

International Journal of Law and Information Technology Vol. 17 No. 1 © Oxford University Press 2008; All rights reserved.
doi:10.1093/ijlit/ean018 Advance Access Published on 29 October 2008

Private Power, Public Interest: An Examination of Search Engine Accountability

EMILY B. LAIDLAW*

Abstract

As information becomes a critical commodity in modern society, the issue is raised whether the entities that manage access to information, that are tools for public discourse and democracy, should be accountable to the public. The Internet has transformed how we communicate, and search engines have emerged as managers of information, organizing and categorizing content in a coherent, accessible manner thereby shaping the Internet user's experience. This article examines whether search engines should have public interest obligations. In order to answer this question, this article first examines comparative public interest regulatory structures, and the growing importance of the Internet to public discourse. Then examined is how the algorithmic designs and manual manipulation of rankings by search engines affects the public interest without a sufficient accountability structure. Finally, the values necessary to a public interest framework are suggested.

1 Introduction

We are living in an 'Information Society'¹ wherein information is now a critical commodity, and those that control this information, whether access to or delivery of it or its content, are in key positions of power. As the Internet infiltrates the very nature of how people relate and communicate, a

* PhD Candidate, Department of Law, The London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, e.b.laidlaw@lse.ac.uk

¹ See discussion in Webster, *Theories of the Information Society* (London: Routledge, 2nd ed, 2002), chapter 2.

spotlight shines on the companies that make these communications possible. By controlling the communication infrastructure of the Internet, they have become information gatekeepers. What they do increasingly embraces principles central to our democracy such as discourse, freedom of expression and public knowledge.² This has attracted the attention of policy makers who question whether or not the public should be protected from unrestricted and unaccountable private power over such an important communications resource. Search engines, it will be argued, are the new information gatekeepers, and have so far slipped through the regulatory net, operating without burdens and without regulation.

In his book *The Control Revolution*, Andrew Shapiro calls for 'collective public action' to create a balance between the market and government,³ identifying control of communication resources as a significant battleground between the market and government.⁴ He warned that if the Internet's communication resources are unregulated, an oligopoly of private power would result, thus 'dashing' the freedoms promised by these communication resources.⁵ He suggested the following guiding principle:

In a democratic society, those who control access to information have a responsibility to support the public interest. By dint of their power over such an important resource, these gatekeepers must assume an obligation as trustees of the greater good. Indeed, barring some clear showing that they are bearing this burden voluntarily, government should impose it upon them.⁶

While this author does not agree that a public interest obligation necessitates government interference, Shapiro raises interesting questions that have become central to this article. Before assessing who should regulate the Internet and in what contexts, certain critical questions must be asked. We must step back for a moment and examine the relationship between private companies and information, which will inform the nature of any regulatory constructs that may be imposed. One must ask: Where do search engines fit in the idea of the democratizing force of the Internet? Do search engines carry out a public function? Is there a shortfall in their accountability that can harm users? To that end, this article examines whether search engines should have public interest obligations, and identifies the values that a public interest framework should exhibit.

² Phrasing is from Shapiro, *The Control Revolution: How the Internet is Putting Individuals in Charge and Changing the World We Know* (New York: Public Affairs 1999) at p 224.

³ *Ibid.* at p 222.

⁴ *Ibid.* at p 224. This is the famous 'east coast code versus west coast code' articulated by Lawrence Lessig: Lessig, *Code and Other Laws of Cyberspace* (New York: Basic Books 1999). On the one hand, Congress enacts laws, while West Coast code writers, through design of software and hardware, determine how the Internet works: how information is to be accessed, transmitted, and stored thereby defining the user experience. Code, Lessig proposes, is law, because by building the environment of the 'social life', it constrains behaviour: *ibid.* at p 53.

⁵ n 2 above, at p 224.

⁶ *Ibid.* at p 225.

2 Definitions

This section contextualises search engines within the notion of online intermediaries and gatekeepers. Search engines are not the only gatekeepers of the Internet, and a public interest duty can also be justified for various other online gatekeepers.

An information or 'online' intermediary is a middleman between the user and the provider of information.⁷ In this role, the intermediary facilitates access to information (or realizes prevention of access), organizes, retrieves, presents, and guides consumers in their information experience. Depending on the function of the intermediary, it may also be a gatekeeper.

Gatekeeping connotes more control than the term intermediary; an additional, more specific, function must be carried out. In sociological terms, it refers to 'an individual who occupies a position that allows him or her to control access to goods, information, and services.'⁸ In traditional media, this can be defined as those who select 'who and what should be given access to channels of publicity or made visible in the public arena.'⁹ Europe's Electronic Commerce Directive¹⁰ implicitly recognizes these differences in assigning varying liability to service providers depending on their functions as mere conduits, caching, or hosts.¹¹

Broadly, intermediary functions can be categorized in terms of their communication services:¹²

Internet Service Providers ('ISP'): ISPs act as conduits between the information source and the recipient. Early in the Internet's commercialization, ISPs were mainly telephone companies, but now other players also offer this service, including cable companies and others with the money and resources to do so. Ronald Mann and Seth Belzley articulate three roles that ISPs play: backbone providers, source ISPs, and destination ISPs.¹³ Backbone ISPs transmit data to the endpoints.¹⁴ Destination ISPs connect the user to the Internet and the World Wide Web, while Source ISPs provide publishing services and online availability to the content

⁷ See Hamelink, *World Communication: Disempowerment & Self-Empowerment* (London: Zed Books 1995) at p 5.

⁸ Calhoun (ed), *Dictionary of Social Sciences* (Oxford: Oxford University Press 2002).

⁹ McQuail, *Media Accountability and Freedom of Publication* (Oxford: Oxford University Press 2003) at p 128. Jonathan Zittrain does not differentiate between these terms as starkly as this author, preferring to define a gatekeeper as being 'intermediaries of various kinds' such as businesses that host, index, and carry others' content: Zittrain, 'A History of Online Gatekeeping' (2006) 19(2) *Harvard Journal of Law & Technology* at p 253-254.

¹⁰ Directive 2000/31/EC (Directive on Electronic Commerce).

¹¹ *Ibid.*, articles 12-15.

¹² There are various ways to categorise online intermediaries. Mann and Belzley prefer to differentiate on the basis of the service provided: ISPs, Payment Intermediaries (i.e. Paypal) and Auction Intermediaries (i.e. eBay): Mann & Belzley, 'The Promise of Internet Intermediary Liability' (2005) 47 *William & Mary Law Review* 239 at p 246.

¹³ *Ibid.* at pp 255-257.

¹⁴ *Ibid.*

provider.¹⁵ ISPs are more appropriately described as intermediaries than gatekeepers, although they can take on gatekeeping functions when a more specific step, such as filtering content, is taken.

Search Engines: At its most basic level, search engines are websites that help users find information on other websites.¹⁶ They have been variously described as akin to library catalogues or the Yellow Pages index.¹⁷ As the number of websites multiplies, search engines become increasingly important resources for organizing the clutter of information on the Internet.¹⁸ They have become ‘the new linchpins on the Internet.’¹⁹ Although every search engine functions differently, certain common functionalities can be identified. Computer robots called spiders crawl the Internet for information, index key words, and allow users to search for words on this index.²⁰ They range from meta-search engines that aggregate the results from several search engines, to general search engines, to issue-specific search engines for scholarship, travel and various other themes.

The current leading brands on the Internet according to Nielson NetRatings all have a search engine as part of their business portfolio: Yahoo!, Microsoft, MSN (which is owned by Microsoft), Google, AOL and eBay (which although it is an auction business, it is operated by a search engine).²¹ However, despite the plethora of search engines,²² there are arguably only two major search engines: Google and Yahoo!.²³ Although this paper will argue for public interest regulation of all search engines, the focus will be on Google and Yahoo! because of their dominance of the search market.

Content Providers and Hosts: Hosts are the systems that store data²⁴ in any form capable of proprietary protection such as text, graphic, audio and video. We are concerned here, in particular, with companies that host sites

¹⁵ Ibid.

¹⁶ See <http://www.howstuffworks.com> (last visited 25 June 2008), and Grimmelmann, ‘The Structure of Search Engine law’ (2007) 93 *Iowa Law Review* 1 at p 6.

¹⁷ Van Couvering, ‘New Media? The Political Economy of Internet Search Engines’, presented to the Communication Technology Policy section at the 2004 Conference of the International Association of Media & Communications Researchers at p 3.

¹⁸ Zarsky, ‘Search Engines, Media Concentration and the Marketplace of Ideas’ (Draft), position paper presented at the Regulating Search Symposium, Yale Law School, 2005 at p 3, at <http://islandia.law.yale.edu/isp/regulatingsearch.html> (last visited 10 July 2007).

¹⁹ Grimmelmann, n 16 above, at p 3.

²⁰ <http://www.howstuffworks.com>, above n 16; <http://www.searchenginehistory.com> (last visited 25 June 2008). Some search engines operate like a directory where inclusion on the index is paid for: see discussion below 4.3.1.

²¹ ‘Top Ten Brands in Internet’ at <http://www.internetworldstats.com/top10.htm> (last visited 25 June 2008).

²² See ‘LM Search and Web Searches’ at <http://www.jus.uio.no/lm/reference.and.resource.miscellany/web.search.html> (last visited 25 June 2008).

²³ n 17 above, at p 24. Google Watch adds Microsoft and AskJeeves to this list: ‘And Then There Were Four’ at <http://www.google-watch.org/bigbro.html> (last visited 25 June 2008). For further discussion, see below 4.3.1.

²⁴ See, for example, <http://www.webopedia.com/TERM/H/host.htm> (last visited 25 June 2008).

for blogs, bulletin boards, video sharing, file sharing, and so on. A content provider, borrowing from the Communications Decency Act²⁵ section 230(f) (3), is 'any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service.'²⁶ The role of content providers and hosts as gatekeepers, both commercial and private, raises several issues worthy of further research.

Any understanding of the meaning of 'information intermediary' must be accompanied by a warning that its meaning is protean in its manifestation. This is largely for two reasons. First, the leading Internet businesses cannot be neatly categorised as content providers, ISPs, or search engines. Most cannot agree on the categories. In addition, the concentration of the market, and diversification of business interests, has created an oligopoly of private power. For example, although search continues to be Google's primary business,²⁷ its' business has expanded far beyond simple search. This diversification of business interests, largely ignored in its own description of what the company does,²⁸ is a focus of concern for privacy experts and competition experts.²⁹

This concentration of the market combined with the diversity of such businesses services intensifies the importance of determining whether search engines have a public interest function necessitating regulation.

²⁵ (1996) 47 U.S.C.

²⁶ *Ibid.*, s 230(f) (3).

²⁷ Google still treats improving search as its primary mission: 'Google's mission is to organize the world's information and make it universally accessible and useful': 'Investor FAQ' at <http://investor.google.com/faq.html> (last visited 25 June 2008). In terms of search, Google has created specialist sites to search for such things as products, blogs, desktops, scholarly materials, books, news, and videos: <http://www.google.com/intl/en/options/> (last visited 25 June 2008).

²⁸ *Ibid.*, <http://investor.google.com/faq.html>.

²⁹ Google now offers maps and satellite images of the Earth and of city streets, the latter currently heavily criticized for potentially breaching individuals' privacy: Robertson, 'Google's Street View could be unlawful in Europe' at <http://www.out-law.com//page-8116> (last visited 25 June 2008). It has introduced applications such as email ('Gmail'), and Docs and Spreadsheets. It has ventured into the market of content sharing with YouTube (video), Picasa (pictures and graphics), Sketch Up (design), and social networking with Orkut and Blogger: n 27 above, <http://www.google.com/intl/en/options/>, and 2006 Annual Report, 'Letter from the Founders' at http://investor.google.com/2006_founders_letter.html (last visited 25 June 2008). In pursuit of its' goal to create 'a single and complete advertising system' (*ibid.*) Google bought DoubleClick, which raised both anti-competitive concerns (for which it was investigated and cleared by the U.S. Federal Trade Commission: 'Federal Trade Commission Closes Google/DoubleClick Investigation' at <http://www.ftc.gov/opa/2007/12/googledc.shtm> (last visited 24 June 2008)) and privacy concerns: 'Consumers groups alarmed by Google online ad merger' at <http://www.euractiv.com> (last visited 18 July 2007). Most recently, Google is moving into the online security market with its agreement to purchase Postini: <http://googleinvestors.blogspot.com/> (last visited 25 June 2008). For further information on Google's business plans see this blog.

Like Google, Yahoo! has diversified its business from its early days as an online directory. It describes itself as 'the No. 1 Internet brand globally': <http://yhoo.client.shareholder.com/faq.cfm> (last visited 25 June 2008). Most notably, it quickly diversified into a web portal, presenting information from various sources regarding, *inter alia*, entertainment, news, and business investments. In addition, it now offers applications such as Yahoo!mail, Del.icio.us (a social bookmark manager), Flickr (a photo host) and MyBlogLog: E. Sokullu, 'Yahoo! 2.0: Its Reorganization and Future' at www.readwriteweb.com/archives/yahoo_reorganization_future.php (last visited 25 June 2008).

3 The Internet as a Public Interest Medium

In crafting a public interest obligation for search engine providers, analogy will be made from the recognition that traditional media carries out a business of importance to society. The public interest notion for both the media and the Internet is rooted in their shared role as conveyors of information and facilitators of public discourse. This will be used as a platform to launch an inquiry into the unique public interest of search engines. It is an issue of accountability, a more accessible word for our purposes, particularly in the context of online gatekeepers, which has been defined as, to explain one's actions and justify them on normative grounds.³⁰

3.1 *Theories of Public Interest in Traditional Media*

For the purposes of this article, public interest is not intended to connote the more limited definition employed in the media context for defences to defamation actions or public service broadcasting. Rather, this term is used in the broad form embraced by theorists for the central role of certain industries in the functioning of democracy. In this context, the definition of public interest articulated by Denis McQuail shall be used:

[T]hey carry out a number of important, even essential, informational and cultural tasks and it is in the general interest (or good of the majority) that these are carried out well and according to principles of efficiency, justice, fairness, and respect for current social and cultural values.³¹

The origins of public interest regulation are with public utilities, such as transport and electricity.³² They were perceived as businesses 'affected with a public interest',³³ justifying imposition of regulation to ensure adequacy of service (assurance of equity of access and efficiency) and to control monopolies.³⁴

Although public interest regulation of the media grew out of this arena, the media is different. Justification is largely found in the freedoms inspired and facilitated by the media.³⁵ There is little disagreement now as to the idea of public interest in the media, but consensus is lost when

³⁰ McQuail, n 9 above, at p 15. In McQuail's view, the basic values of media accountability are truth, freedom, order and cohesion, solidarity and equality, right purpose and responsibility: *ibid.*, chapter 4.

³¹ *Ibid.* at p 47.

³² McQuail, *Media Performance: Mass Communication and the Public Interest* (London: Sage Publications 1992) ('McQuail2') at p 21. William Melody gives the following examples of businesses affected with a public interest: '[i]nns, wharves, bridges, canals, grain warehouses, railways, electricity, gas, water, telephone and other services': Melody, 'Communication Policy in the Global Information Economy: Whither the Public Interest?' in Ferguson (ed), *Public Communication: the New Imperatives, Future Directions for Media Research* (London: Sage Publications 1990) at p 30.

³³ *Ibid.*

³⁴ McQuail2, n 32 above, at pp 4, 21-22.

³⁵ *Ibid.* at p 21.

trying to formulate the exact qualities of this public interest and how it should be manifested.³⁶

Traditional media is at a crossroads as much as the Internet, grappling with changing technologies, convergence, and concentration of the market, leading some to argue that the notion of ‘public interest’ for traditional media needs to be re-defined.³⁷ What should be the goal of public interest obligations? One view is that a public interest obligation has three functions in a democracy: to facilitate judgment and comment on government actions; for the common good; and as a general check on actions.³⁸ Although the Internet cannot be seamlessly inserted into such a public interest framework, analogy can be drawn from the theoretical purpose of public interest regulation of the media.

Treating the media as having a public function is justified because of its central role in public discourse.³⁹ To achieve a vision of a participatory democracy, citizens must have the opportunity to be informed and to be heard, and the media are the primary source of such news and information.⁴⁰ They organize, filter, and interpret information for the public. In so doing, they embrace a powerful role, because they shape public opinion and meaning. Thus regulation has focused on ensuring ‘media access that reflect the broad interests of the general public’,⁴¹ such as balanced reporting, diversity of content, and rights of reply.⁴²

The media’s activities are central to the theoretical justifications for freedom of expression.⁴³ Eric Barendt identifies four categories of historical justifications for freedom of expression: from self-fulfilment, from truth, from democracy, and from suspicion of government.⁴⁴ The media might serve one or more of these categories, but it is the argument from democracy that is most persuasive. The argument from democracy focuses on the idea that participation in discussions and debate are central to the functioning of democracy.⁴⁵ As a self-standing theory of freedom of expression it can be somewhat limiting as such expression is only protected to the extent that it serves democracy. However, in the context of the media, it heightens its power and responsibilities as servants of democracy. The media is arguably ‘the main cultural institution and the principal means of public expression in contemporary society’,⁴⁶ whose main

³⁶ Ibid. at p 3-4.

³⁷ See Feintuck & Varney, *Media Regulation, Public Interest and the Law* (Edinburgh: Edinburgh University Press, 2nd ed, 2006).

³⁸ McQuail2, n32 above, at p 20.

³⁹ n 37 above.

⁴⁰ Melody, n 32 above, at pp 18-19.

⁴¹ Ibid. at p 19.

⁴² Ibid.

⁴³ n 37 above, at p 15.

⁴⁴ Barendt, *Freedom of Expression* (Oxford: Oxford University Press, 2nd ed, 2005), chapter 1.

⁴⁵ Ibid. at p 18.

⁴⁶ McQuail, n 9 above, at p 4.

role is to 'promote an effective public space'.⁴⁷ As such, they have a crucial job in making sure that citizens are informed in order to further their role in democracy.⁴⁸

This notion of the media as a force for freedom of expression, and a servant of democracy, is the basis for the argument that the media is within the public sphere.⁴⁹ Termed by Jurgen Habermas in *The Structural Transformation of the Public Sphere*,⁵⁰ he used it to describe a metaphorical space protected for rational public discourse on issues of common concern.⁵¹ Since few citizens can gather at once for such discourse, 'the mass media have become the chief institutions of the public sphere.'⁵² The media's role is to promote this public space so that diverse viewpoints can be voiced, debated and resolved and public opinion formed.⁵³ Habermas' view has been criticized as naive, idealistic and undemocratic, in addition to underplaying the media's power for negative influence.⁵⁴ Notwithstanding, it is the crucial role of the media in enhancing democracy that resonates regarding the Internet, and search engines in particular.

3.2 *The Democratizing Force of the Internet*

In order to impose any public interest obligations on Internet businesses, these businesses must carry out activities that are important in the functioning of democracy, which raises the initial question of the meta-level importance of the Internet to democracy.

Information societies are so defined for having 'become dependent upon complex electronic information and communication networks'⁵⁵ in order to function efficiently. Information is arguably 'a distinguishing feature of the modern world.'⁵⁶ It infiltrates all aspects of daily life from 'public life, to work, leisure, education, and consumption'.⁵⁷ Although information has historically been recognized as integral to the effective functioning of society, its role is now different and more complex due largely to the technological communication networks that have been developed changing not only how information is accessed, communicated, retrieved, and interpreted,⁵⁸ but also the speed with which this is done. Manuel Castells frames information and

⁴⁷ Ibid. at p 62.

⁴⁸ n 37 above, at p 5.

⁴⁹ McQuail, n 9 above, at p 4; n 37 above, at p 15.

⁵⁰ Habermas, *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society* (Translated Cambridge: Polity 1992).

⁵¹ McQuail, n 32 above, at p 6; McQuail, n 9 above, at p 61; Papacharissi, 'The Virtual Sphere: the internet as a public sphere' (2002) 4(1) *New Media & Society* 9 at pp 10-11.

⁵² n 37 above, at p 15 (quoting Dahlgren).

⁵³ McQuail, n 9 above, at p 61-62.

⁵⁴ n 37 above, at p 15-16; Papacharissi, n 51 above, at p 11-12.

⁵⁵ Melody, n 32 above, at p 26-7.

⁵⁶ n 1 above, at p 1.

⁵⁷ McQuail, n 9 above, at p 4.

⁵⁸ See Melody, n 32 above, at p 30-31.

communication technologies as giving priority to information flows, spurring an era of 'information capitalism'.⁵⁹ Now, 'access to information and communication would appear to be the most essentially public utility.'⁶⁰

Urs Gasser suggests three core democratic values to the Internet. First, the value of informational autonomy in the sense that an individual has the right to choose between various sets of information, but also the right to express his or her opinions, and create sources of information.⁶¹ Second, he promotes diversity in the information sources and distribution.⁶² Third, since we are increasingly dependent on the information we access on the Internet, the quality of this information is a value.⁶³ Such formulations derive from a recognition that the Internet's primary function is as 'a conveyor of information.'⁶⁴ In contrast to traditional media, the costs to becoming a speaker in cyberspace are low, and the distributed architecture of the Internet have 'fundamentally altered' the capacity of individuals to engage in the public sphere.⁶⁵

The Internet has emerged as an essential source for access to information and communication. It is rapidly approaching television with its market penetration. The Internet and Multimedia 2007 report states that 33 per cent of surveyed consumers view the Internet as the most essential medium, trailing television by only three per cent.⁶⁶ Over one billion people in the world use the Internet.⁶⁷ The Internet is also closing in on other mediums with respect to time individuals devote to its use. On average, users spend one quarter of their weekly media time on the Internet.⁶⁸ Yet what is most compelling is the growing role of the Internet in self-development.⁶⁹ In a study by the Pew Internet & American Life Project, 45 per cent of American Internet users said that 'the [I]nternet helped them to make big decisions or negotiate their way through major episodes in their lives in the previous two years.'⁷⁰ The Internet played a sometimes major

⁵⁹ n 1 above, at p 100.

⁶⁰ Melody, n 32 above, at p 31.

⁶¹ Gasser, 'Regulating Search Engines: Taking Stock and Looking Ahead' (2006) 9 *Yale Journal of Law & Technology* 124 at p 150.

⁶² *Ibid.* at p 151.

⁶³ *Ibid.* at pp 152-153.

⁶⁴ Introna & Nissenbaum, 'Shaping the Web: Why the politics of search engines matters' (2000) 16(3) *The Information Society* 169 at p 179.

⁶⁵ Benkler, *The Wealth of Networks: How Social Production Transforms Markets and Freedom* (New Haven: Yale University Press 2006) at p 212.

⁶⁶ Needle, 'Internet Closing in on TV as 'Most Essential' Medium' at www.internetnews.com/bus-news/print.php/3685711 (last visited 25 June 2008).

⁶⁷ 'Sample Internet Statistics Compendium July 2007' at <http://www.e-consultancy.com/publications/internet-stats-compendium/> (last visited 25 June 2008).

⁶⁸ *Ibid.*

⁶⁹ See for example, discussion of the cultural elements of breast cancer patients' use of internet spaces by Orgad, 'The cultural dimensions of online communication: a study of breast cancer patients' internet spaces' (2006) 8(6) *New Media & Society* 877.

⁷⁰ Horrigan & Rainie, 'The Internet's Growing Role in Life's Major Moments' Pew Internet & American Life Project (2006) at http://www.pewinternet.org/pdfs/PIP_Major_per_cent20Moments_2006.pdf (last visited 25 June 2008).

role for American users regarding such critical life experiences as illness, careers, home and finance.⁷¹ The democratizing force of the Internet is taking shape, becoming a tool for speech as a component of one's self-fulfilment.⁷²

3.3 *Public Interest Regulatory Structures*

The media's core role is to publish.⁷³ They have a direct relationship with the consumer as an information source, which informs the regulatory approaches between the three different kinds of media. First, under the print media model, the least degree of regulation is imposed both in the United Kingdom and North America.⁷⁴ It is consultative in the sense that the user chooses what newspaper to read, if any, and what articles to read therein. This may be described as the pull model and includes other media such as film and music.⁷⁵ The second model is for common carriers such as mail, telegraph and telephone, which are interactive in nature and facilitate point-to-point individual interaction.⁷⁶ Such media are historically considered public resources. The third model is the broadcasting or 'push' model, wherein publishing is from one to many.⁷⁷ This industry is heavily regulated on the basis of spectrum scarcity, and due to the potentially invasive nature of broadcasting in a person's private home.

The Internet is arguably part of the mass media⁷⁸ as a giant, interactive publishing house. Yet, the businesses that facilitate the functioning of the Internet and access to the information thereon are not necessarily publishers. Unless these businesses are content providers, they are one or more steps removed from the publishing process, and are often more akin to the common carrier analogy with the telecommunications industry. Even hosts of websites, such as youtube.com, do not create content but rather provide space and storage for others to publish videos. Editorial control is generally avoided, if possible, by hosts of any interactive sites in order to avoid incurring liability for the contents therein.⁷⁹ Search engines are a hybrid. While they are in essence websites that index other sites on the Internet, they have reached a level of power and influence akin to television networks and Hollywood studios.⁸⁰ Recently, a court in the United

⁷¹ Ibid. summary at http://www.pewinternet.org/PPF/r/181/report_display.asp (last visited 25 June 2008).

⁷² n 44 above, at p 13-18.

⁷³ McQuail, n 9 above, at p 4.

⁷⁴ Ibid. at p 108.

⁷⁵ Ibid. at pp 108-110.

⁷⁶ Ibid.

⁷⁷ Ibid.

⁷⁸ See, for example, Orgad, 'The internet as a moral space: the legacy of Roger Silverstone' (2007) 9(1) *New Media & Society* 33 at p 34, and n 64 above. But see n 17 above.

⁷⁹ This is particularly so in Europe where liability depends, to an extent, on knowledge and control: see Electronic Commerce Directive, n 10 above.

⁸⁰ n 17 above, at p 3.

States went so far as to characterize Google's search results as 'opinion' and protected by the First Amendment.⁸¹

Within the regulatory models in traditional media, the Internet adopts some of all three functions. To describe the Internet as a pull technology would be to oversimplify the infrastructure of the Internet and the businesses that exploit it. Roy Rosenzweig describes the Internet as a 'meta-medium' 'that combines aspects of the telephone, post office, movie theatre, television, newspaper, shopping mall, street corner, and a great deal more.'⁸² It is a platform for broadcasting, group discussion, pulpit preaching and researching, and for a variety of content types such as text, video, audio, and image.⁸³ While others caution that incorporating the Internet as a whole into the media category might be inappropriate, describing it rather as 'a technological infrastructure that affords a range of uses',⁸⁴ such a characterization fails to recognize the Internet's democratic force. It is indeed a meta-medium, with the implication that it cannot be pigeonholed into an existing communication model because technological developments and convergence have not only blurred the boundaries of these models but also entirely collapsed them. It is necessary then to view the Internet, and the businesses that operate it, as unique.

If the Internet is an important democratic force, then the institutions that facilitate this public discourse become the target for any public interest obligations. One such key industry is search engines.

4 The Public Interest Duty of Search Engines

4.1 *Why Search Engines owe a Public Interest Duty: They Control Our Informational Experience*

Search engines are more dynamic than simple indices of websites. This is reflected in James Grimmelmann's definition that a search engine 'combines its own knowledge of available content with user queries to provide recommendations to its users'.⁸⁵ It tips its hat at the power of search engines to channel users' attentions to certain information and sites. There is an informational flow to the functioning of search engines,⁸⁶ and

⁸¹ *Search King v Google Technology, Inc.* 2003 WL 21464568 (W.D.Okla.). See discussion below 4.3.2 sub-heading 'Removal at Search Engine Behest'.

⁸² Rosenzweig, 'How Will the Net's History Be Written? Historians and the Internet' in Nissenbaum & Price (eds), *Academy & the Internet* (New York: Peter Lang, 2004) at p 26. See also DiMaggio *et al.*, 'Social Implications of the Internet', *Ibid.* at p 36.

⁸³ *Ibid.*

⁸⁴ n 17 above, at p 6.

⁸⁵ Grimmelmann, n 16 above, at p 4.

⁸⁶ *Ibid.* at p 7.

by controlling this flow search engines have set themselves up as gatekeepers and gateways to the informational experience:

1. [T]he search engine gathers content;
2. a user queries the search engine;
3. the search engine provides the user with results;
4. the user obtains the content.⁸⁷

These flows involve several players: the search engines, content providers, users, government, intellectual property owners and other concerned parties.⁸⁸ By controlling the structure of how information is accessed, search engines control the information flow. Without more, this might not be as consequential, however, search engines are now the portals through which the information on the Internet is experienced. They are seen as authoritative and reliable, and shape public opinion and meaning.

Search engines are some of the most commonly viewed websites on the Internet, almost equalling email for their commonality of use.⁸⁹ As stated by Helen Nissenbaum and Lucas D. Introna, '[t]o exist is to be indexed by a search engine.'⁹⁰ Their role is one of facilitation of access to information rather than direct access to information.⁹¹ As facilitators, they have grown into a critical role on which users depend to make information more easily accessible.⁹² Contrary to the problems of bandwidth scarcity that motivated much of current broadcasting regulation, the Internet struggles to manage information clutter.⁹³ This plethora of information is combined with a lack of knowledge regarding how to find it. Without search engines a user must know the URL (uniform resource locator, or web page address). In Germany, a 2004 survey indicated that 75 per cent of users rely on search engines 'as their principal means of finding web pages.'⁹⁴ As a result, search engines are the principal figures in sorting through and organizing information.⁹⁵

⁸⁷ Ibid.

⁸⁸ Grimmelman sees the information flow as only involving four players: search engines, content providers, users and concerned third parties: *ibid.* In this authors view, governments and intellectual property owners are significant stakeholders in the information flows and must be identified separately from general concerned third parties.

⁸⁹ Hargittai, 'The Social, Political, Economic, and Cultural Dimensions of Search Engines: An Introduction' (2007) 12 *Journal of Computer-Mediated Communication* 3 at p 1.

⁹⁰ n 64 above, at p 171.

⁹¹ Van Eijk, 'Search Engines: Seek and Ye Shall Find? The Position of Search Engines in Law' *Iris plus*, Legal Observations of the European Audiovisual Observatory, Issue 2006-02 at p 5.

⁹² Ibid.; DiMaggio, n 82 above, at p 42.

⁹³ n 18 above, at p 3; Elkin-Koren, 'Let the Crawlers Crawl: on Virtual Gatekeepers and the Right to Exclude Indexing' (2000-01) 26 *University of Dayton Law Review* 179 at p 180.

⁹⁴ Schulz *et al.*, 'Search Engines as Gatekeepers of Public Communication: Analysis of the German framework applicable to internet search engines including media law and anti trust law' (2005) 6(1) *German Law Journal* 1419 at p 1421.

⁹⁵ n 18 above, at p 3.

Yet, 80 per cent of sites visited are just 0.5 per cent of the available websites,⁹⁶ and one-quarter of the most visited websites are ‘portal’ sites, such as Yahoo! and AOL, which might provide, *inter alia*, search engines, category guides, shopping, and information services.⁹⁷ Additionally, most users do not link to sites listed on these indices beyond the first or second page of the search results.⁹⁸ One study states that 80.6 per cent of users review the first page of search results, while only 13.2 per cent view the second page of results.⁹⁹ Thus to be seen is to not only be indexed, but to be highly ranked in the search results.

The market for search engines also raises further concern. While there are several search engines available to a user, the market is significantly more concentrated than just a few years ago. A June 2007 study by Nielsen/Net Ratings, shows that Google holds 52.7 per cent of total searches, followed by Yahoo! at 20.2 per cent, and MSN at 13.3 per cent.¹⁰⁰ Arguably, Yahoo! and Google are the only major search engines.¹⁰¹

Search engines do more than just rank and organize information. They represent a broader struggle ‘to sustain the democratic potential of traditional media, the Internet, and the World Wide Web in particular.’¹⁰² They are not simply directories, as has been asserted by one provider,¹⁰³ but have a political dimension to their business as well, because they affect meaning and thus shape public opinion.¹⁰⁴ Niva Elkin-Koren comments:

They structure categories in response to users’ queries, and thereby have the capacity of creating categories for grasping the world. By defining which information becomes available for each query, search engines may shape positions, concepts and ideas.¹⁰⁵

Their capability to shape public opinion is magnified by user expectations that the search engines will provide relevant and reliable search results. This is a question of quality, on which there is no consensus of meaning.¹⁰⁶ In an interview of leading former or current employees of major search engine providers, Elizabeth Van Couvering concluded that conceptions of quality followed two lines of reasoning: the market schema, which is interested in the business aspect of search engines and judges

⁹⁶ DiMaggio, n 82 above, at p 42.

⁹⁷ *Ibid.*

⁹⁸ n 17 above, at p 17.

⁹⁹ n 94 above, at p 1421.

¹⁰⁰ Burns, ‘Top 10 Search Providers, June, 2007’ at <http://searchenginewatch.com/3626726/print> (last visited 5 June 2008).

¹⁰¹ n 17 above, at p 24.

¹⁰² n 64 above, at p 170.

¹⁰³ See *KinderStart.com, LLC v Google, Inc.*, Case 5:06-cv-02057-JF (2007) (DC N.Cali), discussion below 4.3.2 sub-heading ‘Removal at Search Engine Behest’.

¹⁰⁴ Elkin-Koren, n 93 above, at p 185.

¹⁰⁵ *Ibid.* at pp 185-186.

¹⁰⁶ Van Couvering, ‘Is Relevance Relevant? Market, Science, and War: Discourses of Search Engine Quality’ (2007) 12(3) *Journal of Computer-Mediated Communication* 866.

quality on customer satisfaction, and the science/technology schema from an engineering viewpoint that judges quality based on the relevance of search results.¹⁰⁷ Regardless of the approach, she concluded that they leave little room for issues of public interest such as fairness or bias.¹⁰⁸ Reliability does not go to the authenticity and credibility of a site per se, but is directed rather at the expectation of unbiased search results. As an example, Google has overtly encouraged users to rely on it to assess the value of websites, and to depend on its rankings as a reliable indicator of quality, worth and relevance. A user can download PageRank to his or her toolbar, which will then advise the user, on a scale of 0 to 10, with 10 being the highest, of the importance of the page being viewed. Google advertises the service by asking '[w]ondering whether a new website is worth your time?'¹⁰⁹

The act of framing a user's information experience makes search engines indispensable to access to information on the Internet thus elevating their product to public good status.¹¹⁰

4.2 *The Difficulty with imposing a Public Interest Duty*

Mann and Belzley rightly warn that there is a potential chilling effect in imposing any liability on gatekeepers for the services they provide,¹¹¹ as this might upset the market balance and chill the provision of goods and services. Yet this may be justified when there are social benefits to imposing responsibility that outweighs the resulting burdens.¹¹² Imposing responsibility on search engines for the purportedly automatic rankings might also chill innovation, and suggested solutions to the rankings issue often involve a form of imposed innovation (i.e. modifying algorithms for randomized ranking¹¹³).

Further, manual manipulations can play an important role in protecting the relevancy of search results in the increasingly common situation of link farms and search engine optimisation.¹¹⁴ In addition, some assert that the notion of search engine neutrality is a 'myth', because there can be no objectivity when editorial control determines the results.¹¹⁵ This focuses on the wrong issue, which is, rather, the importance of search engines to information access and public discourse.

¹⁰⁷ Ibid.

¹⁰⁸ Ibid. at p 17.

¹⁰⁹ 'Toolbar for Internet Explorer Help Center' at <http://www.google.com/support/toolbar/bin/static.py?page=features.html> (last visited 25 June 2008).

¹¹⁰ See discussion, n 64 above, at p 178-181.

¹¹¹ n 12 above, at pp 273-274.

¹¹² Ibid. For further discussion on market solutions see Elkin-Koren, n 93 above, at p 103.

¹¹³ Pandey *et al.*, 'Shuffling a Stacked Deck: The Case for Partially Randomized Ranking of Search Engine Results' (2005) at <http://oak.cs.ucla.edu/~cho/papers/cho-shuffle.pdf> (last visited 25 June 2008).

¹¹⁴ See discussion below 4.3.1.

¹¹⁵ Goldman, 'Search Engine Bias and the Demise of Search Engine Utopianism' (Draft), position paper presented at the Regulating Search Symposium, Yale Law School, 2005 at <http://islandia.law.yale.edu/isp/regulatingsearch.html> (last visited 10 July 2007).

Search engine providers should become responsible when ‘publishing and disseminating the information they gather.’¹¹⁶ However, this responsibility is not without drawbacks. Authoritativeness invites certain negative consequences, in that authority lends credibility¹¹⁷ to, for example, statements by women about their ex-boyfriends on www.dontdatehimgirl.com. With responsibility comes an increased cost of negotiating this responsibility, which limits entrance to an already concentrated market.¹¹⁸ In addition, some search engines such as Google are vulnerable if accountability is imposed because users and businesses engage in their own manipulation of rankings.

Online intermediaries are recognizing the important ethical role they play in their dealings with oppressive regimes in Asia and the Middle East.¹¹⁹ Google, Microsoft, Yahoo! and Vodafone are working with the Berkman Center for Internet and Society at Harvard University, Human Rights Watch and the Center for Democracy and Technology to frame a code of conduct.¹²⁰ Their purpose is to commit to principles for the protection of human rights guiding businesses in the challenges faced by doing business internationally.¹²¹ More recently, Google has gone further in its efforts to fight government censorship of the Internet by asking US trade officials to treat censorship of the Internet as a barrier to international trade.¹²²

These businesses should be encouraged to address their role as information gateways and gatekeepers closer to home. Although the Internet as a whole reflects a mixture of regulatory models from the traditional media, at the moment search engines are ‘lightly regulated’ with a regime more akin to print media as it is ‘without guarantees for the citizen that are embodied in the public service model of broadcasting’.¹²³ Yet the press have long acknowledged their central role in democracy, and have adopted self-regulatory regimes to ensure that the public interest is met.

So far, search engines and online intermediaries generally have not collaborated to address whether it owes the public an obligation, and if so, what principles a self-regulatory regime might espouse. Behaviour is currently competitive and proprietary, with each company declaring its priority as the consumer, but at the same time shirking any notion of obligation

¹¹⁶ Pasquale, ‘Rankings, Reductionalism, and Responsibility’ (Rough Draft) (2006) at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=888327 (last visited 25 June 2008) at p 18.

¹¹⁷ *Ibid.* at p 19.

¹¹⁸ See discussion, *ibid.* at p 19-27.

¹¹⁹ Google recognises its social importance to an extent. In 2004 it was reported that Google has an Ethics Committee that assesses changes to PageRank but no further information is available about or from this Committee: Orłowski, ‘Google’s Ethics Committee Revealed’ at http://www.theregister.co.uk/2004/05/17/google_ethics_committee/print.html (last visited 25 June 2008).

¹²⁰ Out-law.com, ‘Google, Yahoo!, Commit to Ethical Code’ at <http://www.out-law.com/page-7682> (last visited 25 June 2008).

¹²¹ *Ibid.*

¹²² Rugaber, ‘Google Looks to U.S. to curb censorship’ at <http://www.msnbc.msn.com/id/19372772/print/1/displaymode/1098/> (last visited 25 June 2008).

¹²³ n 17 above, at p 23.

regarding its power to shape meaning for the public. Google recently took a positive step with the launch of a public policy blog open to the public to contribute their opinions on policy issues.¹²⁴ As Frank Pasquale notes, search engines have become ‘the chief *organizer* and *forum* for research, public discussion, and commercial competition among internet users.’¹²⁵ They are a central figure in facilitating the democratic potential of the Internet. Without recognition by the search engine businesses of this role, and collaboration to frame an accountability regime, they might be faced with government regulation to protect the public interest.

4.3 *Search Engine Practices Affecting the Public Interest*

The goal to protecting the public interest is the creation of a framework that achieves accountability. This author proposes four priorities to achieving accountability: relevant and unbiased search results, a degree of transparency, respect for user dignity, and the implementation of an independent complaints mechanism. The lack of accountability is due to search engine practices associated with how they organize and rank information on indices. There are two levels to this lack of accountability. First, the design of the search algorithms fails to satisfy the above values. Second, the manual manipulation of rankings lacks transparency, consistency, or explanation, which also fails to satisfy these values.

4.3.1 *Algorithm Design*

Each search engine operates differently, and is aimed at varying information sources. Some search engines aggregate results from several search engines (meta-search engines such as InfoSeek),¹²⁶ while others target specific information such as scholarship or travel. The algorithms can be broken down into three categories: the spider approach, the directory approach and the popularity approach.¹²⁷

Search Engines function through the use of spiders or ‘bots’ that crawl the hundreds and millions of pages on the Internet for information and record key words from these sites as well as their links.¹²⁸ A searchable index is then created of the words, and the spider will return to the sites regularly to look for changes. A user submits keywords for his or her search and is provided with a list of websites.¹²⁹

¹²⁴ Metz, ‘Google launches self-protection blog’ at http://www.theregister.co.uk/2007/06/19/google_public_policy_blog/print.html (last visited 5 June 2008).

¹²⁵ n 116 above, at pp 33-34.

¹²⁶ See discussion by Elkin-Koren, n 93 above, at pp 192-194.

¹²⁷ See n 64 above, at pp 171-173.

¹²⁸ See <http://www.searchenginehistory.com>, n 20 above, ‘A History of Search Engines’ at <http://www.wiley.com/legacy/compbooks/sonnenreich/history> (last visited 25 June 2008), and <http://www.howstuffworks.com>, n 16 above.

¹²⁹ The term ‘keyword’ is not used in its normative sense, but rather is “deduced” from the webpages themselves in the process of indexing’: n 64 above, at p 171.

With the directory approach a website owner submits its website for approval to a human editor to be included on a list.¹³⁰ A user then searches the directory and is provided with matches based on descriptions of the websites submitted by the owners.¹³¹ The most notable such directory is the original Yahoo! search engine, which started out simply as a list of the founders' favourite web sites, but grew quickly into a general directory of websites for public use.¹³² Since 2002, Yahoo! has shifted to a crawler-based approach,¹³³ however the directory still exists.¹³⁴

The popularity approach used by Google, involves a two-stage approach. First, 'PageRank' assesses the importance of a web page based on popularity. It counts 'votes' of a page based on the number of sites linking to it, and the importance of the voting sites. Then the search engine analyzes the content of a web page. It used to rely on meta-tags,¹³⁵ but their susceptibility to manipulation by owners in order to boost their rankings on the Google index led Google to move away from meta-tagging. Now, it assesses the content of a page based on the 'fonts, subdivisions and the precise location of each word.'¹³⁶

These algorithms, particularly with respect to the Google approach, can create biased search results. First, popularity is assumed to equal importance, relevance and quality.¹³⁷ This results in a perpetuation of majority interests to the detriment of less 'popular, wealthy and powerful sites'.¹³⁸ Large, economically powerful companies surge to the top of the rankings, while smaller companies that may not have the resources to invest in professionals with the know-how for search engine optimisation are ranked lower, if at all. In addition, more established websites become more entrenched at the head of the rankings because they have had more time to garner popular links.¹³⁹ This *Googlearchy*¹⁴⁰ has the unfortunate effect of delaying recognition of, or excluding entirely, new, quality sites.¹⁴¹ Linking

¹³⁰ Sullivan, 'How Search Engines Work' at <http://searchenginewatch.com/2168031/print> (last visited 25 June 2008).

¹³¹ *Ibid.*

¹³² 'The History of Yahoo! – How It All Started' at <http://docs.yahoo.com/info/misc/history.html> (last visited 25 June 2008), and see <http://www.searchenginehistory.com>, n 20 above.

¹³³ Sullivan, 'Major Search Engines and Directories' at <http://searchenginewatch.com/2156221/print> (last visited 25 June 2008). Yahoo! used Google technology to crawl the web, but in 2004 began using its own technology.

¹³⁴ <http://dir.yahoo.com/> (last visited 25 June 2008).

¹³⁵ Meta tags were created by web page owners to specify the keywords and concepts under which a page was indexed: see <http://www.howstuffworks.com>, n 16 above.

¹³⁶ 'Technology Overview' at <http://www.google.co.uk/intl/en/corporate/techn.htm> (last visited 27 June 2007).

¹³⁷ n 64 above, at p 181 See also n 113 above.

¹³⁸ n 64 above, at p 31; n 115 above, at p 3.

¹³⁹ *Ibid.*

¹⁴⁰ n 113 above, at p 1.

¹⁴¹ *Ibid.* The authors noted that popularity ranking 'can delay widespread awareness of a high-quality page by a factor of over 60': *ibid* at p 1. Pasquale notes that empirical research supports and dismisses this theory: n 116 above, at p 18.

popularity with quality creates an undesirable cycle of exclusivity. The very process of linking users to content providers sets up winners and losers.¹⁴² Some view this bias as merely a gap between the actual results and optimal results,¹⁴³ but this fails to take account of the entrenching effect of popularity ranking.

Second, there is the problem of being highly ranked when one does not want to be. Individuals identified in websites such as www.dontdatehimgirl.com, a website wherein women can warn others against dating their ex-boyfriends, or www.iknowwhatyoudidlastnight.com showing images of individuals from parties, likely do not wish those websites to emerge high on the rankings when a personal name search is conducted. Google's new Street View, recently exposed a lawyer from the Electronic Frontier Foundation as a 'secret smoker'.¹⁴⁴

The practice of Google is to not intervene on the basis that the information is stored on a third-party site.¹⁴⁵ Thus, when it comes to choosing not to be ranked, Google refuses to allow users to exercise this choice. This forces the user to contact the website owner directly. However, due to the search engine practice of caching,¹⁴⁶ removal of the disputed information by the content provider does not mean that the information is then removed from a search engine's search results, which might include a snippet of the relevant material. Nor is there the opportunity to provide a response on the search results page to what appears about you.¹⁴⁷ Yahoo!, on the other hand, does not advise of any practice regarding the removal of sites in its Terms of Service.¹⁴⁸ This is equally detrimental to users, because it lacks transparency, and lacks a standard against which to judge accountability.

Italy recently enacted the 'right to be forgotten'. This law arose from a complaint made to *Garante Per La Protezione Dei Dati Personali* (the privacy authority) wherein an administrative violation against a company that occurred several years before continued to appear first on the indices of a search engine, and no longer reflected the company's current information.¹⁴⁹ It held that the institution should have done something, but did not elaborate on the nature of this obligation.¹⁵⁰ The resulting 'right to be forgotten' was implemented in Italy's data protection code, binding

¹⁴² Grimmelmann, n 16 above, at p 15.

¹⁴³ n 116 above, at p 17.

¹⁴⁴ Holahan, 'Google Is Watching You' at http://www.businessweek.com/print/technology/content/jun2007/tc20070622_338015 (last visited 26 June 2007).

¹⁴⁵ n 116 above, at p 7.

¹⁴⁶ Caching refers to the temporary storing of websites in order to make the Internet work more efficiently.

¹⁴⁷ n 116 above, at p 17.

¹⁴⁸ 'Yahoo! Terms of Service' at <http://info.yahoo.com/legal/us/yahoo/utos/utos-173.html> (last visited 25 June 2008). <http://www.garanteprivacy.it/garante/doc.jsp?ID=1113806> (last visited 25 June 2008) (translated orally by Andrea Glorioso).

¹⁵⁰ *Ibid.*

institutions, rather than search engines, to remove decisions regarding sanctions from its website after a reasonable amount of time.¹⁵¹ Although the page may still be accessible and searchable from the institution's website, it must be excluded from general-purpose search engines.¹⁵² Italy's Data Protection Authority recommends that public bodies advise search engines of an expiry date for the retrieval of pages with personal information.¹⁵³ At the Spring Conference of European Data Protection Commissioners, Mauro Paissan reported that Google is working jointly with the Data Protection Authority to address these issues where out-of-date material is ranked high on search results advising of, for example, an arrest, but failing to rank the subsequent acquittal.¹⁵⁴

Third, search engines, website owners, and users are grappling with the desire of many site owners to be highly ranked. Companies recognize the importance of being ranked highly and compete fiercely for the top ten slots on the first search result page. The primary means of improving ranking is search engine optimisation (SEO) and the purchasing of top slots. Thus experts in the field of 'search engine design' have emerged, teaching companies how to optimise their rankings by anticipating the algorithms,¹⁵⁵ making SEO an industry unto itself.¹⁵⁶ Rankings are sometimes manipulated for sheer humour or political message. Known as 'Google bombing', the search term 'miserable failure' famously returned President George W. Bush's biography as the top search result until it was changed by Google two years later.¹⁵⁷

The first form of search engine optimisation was through manipulation of a company's meta-tags, the key words the search engines read for indexing.¹⁵⁸ Search engines quickly changed their search algorithms to counteract this manipulation, so SEO now uses more sophisticated forms of manipulation such as creating websites that link to the pages in another group of websites in order to improve their popularity ranking with Google, known as link farming.¹⁵⁹ SEO has also been used for other detrimental purposes, to push unflattering company information further down the rankings, and even to prevent competitors from being highly ranked.¹⁶⁰ While search engines are criticized for protecting their

¹⁵¹ See <http://www.garanteprivacy.it/garante/navig/jsp/index.jsp?solotesto=N> (last visited 25 June 2008).

¹⁵² n 150 above, at p 161.

¹⁵³ Paissan, 'Privacy Protection and Right to Know, Striking a Difficult Balance' at <http://www.garanteprivacy.it/garante/doc.jsp?ID=1408388> (last visited 25 June 2008).

¹⁵⁴ *Ibid.*

¹⁵⁵ n 91 above, at p 3.

¹⁵⁶ For one such company, see: <http://www.searchengineoptimising.com/> (last visited 25 June 2008).

¹⁵⁷ Sullivan, 'Google Kill's Bush's Miserable Failure Search & Other Google Bombs' at <http://searchengineland.com/070125-230048.php> (last visited 25 June 2008).

¹⁵⁸ n 91 above, at p 3.

¹⁵⁹ *Ibid.*

¹⁶⁰ See Grimmelman, n 16 above, at p 39.

algorithms as trade secrets, such transparency would arm SEOs with the information needed to further manipulate the rankings leading to stacked, largely commercial and not necessarily relevant search results.¹⁶¹ Although such practices skew the relevancy of search results, a degree of SEO is necessary to allow companies to design their websites appropriately for the algorithms so that the search results best reflect relevancy.¹⁶² Even without transparency, search results are visible, so companies have been relatively successful at reverse engineering to manipulate their placement in the rankings.¹⁶³ What has resulted is a constant tug-o-war between companies and search providers, affecting the relevance of search results, and hindering the ability to frame a relevance principle that a search engine provider could be accountable for.

In addition, companies commonly enhance their position in the rankings by purchasing the top slots or key words.¹⁶⁴ The visibility of paid sponsors has been the subject of investigation by the US Federal Trade Commission demanding that they be identified separate from normal search results.¹⁶⁵ Google has always separated normal from paid search results, but this does not translate into user knowledge of what this separation means. A study by the Pew Internet & American Life Project found that only 38 per cent of American users are aware of the difference between paid and unpaid results.¹⁶⁶ Without users understanding what a paid sponsor is, the effort to alert consumers requires something more robust. The lack of sophistication in the average user combined with the lack of transparency of the companies that deliver search results creates an aura of authority of search results without any corollary accountability by providers.

There are further issues that shall not be developed here, but which should be noted, such as the effect of filtering mechanisms on ranking. In an 'Empirical Analysis of Google SafeSearch', Benjamin Edelman found that Google's SafeSearch, a voluntary filter for sexually explicit material, also filtered innocent sites such as the US Library of Congress and the National Middle School Association.¹⁶⁷ In addition, one can argue that the business model of search providers affects the search results,¹⁶⁸ because revenue is generated largely through user fees, advertiser fees or website fees.¹⁶⁹ Further, search engines have been sued for selling third

¹⁶¹ n 64 above, at p 174; n 116 above, at p 36.

¹⁶² n 64 above, at p 174.

¹⁶³ Grimmelman, n 16 above, at p 44.

¹⁶⁴ n 64 above, at pp 174-175.

¹⁶⁵ A letter from the FTC regarding this matter is available at <http://www.yahoo-watch.org/ftc1.html> (last visited 25 June 2008).

¹⁶⁶ Fallows, 'Internet searchers are confident, satisfied and trusting – but they are also unaware and naive' at http://www.pewinternet.org/PPF/r/146/report_display.asp (last visited 25 June 2008).

¹⁶⁷ Edelman, 'Empirical Analysis of Google SafeSearch' at <http://cyber.law.harvard.edu/people/edelman/google-safesearch> (last visited 25 June 2008).

¹⁶⁸ n 91 above, at p 4.

¹⁶⁹ *Ibid.*

party trademarks to be used in sponsored links and advertisements. This might be misleading because, for example, a user searching for ‘playboy’ on Netscape or Excite at the turn of the century would have found other, likely topically related, sites that were neither owned nor sponsored by Playboy.¹⁷⁰

4.3.2 *Manual Manipulation of Rankings*

Search engines have intervened to manipulate rankings in two respects.¹⁷¹ First, they have removed links from indices in response to complaints concerning content.¹⁷² Second, they have removed links at their own behest.¹⁷³ There is a lack of consistency between and within these categories regarding when and why a search engine will manipulate rankings, and policy is required to fill this gap. Google, being the market leader, is the target of most news and lawsuits in this respect, and therefore will be discussed extensively in this section.

Complaints and Removal As a general rule, Google will not remove a link from its indices; since a third party owns the site Google advises its users to contact the third party.¹⁷⁴ If the web site owner restricts access to his or her site or removes it entirely from the web, ‘Google would consider on a case-by-case basis requests to remove the link to that site from its indices.’¹⁷⁵

Google has removed sites that it deems offensive, however. Last year, Google received complaints about killbattyn.com, a blog hosted on its blogger.com service, which advocated the killing of gays and lesbians.¹⁷⁶ Initially Google refused to remove the site on First Amendment grounds, stating that ‘[t]here are many things on the Web which groups find upsetting or distasteful...It is up to governments to decide at the end of the day where freedom of speech begins and ends.’¹⁷⁷ Effectively denying any role in disseminating the hate material, it ‘passed the buck’ to government. Google finally removed the site relying on blogger.com’s terms of service, but refused to elaborate further on the specific terms that it deemed were breached.¹⁷⁸ This, however, is an instance where it acted as a host of the content provider rather than as a search engine. With regards to

¹⁷⁰ *Playboy Enterprises, Inc. v Netscape Communications Corp.* 354 F.3d 1020 (9th Cir. 2004). For further discussion of trademarks issues see n 61 above.

¹⁷¹ There is a third category for removals required by law, but it will not be discussed herein.

¹⁷² See ‘An explanation of our search results’ at <http://www.google.com/explanation.html> (last visited 25 June 2008), and D. Goodin, ‘Google removes gaybashing website’ at http://www.theregister.co.uk/2007/03/02/google_removes_site/ (last visited 25 June 2008).

¹⁷³ n 116 above, at p 23.

¹⁷⁴ *Ibid.* at p 7.

¹⁷⁵ ‘Google Terms of Service For Your Personal Use’ at <http://text.usg.edu:8080/tt/wifi.google.com/ggterms.html> (last visited 25 June 2008). See discussion by Pasquale, n 116 above, at p 6.

¹⁷⁶ Goodin, n 172 above.

¹⁷⁷ *Ibid.*

¹⁷⁸ *Ibid.*

search results, Google was prompted to issue a statement regarding anti-Semitic sites topping the rankings when the search term 'Jew' was used, explaining why this might occur and expressing condemnation, but stopping short of removing the site.¹⁷⁹ Pasquale commented that the fact that Google responded raised questions about whether 'certain types or [sic] sites or particularly sensitive information *by their very nature* [are] underserving of the type of publicity high-ranked results provide'.¹⁸⁰ Google used the opportunity to elaborate further on its policy regarding removal of sites stating:

Individual citizens and public interest groups do periodically urge us to remove particular links or otherwise adjust search results. Although Google reserves the right to address such requests individually, Google views the comprehensiveness of our search results as an extremely important priority. Accordingly, we do not remove a page from our search results simply because its content is unpopular or because we receive complaints concerning it. We will, however, remove pages from our results if we believe the page (or its site) violates our Webmaster Guidelines, if we believe we are required to do so by law, or at the request of the webmaster who is responsible for the page.¹⁸¹

This response evidences Google's perspective that intervention would taint the value of the search results, but closer examination reveals a form of double-speak. On the one hand, Google will only remove pages in three circumstances: where it is required by law, at the request of the Webmaster, or due to a violation of their Webmaster Guidelines. With regard to the latter, these Guidelines are only concerned with breaches caused by efforts to manipulate rankings such as 'link schemes' and 'sneaky redirects'.¹⁸² On the other hand, Google reserves the discretionary right to act on individual requests. Google's Code and Yahoo! Terms of Service employ even broader language denying any responsibility but also any accountability. Google reserves the right, but not the obligation 'to pre-screen, flag, filter, refuse, modify or remove Content'.¹⁸³ Similarly Yahoo! reserves the right to pre-screen, refuse or remove any content based on violations of the Terms of Service or if it is deemed 'otherwise objectionable'.¹⁸⁴ This behaviour represents search engines' broader attempt to restrict their

¹⁷⁹ n 172 above, <http://www.google.com/explanation.html>. The explanation is currently ranked 7th on the search results for the term 'Jew' (as at 24 June 2008).

¹⁸⁰ n 116 above, at p 9.

¹⁸¹ n 172 above, <http://www.google.com/explanation.html>.

¹⁸² 'Webmaster Guidelines' at <http://www.google.com/support/webmasters/bin/answer.py?hl=en&answer=35769> (last visited 25 June 2008).

¹⁸³ 'Google Code' at <http://text.usg.edu:8080/tt/wifi.google.com/ggterms.html> (last visited 25 June 2008).

¹⁸⁴ 'Yahoo! Terms of Service' at <http://info.yahoo.com/legal/us/yahoo/utos/utos-173.html> (last visited 25 June 2008).

roles to that of mere directory or conduit, while simultaneously empowering themselves to censor without explanation or accountability.

Removal at Search Engine Behest The power of search engines to punish website owners for behaviour it deems objectionable is evident in cases such as *Search King Inc. v. Google Technology Inc.*¹⁸⁵ Search King introduced PR Ad Network (PRAN) in 2002 to act as a middleman between clients and third party sites highly ranked by Google's algorithm. The purpose was to improve their clients' rankings in Google search results. PRAN would arrange for the advertisement of the clients on these third party sites.¹⁸⁶ This was effectively a link farm. Google ranks websites between 1 and 10, with 10 being the best ranking. In 2002, Search King's ranking decreased from 8 to 4, while PRAN's ranking went from 2 to being eliminated completely from the ranking.¹⁸⁷ Search King soon after filed a lawsuit alleging tortious interference with contractual relations. Google argued that its PageRank is protected speech under the First Amendment. US District Court Judge Miles La Grange reasoned that PageRank relates to matters of public concern, but that it is impossible to prove that the ranking given to a web site is false.¹⁸⁸ She dismissed the action, concluding that the search results are opinions and accordingly protected speech.¹⁸⁹ The opinion Google was expressing was as to 'the significance of particular web sites as they correspond to a search query.'¹⁹⁰ Without access to the algorithms and reasons for manual manipulation decisions, plaintiffs cannot evidence that a drop in rankings is untoward or illegitimate. Categorizing search results as expressions of opinion exacerbates this, because it further insulates search engine providers from accountability for their products. Arguably, search engines can now engage in openly discriminatory manipulations under the protection of it being opinion, although this would be a public relations nightmare.

Subsequently, in *Kinderstart.com LLC et al v. Google Inc.*,¹⁹¹ Google relied on *Search King's* finding that its search engine results were opinion. Since 2000 KinderStart has operated a website dedicated to information related to small children.¹⁹² In 2005, Kinderstart alleged that its site's traffic dropped by 70 per cent, and its AdSense revenue dropped by 80 per cent.¹⁹³ KinderStart 'realized' that the website was no longer listed with the same visibility in Google's search results (Google apparently advised that this would occur).¹⁹⁴ In addition, KinderStart's website received a

¹⁸⁵ n 81 above.

¹⁸⁶ *Ibid.* at p 2.

¹⁸⁷ *Ibid.*

¹⁸⁸ *Ibid.* at p 4.

¹⁸⁹ *Ibid.*

¹⁹⁰ *Ibid.*

¹⁹¹ n 103 above.

¹⁹² *Ibid.* at p 3.

¹⁹³ *Ibid.* at p 4.

¹⁹⁴ *Ibid.*

PageRank of '0'.¹⁹⁵ KinderStart claimed various causes of action such as violation of free speech, attempted monopolization and unfair competition.¹⁹⁶ US District Court Judge Jeremy Fogel granted Google's motion to dismiss the case in a strongly worded judgment highly critical of KinderStart.com. It is not the purpose herein to assess the reasonableness of the Judge's dismissal of the case,¹⁹⁷ but rather to use the case to highlight Google's view of its function as articulated in its Brief to the Court.

Google characterized its search engine as (a) an expression of opinion (b) by a private business (c) about the importance of websites.¹⁹⁸ The value to websites, it asserted, was 'promotion of [the website's] message via Google.'¹⁹⁹ Drawing an analogy to other lists or indices that serve a promotional purpose, it argued:

Over the years, authors who felt their books belonged on bestseller lists, airlines who thought their flights should be featured more prominently in airline flight listings, bond issuers dissatisfied with their ratings, and even website owners angry about Google's ranking of their sites, have turned to litigation seeking to override such judgments. Each time, the courts have rejected such claims, recognizing that private businesses have a right to express these opinions freely.²⁰⁰

It later emphasized: 'Google's index and its search results are a private forum for Google's speech.'²⁰¹ With reference to the zero-rating received by KinderStart, Google characterized this simply as a difference of opinion regarding the site's quality.²⁰²

Search King and *KinderStart* highlight the mismatch between the critical importance of search engine results to companies' viabilities, and the immunity of search engines from any form of liability or responsibility for the indices. *KinderStart* highlights the impossibility of succeeding at a claim against a search engine provider for changes in a company's rankings on a search index, because how or why the change occurred cannot be quantitatively proven. Judge Fogel stated that '[t]he Court concludes

¹⁹⁵ Ibid. at pp 4-5.

¹⁹⁶ Ibid. at p 2.

¹⁹⁷ Judge Fogel reprimanded KinderStart's lawyer for making unsupported claims regarding the companies' treatment of other companies, relying on double hearsay or hearsay speculation, and concluded that the behaviour was sanctionable against the lawyer: see Williams, 'Judge boots out Google delisting suit' at http://www.theregister.co.uk/2007/03/21/kinderstart_thrown_out/print.html (last visited 25 June 2008), and Out-law.com, 'Google search rank claim thrown out for second time' <http://www.out-law.com/default.aspx?page=7901> (last visited 25 June 2008).

¹⁹⁸ Google's Notion of Motion and Motion to Dismiss the First Amended Complaint, and Memorandum of Points, 2006 WL 1232481 at p 8.

¹⁹⁹ Ibid. at p 9.

²⁰⁰ Ibid. at p 6.

²⁰¹ Ibid. at p 10.

²⁰² Google's Reply in Support of Its Motion to Dismiss Plaintiff's Second Amended Complaint, 2006 WL 3619654 at p 7.

that the allegation that Google sell priority placement in its results should not have been made based upon the limited information identified by [the plaintiff's lawyer]'.²⁰³ The search algorithms are protected as trade secrets, and the reasons for manual manipulation of rankings, particularly in any given case, are not publicly or privately revealed. This results in a quagmire wherein businesses rely on the search results, yet have no access to understanding changes in such results, even when the changes have a marked impact on the company's sustainability.

5 Public Interest Framework

Search engines are indispensable tools for sorting through the overwhelming amount of information available on the Internet, and organizing it in a coherent, accessible manner. Search providers might bitterly reflect that public interest-type obligations are punishing them for being good at what they do. However, the idea that search engines are merely 'businesses' or 'promotional services for website owners' is untenable. If information is becoming a critical commodity in modern society, then such bodies that manage access to information, that are tools for public discourse and democracy, should be accountable to the public. When the structure for access shapes meaning for its users, and influences public opinion, this need for a public-interest obligation is magnified.

What would be the goal for a framework of public interest obligation? In short, the goal is to create accountability. Such a proposition seems straightforward, but like disputes in traditional media, it becomes murky when attempting to define the qualities of this accountability and how it should be implemented. The goal, guided by traditional media is 'that these are carried out well and according to principles of efficiency, justice, fairness, and respect'.²⁰⁴ Three values should be prioritised: relevant and unbiased search results, a certain degree of transparency, and respect for user dignity. In addition, for any of these values to have meaning, there must be a check on their power and practices, and therefore an independent complaints mechanism should be implemented.

5.1 *Relevant and Unbiased Search Results*

Users want relevant and reliable search results. Pasquale describes having two mutually reinforcing goals: authoritative search results, and responsibility as information gateways.²⁰⁵ Treating relevancy as a matter of public interest raises complex issues. How can we determine relevancy when there is no transparency on the part of the businesses involved? How can

²⁰³ n 197 above, <http://www.out-law.com/default.aspx?page=7901>.

²⁰⁴ McQuail2, n 32 above, at p 47.

²⁰⁵ n 116 above, at p 14.

we even expect reliable and authoritative search results if the indices are, as argued by the providers themselves, merely directories or indices of third party information? Should reliability be a factor when dealing with the dynamic and participatory nature of the Internet?

Several ideas have been proposed to achieve more relevant, unbiased search results. Partially randomized ranking has been suggested to counter-balance the problem of entrenchment. The suggestion is to 'promote a small fraction of unexplored pages up in the result list'²⁰⁶ to randomly chosen rank positions. Publicly funded search engines have also been suggested.²⁰⁷ France and Germany have cooperated to fund the Quaero Project, led by Thomson.net and involving several other companies, with the goal of, *inter alia*, creating a search competitor to the American companies.²⁰⁸ Further, the development of Wikia should be followed closely. Created by the founder of Wikipedia, Wikia is a for-profit search engine with one distinguishing feature: it is open source,²⁰⁹ meaning that the search algorithms will be accessible to anyone to review, modify or use on other websites.²¹⁰

Some posit that relevancy will take care of itself. The increase in personalized ranking algorithms, it is argued, will render bias irrelevant, because results are tailored to the interests of the user.²¹¹ Search engines are beginning to move in this direction with Yahoo!'s 'Yahoo! Mindset' and Google's option to order results based on past searches.²¹² This shift of reliance to personalized searches as the panacea to the current problems should be cautioned against. It underplays the values at stake, is a privacy concern, and encourages balkanization of knowledge and experience.

Although users are becoming savvy and might 'mix and match' searching on various search sites,²¹³ the availability of alternative search engines does not adequately address user dependency on search engines, nor satisfies the public's interest in relevant search results. 44 per cent of American searchers regularly use only one search engine, while 48 per cent use only two or three.²¹⁴ Further, concentration of the market, and domination by

²⁰⁶ n 113 above, at p 2.

²⁰⁷ See, for example, n 115 above, at p 4.

²⁰⁸ Baker, 'Quaero European Search Engine Goals and Plans' at <http://www.searchenginejournal.com/quaero-european-search-engine-goals-and-plans/2766/#more-2766> (last visited 25 June 2008). See also <http://en.wikipedia.org/wiki/Quaero> (last visited 25 June 2008). It has been reported that Germany is no longer a part of the Quaero Project, and started project Theseus to develop a text-based search engine, whereas Quaero is focused on multimedia search: *ibid.*

²⁰⁹ MacKinnon, 'Jimbo Wales: Google's China mistake' at <http://rconversation.blogs.com/rconversation/2007/08/jimmy-wales-goo.html> (last visited 25 June 2008).

²¹⁰ Auchard, 'Wikia Plans Community-Based Search' at <http://www.pcworld.com/article/id,135275-c,companynews/article.html> (last visited 20 August 2007).

²¹¹ n 115 above, at p 6.

²¹² *Ibid.* at pp 6-7.

²¹³ See discussion of the quality of search engines other than Google in Beam, 'FrankenGoogle' at <http://www.slate.com/toolbar.aspx?action=print&id=2172222> (last visited 25 June 2008).

²¹⁴ n 166 above.

Google followed not-so-closely by Yahoo!, undermines the argument that there is sufficient diversity in the market to address issues of relevancy. Even if users had a diversity of competitive search engines, this does not negate the need for a public interest obligation, nor is an answer to it.

Any public interest in relevancy must be realistic. Businesses must expect fluctuations in rankings, and not every website can be highly ranked, nor is every undesirable link going to be caught and removed. Relevance is hard to measure, because it is as much dependent on the users search terms as it is on the algorithm of the provider, and results can be manipulated by google bombing, search engine optimisation and the like. For relevance to have meaning, the key is consistency: consistency in algorithms, consistency in decision-making regarding any manipulations of search results, and consistency in the values that drive such manual manipulation.

5.2 *Transparency*

For relevance to be of any value, there must be openness about how the algorithms work and when and why the providers make certain decisions. This author hesitates to go so far as call it transparency, because full transparency is not promoted here.

Several authors have argued for transparency of how search results are reached, both automated and manual.²¹⁵ Nissenbaum and Introna comment:

As a first step we would demand full and truthful disclosure of the underlying rules (or algorithms) governing indexing, searching, and prioritizing, stated in a way that is meaningful to the majority of web users.²¹⁶

Nissenbaum and Introna see transparency as the optimal solution, because it enables users to make more informed decisions about which search engines to use and for what.²¹⁷ Without transparency, how is a user to assess relevance?²¹⁸ There are difficulties with full transparency, however, including impacting competition among providers, stymieing incentives to innovate, arming users with knowledge to manipulate the rankings, and risking user privacy if transparency requires disclosure of user queries.²¹⁹

This author views the negative consequences of disclosure of algorithms on the market as sufficiently concerning that blanket transparency should not be the solution. Further investigation is required to determine the potential market fallout of full algorithmic transparency. However, disclosure should be required regarding manual manipulations. Such

²¹⁵ For example, see n 115 above, at p 74, and Grimmelman, n 16 above.

²¹⁶ n 64 above, at p 181.

²¹⁷ *Ibid.*

²¹⁸ Elkin-Koren, n 93 above, at p 191.

²¹⁹ See discussion by Grimmelman, n 16 above, at p 44-50.

openness would alleviate user concerns to understand how a sudden shift in ranking occurs, and achieve accountability from search providers now aware that their decisions are under scrutiny.

5.3 *Respect for User Dignity*

User dignity must be addressed separately, because the participatory nature of the Internet, and the speed with which information is spread, increases the power and incidence of unsavoury and defamatory information circulating. This concern is compounded when potential employers, businesses partners, or curiosity seekers perform a name search and such information appears high on the rankings.²²⁰ Several incidents have arisen lately. Two Yale law students have sued one of the administrators of www.autoadmit.com, a bulletin board for discussion of American colleges, and several 'John Doe' authors, for defamatory comments posted by anonymous users on the law forum.²²¹ One Plaintiff insists that her inability to find a job after graduation was in part caused by prospective employers 'googling' her name and finding discussion threads concerning her from AutoAdmit high on Google's search results. Whether this allegation has any merit is yet unclear, but it serves to highlight the role of search engine results in finding such information. It is increasingly standard procedure to vet a prospective employee by searching social networking sites and search engines. A recent survey suggested that 35 per cent of American employers conducted a name search of potential employees.²²² Should individuals have control over access to such information? Pasquale comments that if individuals may control the transfer of personal information in other arenas such as banking and medicine, then perhaps similar controls should be implemented regarding such information sought via search engines.²²³

Search engine providers advise users to contact the content providers directly, and this approach is appropriate but does not provide a complete solution. Search engines' caching function allows the information to survive longer than the content provider might intend. The Internet Archive WayBackMachine service allows user to search 'archived versions of websites' advertising '[i]magine surfing circa 1999 and looking at all the Y2K hype'.²²⁴ Both Yahoo! and Google allow users to click on a 'cached' function taking them to a cached version of a website if it is available.²²⁵ Once on the

²²⁰ See, for example, ReputationDefender, a business set up to search for and remove undesirable personal information from the Internet: <http://www.reputationdefender.com> (last visited 25 June 2008).

²²¹ *Doe I and Doe II v Anthony Ciolli et al*, http://www.concurringopinions.com/archives/2007/06/the_autoadmit_1.html (last visited 25 June 2008).

²²² Du, 'Job candidates get tripped up by Facebook' at <http://www.msnbc.msn.com/id/20202935/print/1/displaymode/1098> (last visited 25 June 2008).

²²³ n 116 above, at p 16-17.

²²⁴ Internet Archive 'Frequently Asked Questions' at http://www.archive.org/about/faqs.php#The_Wayback_Machine (last visited 25 June 2008).

²²⁵ This author searched 'Y2K' on both search engines showing this result.

cached site, Yahoo! also provides direct links to the WayBackMachine for further searching. While website owners may request that their content not be cached,²²⁶ they cannot be a full-service outfit to respond to complaints regarding content. Collaboration of search engines is necessary to successfully remove the content.

In circumstances where a person wants content to be removed, or to no longer appear on the rankings, Pasquale suggests that in meritorious circumstances, where a complainant alleges a false or misleading search result, an asterisk should be placed next to the result linking to the complainant's website where he or she can tell his or her side of the story.²²⁷ He defends this suggestion not as imposed free speech but rather as 'a way of restructuring a forum with deep roots in our First Amendment tradition.'²²⁸ In Italy, the 'right to be forgotten' has been implemented, and although it is directed at content providers rather than search engines, Google has been collaborating to make this right effective.²²⁹ As in the cases of *Search King* and *KinderStart*, many website owners want their websites to be highly ranked. It has been suggested that when a website drops suddenly in the rankings, the solution is for these owners to be informed of what caused the sudden change.²³⁰

Should search engines have a takedown obligation akin to ISPs under the Electronic Commerce Directive? The latter has been criticized for over-censoring content because a mere complaint without proof of the contents defamatory nature justifies taking down of the content. If applied to search engines it would fundamentally undermine freedom of expression to an extent far more invasive than an obligation on ISPs, ultimately failing to serve the public. The above suggestions provide possible solutions to some of the rankings controversies, but they address small portions of a much bigger problem. For a user's dignity to be respected, a one-size-fits-all approach will not work. Owners, users, and concerned citizens need a place where their complaints can be heard. For relevance, transparency, and user dignity to have any real value, there must be a body to deal with complaints.

5.3 *Independent Complaints Mechanism*

Cases such as *KinderStart* and *Search King* illustrate the need for a complaints mechanism for website owners concerned about sudden shifts in their position on the rankings, or high rankings of undesirable information. The responses of search engines thus far have been unsatisfactory and makes clear that any such complaints mechanism must be independent. For such a mechanism to be effective there must be a code of practice to ground the

²²⁶ Google Help Center at <http://www.google.com/help/features.html#cached> (last visited 25 June 2008).

²²⁷ n 116 above, at p 28.

²²⁸ *Ibid.* at p 30.

²²⁹ n 149 above.

²³⁰ n 116 above, at p 32.

nature of search engine obligations, and a standard against which a decision maker can determine if the code has been breached.

In Germany, the Association of Voluntary Self-Regulating Multimedia Service Providers ('FSM')²³¹ implemented a Subcode of Conduct for Search Engine Providers for the purpose of consumer protection and protection of children. It states that although information intermediaries 'do not make their own content available... [t]hey are, however, aware of their particular Role in making information available on the Internet.'²³² Thus, this Code is an expression of commitment to abide by three main rules: disclosure of how the search engine functions; identifiable separation of normal search results from paid placements; and the enabling of technical precautionary measures for the protection of children from harmful content.²³³ Members of FSM are only bound by the Code if they are signatories to it, and it is as yet unclear who has signed on to it, although members of FSM include Google, Lycos, and MSN/Microsoft.²³⁴ Although this Code addresses issues critical to any code governing search engine conduct, something more comprehensive is needed to address user complaints regarding undesirably high or low rankings.

What would be the characteristics of such a complaints mechanism? A complete discussion of such a mechanism would require a separate article. Therefore, this article will offer some thoughts on comparative regulatory structures and the values an optimal regime might possess.

Guidance can be found from the media where various regulatory structures are used. Any structure will be exposed to the many criticisms made of media regulators; that they are ineffective, merely 'public relations operations',²³⁵ and might only serve to stave off legally enforceable safeguards.²³⁶ However, cases such as *KinderStart* and *Search King* show the inability of the courts to currently manage such claims, and legal regulation might be undesirable, and perhaps would upset the market balance and chill innovation.

On a national level, there are two options. First, the preferable route is to adopt a self-regulatory regime akin to the Press Complaints Commission (PCC)²³⁷ in the United Kingdom. Although newspaper proprietors fund the PCC,²³⁷ it is an independent body established for the purpose of resolving complaints from the public concerning content in newspapers or magazines.²³⁸ The PCC's main role is to enforce the Code of Practice,²³⁹ drafted

²³¹ Feiwillige Selbstkontrolle Multimedia-Diensteanbieter at http://www.fsm.de/e/SubCoC_Search_Engines (last visited 22 July 2007).

²³² *Ibid.*

²³³ *Ibid.*

²³⁴ *Ibid.*

²³⁵ Robertson & Nicol, *Media Law* (London: Penguin Books, 4th ed, 2002) at p 675.

²³⁶ *Ibid.*

²³⁷ *Ibid.*

²³⁸ PCC, 'What is the PCC?' at <http://www.pcc.org.uk/about/whatispcc.html> (last visited 25 June 2008).

²³⁹ PCC, 'Editors' Code of Practice' at http://www.pcc.org.uk/assets/111/Code_Aug_2007.pdf (last visited 25 June 2008).

and updated by the Editor's Commission whose members are from the industry.²⁴⁰ The Code sets the ethical standards for journalists regarding, *inter alia*, accuracy of publications, protection of children, discrimination, and the use of hidden cameras and listening devices.²⁴¹ The PCC is not funded by government, nor is it under statutory control,²⁴² and the service is free to the public.²⁴³ The Commission members are laypersons, and currently include only one person from the industry.²⁴⁴ The Commission cannot sanction a newspaper or editor financially, but may only require that the Commission's decision be published by the newspaper or magazine found to be in breach of its Code of Practice. Such a regime is advantageous because it is voluntary, created by members of the industry, independent from government, and relies on the commitment of the industry to abide by decisions of a tribunal rather than through penalties and fines. Of course, it is this very lack of legally enforceable power that exposes the PCC to criticism,²⁴⁵ and raises doubts as to its compliance with human rights laws.²⁴⁶ The roots of press freedom in freedom of expression resonate with search engines, and a light regulatory approach may be optimal to create the consistency and openness sought from search engines.

Second, a quasi-governmental regulatory regime might be adopted. The UK's Office of Communications (Ofcom) operates as such a regulator for the broadcasting industry.²⁴⁷ The British Government set up Ofcom under the Communications Act, 2003.²⁴⁸ Its duties are far more extensive than the PCC, ranging from promotion of competition, to diversity of broadcasting content, and to protection of consumer privacy.²⁴⁹ The Communications Act and the Broadcasting Act 1996 required that a code be drafted to set the standards for radio and television,²⁵⁰ which 'Broadcasting Code' came into

²⁴⁰ *Ibid.*

²⁴¹ *Ibid.*

²⁴² PCC, 'Governance and Accountability' at <http://www.pcc.org.uk/about/governance.html> (last visited 25 June 2008).

²⁴³ PCC, 'Key Benefits of the System of Self Regulation' at <http://www.pcc.org.uk/about/benefits.html> (last visited 25 June 2008).

²⁴⁴ PCC, 'Frequently Asked Questions' at http://www.pcc.org.uk/faqs/index.html#faq1_8 (last visited 25 June 2008).

²⁴⁵ Further examination of such a regime requires consideration of the weaknesses of self-regulation, as discussed by scholars such as Julia Black in, for example, 'Decentering Regulation: Understanding the Role of Regulation and Self-Regulation in a "Post-Regulatory" World' (2001) 54 *Current Legal Problems* 103, in which she describes self-regulation as a 'normatively loaded term': at p 3.

²⁴⁶ In *Peck v United Kingdom* (2003) 36 EHRR 41, the European Court of Human Rights held that since the media commissions lacked the legal power to award damages, they could not provide an effective remedy to Peck under article 13 of the European Convention on Human Rights.

²⁴⁷ See <http://www.ofcom.org.uk/> (last visited 25 June 2008). For further discussion of Ofcom see n 37 above.

²⁴⁸ 2003 Chapter 21.

²⁴⁹ Ofcom, 'Statutory Duties and Regulatory Principles' at <http://www.ofcom.org.uk/about/sdrp/> (last visited 25 June 2008).

²⁵⁰ Ofcom, 'Legislative Background to the Code' at http://www.ofcom.org.uk/tv/ifi/codes/bcode/tvcode_b/ (last visited 25 June 2008).

effect on July 25, 2005.²⁵¹ Unlike the PCC, the Code is enforced through a condition in every licence that it be observed,²⁵² and failure to adhere to the Code may result in fines or forfeiture of the licence.²⁵³

The complexity and concerns regarding an OFCOM structure are far more than indicated herein.²⁵⁴ Such complexities are exacerbated if one attempts to apply it to the Internet, because of the political ramifications of attempting public regulation of a transnational communications network. However, at this stage a fundamental question is whether a complaints mechanism is better run by the industry, with government involvement, or by international collaboration and so on. OFCOM is merely one example of a quasi-governmental approach. It is advantageous in binding the industry to the Code, and the threat of licence removal is significant to deter certain behaviour. Whether it is effective, or used for that matter, is uncertain as OFCOM is still in its infancy.

The principles for government intervention in broadcasting do not resonate as acutely with search engines as the self-regulatory regime of the PCC. The concentration of the market faced by both industries is parallel, but broadcasting invests in further public interest regulations such as diversity of programming and public service broadcasting, and it is not clear that search engines require such a re-structuring of the forum. Randomized ranking may arguable be a form of diversity imposition, but this author hesitates to conclude that government involvement will necessarily protect the public interest without further investigation of the potential stifling of market innovation.

One significant hurdle to a complaints mechanism is the transnational nature of Internet communications and of the search engine providers themselves. How can such a system be implemented when varying laws and values bind each nation? This terrain, however, is not exclusive to search engine providers, nor the Internet. Efforts to tackle issues concerning, *inter alia*, domain names, hate speech, defamation, and electronic commerce have been addressed at a national and international level since the Internet's inception. Recognition of search engines as affected with a public interest and requiring regulation is simply the first step, and another important issue to be added to burgeoning stack of quandaries we are encountering.

6 Conclusion

Access to information has become one of the most important features of modern society. The Internet has fundamentally changed how we

²⁵¹ *Ibid.*

²⁵² *Ibid.*

²⁵³ Carey & Sanders, *Media Law* (London: Sweet & Maxwell, 3rd ed, 2004) at p 241.

²⁵⁴ For a discussion of the evolution of media policy and the changes brought about by the *Communications Act*, n 248 above, see Vick, 'Regulatory convergence?' (2006) (26)1 *Legal Studies* 26.

communicate and our relationship to information. It is a discursive platform, where the costs of contributing are low, and the opportunity to express opinions from the insipid to the profound are available at our fingertips. It is, indeed, a democratizing force, and its' increasing penetration rate as a key medium for users propels its force forward. It has moved beyond a mere information portal to one that shapes meaning and self-development.

Yet with this growth has come information clutter. Over one billion people use the Internet, creating millions upon millions of pages for consumption. Search engines have emerged as managers of information, organizing and categorizing content in a coherent, accessible manner, shaping public opinion and the user's information experience. Users have become dependent on search engines, viewing them as authoritative and reliable. Search engines have become the tools through which the democratic potential of the Internet can be advanced or hindered.

The way that information is organized and ranked on search engine results, the story of the searched that is created, is capable of harm. Sudden shifts in rankings can cripple a company, while individuals might find that scurrilous commentary or defamatory statements are the foremost links concerning them on search indices. Where can such companies or individuals turn? With regards to the latter, individuals might contact the website owner or administrator, but removal from the website does not ensure removal from the Internet or search engine results. Otherwise, they have no redress. Search engine practice is, generally, not to intervene, nor to provide information concerning the cause of shifts in rankings. Yet, search engines, such as Google, do intervene to manipulate rankings, but the public is not privy to information regarding when it occurs, for which sites, or why.

There is a chasm between the role of search engines to democracy, and the lack of any framework of accountability for its practices. A public interest framework is needed to bridge this gap and protect consumers in order to further their ability to engage in public discourse and access information.

Three values should be present in such a framework: the value of relevant and unbiased search results; the value of a degree of transparency concerning algorithms and reasons for manual manipulation, and respect for the dignity of the users recognizing that how information is presented on indices can cause harm.

For any of these values to have meaning, for these values to be considered standards in the industry, there must be an independent complaints mechanism for individuals and companies to turn to for resolution of conflicts. Search engine providers might form such a mechanism themselves, but they must first recognize their role as more than mere directories, or expressions of opinion, and embrace their position as a critical democratic force for the Internet.