

University of Maryland Francis King Carey School of Law

DigitalCommons@UM Carey Law

Faculty Scholarship

Francis King Carey School of Law Faculty

2010

Trusting (and Verifying) Online Intermediaries' Policing

Frank A. Pasquale

University of Maryland Francis King Carey School of Law, fpasquale@law.umaryland.edu

Follow this and additional works at: https://digitalcommons.law.umaryland.edu/fac_pubs



Part of the [Internet Law Commons](#)

Digital Commons Citation

Pasquale, Frank A., Trusting (and Verifying) Online Intermediaries' Policing, in *The Next Digital Decade, Essay on the Future of the Internet*, 347 (Berin Szoka, Adam Marcus, eds., 2010).

This Book Chapter is brought to you for free and open access by the Francis King Carey School of Law Faculty at DigitalCommons@UM Carey Law. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of DigitalCommons@UM Carey Law. For more information, please contact smccarty@law.umaryland.edu.

CHAPTER 6

SHOULD ONLINE INTERMEDIARIES BE REQUIRED TO POLICE MORE?

Trusting (and Verifying) Online Intermediaries' Policing 347

Frank Pasquale

Online Liability for Payment Systems 365

Mark MacCarthy

Fuzzy Boundaries: The Potential Impact
of Vague Secondary Liability Doctrines
on Technology Innovation 393

Paul Szynol

Trusting (and Verifying) Online Intermediaries' Policing

By Frank Pasquale*

Introduction

Internet Service Providers (ISPs) and search engines have mapped the Web, accelerated e-commerce, and empowered new communities. They can also enable intellectual property infringement, harassment, stealth marketing, and frightening levels of surveillance. As a result, individuals are rapidly losing the ability to control their own image on the web, or even to know what data others are presented with regarding them. When Web users attempt to find information or entertainment, they have little assurance that a carrier or search engine is not subtly biasing the presentation of results in accordance with its own commercial interests.¹

None of these problems is readily susceptible to swift legal intervention. Instead, intermediaries themselves have begun policing their own virtual premises. eBay makes it easy for intellectual property owners to report infringing merchandise. A carrier like Comcast has the technical power to slow or block traffic to and from a site like BitTorrent, which is often accused of infringement.² Google's StopBadware program tries to alert searchers about malware-ridden websites,³ and YouTube employs an indeterminate number of people to police copyright infringement, illegal obscenity, and even many grotesque or humiliating videos.⁴ Reputable social networks do the same for their own content.

* Professor of Law, Seton Hall Law School; Visiting Fellow, Princeton Center for Information Technology Policy.

¹ Benjamin Edelman, *Hard-Coding Bias in Google "Algorithmic" Search Results*, Nov. 15, 2010, available at <http://www.benedelman.org/hardcoding/> ("I present categories of searches for which available evidence indicates Google has "hard-coded" its own links to appear at the top of algorithmic search results, and I offer a **methodology** for detecting certain kinds of tampering by comparing Google results for similar searches. I compare Google's hard-coded results with Google's public **statements** and promises, including a dozen denials but at least one admission.").

² See, e.g., *Comcast Corp. v. FCC*, No. 08-1291, 2010 U.S. App. LEXIS 7039 (D.C. Cir. April 6, 2010).

³ For more information, visit <http://stopbadware.org/>.

⁴ YouTube, *YouTube Community Guidelines*, http://www.youtube.com/t/community_guidelines ("YouTube staff review flagged

Yet all is not well in the land of online self-regulation. However competently they police their sites, nagging questions will remain about their fairness and objectivity in doing so. Is Comcast blocking BitTorrent to stop infringement, or to decrease access to content that competes with its own for viewers? How much digital due process does Google need to give a site it accuses of harboring malware? If Facebook eliminates a video of war carnage, is that a token of respect for the wounded or one more reflexive effort of a major company to ingratiate itself with a Washington establishment currently committed to indefinite military engagement in the Middle East?

Questions like these will persist, and erode the legitimacy of intermediary self-policing, as long as key operations of leading companies are shrouded in secrecy. Administrators must develop an institutional competence for continually monitoring rapidly-changing business practices. A trusted advisory council charged with assisting the Federal Trade Commission (FTC) and Federal Communications Commission (FCC) could help courts and agencies adjudicate controversies concerning intermediary practices. An Internet Intermediary Regulatory Council (IIRC) would spur the development of what Christopher Kelty calls a “recursive public”—one that is “vitaly concerned with the material and practical maintenance and modification of the technical, legal, practical, and conceptual means of its own existence as a public.”⁵ Questioning the power of a dominant intermediary is not just a preoccupation of the anxious. Rather, monitoring is a prerequisite for assuring a level playing field online.

Understanding Intermediaries’ Power

Internet intermediaries govern online life.⁶ ISPs and search engines are particularly central to the web’s ecology. Users rely on search services to map the web for them and use ISPs to connect to one another. Economic sociologist David Stark has observed that “search is the watchword of the information age.”⁷ ISPs are often called “carriers” to reflect the parallel

videos 24 hours a day, seven days a week to determine whether they violate our Community Guidelines. When they do, we remove them.”)

⁵ CHRISTOPHER M. KELTY, *TWO BITS: THE CULTURAL SIGNIFICANCE OF FREE SOFTWARE* 3 (Duke Univ. Press 2007).

⁶ For a definition of intermediary, see Thomas F. Cotter, *Some Observations on the Law and Economics of Intermediaries*, 2006 MICH. ST. L. REV. 67, 68–71 (“[A]n ‘intermediary’ can be any entity that enables the communication of information from one party to another. On the basis of this definition, any provider of communications services (including telephone companies, cable companies, and Internet service providers) qualify as intermediaries.”).

⁷ DAVID STARK, *THE SENSE OF DISSONANCE: ACCOUNTS OF WORTH IN ECONOMIC LIFE* 1 (Princeton Univ. Press 2009) (“Among the many new information technologies that are reshaping work and daily life, perhaps none are more empowering than the new technologies of search.”).

between their own services in the new economy and transportation infrastructure. Online intermediaries organize and control access to an extraordinary variety of digitized content. Content providers aim to be at the top of Google Search or Google News results.⁸ Services like iTunes, Hulu, and YouTube offer audio and video content. Social networks are extending their reach into each of these areas. Cable-based ISPs like Comcast have their own relationships with content providers.⁹

When an Internet connection is dropped, or a search engine fails to produce a result the searcher knows exists somewhere on the web, such failures are obvious. However, most web experiences do not unfold in such a binary, pass–fail manner. An ISP or search engine can slow down the speed or reduce the ranking of a website in ways that are very hard for users to detect. Moreover, there are many points of control, or layers, of the Web.¹⁰ Even when users' experience with one layer causes suspicion, it can blame others for the problem.

The new power of intermediaries over reputation and visibility implicates several traditional concerns of the American legal system.¹¹ Unfortunately, Internet intermediaries are presently bound only by weak and inadequate enforcement of consumer protection and false advertising statutes, which were designed for very different digital infrastructures.

-
- ⁸ See Deborah Fallows & Lee Rainie, Pew Internet & Am. Life Project, *Data Memo: The Popularity and Importance of Search Engines 2* (Aug. 2004), http://www.pewinternet.org/pdfs/PIP_Data_Memo_Searchengines.pdf (“The average visitor scrolled through 1.8 result pages during a typical search.”); Leslie Marable, *False Oracles: Consumer Reaction to Learning the Truth About How Search Engines Work: Results of an Ethnographic Study*, CONSUMER WEBWATCH, June 30, 2003, at 5, available at <http://www.consumerwebwatch.org/pdfs/false-oracles.pdf> (“The majority of participants never clicked beyond the first page of search results. They trusted search engines to present only the best or most accurate, unbiased results on the first page.”).
- ⁹ ROBERT W. MCCHESENEY, RICH MEDIA, POOR DEMOCRACY: COMMUNICATION POLITICS IN DUBIOUS TIMES 123 (2000) (describing how convergence of digital technology “eliminates the traditional distinctions between media and communications sectors”).
- ¹⁰ JONATHAN ZITTRAIN, THE FUTURE OF THE INTERNET—AND HOW TO STOP IT 67 (2008) (describing a physical layer, the “actual wires or airwaves over which data will flow;” an application layer, “representing the tasks people might want to perform on the network;” a content layer, “containing actual information exchanged among the network’s users;” and a social layer, “where new behaviors and interactions among people are enabled by the technologies underneath”).
- ¹¹ Yochai Benkler, *Communications Infrastructure Regulation and the Distribution of Control over Content*, 22 TELECOMM. POL’Y 183, 185–86 (1998) (describing the power of intermediaries over information flow: “technology, institutional framework, and organizational adaptation . . . determine . . . who can produce information, and who may or must consume, what type of information, under what conditions, and to what effect”); Cotter, *supra* note 6, at 69–71 (discussing some of the functions of technological intermediaries, including their control of information flow from suppliers to consumers).

In the space of a brief essay, I cannot survey the entire range of intermediary policing practices. But it is worthwhile to drill down a bit into the tough questions raised by one intermediary—the dominant search engine, Google—as it decides what is and is not an acceptable practice for search engine optimizers who want their clients’ sites to appear higher in the rankings for given queries.

Search engineers tend to divide the search engine optimization (SEO) business into “good guys” and “bad guys,” often calling the former “white hat SEO” and the latter “black hat SEO.”¹² Some degree of transparency regarding the search engine’s algorithm is required to permit white hat SEO. These rules are generally agreed upon as practices that “make the web better;” *i.e.*, have fresh content, don’t sell links, don’t “stuff metatags” with extraneous information just to get attention. However, if there were complete transparency, “black hat” SEOs could unfairly elevate the visibility of their clients’ sites—and even if this were only done temporarily, the resulting churn and chaos could severely reduce the utility of search results. Moreover, a search engine’s competitors could use the trade secrets to enhance its own services.

This secrecy has led to a growing gray zone of Internet practices with uncertain effect on sites’ rankings. Consider some of the distinctions below, based on search engine optimization literature:

White Hat (acceptable) ¹³	Gray Area (unclear how these are treated) ¹⁴	Black Hat (unacceptable; can lead to down-ranking in Google results or even the “Google Death Penalty” of De-Indexing)
Asking blogs you like to link to you, or engaging in reciprocal linking between your site and other sites in a legitimate dialogue. ¹⁵	Paying a blogger or site to link to your blog in order to boost search results and not just to increase traffic.	Creating a “link farm” of spam blogs (splogs) to link to you, or linking between multiple sites you created (known as link farms) to boost search results. ¹⁶

¹² Elizabeth van Couvering, *Is Relevance Relevant?*, <http://jcmc.indiana.edu/vol12/issue3/vancouvering.html> (search engineers’ “animosity towards the ... guerilla fighters of spamming and hacking, is more direct” than their hostility toward direct business competitors); Aaron Wall, *Google Thinks YOU Are a Black Hat SEO. Should You Trust Them?*, SEOBOK, Apr. 17, 2008, <http://www.seobook.com/to-google-you-are-a-spammer> (claiming that Google discriminates against self-identified SEOs).

¹³ Phil Craven, ‘Ethical’ Search Engine Optimization Exposed!, WebWorkshop, <http://www.webworkshop.net/ethical-search-engine-optimization.html> (last visited Jun. 8, 2009).

¹⁴ Grey Hat SEO, <http://greyhatseo.com/> (last visited Jun. 5, 2006) (claiming a Grey Hat SEO is someone who uses black hat techniques in an ethical way.)

¹⁵ *Link Schemes*, GOOGLE WEBMASTER CENTRAL, <http://www.google.com/support/webmasters/bin/answer.py?answer=66356> (“The

White Hat (acceptable) ¹³	Gray Area (unclear how these are treated) ¹⁴	Black Hat (unacceptable; can lead to down-ranking in Google results or even the “Google Death Penalty” of De-Indexing)
Running human-conducted tests of search inquiries with permission from the search engine.	Doing a few queries to do elementary reverse engineering. (This may not be permitted under the Terms of Service).	Using computer programs to send automated search queries to gauge the page rank generated from various search terms (Terms of Service prohibit this) ¹⁷
Creating non-intentional duplicate content (through printer-friendly versions, pages aimed at mobile devices, <i>etc.</i>) ¹⁸	Intentionally creating permitted duplicate content to boost search results	Intentionally creating unnecessary duplicate content on many pages and domains to boost results
Generating a coherent site with original and informative material aimed at the user	Creating content or additional pages that walk the line between useful information and “doorway pages”	Creating “doorway pages” that are geared towards popular keywords but that redirect to a largely unrelated main site. ¹⁹

best way to get other sites to create relevant links to yours is to create unique, relevant content that can quickly gain popularity in the Internet community. The more useful content you have, the greater the chances someone else will find that content valuable to their readers and link to it.”)

¹⁶ Duncan Riley, *Google Declares Jihad On Blog Link Farms*, TECHCRUNCH, Oct. 24, 2007, <http://www.techcrunch.com/2007/10/24/google-declares-jihad-on-blog-link-farms/>.

¹⁷ *Automated Queries*, GOOGLE WEBMASTER CENTRAL, <http://www.google.com/support/webmasters/bin/answer.py?answer=66357> (“Google’s Terms of Service do not allow the sending of automated queries of any sort to our system without express permission in advance from Google.”); Google Terms of Service: Use of the Services by you, <http://www.google.com/accounts/TOS> (last visited Jun. 4, 2009) (“You agree not to access (or attempt to access) any of the Services by any means other than through the interface that is provided by Google, unless you have been specifically allowed to do so in a separate agreement with Google.”).

¹⁸ *Duplicate Content*, GOOGLE WEBMASTER CENTRAL, <http://www.google.com/support/webmasters/bin/answer.py?answer=66359> (“Examples of non-malicious duplicate content could include: Discussion forums that can generate both regular and stripped-down pages targeted at mobile devices, Store items shown or linked via multiple distinct URLs, Printer-only versions of web pages”).

¹⁹ Google Blogscoped, *German BMW Banned From Google*, Feb. 4, 2006, <http://blogscoped.com/archive/2006-02-04-n60.html>; Matt Cutts, *Ramping up on International Webspam*, MATT CUTTS: GADGETS, GOOGLE, AND SEO, Feb. 4, 2006, <http://www.mattcutts.com/blog/ramping-up-on-international-webspam/> (Google employee confirming BMW’s ban).

White Hat (acceptable) ¹³	Gray Area (unclear how these are treated) ¹⁴	Black Hat (unacceptable; can lead to down-ranking in Google results or even the “Google Death Penalty” of De-Indexing)
Targeting an appreciative audience ²⁰	Putting random references to salacious or celebrity topics on a blog primarily devoted to discussing current affairs ²¹	Distracting an involuntary audience with completely misleading indexed content (akin to “initial interest confusion” in Internet trademark law) ²²
Influencing search engine by making pages easier to scan by automated bots ²³	Creating “hidden pages” when there may be a logical reason to show one page to search engine bots and another page to users who type in the page’s URL	Using “hidden pages” to show a misleading page to search engine bots, and another page to users who type in the page’s URL.

As these practices show, search engines are referees in the millions of contests for attention that take place on the web each day. There are hundreds of entities that want to be the top result in response to a query like “sneakers,” “restaurant in New York City,” or “best employer to work for.” Any academic who writes on an obscure subject wants to be the “go-to” authority when it is Googled—and for consultants, a top or tenth-ranked result could be the

²⁰ *Webmaster Guidelines: Design and content guidelines*, GOOGLE WEBMASTER CENTRAL, <http://www.google.com/support/webmasters/bin/answer.py?answer=35769> (last visited Jun. 4, 2009) (“Think about the words users would type to find your pages, and make sure that your site actually includes those words within it.”).

²¹ Daniel Solove, *Thanks, Jennifer Aniston (or the Manifold Ways to Do the Same Search)*, CONCURRING OPINIONS, http://www.concurringopinions.com/archives/2006/01/thanks_jennifer.html (“One of my more popular posts is one entitled *Jennifer Aniston Nude Photos and the Anti-Paparazzi Act*. It seems to be getting a lot of readers interested in learning about the workings of the Anti-Paparazzi Act and the law of information privacy. It sure is surprising that so many readers are eager to understand this rather technical statute. Anyway, for the small part that Jennifer Aniston plays in this, we thank her for the traffic.”); Dan Filler, *Coffee Or Nude Celebrity Photos: A Tale Of Two Evergreen Posts*, THE FACULTY LOUNGE, <http://www.thefacultylounge.org/2008/04/coffee-or-nude.html> (“significant amounts of traffic arrived in the form of web surfers seeking out pictures of Jennifer Aniston”).

²² Jason Preston, *Google punishes Squidoo for having too much Spam*, BLOG BUSINESS SUMMIT, Jul. 11, 2007, <http://blogbusinesssummit.com/2007/07/google-punishes-squidoo-for-having-too-much-spam.htm>.

²³ *Webmaster Guidelines: Design and Content Guidelines*, GOOGLE WEBMASTER CENTRAL, <http://www.google.com/support/webmasters/bin/answer.py?answer=35769> (last visited Jun. 4, 2009) (“Create a useful, information-rich site, and write pages that clearly and accurately describe your content.”); *Id.* (“Try to use text instead of images to display important names, content, or links. The Google crawler doesn’t recognize text contained in images.”).

difference between lucrative gigs and obscurity. The top and right hand sides of many search engine pages are open for paid placement; but even there the highest bidder may not get a prime spot because a good search engine strives to keep even these sections very relevant to searchers.²⁴ The organic results are determined by search engines' proprietary algorithms, and preliminary evidence indicates that searchers (and particularly educated searchers) concentrate attention there. Businesses can grow reliant on good Google rankings as a way of attracting and keeping customers.

For example, John Battelle tells the story of the owner of 2bigfeet.com (a seller of large-sized men's shoes), whose site was knocked off the first page of Google's results for terms like "big shoes" by a sudden algorithm shift in November 2003, right before the Christmas shopping season. The owner attempted to contact Google several times, but said he "never got a response." Google claimed the owner may have hired a search engine optimizer who ran afoul of its rules—but it would not say precisely what those rules were.²⁵ Like the IRS's unwillingness to disclose all of its "audit flags," the company did not

²⁴ Steven Levy, *Secret of Googlenomics: Data-Fueled Recipe Brews Profitability*, WIRED, May 2, 2009, http://www.wired.com/culture/culturereviews/magazine/17-06/nep_googlenomics (in Google's AdWords program, "The bids themselves are only a part of what ultimately determines the auction winners. The other major determinant is something called **the quality score**. This metric strives to ensure that the ads Google shows on its results page are true, high-caliber matches for what users are querying. If they aren't, the whole system suffers and Google makes less money."); see also Google, *What is the Quality Score and How is it Calculated*, <http://adwords.google.com/support/aw/bin/answer.py?hl=en&answer=10215> (last visited Sept. 1, 2009) ("The AdWords system works best for everybody—advertisers, users, publishers, and Google too—when the ads we display match our users' needs as closely as possible.").

²⁵ JOHN BATTELLE, *THE SEARCH: HOW GOOGLE AND ITS RIVALS REWROTE THE RULES OF BUSINESS AND TRANSFORMED OUR CULTURE* (Portfolio Trade 2005). See also Joe Nocera, *Stuck in Google's Doghouse*, N.Y. TIMES, Sept. 13, 2008, <http://www.nytimes.com/2008/09/13/technology/13nocera.html> ("In the summer of 2006 ... Google pulled the rug out from under [web business owner Dan Savage, who had come to rely on its referrals to his page, Sourcedool]... . When Mr. Savage asked Google executives what the problem was, he was told that Sourcedool's "landing page quality" was low. Google had recently changed the algorithm for choosing advertisements for prominent positions on Google search pages, and Mr. Savage's site had been identified as one that didn't meet the algorithm's new standards... . Although the company never told Mr. Savage what, precisely, was wrong with his landing page quality, it offered some suggestions for improvement, including running fewer AdSense ads and manually typing in the addresses and phone numbers of the 600,000 companies in his directory, even though their Web sites were just a click away. At a cost of several hundred thousand dollars, he made some of the changes Google suggested. No improvement."). Savage filed suit against Google on an antitrust theory, which was dismissed in March 2010. See *TradeComet, LLC v. Google, Inc.*, 2010 U.S. Dist. LEXIS 20154 (S.D. N.Y. March 5, 2010), <http://www.courthousenews.com/2010/03/08/Google%20opinion.pdf>.

want to permit manipulators to gain too great an understanding of how it detected their tactics.

So far, claims like 2bigfeet.com's have not been fully examined in the judicial system, largely because Google has successfully deflected them by claiming that its search results embody opinions protected by the First Amendment. Several articles have questioned whether blanket First Amendment protection covers all search engine actions, and that conclusion has not yet been embraced on the appellate level in the United States.²⁶ The FTC's guidance to search engines, promoting the clear separation of organic and paid results, suggests that search engines' First Amendment shield is not insurmountable.²⁷ While a creative or opportunistic litigant could conceivably advance a First Amendment right to promote products or positions without indicating that the promotion has been paid for, such a challenge has not yet eliminated false advertising law, and even political speakers have been required to reveal their funding sources.²⁸

Qualified Transparency for Carrier & Search Engine Practices

Both search engines' ranking practices and carriers' network management should be transparent to some entity capable of detecting biased policing by these intermediaries.²⁹ There are some institutional precedents for the kind of monitoring that would be necessary to accomplish these goals. For example, the French Commission Nationale De L'Informatique et des Libertes (CNIL) has several prerogatives designed to protect the privacy and reputation of

²⁶ Frank Pasquale, *Rankings, Reductionism, and Responsibility*, CLEVELAND ST. L. REV. (2006); Frank Pasquale & Oren Bracha, *Federal Search Commission*, 93 CORNELL L. REV. 1149 (2008); Jennifer A. Chandler, *A Right to Reach an Audience: An Approach to Intermediary Bias on the Internet*, 35 HOFSTRA L. REV. 1095, 1109 (2007).

²⁷ See Bracha & Pasquale, *Federal Search Commission*, *supra* note 26 (discussing the implications of Ellen Goodman's work on "stealth marketing" for search engines, and how the Hipsley Letter of 2002 inadequately addressed such concerns in the industry).

²⁸ In early cases alleging an array of unfair competition and business torts claims against search engines, the First Amendment has proven a formidable shield against liability. Search engines characterize their results as opinion, and lower courts have been reluctant to penalize them for these forms of expression. In other work, I have described why this First Amendment barrier to accountability should not be insurmountable. Search engines take advantage of a web of governmental immunities that they would be loath to surrender. *FAIR v. Rumsfeld*, 547 U.S. 47 (2006) and cognate cases stand for the proposition that such immunities can be conditioned on agreement to certain conditions on an entity's speech. Whatever the federal government's will, it is within its power to regulate ranking and rating entities in some way when they are so deeply dependent on governmental action. Frank Pasquale, *Asterisk Revisited*, 3 J. BUS. & TECH. LAW 61 (2008).

²⁹ I mean partial in two senses of the word—unduly self-interested, or only partly solving problems they claim to be solving.

French citizens, and to enforce standards of fair data practices.³⁰ CNIL “ensure[s] that citizens are in a position to exercise their rights through information” by requiring data controllers to “ensure data security and confidentiality,” to “accept on-site inspections by the CNIL,” and to “reply to any request for information.”³¹ CNIL also grants individual persons the right to obtain information about the digital dossiers kept on them and the use of this information. For example, CNIL explains that French law provides that:

Every person may, on simple request addressed to the organisation in question, have free access to all the information concerning him in clear language.

Every person may directly require from an organisation holding information about him that the data be corrected (if they are wrong), completed or clarified (if they are incomplete or equivocal), or erased (if this information could not legally be collected).

³⁰ Law No. 78-17 of January 6, 1978, J.C.P. 1978, III, No. 44692. English translation of law as amended by law of August 6, 2004, and by Law of May 12, 2009, <http://www.cnil.fr/fileadmin/documents/en/Act78-17VA.pdf>; French language text modified through Law No. 2009-526 of May 12, 2009, J.O., May 13, 2009, <http://www.cnil.fr/la-cnil/qui-sommes-nous/>; French language consolidated version as of May 14, 2009, <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000886460&fastPos=1&fastReqId=826368234&categorieLien=cid&oldAction=rechTexte>. Commission Nationale de l’Informatique et des Libertés (CNIL), founded by Law No. 78-17 of January 6, 1978, *supra*, is an independent administrative French authority protecting privacy and personal data held by government agencies and private entities. Specifically, CNIL’s general mission consists of ensuring that the development of information technology remains at the service of citizens and does not breach human identity, human rights, privacy, or personal or public liberties.

³¹ CNIL, *Rights and Obligations*, <http://www.cnil.fr/english/the-cnil/rights-and-obligations/> (last visited Mar. 12, 2010). Specifically, Chapter 6, Article 44, of the CNIL-creating Act provides:

The members of the “Commission nationale de l’informatique et des libertés” as well as those officers of the Commission’s operational services accredited in accordance with the conditions defined by the last paragraph of Article 19 (accreditation by the commission), have access, from 6 a.m to 9 p.m, for the exercise of their functions, to the places, premises, surroundings, equipment or buildings used for the processing of personal data for professional purposes, with the exception of the parts of the places, premises, surroundings, equipment or buildings used for private purposes.

Law No. 78-17 of January 6, 1978, J.C.P. 1978, III, No. 44692, ch. 6, art. 44, at 30, <http://www.cnil.fr/fileadmin/documents/en/Act78-17VA.pdf>.

Every person may oppose that information about him is used for advertising purposes or for commercial purposes.³²

While the United States does not have the same tradition of protecting privacy prevalent in Europe,³³ CNIL's aims and commitments could prove worthwhile models for U.S. agencies.

U.S. policymakers may also continue to experiment with public–private partnerships to monitor problematic behavior at search engines and carriers. For instance, the National Advertising Division (NAD) of the Council of Better Business Bureaus is a “voluntary, self-regulating body” that fields complaints about allegedly untruthful advertising.³⁴ The vast majority of companies investigated by NAD comply with its recommendations, but can also resist its authority and resolve the dispute before the FTC.³⁵ Rather than overwhelming the agency with adjudications, the NAD process provides an initial forum for advertisers and their critics to contest the validity of statements.³⁶ NAD is part of a larger association called the National Advertising Review Council (NARC), which promulgates procedures for NAD, the Children's Advertising Review Unit (CARU), and the National Advertising Review Board (NARB).³⁷

³² CNIL, Rights and Obligations, *supra* note 31.

³³ James Whitman, *The Two Western Cultures of Privacy: Dignity Versus Liberty*, 113 YALE L.J. 1151, 1155 (2004) (comparing U.S. and European privacy law).

³⁴ Seth Stevenson, *How New Is New? How Improved Is Improved? The People Who Keep Advertisers Honest*, SLATE, July 13, 2009, <http://www.slate.com/id/2221968>.

³⁵ *Id.* (“When an ad is brought to their attention, the NAD’s lawyers review the specific claims at issue. The rule is that the advertiser must have substantiated any claims before the ad was put on the air, so the NAD will first ask for any substantiating materials the advertiser can provide. If the NAD lawyers determine that the claims aren’t valid, they’ll recommend that the ad be altered. The compliance rate on this is more than 95 percent. But if the advertiser refuses to modify the ad (this is a voluntary, self-regulating body, not a court of law), the NAD will refer the matter to the Federal Trade Commission. One such FTC referral resulted in an \$83 million judgment against a weight-loss company.”).

³⁶ *Id.*

³⁷ NATIONAL ADVERTISING REVIEW COUNCIL, THE ADVERTISING INDUSTRY’S PROCESS OF VOLUNTARY SELF-REGULATION: POLICIES AND PROCEDURES § 2.1(a) (July 27, 2009) (“The National Advertising Division of the Council of Better Business Bureaus (hereinafter NAD), and the Children’s Advertising Review Unit (CARU), shall be responsible for receiving or initiating, evaluating, investigating, analyzing (in conjunction with outside experts, if warranted, and upon notice to the parties), and holding negotiations with an advertiser, and resolving complaints or questions from any source involving the truth or accuracy of national advertising.”). Though billed as “self-regulation,” it is difficult to see how the policy would have teeth were it not self-regulation in the shadow of an FTC empowered by the Lanham Act to aggressively police false advertising. The FTC has several mechanisms by which to regulate unfair business practices in commerce. *See, e.g.*, 15 U.S.C. § 45(b) (2006)

Instead of an “Innovation Environment Protection Agency (iEPA)” (the agency Lawrence Lessig proposed to supplant the FCC), I would recommend the formation of an Internet Intermediary Regulatory Council (IIRC), which would assist both the FCC and FTC in carrying out their present missions.³⁸ Like the NARC, the IIRC would follow up on complaints made by competitors, the public, or when it determines that a practice deserves investigation. If the self-regulatory council failed to reconcile conflicting claims, it could refer complaints to the FTC (in the case of search engines, which implicate the FTC’s extant expertise in both privacy and advertising) or the FCC (in the case of carriers). In either context, an IIRC would need not only lawyers, but also engineers and programmers who could fully understand the technology affecting data, ranking, and traffic management practices.

An IIRC would research and issue reports on suspect practices by Internet intermediaries, while respecting the intellectual property of the companies it investigated. An IIRC could generate official and even public understanding of intermediary practices, while keeping crucial proprietary information under the control of the companies it monitors. An IIRC could develop a detailed description of safeguards for trade secrets, which would prevent anyone outside its offices from accessing the information.³⁹ Another option would be to allow IIRC agents to inspect such information without actually obtaining it. An IIRC could create “reading rooms” for use by its experts, just as some courts allow restrictive protective orders to govern discovery in disputes involving trade secrets. The experts would review the information in a group setting (possibly during a period of days) to determine whether a given intermediary had engaged in practices that could constitute a violation of privacy or consumer protection laws. Such review would not require any outside access to sensitive information.

I prefer not to specify at this time whether an IIRC would be a private or public entity. Either approach would have distinct costs and benefits explored (in part) by a well-developed literature on the role of private entities in Internet

(giving the commission the authority to register an official complaint against an entity engaged in unfair business methods).

³⁸ It could include a search engine division, an ISP division focusing on carriers, and eventually divisions related to social networks or auction sites if their practices begin to raise commensurate concerns.

³⁹ This is the way that the NAD proceeds. It provides specific procedures under which the participants can request that certain sensitive information be protected. *See* NAT’L ADVERTISING REVIEW COUNCIL, *THE ADVERTISING INDUSTRY’S PROCESS OF VOLUNTARY SELF-REGULATION* § 2.4(d)–(e), at 4–5 (2009), http://www.nadreview.org/07_Procedures.pdf (discussing procedure for confidential submission of trade secrets).

governance.⁴⁰ Regardless of whether monitoring is done by a governmental entity (like CNIL) or an NGO (like NARC), we must begin developing the institutional capacity to permit a more rapid understanding of intermediary actions than traditional litigation permits.⁴¹

It is not merely markets and antitrust enforcement that are insufficient to constrain problematic intermediary behavior—the common law is also likely to fall short. It is hard to imagine any but the wealthiest and most sophisticated plaintiffs’ attorneys attempting to understand the tweaks to the Google algorithm that might have unfairly diminished their clients’ sites’ salience. Trade secrets have been deployed in the context of other litigation to frustrate investigations of black box algorithms.⁴² Examination of Google’s algorithms subject to very restrictive protective orders would amount to a similar barrier to accountability. Given its recent string of litigation victories, it is hard to imagine rational litigants continuing to take on that risk. Moreover, it makes little sense for a court to start from scratch in understanding the complex practices of intermediaries when an entity like the IIRC could develop lasting expertise in interpreting their actions.

A status quo of unmonitored intermediary operations is a veritable “ring of Gyges,”⁴³ tempting them to push the envelope with policing practices which

⁴⁰ See, e.g., Philip J. Weiser, *Internet Governance, Standard Setting, and Self-Regulation*, 28 N. KY. L. REV. 822, 822 (2001) (examining “in particular the nature and limits of a key private regulator of the Internet: standard-setting organizations and their institution of open, interoperable standards”).

⁴¹ Google has already recognized the need for some kind of due process in response to complaints about its labeling of certain websites as “harmful” (due to the presence of viruses or other security threats at the sites) via the StopBadware program. See ZITTRAIN, *FUTURE OF THE INTERNET*, *supra* note 10, at 171 (“Requests for review—which included pleas for help in understanding the problem to begin with—inundated StopBadware researchers, who found themselves overwhelmed in a matter of days by appeals from thousands of Web sites listed. Until StopBadware could check each site and verify it had been cleaned of bad code, the warning page stayed up.”). Google’s cooperation with the Harvard Berkman Center for Internet Research to run the StopBadware program could prefigure future intermediary cooperation with NGOs to provide “rough justice” to those disadvantaged by certain intermediary practices.

⁴² See Jessica Ring Amunson & Sam Hirsch, *The Case of the Disappearing Votes: Lessons from the Jennings v. Buchanan Congressional Election Contest*, 17 WM. & MARY BILL RTS. J. 397, 397–98 (2008) (“[T]he litigation ultimately was utterly inconclusive as to the reason for the 18,000 electronic undervotes because discovery targeting the defective voting system was thwarted when the voting machines’ manufacturer successfully invoked the trade-secret privilege to block any investigation of the machines or their software by the litigants.”).

⁴³ “The Ring of Gyges is a mythical magical artifact mentioned by the philosopher Plato in book 2 of his Republic (2.359a–2.360d). It granted its owner the power to become invisible at will. Through the story of the ring, Republic discusses whether a typical person would be moral if he did not have to fear the consequences of his actions.” Wikipedia, *Ring of Gyges*, http://en.wikipedia.org/wiki/Ring_of_Gyges (last accessed Dec. 1, 2010).

cannot be scrutinized or challenged. Distortions of the public sphere are also likely. While a commercially-influenced “fast-tracking” or “up-ranking” of some content past others might raise suspicions among its direct (but dispersed) victims, the real issues it raises are far broader. If an online ecology of information that purports to be based on one mode of ordering is actually based on another, it sets an unfair playing field whose biases are largely undetectable by lay observers. Stealth marketing generates serious negative externalities that menace personal autonomy and cultural authenticity. Moreover, the degree of expertise necessary to recognize these externalities in the new online environment is likely to be possessed by only the most committed observers.

This potent combination of expertise and externalities is a classic rationale for regulation. As Danny Weitzner’s proposal for “extreme factfinding” (in the context of the Google–DoubleClick merger review) recognized, only a dedicated group of engineers, social scientists, attorneys, and computer scientists are likely to be adept enough at understanding search engine decisions as a whole to understand particular complaints about them.⁴⁴ Someone needs to be able to examine the finer details of the publicly undisclosed operation of culturally significant automated ranking systems—that is, to watch those who watch and influence us.⁴⁵

⁴⁴ See generally, Danny Weitzner, *What to Do About Google and Doubleclick? Hold Google to It’s Word With Some Extreme Factfinding About Privacy Practices*, GOOGLE OPEN INTERNET POLICY BLOG, Oct. 8, 2007, <http://dig.csail.mit.edu/breadcrumbs/node/203>:

In the 1990s, the FTC under Christine Varney’s leadership pushed operators of commercial websites to post policies stating how they handle personal information. That was an innovative idea at the time, but the power of personal information processing has swamped the ability of a static statement to capture the privacy impact of sophisticated services, and the level of generality at which these policies tend to be written often obscure the real privacy impact of the practices described. It’s time for regulators to take the next step and assure that both individuals and policy makers have information they need.

Weitzner proposes that “[r]egulators should appoint an independent panel of technical, legal and business experts to help them review, on an ongoing basis the privacy practices of Google.” *Id.* The panel would be “made up of those with technical, legal and business expertise from around the world.” *Id.* It would hold “public hearings at which Google technical experts are available to answer questions about operational details of personal data handling.” *Id.* There would be “staff support for the panel from participating regulatory agencies,” “real-time publication of questions and answers,” and “[a]n annual report summarizing what the panel has learned.” *Id.*

⁴⁵ In the meantime, Google has been developing a tool that would help consumers detect if their Internet service provider was “running afoul of Net neutrality principles.” Stephanie Condon, *Google-Backed Tool Detects Net Filtering, Blocking*, CNET NEWS, Jan. 28, 2009, http://news.cnet.com/8301-13578_3-10152117-38.html (“[The tool, M-Lab,] is running three diagnostic tools for consumers: one to determine whether BitTorrent is being blocked or throttled, one to diagnose problems that affect last-mile broadband networks, and one to

Why Dominant Search Engines & Carriers Deserve More Scrutiny than Dominant Auction Sites & Social Networks

Those skeptical of the administrative state may find this proposal to “watch the watchers” problematic. They think of intermediaries as primarily market actors, to be disciplined by market constraints. However, the development of dominant Web 2.0 intermediaries was itself a product of particular legal choices about the extent of intellectual property rights and the responsibilities of intermediaries made in legislative and judicial decisions in the 1990s. As intermediaries gained power, various entities tried to bring them to heel—including content providers, search engine optimizers, trademark owners, and consumer advocates. In traditional information law, claims under trademark, defamation, and copyright law might have posed serious worries for intermediaries. However, revisions of communications and intellectual property law in the late 1990s provided safe harbors that can trump legal claims sounding in each of these other areas.⁴⁶ Some basic reporting responsibilities are a small price to pay for continuing enjoyment of such immunities.

An argument for treating internet intermediaries more like regulated entities owes much to the trail-blazing work of legal realists. Among these, Robert Hale’s work on utilities remains especially inspirational.⁴⁷ Hale developed many of the theoretical foundations of the New Deal, focusing on the ways in which the common law became inadequate as large business entities began ordering

diagnose problems limiting speeds.”). It remains to be seen whether Google itself would submit to a similar inspection to determine whether it was engaging in stealth marketing or other problematic practices.

⁴⁶ 17 U.S.C. § 512(d) (2000) (Digital Millennium Copyright Act of 1998 safe harbor); 47 U.S.C. § 230(c)(1) (2000) (Communications Decency Act of 1997 safe harbor for intermediaries). For critical commentary on the latter, see Michael L. Rustad & Thomas H. Koenig, *Rebooting Cybertort Law*, 80 WASH. L. REV. 335, 371 (2005) (“An activist judiciary, however, has radically expanded § 230 by conferring immunity on distributors. Section 230(c)(1) has been interpreted to preclude all tort lawsuits against ISPs, websites, and search engines. Courts have ... haphazardly lump[ed] together web hosts, websites, search engines, and content creators into this amorphous category.”).

⁴⁷ Ilana Waxman, Note, *Hale’s Legacy: Why Private Property is Not a Synonym for Liberty*, 57 HASTINGS L.J. 1009, 1019 (“Hale’s most fundamental insight was that the coercive power exerted by private property owners is itself a creature of state power... . By protecting the owner’s property right ... ‘the government’s function of protecting property serves to delegate power to the owners’ over non-owners, so that ‘when the owners are in a position to require non-owners to accept conditions as the price of obtaining permission to use the property in question, it is the state that is enforcing compliance, by threatening to forbid the use of the property unless the owner’s terms are met.’ ... [A]ll property essentially constitutes a delegation of state power to the property owner...”). For a powerful application of these ideas to Internet law, see Julie Cohen, *Lochner in Cyberspace: The New Economic Orthodoxy of ‘Rights Management,’* 97 MICH. L. REV. 462 (1998).

increasing proportions of the national economy.⁴⁸ Hale's crucial insight was that many of the leading businesses of his day were not extraordinary innovators that "deserved" all the profits they made; rather, their success was dependent on a network of laws and regulation that could easily shift favor from one corporate player to another.⁴⁹ Hale focused his theoretical work on the utilities of his time, expounding an economic and philosophical justification for imposing public service obligations on them. Regulatory bodies like state utility commissions and the FCC all learned from his work, which showed the inadequacy of private law for handling disputes over infrastructural utilities.

Market advocates may worry that monitoring of search engines and carriers will lead to more extensive surveillance of the affairs of other intermediaries, like social networks and auction platforms. They may feel that competition is working in each of those areas, and should be the foundation of all intermediary policy. However, competition is only one of many tools we can use to encourage responsible and useful intermediaries. We should rely on competition-promotion via markets and antitrust only to the extent that (a) the intermediary in question is an economic (as opposed to cultural or political) force; (b) the "voice" of the intermediary's user community is strong;⁵⁰ and (c) competition is likely to be genuine and not contrived. These criteria help us map older debates about platforms onto newer entities.

For search engines and carriers, each of these factors strongly militates in favor of regulatory intervention. Broadband competition has failed to materialize beyond duopoly service for most Americans. There are several reasons to suspect that Google's dominance of the general purpose search market will continue to grow.⁵¹ Just as past policymakers recognized the need for common

⁴⁸ Duncan Kennedy, *The Stakes of Law, or, Hale and Foucault*, 15 LEGAL STUDIES FORUM (4) (1991).

⁴⁹ BARBARA FRIED, THE PROGRESSIVE ASSAULT ON LAISSEZ FAIRE: ROBERT HALE AND THE FIRST LAW AND ECONOMICS MOVEMENT (1998), available at <http://www.hup.harvard.edu/catalog/FRIPRA.html>.

⁵⁰ Competition is designed to provide users an "exit" option; regulation is designed to give them more of a "voice" in its governance. Hirschman ALBERT O. HIRSCHMAN, *Exit and Voice: An Expanding Sphere of Influence*, in RIVAL VIEWS OF MARKET SOCIETY AND OTHER RECENT ESSAYS 78–80 (1986) (describing "exit" and "voice" as two classic options of reform or protest). To the extent exit is unavailable, voice (influence) within the relevant intermediary becomes less necessary; to the extent voice is available, exit becomes less necessary.

⁵¹ Bracha & Pasquale, *Federal Search Commission*, *supra* note 26, at 1179. Section III of the article, "Why Can't Non-Regulatory Alternatives Solve the Problem?," addresses the many factors impeding competition in the search market. Present dominance entrenches future dominance as the leading search engine's expertise on user habits grows to the extent that no competitor can match its understanding of how to target ads well. *Id.* Since that article was published, Harvard Business School Professor Ben Edelman has investigated another self-reinforcing aspect of Google's market power: the non-portability of AdSense data, which

carrier obligations for concentrated communications industries, present ones will need to recognize carriers' and search engines' status as increasingly essential facilities for researchers, advertisers, and media outlets.⁵²

The parallel is apt because, to use the three dimensions discussed above, carriers and dominant general-purpose search engines a) are just as important to culture and politics as they are to economic life, b) conceal key aspects of their operations, and are essentially credence goods, vitiating user community influence, and c) do not presently face many strong competitors, and are unlikely to do so in the immediate future. The first point—regarding cultural power—should lead scholars away from merely considering economies of scale and scope and network effects in evaluating search engines. We need to consider all dimensions of network power—the full range of cultural, political, and social obstacles to competition that a dominant standard can generate.⁵³ Moreover, policymakers must acknowledge that competition itself can drive practices with many negative externalities. The bottom line here is that someone needs to be able to “look under the hood” of culturally significant automated ranking systems.

What about auction platforms, another important online intermediary?⁵⁴ Here, a purely economic, antitrust-driven approach to possible problems is more appropriate. To use the criteria mentioned above: (a) a site like eBay is a very important online marketplace, but has little cultural or political impact and (b) the user community at eBay understands its reputation rankings very well, and has shown remarkable capacities for cohesion and self-organization to protest

makes it difficult for Google customers to apply what they have learned about their Internet customers to ad campaigns designed for other search engines. Ben Edelman, *PPC Platform Competition and Google's 'May Not Copy' Restriction*, June 27, 2008, <http://www.benedelman.org/news/062708-1.html>. As Edelman shows, Google has tried to make the data it gathers for companies “sticky,” inextricable from its own proprietary data structures.

⁵² TIM WU, *THE MASTER SWITCH* (Knopf, 2010) (promoting “separations principle” in the digital landscape.).

⁵³ DAVID GREWAL, *NETWORK POWER: THE SOCIAL DYNAMICS OF GLOBALIZATION* 45 (Yale Univ. Press 2008) (“[T]he network power of English isn’t the result of any intrinsic features of English (for example, ‘it’s easy to learn’): it’s purely a result of the number of other people and other networks you can use it to reach... . The idea of network power ... explains how the convergence on a set of common global standards is driven by the accretion of individual choices that are free and forced at the same time.”).

⁵⁴ David S. Evans, *Antitrust Issues Raised by the Emerging Global Internet Economy*, 102 Nw. U. L. REV. COLLOQUY 285, 291 (2008) (“European Community law and decisional practice ... impose special obligations and significant scrutiny on firms that have market shares as low as 40 percent.”). Evans compiles data demonstrating that some leading auction platforms (such as eBay) are well above this market share in Europe and the U.S. *Id.* (citing comScore, *MyMetrix qSearch 2.0 Key Measures Report*, Dec. 2007, <http://www.comscore.com/method/method.asp>).

(and occasionally overturn) policies it dislikes. These factors overwhelm the possibility that (c) competition in the general auction market (as opposed to niche auctions) may be unlikely to develop. If real competitors fail to materialize due to illicit monopolization, antitrust judgments against Microsoft (and parallel requirements of some forms of “operating system neutrality”) can guide future litigants seeking online auction platform neutrality. While eBay’s user community successfully pressured Disney to end its 2000 special-preference deal with eBay, in the future antitrust judgments or settlements might require the full disclosure of (and perhaps put conditions on) such deals.⁵⁵

In social networks, another area where tipping can quickly lead to one or a few players’ dominance,⁵⁶ the situation is more mixed. While Rebecca Mackinnon and danah boyd have compared Facebook to a utility, the famously market-oriented Economist magazine has compared it to a country, possibly in need of a constitution and formal input from users. Social networks are closer to search engines than auction sites with respect to factor a: they are becoming crucial hubs of social interactions, cultural distribution and promotion, and political organizing.⁵⁷

On the other hand, social networks provide a some leverage to their members to police bad behavior, opening up “voice” options, with respect to factor b, far more potent than those available to the scattered searchers of Google. A group named “Facebook: Stop Invading My Privacy” became very popular within Facebook itself, catalyzing opposition to some proposed features of its Beacon program in 2008.⁵⁸ Facebook’s privacy snafus in early 2009 led the company to organize formal user community input on future alterations to the company’s terms of service. On the final factor, competitive dynamics, it appears that competition is more likely to develop in the social network space than in the broadband, search engine, or auction platform industries. There is a more

⁵⁵ In 2000, eBay granted special perks to Disney on a platform within its auction site. After protest from “the eBay community,” the perks ceased. eBay CEO Meg Whitman said of the special Disney deal: “We’ve concluded that eBay has to be a level playing field. That is a core part of our DNA, and it has to be going forward.” ADAM COHEN, *THE PERFECT STORE: INSIDE EBAY* 292 (Back Bay Books 2006).

⁵⁶ In early 2008, 98% of Brazilian social networkers used Google’s Orkut; 97% of South Korean social networkers used CyWorld, and 83% of American social networkers used MySpace or Facebook. Evans, *supra* note 54 at 292.

⁵⁷ James Grimmelman, *Saving Facebook*, 94 IOWA L. REV. 1137, 52-59 (2009), http://works.bepress.com/james_grimmelman/20/.

⁵⁸ William McGeveran, *Disclosure, Endorsement, and Identity in Social Marketing*, 2009 ILL. L. REV. 1105, 1120 (2009), http://www.law.uiuc.edu/lrev/publications/2000s/2009/2009_4/McGeveran.pdf.

diverse playing field here than in the carrier or search space, with more than 4,000 social networks in the United States.⁵⁹

Any policy analysis of dominant intermediaries should recognize the sensitive cultural and political issues raised by them. The cultural, communal, and competitive dynamics surrounding dominant search engines and carriers defy easy or stereotyped responses. Qualified transparency will assist policymakers and courts that seek to address the cultural, reputational, and political impact of dominant intermediaries.

Conclusion

As David Brin predicted in *The Transparent Society*, further disclosure from corporate entities needs to accompany the scrutiny we all increasingly suffer as individuals.⁶⁰ While the FTC and the FCC have articulated principles for protecting privacy, they have not engaged in the monitoring necessary to enforce these guidelines. This essay promotes institutions designed to develop better agency understanding of privacy-eroding practices. Whether public or private, such institutions would respect legitimate needs for business confidentiality while promoting individuals' capacity to understand how their reputations are shaped by dominant intermediaries.

⁵⁹ Evans, *supra* note 54, at 290.

⁶⁰ DAVID BRIN, *THE TRANSPARENT SOCIETY: WILL TECHNOLOGY FORCE US TO CHOOSE BETWEEN PRIVACY AND FREEDOM?* (Basic Books 1999).