holding that the defendant, who specialized in repairing Volkswagens, could use the plaintiff's trademarks as long as the use did not create a likelihood of confusion). There is also a statutory defense against incontestability status. See 15 U.S.C. § 1115(b)(4) (1994); 1 MCCARTHY, *supra* note 74, § 11:45. Some of the "fair use" defenses are not necessarily affirmative defenses, because they cannot coexist with a likelihood of confusion, which is a prerequisite for liability. See *id.* § 11:47. Regarding dilution, the requirement that defendant's use be that of a "trademark or trade name" suggests that descriptive uses (i.e., non-trademark uses) are also free from liability. See RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 25(2), cmt. I (1995); 3 MCCARTHY, *supra* note 74, § 24:103.

A variant is the "nominative" use defense, which is where the defendant uses the plaintiff's trademark to refer to the plaintiff. See *New Kids*, 971 F.2d at 308; see also Derek J. Westberg, Note, *New Kids on the Block v. News America Publishing, Inc.: New Nominative Use Defense Increases the Likelihood of Confusion Surrounding the Fair Use Defense to Trademark Infringement*, 24 GOLDEN GATE U. L. REV. 685 (1994). Defenses to dilution are somewhat similar, including noncommercial use and fair use in comparative commercial advertising or promotion. See 15 U.S.C. § 1123(c)(4)(A)–(B) (Supp. I 1995).

382. The Ninth Circuit's nominative defense provides a good example. In *New Kids*, the Ninth Circuit held that the use of the "New Kids" name in newspaper ads for 900-numbers to vote for one's favorite "New Kid" was fair nominative use of the trademarks, even if the purpose was money-making. See *New Kids*, 971 F.2d at 309. Having a trademark does not give the plaintiff complete control over how others may profit by making competitive or derivative products: "[T]he trademark laws do not give the New Kids the right to channel their fans' enthusiasm (and dollars) only into items licensed or authorized by them." *Id.*

At least one court addressing meta tag abuse has used the nominative defense to dismiss infringement and dilution charges. See *Patmont Motor Werks, Inc. v. Gateway Marine, Inc.*, No. C 96-2703, 1997 WL 811770, at \*3–\*4 (N.D. Cal. Dec. 18, 1997). Plaintiffs claimed that the defendant "used [plaintiff's federally registered trademark] 'in such a way' that websurfers were lured away" by introducing plaintiff's mark into search engines. *Id.* at \*4 n.7. The court chided the plaintiffs for failing to "specify how" this was done, see *id.*, but it is likely that defendant used the trademark as a meta tag.

*Enterprises, Inc. v. Welles*<sup>383</sup> case is a paradigmatic example. Although "Playmate" was a federally registered trademark belonging to plaintiff Playboy Enterprises, Inc. ("PEI"), it was also a "title[] which *Playboy* magazine awards to certain *Playboy* models, who then use the title to describe themselves . . . . [These titles] become[] part of their identity and add[] value to their name[s]."<sup>384</sup> Here, Welles, as a former Playmate of the Year, "used the trademark [sic] term *Playmate of the Year* to identi[f]y and describe herself."<sup>385</sup> The court refused to allow PEI to overreach with its trademark registration to stifle even direct competition by Welles: "The 'fair use' defense, in essence, forbids a trademark registrant to appropriate a descriptive term for his exclusive use and so prevent others from accurately describing a characteristic of their goods."<sup>386</sup>

The court looked to actual content to determine whether the use of meta tags was infringing. The key question was whether the

---

The nominative defense should be applied to dilution as well as infringement claims. There is support for this assertion. *Welles* quotes *New Kids* as support for fair use as a defense to trademark dilution. See *Playboy Enters., Inc. v. Welles*, 7 F. Supp. 2d 1098, 1105 (S.D. Cal.), *aff'd*, No. 98-55911, 1998 U.S. App. LEXIS 27739 (9th Cir. Oct. 27, 1998). Indeed, the rationale behind *New Kids* may apply even better in the dilution context, due to the danger of the broad "property" rights implicated by dilution protection. In the spamdexing context, the nominative defense might apply to sites that offer products that replace Microsoft Windows 98 or Netscape Navigator, and use those names as meta tags. Here, as long as there was no likelihood of confusion over source or affiliation, webmasters should be free to advertise their products as being compatible with these other products. Thus, indexing by the terms "Microsoft," "Microsoft Windows," and "Netscape" would be relevant. The fact that these terms make it more difficult for the public to find the real Microsoft does not change matters. Here, the use is relevant, and fair under trademark law; the fact that it makes the public's search more difficult only suggests that the consumer should use a more focused search.

383. 7 F. Supp. 2d 1098 (S.D. Cal.), *aff'd*, No. 98-55911, 1998 U.S. App. LEXIS 27739 (9th Cir. Oct. 27, 1998); see also *supra* notes 132-40 and accompanying text.

384. *Welles*, 7 F. Supp. 2d at 1102.

385. *Id.* at 1103.

386. *Id.* (quoting *New Kids*, 971 F.2d at 306 (quoting *Soweco, Inc. v. Shell Oil Co.*, 617 F.2d 1178, 1185 (5th Cir. 1980))). The court cited 15 U.S.C. § 1115(b)(4) (1994). See *id.* As the Court of Appeals for the Ninth Circuit put it in another case:

It is not disputed that [the defendant, an auto repairman specializing in Volkswagens,] may specialize in the repair of Volkswagen vehicles. He may also advertise to the effect that he does so, and in such advertising it would be difficult, if not impossible, for him to avoid altogether the use of the word 'Volkswagen' or its abbreviation 'VW,' which are the normal terms which, to the public at large, signify appellant's cars.

*Volkswagenwerk Aktiengesellschaft*, 411 F.2d at 352. The *Welles* court cites this case with approval. See *Welles*, 7 F. Supp. 2d at 1104.



“defendant . . . used plaintiff’s trademarks in good faith to index the content of her website.”<sup>387</sup> Therefore, using “Playmate” as a meta tag was fair use under 15 U.S.C. § 1115(b)(4) because it was an accurate description of the defendant: “It is clear that defendant is selling Terri Welles and only Terri Welles on the website.”<sup>388</sup> Further, the use of “Playboy” was also acceptable to the court because (1) it referred to Welles’ identity as a “*Playboy* Playmate;” and (2) because “it may also reference the legitimate editorial uses of the term *Playboy* contained in the text of defendant’s website.”<sup>389</sup> These same factors raised a fair use defense against trademark dilution.<sup>390</sup>

What is especially interesting about *Welles* are the factors that help separate this case from others that may not merit a fair use defense. The court was impressed by steps that Welles took, and refrained from taking, that helped to cut against a finding that she was trying “to trick consumers into believing that they are viewing a *Playboy*-endorsed website.”<sup>391</sup> The court found that Welles:

- Did not create a *Playboy*-related website;
- Did not use “Playboy” or “Playmate” in her domain name;
- Did not use the classic *Playboy* bunny logo;
- Inserted disclaimers to clearly indicate that her website was not endorsed by PEI; and
- Did not mimic the look and feel of *Playboy* (specifically, the font of the title “Playmate of the Year 1981” was not “recognizable as a *Playboy* magazine font”).<sup>392</sup>

The court was also impressed by factors that cut against a likelihood of confusion:

- A lack of empirical data showing actual confusion;
- Evidence suggesting good faith use; and
- The fact that the use of the meta tags was descriptive of (or, in the terms used by this Article, *relevant to*) the plaintiff’s

---

387. *Welles*, 7 F. Supp. 2d at 1104.

388. *Id.* Note that the defense under 15 U.S.C. § 1115(b)(4) (1994) applied to both the trademark infringement and false designation of origin claim. *See id.*

389. *Id.*

390. *See id.* at 1105.

391. *Id.* at 1103–04.

392. *Id.* at 1104.

products or services, or referenced (again, *relevant to*) actual content at defendant's site.<sup>393</sup>

The finding of good faith was inferred from factors that suggested fair use.<sup>394</sup> Put differently, a finding of descriptive or nominative fair use is essentially the same as finding good faith non-confusing use.<sup>395</sup> In any case, these factors, especially those that lean toward a finding of good faith, help to separate fair use cases from others such as *Playboy Enterprises, Inc. v. AsiaFocus International, Inc.*<sup>396</sup>

Although the *Welles* and *AsiaFocus* defendants offered similar services, many of the *Welles* factors were absent in *AsiaFocus*. This absence eliminated a fair use defense, created a likelihood of confusion, and suggested actual dilution. Thus, *AsiaFocus* and *Welles* might be considered to be at the opposite ends of the fair use spectrum. First, unlike Terri Welles, the *AsiaFocus* defendants registered the domain names "asian-playmates.com" and "playmates-asian.com."<sup>397</sup> Second, they apparently did not insert disclaimers of sponsorship, endorsement, or affiliation with PEI.<sup>398</sup> Third, they offered merchandise, such as key chains, calendars, wrist watches, and playing cards that used PEI trademarks in part.<sup>399</sup> Fourth, there was evidence of actual confusion.<sup>400</sup> Finally, unlike *Welles*, they did not have a special need or good faith reason to use PEI trademarks as indexing terms. Thus, the court found that *AsiaFocus*'s use was willful: "No other purpose appears for choosing PLAYMATE and PLAYBOY but to create that false association in the mind of the consuming public."<sup>401</sup> The court was especially upset by the use of the meta tags, finding it to be further

---

393. *See id.* at 1104–05.

394. *See id.* at 1104. Indeed, the successful use of a descriptive fair use defense suggests that a likelihood of confusion is not possible. Nevertheless, the court considered in the alternative whether there was a likelihood of confusion, and determined that there was not. *See id.* Note, however, that a finding of fair use is not necessarily equivalent to a finding of a lack of dilution. After all, if every former Playmate were to establish a website using PEI's trademarks as meta tags, the ability of those marks to serve as keywords could be diluted, however fairly.

395. *See id.*

396. No. 97-734-A, 1998 U.S. Dist. LEXIS 10359 (E.D. Va. Feb. 2, 1998) (Mag. J.), *adopted by* 1998 U.S. Dist. LEXIS 10459 (E.D. Va. Apr. 10, 1998); *see also supra* notes 141–45 and accompanying text.

397. *See id.* at \*6–\*7.

398. The decision did not state whether disclaimers were used.

399. *See AsiaFocus*, 1998 U.S. Dist. LEXIS 10359, at \*8.

400. *See id.* at \*9. PEI "presented evidence of at least one instance of actual consumer confusion." *Id.* at \*18.

401. *Id.* at \*21.



evidence of defendant's willfulness.<sup>402</sup> The finding of willfulness was extremely damaging to defendants — the magistrate judge recommended awarding PEI maximum statutory damages under 15 U.S.C. § 1117(c)(2), as well as costs and reasonable attorneys' fees, for a total of \$3 million.<sup>403</sup>

Thus, in the spamdexing context, the fair use defense seems to implicitly incorporate an element of necessity, or at least a requirement of a high degree of relevance between actual content and meta tag. Terri Welles could not effectively market herself if she could not assert her status as a former Playmate of the Year. The *AsiaFocus* defendants, however, could market themselves without the use of PEI's trademarks. Certainly the terms "Playboy" and "Playmate" are arguably indirectly relevant as terms meaning "adult content." Unlike in *Welles*, however, these terms are not necessary for public understanding of who AsiaFocus is. Further, the index terms are not highly relevant.

Under this rationale, the *Advanced Concepts* defendants would also be hard-pressed to show a necessity to use the plaintiff's names as index terms. The fact that their spamdexing makes them easier to find is insufficient. Thus, in the case of a competitive injury, if infringement or dilution is shown, the meta tag must have an element of necessity or a high degree of relevance to actual content to establish fair use. On the other hand, if the *Advanced Concepts* defendants had good faith editorial content about the plaintiffs, they would have a much stronger case of fair use.<sup>404</sup> Finally, even if a fair use defense is otherwise applicable, it is still possible to imagine scenarios where the use of the meta tag would nonetheless be actionable — such as by stuffing, or by using overly remote meta tags.<sup>405</sup>

---

402. *See id.*

403. *See id.* at \*22–\*25.

404. Contrast a good faith use, such as Pepsi using meta tags on a page that contains the results of a Pepsi-Coke taste test, with that of CityAuction, which used text referring to Monica Lewinsky and attached meta tags in what appears to be a somewhat questionable attempt to appear in queries for her. *See supra* note 5 and accompanying text; *see also* Search Engine Watch, *supra* note 196.

405. Continuing the hypothetical posed in the previous footnote, suppose Pepsi placed the results of this hypothetical study on a sub-page at its World Wide Web site. *See* Search Engine Watch, *supra* note 196. Under the analysis of this Article, Pepsi may have the right to use "Coke" and "Coca-Cola" as indexing terms *at that page*. However, this does not give it the right to stuff these terms repeatedly. *Cf. supra* note 102 and accompanying text. Further, whatever relevance that use of "Coke" may have does not extend to every page at the Pepsi website. Pepsi would also have a much harder time showing that use of these terms at its main page would bear the same relevance.

First Amendment concerns supply a second defense to trademark infringement and dilution.<sup>406</sup> The Internet arguably deserves greater protection than traditional media.<sup>407</sup> Courts should think twice before precluding a First Amendment defense in cases where the spamdexer uses meta tags that are directly or indirectly related to actual site content.<sup>408</sup> It bears noting, however, that First Amendment protection, even when applicable, will extend only to relevant uses of meta tags; irrelevant uses, as deceptive speech, do not merit free speech protection.<sup>409</sup>

An extended discussion of the First Amendment is beyond the scope of this Article. However, the history and rationale behind it may well be intertwined with the progressing history of information technology. As we are faced with new media, we must realize that as technology shapes what kinds of information are possible, these new kinds of information in turn may call into question the values that inform what we characterize as "First Amendment" concerns. That which may be framed in First Amendment terms also implicates a broader range of concerns:

---

406. See 15 U.S.C. § 1125(c)(4) (Supp. I. 1995); see also Filippo M. Cinotti, "Fair Use" of Comparative Advertising Under the 1995 Federal Dilution Act, 37 IDEA 133 (1996).

407. See *Reno v. ACLU*, 117 S. Ct. 2329, 2343 (1997). The Communications Decency Act, found unconstitutional in *Reno*, "cast[] a . . . dark[] shadow over free speech [and] threaten[ed] to torch a large segment of the Internet community." *Id.* at 2350; see also Keith Aoki, *How the World Dreams Itself to be American: Reflections on the Relationship Between the Expanding Scope of Trademark Protection and Free Speech Norms*, 17 LOY. L.A. ENT. L.J. 523, 546-47 (1997) (arguing that overly broad protection of trademarks violates the First Amendment by leaving Corporate America as "the only institute that is allowed to impregnate text with meaning").

408. James Boyle criticizes courts for being "traditionally much less sensitive to First Amendment, free speech and other 'free flow of information' arguments when the context is viewed as private rather than public, or property rather than censorship." Boyle, *Politics*, *supra* note 27, at 98; see also Keith Aoki, *Adrift in the Intertext: Authorship and Audience "Recoding" Rights — Comment on Robert H. Rotstein, "Beyond Metaphor: Copyright Infringement and the Fiction of the Work"*, 68 CHI.-KENT L. REV. 805, 832 (1993) (stating that "a trademark owner's increasingly concretized property-like rights in a mark have frequently trumped free speech concerns in several state law anti-dilution cases which have ruled against 'recodings,' or subsequent unauthorized uses of marks").

409. See David J. Goldstone, *A Funny Thing Happened on the Way to the Cyber Forum: Public vs. Private in Cyberspace Speech*, 69 U. COLO. L. REV. 1, 56 (1997); Rosalie Berger Levinson, *State and Federal Constitutional Law Developments*, 30 IND. L. REV. 965, 994 (1997).



In addition (and sometimes in opposition) to First Amendment values, our information policy incorporates concerns ranging from protecting children to vindicating injured reputations, defending individual privacy, preserving the value of business property, preventing misleading advertisements, ensuring the fairness of elections, and guarding our national security.<sup>410</sup>

Indeed, James Boyle laments the lack of a politics of intellectual property, the “legal form of the Information Age.”<sup>411</sup> Part V attempts to supply a tapestry for these difficult policy choices by examining the broader concerns of information dissemination and retrieval. It examines the infoglut and spamdexing in light of the historical context of information dissemination and retrieval as discussed in Part II, supplying a landscape to help delineate the difficult border-drawing questions that arise in the context of spamdexing liability.

## V. WHERE DO YOU WANT TO GO *TOMORROW*?: TAMING THE INTERNET INFOGLUT

*With over 100 million World Wide Web pages on-line, the Internet is a virtual Library of Congress. Now it's about to get a few more librarians.*

— *Wall Street Journal*, Nov. 6, 1997<sup>412</sup>

By introducing a tremendous amount of noise into the Internet, spamdexing harms many. By externalizing a webmaster's costs of getting noticed, spamdexing harms everyone by making information retrieval more difficult. Ironically, because so many webmasters spamdex, they collectively raise the information noise to a level where spamdexing helps few. Essentially, spamdexers externalize their advertising costs over the short run, only to re-internalize those costs — and more — in the end.

These costs extend far beyond the pecuniary. It is essential that anybody with a basic level of education be able to navigate the Internet.

---

410. Jessica Litman, *Reforming Information Law in Copyright's Image*, 22 U. DAYTON L. REV. 587, 591–92 (1997) (footnotes omitted).

411. Boyle, *Politics*, *supra* note 27, at 87–90.

412. Rebecca Quick, *New “Filters” Could Help Narrow Searches*, WALL ST. J., Nov. 6, 1997, at B6.

Although the digital revolution promises great benefit to society, history has shown that technological revolutions reverberate with the potential for conflict between the “deserving” and the disenfranchised, all too often leading to human tragedy on a grand scale.<sup>413</sup> Spamdexing, by making it more difficult to navigate this expanded world of information, widens the gap between the technological elite and “technofailures,” fueling the uncertainty that human revolution inevitably brings.<sup>414</sup> Accessibility, however, has a democratizing effect.<sup>415</sup> As we become more accustomed to navigating the infoglut, ease of access will become more important than ownership.<sup>416</sup> It is important that we not widen the gap between the digital savvy and technofailures, lest we contribute toward creating a new information underclass.

Katsh characterizes the problem of accessibility as one of “informational distance,” something inherent to any medium.<sup>417</sup> He suggests that if information is too difficult to access, it will not be found, or will be of little value because it will not be understood.<sup>418</sup> On the other hand, if one can locate information with relative ease, value is added.<sup>419</sup> On its face, one might take Katsh’s characterization to support

---

413. Michael S. Malone, *Forget Digital Utopia: We Could Be Headed for Technofascism*, UPSIDE, Aug. 1998, at 78, 136. Malone fears the social upheaval that will be caused by the moral arrogance that underlies the digital revolution:

[W]hat human revolutions give, they also take away. It’s polite to ignore the fact that the industrial revolution also killed the Enlightenment, set off the destructive and narcissistic counterforce called romanticism, buried us in soul-killing bureaucracies and, worst of all, gave us machine-age “total war.” It’s a single developmental thread that runs from Cold Harbor to Verdun to Stalingrad to Hue.

Here at the end of the most homicidal century in human history, the memory of millions of murdered innocents ought to be more than enough to make us wary of all the talk about the New Digital Man, Homo computatis.

*Id.* at 82.

414. Malone characterizes those who fall by the wayside in the so-called digital utopia “technofailures.” *See id.* at 136.

415. “We can obtain information that was previously distant, not only in the sense that it was far away but also in that it was inaccessible because special skills were required in order to access it.” KATSH, DIGITAL WORLD, *supra* note 27, at 14.

416. *See id.* at 208–09 (“Unlike in the past, access to some kinds of electronic information may provide more value than ownership and, therefore, may be more desired in the future than ownership or possession. For other kinds of information, access may be indistinguishable from ownership.”).

417. *See id.* at 51.

418. *See id.* at 52.

419. *See id.* at 54.



the hacker mantra “information wants to be free.”<sup>420</sup> Narrowing informational distance, however, is not a function of making information freely available without restraint. If this Article has shown anything at all, it is that having access to too much information is just as paralyzing as having access to too little. Freedom of information is of little use if one cannot find it. Thus, the process of filtering is as essential to adding value as is the process of dissemination.

The tension between filtering and freedom is exemplified by intellectual property law. As a society, we have long assumed that limited intellectual property protection — a form of filtering — will increase incentives, and therefore output.<sup>421</sup> Ironically, this assumption, which is the cornerstone of intellectual property law, is empirically unprovable.<sup>422</sup> Katsh makes the seemingly contrary, yet similar assumption that if “a greater variety of information becomes available to us without our going anywhere for it, new opportunities and ways to use information arise that were not previously available.”<sup>423</sup> Katsh’s assumption is equally unprovable. This is not to say that either approach is incorrect. Indeed, variants of Katsh’s assumption are now receiving well-deserved scholarly attention.<sup>424</sup> Rather, Katsh’s assumption mirrors the tension implicit in the Copyright Clause and embodied throughout intellectual property law — how to maximize the amount of information available to society.

Encouraging accessibility, Katsh correctly recommends the development of new searching technologies and user skills.<sup>425</sup>

---

420. Stewart Brand is generally credited with this statement, now a mantra of hacker culture. See John Perry Barlow, *The Economy of Ideas*, WIRED, Mar. 1994, at 84, 86, available at <<http://www.wired.com/wired/archive/2.03/economy.ideas.html>>; see also STEWART BRAND, *THE MEDIA LAB: INVENTING THE FUTURE AT M.I.T.* 202 (1987); John Perry Barlow, *A Declaration of the Independence of Cyberspace* (Feb. 9, 1996) <<http://www.atd.ucar.edu/rdp/gfc/barlow.html>>. Brand wryly notes that “[i]nformation also wants to be expensive,” because of its potential value to the holder of proprietary rights. BRAND, *supra*, at 202.

421. The plain language of the Copyright Clause bears this out. See U.S. CONST. art. I, § 8, cl. 8 (securing to Congress the power to “promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries”).

422. See Julie E. Cohen, *Intellectual Privacy and Censorship of the Internet*, 8 SETON HALL CONST. L.J. 693, 693 (1998).

423. KATSH, *DIGITAL WORLD*, *supra* note 27, at 54.

424. See, e.g., Lemley & McGowan, *supra* note 30 (discussing network effects).

425. Despite one’s physical (or virtual) closeness to information, the difficulty of accessing such information without “special skills that only a few may possess” may leave it “difficult to retrieve, use, or understand.” KATSH, *DIGITAL WORLD*, *supra* note 27, at 59 (referring to communications theorist Joshua Meyrowitz) (citation omitted); see

Maximizing the benefit to society requires finding a balance where each party — webmaster, proprietary-rights holder, and the public — internalizes its reasonable share of the costs of information dissemination and retrieval. First, webmasters should not spamdex, because it vastly increases the amount of Internet noise, externalizing the publicity costs of websites.<sup>426</sup> Second, granting overly broad proprietary rights also forces others to internalize the costs of searching and being found. Searchers will have a much harder time finding information due to underinclusive searches, and webmasters will have a difficult time being found if they cannot, when appropriate, use the trademarks and names of others. Third, the public must be expected to develop reasonable searching skills. They must internalize some of the difficulties of finding information by searching in a reasonable manner, such as using multiple terms and Boolean connectors. They must use information-rich terms, not generic terms or terms with multiple meanings. By seeking a proper balance between each party, we maximize the accessibility of information.

To explore the issue of accessibility, it is instructive to look at the historical role of libraries. Libraries have always acted as a form of control over information.<sup>427</sup> Traditionally, they performed the twin goals of preserving and providing access to information.<sup>428</sup> Whereas the

---

*also* Katsh, *New Media*, *supra* note 27, at 1465 (“Print, as compared to writing, provided the reader not only with new resources, but with new skills.”). Well before the advent of the World Wide Web, information professionals noted that for online end-user access to become widespread, (1) end users must become “information savvy;” and (2) online systems must become “user friendly.” FENICHEL & HOGAN, *supra* note 273, at 4–5. If everyone has the basic skills and tools for “efficient information accessibility and management,” then everyone has the opportunity to add value to the pool of information. KATSH, *DIGITAL WORLD*, *supra* note 27, at 211. It is therefore essential to encourage the development of useful tools, habits, and concepts to deal with the electronic culture. *See id.* at 78, 80. “The ability to use information retrieval systems in today’s society is tantamount to literacy; for if you cannot locate what you need you cannot read it, and, from that perspective, it might as well not exist.” Summit, *supra* note 1, at 102.

426. *See* SHENK, *DATA SMOG*, *supra* note 35, at 102.

In the immediate sense, pumping up the volume is an extremely effective solution. More broadly, though, it becomes part of the problem, feeding a vicious spiral in which the data smog gets thicker and thicker and the efforts to cut through the smog get ever more desperate. As the people of earth collectively try to rise above the noise, they unwittingly create more of it.

*Id.*

427. *See* Robert C. Berring, *Future Libraries*, in *FUTURE LIBRARIES* 101 (R. Howard Bloch & Carla Hesse eds., 1995).

428. *See id.* (“Formerly librarians assumed the roles of guardians of books, first gathering the books together, then protecting them by controlling both the books and



traditional librarian was concerned with gathering and archiving information, the modern librarian is more concerned with sorting and filtering it.<sup>429</sup> The librarian-as-filter<sup>430</sup> serves the same public interest discussed throughout this Article, by helping the information consumer find relevant information (maximizing recall by finding the relevant) while filtering out the noise (maximizing precision by eliminating the irrelevant).<sup>431</sup> In short, today's librarian is essentially a highly trained sailor trying to navigate the seas of the infoglut.<sup>432</sup>

---

access to them.”).

429. *See id.* at 98–99.

The great explorer librarians who traveled the world searching out rare collections and finding odd titles are now a tiny band indeed. Each year there is a river of new books. No library can own it all. The function of gathering has been transformed into the challenge of choosing. . . . The problem is no longer finding information, but sorting it, filtering it.

*Id.*; *see also id.* at 99 (noting that traditionally “[t]here were always physical threats to books [such as] . . . vermin, fire, water, even librarians,” but that today’s librarian is a filter, not an archivist).

430. The public’s interest in filtering to enable retrieval of relevant information is the flip side of a contentious issue facing libraries today: filtering out pornography, by targeting information to prevent retrieval. Filtering in libraries has become a politically charged issue. Many decry filtering software as both over- and underinclusive, and a threat to free speech. *See* Noah Robischon, *Porn to Lose — Maybe: Summit on Internet Free Speech*, ENT. WKLY., Jan. 9, 1998, at 81 (“Free-speech groups believe that website-blocking software and self-rating systems could lead to censorship, especially for Net news sites.”); Elizabeth Wasserman, *On-Line Smut Filters Are Not Infallible*, THE RECORD (Bergen Co., N.J.), Jan. 5, 1998, at H9 (“The assortment of software-filtering products and browser-supported ratings systems available for parents to shield their children from Net pornography, hate messages, and chat-room perpetrators are very much like parents themselves: None is perfect.”).

431. *See supra* Part III.A.

432. *See supra* Part II. The tension between librarian-as-filter, librarian-as-preserver, and librarian-as-access provider explains, from a social perspective, the conflict that librarians must feel over whether they should employ filtering software to block access to pornography on library computers. Librarians, by choosing which books to select, have always acted as a filter. Only after a book is in a collection do the preservation and access duties exist. The instant access to anything over the Internet, however, arguably makes the Web already part of all libraries’ collections, which leads to the struggle over whether the duty to provide such access exists in all cases. *See* *Mainstream Loudoun v. Board of Trustees of the Loudoun County Library*, No. Civ.A. 97-2049-A, 1998 WL 822105 (E.D. Va. Nov. 23, 1998).

The Internet has often been characterized as a huge library.<sup>433</sup> Although appealing, this comparison is flawed.<sup>434</sup> If the Internet is a library, then it is one with all the books scattered on the floor.<sup>435</sup> Unbridled chaos could render the World Wide Web useless as a form of discourse<sup>436</sup> unless we can reasonably rely on our retrieval and filtering tools. Search engines, trying to meet this need, are essentially turbo-charged card catalogs, allowing great flexibility in how information may be indexed and retrieved.<sup>437</sup> Like librarians, search engines act to collect, index, and filter information subject to a query. Search engines index the meta tag, title tag, and description tag fields; the URL; links

---

433. See, e.g., Peter Coy et al., *Has the Net Finally Reached the Wall? America Online's Crash May Portend Constant Crises Unless the Internet Is Revamped*, BUS. WK., Aug. 26, 1996, at 62; James S. Derk, *19th Century Nerd Deserves Credit for Refining Net Search*, COMM'L APPEAL (Memphis, TN), Sept. 25, 1998, at C4.

434. Libraries, by filtering, can prioritize information. Search engines try this, but obviously with limited success. Libraries preserve information; search engines do not. Indeed, one often clicks on a link from a search engine only to find that the site no longer exists. One recent and notable exception is Alexa <<http://www.alexa.com>>, which is engaging in the ambitious project of archiving the entire Internet. In theory, if one tries to go to a dead link, one could access an archived version of that page from Alexa's massive database. Alexa's practice of archiving entire sites that no longer exist raise significant and interesting issues of copyright and trademark infringement and fair use — issues that deserve commentary in another article.

435. See Derk, *supra* note 433. The same sentiment has been expressed about libraries themselves. Harvard President Neil Rudenstine notes that after the Civil War, a publishing boom caused similar information overflow problems at Harvard's libraries. See James S. Derk, *Wired For the Future — Taming the Internet May Transform Our World*, ARIZ. REPUBLIC, June 11, 1996, at B5. Interestingly, Rudenstine comments that the real problems of that time "were not those of space and money. They were organizational and conceptual." *Id.* The author of the article in which Rudenstine is quoted concludes that "the Internet is at the same stage of development as Harvard's library in the 1870s." *Id.*

436. Recall the example of Internet newsgroups: overparticipation without filtering made them too chaotic to use efficiently. See Lou Dolinar, *At Their Best, Newsgroups Are a Forum For Exchanging Ideas*, BUFFALO NEWS, Dec. 2, 1997, at D8.

[U]nder comp, we'd see comp.sys (sys for computer systems) running from comp.sys.acorn, comp.sys.ibm.pc, comp.sys.mac, comp.sys.sun and so on. This is an odd taxonomy for the sum total of human existence, but hey, that's the way the Internet works. The indexing system is generally so screwy that it often makes more sense to use a search engine to look up subjects on UseNet that interest you.

*Id.* On the other hand, new technologies that allow filtering, such as DejaNews <<http://www.dejanews.com>>, a newsgroups search engine, can breathe new life into the infoglut.

437. See GILSTER, *supra* note 48, at 165–66.



contained within; and more.<sup>438</sup> Further, unlike a traditional library, many search engines allow full-text searching.<sup>439</sup> Not only do multiple indexing fields allow very powerful search capabilities, but the ability to search full text is unprecedented.<sup>440</sup> Thus, to the extent that search tools are effective, the Internet may function similarly to traditional libraries, plus offer much more.<sup>441</sup>

However, search engines, like card catalogs, are not suited to every searching need.<sup>442</sup> Just as searching a card catalog limits one to looking at title, author, or concept, search engines are also best suited to targeted searches, especially those involving information-rich terms.<sup>443</sup> This is because one has a reasonable expectation that targeted index terms have a limited scope.<sup>444</sup> A library, of course, is much more than its card catalog—it also has trained librarians, “what’s new” shelves, paperback racks, and separate sections for science, social sciences, law, music, and children’s books. Each is appropriate for different searching needs, reminding us that not all forms of organization are appropriate for all kinds of searches.

Similarly, the Internet offers other ways to retrieve information besides search engines.<sup>445</sup> These forms also vary in their appropriateness

438. See *supra* notes 81–83 and accompanying text.

439. See, e.g., Altavista <<http://www.altavista.com>>; Lycos <<http://www.lycos.com>>.

440. See GILSTER, *supra* note 48, at 176.

441. See Carolyn R. Pool, *A New Digital Literacy: A Conversation With Paul Gilster*, EDUCATIONAL LEADERSHIP, Nov. 1, 1997, at 6 (quoting Paul Gilster, author of *Digital Literacy*, *supra* note 48, as saying “the way we find [Internet] information is different from the way we use a card catalog, check out a book, buy a magazine, or sit down to read on a rainy day. [The Internet] enables people to truly construct information . . .”).

442. See *id.* (quoting Gilster as stating that “[t]he Internet provides us something like a library of information online, though I’m hesitant to use the term library”); see also Billy Barron, *Symbiotic Cyberspace Libraries*, in THINKING ROBOTS, *supra* note 33, at 157 (noting that a card catalog is a difficult way to find very specific information, such as how to remove battery acid from a counter-top).

443. See *supra* notes 164–68 and accompanying text and tbls. 1–2.

444. See David Lidsky & Larry Seltzer, *Fuel-Efficient Web Searching*, PCMag., Dec. 16, 1997, at 29 (“The key is to think like the search engine. Avoid words that occur too frequently. Most engines accept quotes for exact phrases (or let you specify an exact search via the interface), and that’s also an effective and simple means of filtering out unwanted junk.”); see also Brock, *supra* note 29.

445. For example:

- “Hot” lists — see *Editorial: Furnishing the Internet*, WASH. POST, Nov. 8, 1997, at A24 (noting that the “Encyclopedia Britannica, the National Endowment for the Humanities and the American Library Association” have published lists of recommended sites);
- “Cool” sites — see, e.g., *Cool Site of the Day* <<http://cool.infi.net>>. This site is “generally acknowledged as the originator of the ‘Cool Site’ genre.”

to different kinds of searches. These alternate forms of dissemination and retrieval are important for several reasons. First, the existence of alternatives undercuts arguments that webmasters need to spamdex to be noticed. Second, it equally undercuts the reasonableness of any public expectation of using information-poor terms or inefficient searching techniques. Third, because proprietary-rights holders have other, more effective ways of being found — such as domain names or print advertising — courts must be wary of overblown claims of spamdexing liability.

As discussed in Part III, search engines are well-suited to finding targeted information with information-rich terms. However, categorical and relational terms are less useful because, due to their vagueness, ambiguity, or multiplicity of meaning, they are often information-poor.<sup>446</sup> Such terms can still be useful if the public narrows its searching with reasonable searching techniques, such as combining multiple terms.<sup>447</sup> For instance, the Boolean search “law OR computers” uses

---

SELLERS, *supra* note 11, at 54;

- “Top Ten” or award sites — *see id.* at 57;
- E-zines — *see, e.g.*, HotWired <<http://www.hotwired.com>>; CNET <<http://www.cnet.com>>;
- Advertising (traditional and Internet-based) — *see* ActivMedia, *supra* note 37;
- Magazines — *see, e.g.*, INTERNET WORLD;
- WebRings — WebRings allow persons who create similar websites to be connected in a “ring” that, in turn, allows users to go from one site in the WebRing to the next by clicking on an arrow. *See* Kathryn Lively, *Ring Around the Internet*, INFO. TODAY, Nov. 1, 1996, at 16. There are more than 10,000 rings comprising more than 80,000 sites so far. *See* Margot Williams, *How to Run Rings 'Round the World Wide Web*, WASH. POST, May 19, 1997, at F19. The WebRing site is at <<http://www.webring.org>>;
- Books — *see, e.g.*, CYBERHOUND’S INTERNET GUIDE TO THE COOLEST STUFF OUT THERE (1996); and
- “Push” content — This includes providers such as PointCast, which provides proprietary software that allows users to automatically receive information from news providers that a user is interested in, such as the New York Times, Ziff-Davis Net, People Magazine, and CNN. *See* PointCast <<http://www.pointcast.com>>. In their latest incarnations, both Microsoft Internet Explorer 4.0 and Netscape Navigator/Communicator also allow content to be downloaded automatically to one’s computer. *See* Microsoft <<http://www.microsoft.com>>; Netscape <<http://www.netscape.com>>.

446. *Cf.* Chaffee, *supra* note 14, at 388–89 (“We find abundant examples in law of the trouble caused by a word which is capable of standing for two or more different objects . . .”).

447. *See* Brock, *supra* note 29 (noting that the term “gates” may bring up a plethora of unrelated topics, and suggesting that “[t]he key is to narrow the search”).



information-poor terms and will certainly result in a low-precision, low-recall search. However, "law AND computers" is better. Using a phrase such as "computer law" is better yet.<sup>448</sup> Even then, terms such as "law" and "computers" are so broad that even the narrowed search of "computer law" will be overly broad and underinclusive, i.e., low in both precision and recall.

The type of information desired dictates the appropriate search strategies.<sup>449</sup> Sometimes, the Internet is literally the last place you want to start your search.<sup>450</sup> After all, you wouldn't expect to find your cousin's home phone number in the yellow pages. Similarly, you wouldn't look for a local pizza parlor by reading the white pages starting on page one. Thus, even when using the Internet, traditional search engines are not always the best place to start. If one is interested in computer law, for example, there are far better ways to research that topic than using a traditional search engine. One can go to FindLaw<sup>451</sup> or Law Journal Extra,<sup>452</sup> and search their collections of articles and links.

Thus, it becomes very important to consider search tools in light of their strengths and weaknesses. Search engines and their brethren are well-suited to targeted searches, which favor discrete information with information-rich terms, or information indexed in an alphanumeric or other linear form. The white pages is an archetypal non-electronic example.<sup>453</sup> Computer programs such as ProCD,<sup>454</sup> and Internet

---

448. One webmaster realized that by using specific indexing terms together, he might attract attention from employers looking for somebody with his particular experience. See Justin Martin, *So, You Want to Work for the Best . . . You'll Have to Sit Through Multiple Interviews, Pass a Battery of Tests, Tell Your Life Story — Maybe Even Come up with a Few Jokes*, FORTUNE, Jan. 12, 1998, at 77.

449. See Suzanne S. Bell, *Net Search Strategies*, INFO. OUTLOOK, Dec. 1997, at 17.

[If] users master various Web conventions, and differentiate amongst the various types of tools — subject lists, word-searchable indexes, searchable newsgroup archives — they will be able to adapt as new tools appear and current tools change. If they can analyze their question in order to make an informed choice about which tool to start with, their searching will be more efficient and, hopefully, successful.

*Id.*; see also ALFRED GLOSSBRENNER, HOW TO LOOK IT UP ONLINE 90–91 (1987) (recommending that those looking for a particular type of information consider "[w]ho would publish this kind of information?"); Brock, *supra* note 29; Lou Dolinar, *Search Engines Need to Know Where to Go*, TULSA WORLD, Nov. 15, 1997, at E8 (suggesting use of phrase searching, Boolean operators, URL "guesstimating," backtracking, etc.).

450. See ROSZAK, *supra* note 35, at 174 ("By virtue of their training and experience, [librarians know] when *not* to use the computer.").

451. <<http://www.findlaw.com>>.

452. <<http://www.ljx.com>>.

453. If looking for a plumber named "K.B. McFee, Jr.," the white pages is an entirely

databases such as WhoWhere<sup>455</sup> present the next generation of targeted search systems. In these systems, input is limited to discrete categories (name, address, city, e-mail address, etc.). Although targeted searching exists along a continuum from narrow to broad,<sup>456</sup> it is most effective when the search seeks one specific item. A name/address/city query in a telephone number database is one example. There is typically only one "John J.J. Schmidt" at "123 Main Street" in "Erie, PA." Domain names, where only one party can own "sony.com," is an excellent example of binary searching.<sup>457</sup> Legal citation is yet another: there is only one "7 F. Supp. 2d 1098."<sup>458</sup>

Categorical searching, however, often requires different techniques and tools than does targeted searching. Because the white pages, a targeted searching tool, are not at all suited to a categorical search, telephone companies developed the yellow pages, which organize businesses by category.<sup>459</sup> The success of yellow pages is rooted in its basic common sense. Thus, it should come as little surprise that the Internet directory Yahoo!, which is evocative of yellow pages, is the most popular and successful Internet search tool.<sup>460</sup> Yahoo!'s categories may not be exhaustive, or even well-maintained, but they allow easy

---

reasonable searching tool.

454. *See ProCD, Inc. v. Zeidenberg*, 86 F.3d 1447, 1449 (7th Cir. 1996).

455. <<http://www.whowhere.lycos.com/>>.

456. *See supra* tbl. 2; *see also supra* Part III.B.1.

457. People engage in binary searches by "guesstimating" domain names via the template [www.fill-in-the-blank.com](http://www.fill-in-the-blank.com). For example, Sony is at "www.sony.com." Domain name guesstimation is by no means perfect. The British Broadcasting Corp. (BBC), for example, is not at "www.bbc.com." *See* Nathenson, *supra* note 210, at 959-60 n.282.

458. *Playboy Enters., Inc. v. Welles*, 7 F. Supp. 2d 1098 (S.D. Cal.), *aff'd*, No. 98-55911, 1998 U.S. App. LEXIS 27739 (9th Cir. Oct. 27, 1998). Of course, more than one case can be on the same page of the Federal Supplement, just as more than one person can live at 123 Main Street.

459. It would be somewhat foolish to use the white pages to find all the plumbers in town, for example. *See* Lake, *supra* note 35 (stating that directories are appropriate for general, topical searches while search engines are more appropriate for specific information).

460. The same rationale applies to those parties, private or institutional, that go to the effort of compiling lists of links. "A number of universities and other agencies patrol the Internet for useful information and provide links. Does this mean that they are infallible? It does not. Does it mean that they will save you time by sifting through the whole thing and picking out the best stuff? It does." Terry Ballard, *Internet Reference: Just the Good Stuff*, INFO. TODAY, Dec. 1997, at 51; *see also* RAWLINS, *supra* note 44, at 56-57.



searching for those who know the categories in which they are interested.<sup>461</sup>

Thus, just as we developed white pages for targeted searching, we created yellow pages for conceptual searching. The same goes for search engines and directories. The development of parallel search tools to handle both targeted and conceptual searches runs through our culture.<sup>462</sup> Although one's interest in performing relational searches with a search engine is low, one's interest in doing the same search on Yahoo! would be significantly higher. A short examination of the Yahoo! concept is therefore in order.

While Yahoo! has its similarities to yellow pages, it is more flexible. In traditional yellow pages, for an HVAC contractor to reach all potential clients, she must advertise under "furnaces," "heating," and "air conditioning." This is expensive to the businessperson and frustrating to the consumer who does not know which categories are most squarely on point. Further, the consumer is unable to narrow his search by neighborhood, products offered, or other categories. Yellow pages would be more useful if the consumer could construct his own categories and conditions.<sup>463</sup> This is possible with Yahoo!, which allows the user to browse by category and subcategory, or to search its database of categories and links.

By allowing the consumer to choose his own filtering constraints, information technology allows new value to be added where it did not before exist.<sup>464</sup> Consumers, having the power to filter and organize information from search engine databases, will act differently. Creative consumers will use these tools not merely to find, but to cull valuable information that did not exist before. This idea is often described as

---

461. In fact, when a searcher has a particular topic or category in mind, a directory may be the most reasonable, if not exhaustive, place to start searching. See Witulski, *supra* note 77, at 89.

462. Just as West has its citation system, a targeted search tool, it also has its Key Number and Digest system, which is a conceptual search system.

463. "To be really useful, Yellow Pages should list all businesses by each street, neighborhood, and mall; by the time needed to get to them from our current location; by whether they're in a safe neighborhood; [and by other factors]." RAWLINS, *supra* note 44, at 56-57.

464. See GILSTER, *supra* note 48, at 177 (stating that "on-line works, as opposed to their physical cousins like books and newspapers, draw added meaning from the universe of documents and other media surrounding them"); ROSZAK, *supra* note 35, at 163 (quoting John Naisbitt as stating that "[i]nformation technology brings order to the chaos of information pollution and therefore gives value to data that would otherwise be useless").

“data mining.”<sup>465</sup> The “reworking of information is a key ingredient of the information society or information economy in which new information, or old information processed into a new form, acquires substantial economic value.”<sup>466</sup> Allowing third parties to data mine may even enhance the value of the proprietary interests of others, by culling new information that the owner might not have otherwise noticed.

Like consumers, those with proprietary rights may also have to adapt. It is true that as new technologies emerge, battles over proprietary rights inevitably follow. Sometimes, courts will protect — or even invent — such rights, and sometimes they will not. When courts refuse to expand proprietary rights to new media contexts, those with traditional proprietary interests must sometimes adapt or risk forfeiting new markets to others. Disney was so fearful of video-cassette recorders that it took Sony all the way to the Supreme Court.<sup>467</sup> It lost the case,<sup>468</sup> but won the war. Today, Disney is the most successful marketer of videos.<sup>469</sup> Those claiming proprietary rights may sometimes prosper more by adapting than by building fences around their property.<sup>470</sup> Further, practical considerations, such as avoiding bad publicity, may suggest that parties with famous names (such as celebrities and sports leagues) should be cautious not to alienate their fans.<sup>471</sup>

---

465. Robert Schwarzwald, *1997: A Quiet Year for Sci/Tech?*, DATABASE, Dec. 1997, at 59 (stating that “[a]s information systems advance, they will use information more as a raw material than as an end-product”).

466. Katsh, *New Media*, *supra* note 27, at 1478; *see also* JOHNSON, *supra* note 55, at 155–59 (describing how high-speed computer textual analysis was able to show which roles in his own plays Shakespeare likely performed as an actor).

467. *See Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 421 (1984) (noting Disney’s presence as a petitioner).

468. *See id.* at 456.

469. *See* John Hartl, ‘Hercules and Xena’ Poised to Fight for Disney Turf, SEATTLE TIMES, Jan. 8, 1998, at E6 (noting that four of the ten top-selling videos of 1997 were from Disney, including the restored 1942 classic *Bambi*).

470. The National Basketball Association was not able to enjoin Motorola from marketing a pager that gave real-time updates of basketball scores. *See National Basketball Ass’n v. Motorola, Inc.*, 105 F.3d 841, 843–44, 855 (2d Cir. 1997). The NBA could sit idly by and let others take advantage of this market. The smarter approach, however, would be to offer better services in competition or cooperation with Motorola. Even Sony — the defendant in the Disney suit — eventually got into the movie studio business. Sony purchased Columbia Tristar pictures for \$3.4 billion in 1989. *See Sony’s Billions Get Little Play at the Box Office*, STAR-LEDGER (Newark, N.J.), Oct. 3, 1996, at F3. Ironically, however, Sony has not been very successful at the box office, causing it to write off \$2.71 billion in fiscal 1995. *See id.*

471. *See* Koster & Shatz-Akin, *supra* note 204, at 21–22. These authors rightly point out that propriety-rights holders should be circumspect in suing those who establish fan sites, even when copyright or trademark infringement is clear. The costs of bad publicity



Webmasters must adapt as well. Because they have alternative ways to advertise, there is little justification for using meta tags that have little or no good faith relevance to actual content.<sup>472</sup> A spamdexer is rarely more than a bad faith advertiser who tries to put his ad for puppies at the beginning of every section of the classifieds for free.<sup>473</sup> Still, spamdexers may succeed at getting more short-term traffic. Sometimes, members of the public may even be pleased to chance upon an interesting, but unexpected result, the same way one might chance upon an interesting book on physics misplaced in the literature section.<sup>474</sup> This remote possibility, however, does not justify a library “with all the books scattered on the floor.”<sup>475</sup>

Long-term Internet success requires sustaining our interest, not just “momentarily grabbing our attention.”<sup>476</sup> Marketing success may depend more on maximizing the frequency of return visits and the length of time a person spends browsing than it might on tricking us into an unplanned visit.<sup>477</sup> The diminishing returns of spamdexing suggest that

---

may significantly outweigh the benefits of deterring otherwise de minimus infringement. Indeed, fan sites keep people happy and serve as an additional way for the trademark holder to maintain and increase its popularity. Avid fans may even feel they have an entitlement to infringe. These sites serve to spread valuable information among their users. Spamdexing the name of a famous party, on the other hand, serves only the spamdexer. Unlike a fan website that serves some public need, no public interest is served by spamdexing. The backlash that one can expect from harshly worded cease and desist letters in the fan site context is much less likely to occur with spamdexers.

472. First Amendment concerns should caution us against too freely applying liability without a substantial imbalance in the interests.

473. See Pridgen, *supra* note 285, at 245 (suggesting that spamdexing is a “variation on the old practice of naming your company the ABC company so it will be the first listing in the Yellow Pages”).

474. Cf. Berring, *supra* note 427, at 101 (“A professional librarian is always amused to hear a researcher talk about how they often [go] to the library shelf looking for one book and then, incredibly find an even better one by miraculous chance.”). Of course, there still has to be a method to the madness. See *id.* (“[L]egions of librarians have spent their lives trying to create arrangements that would allow just this kind of miracle to take place. Think of it as prearranged serendipity.”).

475. See *supra* text accompanying note 435.

476. SCHWARTZ, *supra* note 166, at 22. “Once a site succeeds in hooking thousands, or even millions, of eyeballs, it has to deliver something special. Something that causes people to return to the site again and again. Otherwise, it’s just the digital equivalent of an accident on the side of the road.” *Id.* at 23.

477. See *id.* at 23–26 (noting that one adult website, which is updated more frequently than the *Playboy* website, generates more return visitors than *Playboy* even though the latter generates four times the number of initial hits); see also Jacques Barbey, *Once Upon a Time*, TIME DIGITAL, Nov. 2, 1998, at 37 (noting that Internet book retailer Amazon.com gets 80,000 visitors a day, has three million in all, two million of which “come back for more”).

webmasters' efforts are better spent seeking results in places where a "target audience" is likely to be found.<sup>478</sup>

The search industry has recognized the public's need for a better way to navigate the infoglut. It is in turmoil as it tries to determine what the public values.<sup>479</sup> Search services may need to come up with value-added extras to attract more people, especially if they ever hope to entice

---

478. See SELLERS, *supra* note 11, at 27–28.

Don't become hung up on squeezing the last drop out of the results. . . . [Y]ou will benefit from promoting your site in places where you are more likely to be noticed by your target audience: magazine reviews, links from related sites, or notices in mailing lists and newsgroups. Making the right connection with any of these may result in more referrals than a search engine.

*Id.*

479. See Schwarzwald, *supra* note 465, at 59. Northern Light is an example of the "next generation." See Northern Light Technology LLC <<http://www.northernlight.com>>. It acts as a standard search engine, but also gives links to non-public information that may be accessed for a fee. Thus, one may use it to find both free and premium information. Further, it "reverse-engineers" your search into categories that may be used to refine the search. See Richard Wiggins, *Vendors Future: Northern Light — Delivering High-Quality Content to a Large Internet Audience*, SEARCHER, Nov. 21, 1997, at 47.

Another service, "searchmill.com," offers real-time human assistance. See *Relief for Frustrated Internet Users: New Service Finds It for You*, PR NEWSWIRE, Dec. 2, 1997 (pg. unavail. online) ("SearchMill.com's human searchers act as a help line for the Web, giving information-starved surfers a welcome break from search engines and subject catalogs.").

Another approach, collaborative filtering, "builds profiles of end users' interests and recommends content one user finds appealing to other users who have similar tastes." Justin Hibbard, *Just Add People — Collaborative Filtering Brings Human Input to Information Retrieval in the Enterprise*, INFO. WK., Dec. 22, 1997, at 65; see also Beth Lipton, *Direct Hit Aims to Refine Searches*, CNET NEWS.COM (Aug. 21, 1998) <<http://www.news.com/News/Item/0,4,25533,00.html?dd.ne.htmldisp.hl.ne>>.

"Intelligent agents" might eventually act as independent search engines. See William E. Halal et al., *Emerging Technologies: What's Ahead for 2001–2030: Scholars Assemble a Comprehensive Forecast of Coming Technologies*, FUTURIST, Nov. 21, 1997, at 20; Mohammed Omar, *How to Deal with Information Overload on the Internet?: The Intelligent Agent Concept*, THE STAR, Nov. 27, 1997, available at <<http://star.arabia.com/971127/TE2.html>>; Jan Davis Tudor, *The New Alchemy: Using Droids & Agents to Treat Information Overload*, ONLINE, Nov. 21, 1997, at 50. These agents act on behalf of an individual to seek out and retrieve information. See GILSTER, *supra* note 48, at 234. Intelligent agents may even become more sophisticated as they learn the preferences of their principals. See *id.* ("[T]he agent will adapt by strengthening its searching on [learned] parameters, while abandoning less productive channels of inquiry."). Regardless, increased sophistication is unlikely to guarantee perfect results. See *id.* at 235–36 (stating that the amount of raw data in existence will arguably increase nineteenfold between 1990 and 2000, and that, at best, intelligent agents "can only make informed guesses as to what we need").



them to pay for searching the Web. After all, the public has grown accustomed to searching for free.<sup>480</sup> While the advent of paid searching will undoubtedly bring higher quality service, it will also lead to information stratification — only those who can afford to pay will have access to quality information.<sup>481</sup>

Meta tags are perhaps undeserving of their bad name. They are nothing more than a variant of indexing by keyword, a traditional and powerful method of information organization and retrieval.<sup>482</sup> Reasonable reliance on relevant, good faith meta tags would vastly enhance the Internet-searching experience.<sup>483</sup> Technological advances in filtering and information technology may help to tackle spamdexing. If search engines can better correlate meta tags to actual content, then query output may become more accurate. Some suggest that using “fuzzy” logic will surmount the difficulties involved in the use of conceptual search terms.<sup>484</sup> Such an approach, although currently

---

480. Some argue that there is a danger in the public’s expectation of free searching. See Schwarzwald, *supra* note 465, at 59; Wiggins, *supra* note 479, at 47.

481. One commentator notes the coming schism in quality between paid and free searching tools:

The future seems clear. One set of future information products will be low-end. Raw information will be taken from the Internet, reformatted, and resold. The quality and credibility of these products will be open to question. The other set of information products will be high-end. These high-end products will be easy to use, full of prompts and links to other systems. Researchers will be able to sit down and use them with no training.

Robert Berring, *Chaos, Cyberspace and Tradition: Legal Information Transmogrified*, 12 BERKELEY TECH. L.J., 189, 209 (1997).

482. See Colin Johnson, *Conceptual Filtering “SIG Pidgins” Home in on Web Info*, ELECTRONIC ENGINEERING TIMES, Nov. 17, 1997, at 48 (“If contextual meta-term handles were attached to on-line documents,” [said one expert,] “then any collection of information could be filtered conceptually to a degree not possible today.”).

483. See Andrew Lawrence, *The Race Is on for New Generation of Search Engines*, COMPUTERGRAM, Nov. 7, 1997. Some suggest the use of Resource Description Framework (RDF), a new kind of metadata:

[Proposed specifications for] metadata about a document or book might include its title, author, publication date, publisher and subject. Even the table of contents and index can be thought of as metadata. Other, less common metadata might include pointers to reviews of that document or book, content ratings . . . and signatures or seals asserting its authenticity.

Eamonn Sullivan, *Standards Aim To Tame The Web*, PC WEEK, Nov. 3, 1997, at 41.

484. Fuzzy logic would involve meta tags that themselves indicate their own relationship or degree of relevance to actual content. “For instance, an automotive service might “somewhat” specialize in sports cars, “usually” do body work, and “especially” handle glass, detailing and painting. Those shades of meaning are best

impractical, would allow webmasters to internalize into meta tags the very degrees of relevance discussed at length in this Article.

While meta tags may come and go, searching and retrieval tools will always be with us. As new search tools emerge, we must be willing to look anew at the balance of interests. An unreasonable search today may become reasonable tomorrow. The temptation to abuse, deceive, or spam will also always be with us.<sup>485</sup> While the balance itself may change, the framework detailed in this Article will hopefully remain a viable consideration to courts faced with disputes between those who engage in abuse and those who seek protection from abuse.<sup>486</sup> In any case, keeping the public interest paramount will benefit all in the long run.

## VI. CONCLUSION

Search engines have become an important tool in our navigation of the noisy waters of the infoglut. Spamdexers have cleverly but dangerously hijacked search engines for their own needs. In most cases, spamdexing is abusive, but because many webmasters have an arguably relevant reason to use a term, it is necessary to provide a workable analytic framework. I have suggested looking to the relevance of the search term in light of the reasonable expectations of the public, webmasters, and proprietary-rights holders.

The tension in trying to balance the diverse interests of these parties will not go away. New forms of information dissemination and retrieval, by changing the reasonableness of one's interests, will undoubtedly require the re-examination of legal concepts of liability.<sup>487</sup>

---

handled with fuzzy set-membership functions that quantify the probability of membership." Johnson, *supra* note 482.

485. As noted, there are other forms of indexing abuse, such as bridge pages, frame abuse, and title tag abuse. See *supra* notes 108–11 and accompanying text.

486. After all, while the tools may be refined, our basic desires are not. Those who search want to find things bereft of noise, those who want to be found don't care how "loud" they are, and those who have proprietary interests will always want to maximize the value of their interests.

487. Some interests, such as free speech, should not necessarily be allowed to change just because new technologies change the balance of reasonable interests. Cf. Nicole B. Casarez, *Deconstructing the Fair Use Doctrine: The Cost of Personal and Workplace Copying After American Geophysical Union v. Texaco, Inc.*, 6 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 641, 717–18 (1996). But cf. *American Geophysical Union v. Texaco Inc.*, 60 F.3d 913 (2d Cir. 1994) (new forms of licensing may lessen fair use under the Copyright Act). As noted before, First Amendment considerations are beyond the scope of this Article. Nevertheless, I hope that the framework of interests developed in Parts



The FTC should act against irrelevant spamdexers, and do so publicly through Surf Days. Courts have also already entertained spamdexing suits under trademark law. Courts should not follow the initial interest doctrine, because a bare meta tag cannot infringe without reference to actual content. Dilution is a more appropriate remedy, but unavailable for non-famous marks. In the small class of cases where neither dilution nor infringement apply, courts must be especially hesitant to invent new categories of liability. Fair use considerations should be liberally considered throughout infringement/dilution analysis to ensure that non-trademark, nominative, and first amendment speech is not chilled.

Courts and regulators should also consider the broader concerns explored in Part V. The existence of alternate forms of information dissemination and retrieval serves a triple purpose. First, webmasters have viable alternatives to spamdexing, which makes meta tag abuse even less excusable. Second, because some types of searches are more suited to alternative tools, the public cannot complain when it unsuccessfully uses conceptual search terms. Third, because proprietary-rights holders have a variety of alternative ways to publicize themselves, and because they may benefit by allowing data mining or other good faith use of their names, they should be cautious in seeking legal remedies.

There is no Holy Grail for search engines. It is inevitable that as new forms of information emerge, new forms of deception will “tag” along. When that happens, it may become necessary to re-examine the balance of interests. This flexibility is the power that underlies the law. Hopefully, the framework suggested in this Article will remain a viable tool for future analysis.