

# Information Science and Knowledge Management

---

*Editor-in-Chief*

J. Mackenzie Owen

*Editorial Board*

M. Bates

P. Bruza

R. Capurro

E. Davenport

R. Day

M. Hedstrom

A.M. Paci

C. Tenopir

M. Thelwall

Amanda Spink • Michael Zimmer  
Editors

# Web Search

Multidisciplinary Perspectives

 Springer

*Editors*

Prof. Dr. Amanda Spink  
Queensland University of Technology  
Faculty of Information Technology  
Gardens Point Campus  
2 George St.  
Brisbane, QLD 4001  
Australia  
ah.spink@qut.edu.au

Michael Zimmer  
Information Society Project  
Yale Law School  
127 Wall Street  
New Haven, CT 06520  
USA  
michael.zimmer@yale.edu

ISBN: 978-3-540-75828-0

e-ISBN: 978-3-540-75829-7

Information Science and Knowledge Management ISSN: 1568-1300

Library of Congress Control Number: 2007938412

ACM Codes: H.4, J.5, K.4

© 2008 Springer-Verlag Berlin Heidelberg

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

*Cover Design:* KuenkelLopka, Heidelberg, Germany

Printed on acid-free paper

5 4 3 2 1 0

springer.com

# Preface

The inspiration for this book emerged from the editors' participation in a panel discussion on Web search engines at an annual meeting of the Association of Internet Researchers. This multidisciplinary panel of Web search researchers revealed the diversity of scholars interested in Web searching, coupled with a broad range of questions, attitudes, and approaches. It became clear that more "cross-fertilization" was necessary between the disciplines to ensure Web search engines (the entities) and Web searching (the user behavior) received the thorough scholarly attention they deserved. This book is a result of that realization, and an important first step in achieving new levels of awareness and collaboration across disciplines.

The book represents a core theme within the intellectual pursuits of the editors. The first editor (Spink) is an information scientist who has worked with, taught and has researched the informational dimensions of Web searching since 1997. The second editor (Zimmer) is a scholar of culture and communication who focuses on the political and ethical dimensions of new media and information technologies, and whose dissertation research focused on the value-related consequences of the quest for the "perfect" search engine.

This book is intended as a resource for researchers, educators, students, and practitioners. Researchers in the fields of social sciences, communication studies, cultural studies, information science, and related disciplines will all find the chapters presented here as a valuable source of new ideas on Web search. This book is also an appropriate text for advanced undergraduate, graduate, and doctoral level courses in areas of Web search. In addition, anyone who is interested in understanding Web search behavior and Web search engines will surely find this book a valuable read. Each section contains one or more chapters relating to the broader area of the section. Each chapter has a unique perspective and reference list. The chapters are cross-referenced where appropriate to illustrate how the different topics mesh together to form a broader expanse of Web search.

We greatly thank the chapter authors for their ground breaking and stimulating contributions. Many chapters represent the work of collaborations between researchers. We also thank those who edited sections of the book.

Amanda Spink thanks Michael Zimmer for his hard work and academic excellence during this project.

Michael Zimmer thanks Amanda Spink for her leadership, Helen Nissenbaum for her encouragement, and his wife, Rebecca, for her patience and support over the course of this project.

Amanda Spink  
Professor of Information Technology  
Queensland University of Technology

Michael Zimmer  
Fellow, Information Society Project  
Yale Law School

# Contents

## Part I Introduction

<b>1 Introduction</b> . . . . .	3
A. Spink and M. Zimmer	

## Part II Social, Cultural, and Philosophical Perspectives

<b>2 Through the Google Goggles: Sociopolitical Bias in Search Engine Design</b> . . . . .	11
A. Diaz	

<b>3 Reconsidering the Rhizome: A Textual Analysis of Web Search Engines as Gatekeepers of the Internet</b> . . . . .	35
A. Hess	

<b>4 Exploring Gendered Notions: Gender, Job Hunting and Web Searches</b> . . . . .	51
R.M. Martey	

<b>5 Searching Ethics: The Role of Search Engines in the Construction and Distribution of Knowledge</b> . . . . .	67
L.M. Hinman	

<b>6 The Gaze of the Perfect Search Engine: Google as an Infrastructure of Dataveillance</b> . . . . .	77
M. Zimmer	

## Part III Political, Legal, and Economic Perspectives

<b>7 Search Engine Liability for Copyright Infringement</b> . . . . .	103
B. Fitzgerald, D. O'Brien, and A. Fitzgerald	

**8 Search Engine Bias and the Demise of Search Engine Utopianism** ..... 121  
 E. Goldman

**9 The Democratizing Effects of Search Engine Use: On Chance Exposures and Organizational Hubs** ..... 135  
 A. Lev-On

**10 ‘Googling’ Terrorists: Are Northern Irish Terrorists Visible on Internet Search Engines?** ..... 151  
 P. Reilly

**11 The History of the Internet Search Engine: Navigational Media and the Traffic Commodity** ..... 177  
 E. Van Couvering

**Part IV Information Behavior Perspectives**

**12 Toward a Web Search Information Behavior Model** ..... 209  
 S.A. Knight and A. Spink

**13 Web Searching for Health: Theoretical Foundations and Connections to Health Related Outcomes** ..... 235  
 M.J. Dutta and G.D. Bodie

**14 Search Engines and Expertise about Global Issues: Well-defined Landscape or Undomesticated Wilderness?** ..... 255  
 J. Fry, S. Virkar, and R. Schroeder

**15 Conceptual Models for Search Engines** ..... 277  
 D.G. Hendry and E.N. Efthimiadis

**16 Web Searching: A Quality Measurement Perspective** ..... 309  
 D. Lewandowski and N. Höchstötter

**Part V Conclusion**

**17 Conclusions and Further Research** ..... 343  
 A. Spink and M. Zimmer

**Index** ..... 349

# Contributors

Bodie, Graham D.

Department of Communication, Purdue University,  
Beering Hall of Liberal Arts and Education, Room 2114  
100 North University Street, Lafayette, IN 47901, USA

Diaz, Alejandro

Department of Communication, Stanford University, Building 120,  
Room 110, 450 Serra Mall, Stanford, CA 94305, USA

Dutta, Mohan J., Ph.D.

Department of Communication, Purdue University,  
Beering Hall of Liberal Arts and Education, Room 2114  
100 North University Street, Lafayette, IN 47901, USA

Efthimiadis, Efthimis N., Ph.D.

The Information School, University of Washington, Box 352840,  
Seattle, WA 98195, USA

Fitzgerald, Anne, J.D.

Faculty of Law, Queensland University of Technology Level 1,  
Room 110, 126 Margaret Street, Brisbane, Qld 4000, Australia

Fitzgerald, Brian, Ph.D.

Faculty of Law, Queensland University of Technology, Level 1,  
Room 110, 126 Margaret Street, Brisbane, Qld 4000, Australia

Fry, Jenny, Ph.D.

Department of Information Science, Loughborough University,  
Leicester-shire, LE11 3TU, UK

Goldman, Eric, J.D.

High Tech Law Institute, Santa Clara University School of Law,  
500 El Camino Real, Santa Clara, CA 95053, USA

Hendry, David G., Ph.D.

The Information School, University of Washington, Box 352840, Seattle,  
WA 98195, USA



Hess, Aaron

Hugh Downs School of Human Communication, Arizona State University,  
Stauffer Hall Building A, Room 412, PO Box 871205, Tempe, AZ 85287, USA

Hinman, Lawrence M., Ph.D.

Department of Philosophy, University of San Diego, 5998 Alcalá Park,  
San Diego, CA 92110-2492, USA

Höchstötter, Nadine, Ph.D.

Institute for Decision Theory and Management Science, Universitaet Karlsruhe (TH),  
Kaiserstrasse 12, D - 76128 Karlsruhe, Germany

Knight, Shirlee Ann

Edith Cowan University, Joondalup Campus, 100 Joondalup Drive,  
Joon-dalup, 6027, Australia

Lewandowski, Dirk, Ph.D.

Department Information, Hamburg University of Applied Sciences  
Berliner Tor 5, D - 20099 Hamburg, Germany

Lev-On, Azi, Ph.D.

Philosophy, Politics, and Economics Program, University of Pennsylvania  
311 Logan Hall, 249 S 36th Street, Philadelphia, PA 19104, USA

Martey, Rosa Mikeal, Ph.D.

Department of Journalism & Technical Communication, Colorado State  
University, Clark Building, Fort Collins, CO 80523, USA

O'Brien, Damien

Faculty of Law, Queensland University of Technology, Level 1,  
Room 110, 126 Margaret Street, Brisbane Qld 4000, Australia

Reilly, Paul

Department of Politics, University of Glasgow, Adam Smith Building,  
Bute Gardens, Glasgow

Schroeder, Ralph, Ph.D.

Oxford Internet Institute, University of Oxford, One St. Giles, Oxford,  
OX1 3JS, UK

Spink, Amanda, Ph.D.

Faculty of Information Technology, Queensland University of Technology,  
Gardens Point Campus, GPO Box 2434, Brisbane, Queensland 4001, Australia

Van Couvering, Elizabeth

Department of Media & Communications, London School of Economics  
& Political Science, Houghton Street, London WC2A 2AE, UK

Virkar, Shefali

Oxford Internet Institute, University of Oxford, One St. Giles, Oxford,  
OX1 3JS, UK

Zimmer, Michael, Ph.D.

Information Society Project, Yale Law School, 127 Wall Street,  
New Haven, CT 06520, USA

**Part I**  
**Introduction**

# 1 Introduction

A. Spink and M. Zimmer

## 1.1 Book Synopsis

Web search engines have emerged as one of the dominant technologies of modern, digital life, providing doorways to the universe of information available online. According to the Pew Internet & American Life Project, 84% of American adult Internet users have used a search engine to seek information online (Fallows 2005: 1). On any given day, more than 60 million American adults send over 200 million information requests to Web search engines, making Web searches second most popular online activity (behind using e-mail) (Rainie 2005).

More than just an indispensable tool for finding and accessing information online, Web searching has also become a defining component of the human condition. Web searching can be conceptualized as a complex behavior embedded within an individual's everyday social, cultural, political, and information-seeking activities. Following this broad impact of Web searching on daily life, the scholarly study of Web searching spans a multidisciplinary collection of researchers from the social sciences, media and cultural studies, law, information science and other related disciplines. *Web Search: Multidisciplinary Perspectives* brings together chapters that represent this range of multidisciplinary theories, models, and ideas about Web searching, drawing out and examining the various roles and impacts of Web searching on the social, cultural, political, legal, and informational spheres of our lives, such as the impact on individuals, social groups, modern and postmodern ways of knowing, and public and private life. By critically examining the issues, theories, and formations arising from, and surrounding, Web searching, *Web Search: Multidisciplinary Perspectives* represents an important contribution to the emerging multidisciplinary body of research on Web search engines.

Not surprisingly, some of the earliest research publications on Web search engines were technical in nature. Numerous computer scientists have contributed not only valuable research on improving and enhancing the underlying Web search engine technology (Brin and Page 1998; Heydon and Najork 1999; Page et al. 1998), but also technical analyses of the extent of coverage achieved by search engine products and how it relates to information access (Kleinberg and Lawrence 2001; Lawrence and Giles 1998, 2000).

Social studies of Web search engines quickly emerged, typically by information scientists attempting to isolate the habits and characteristics of search engine users through the analysis of transaction log data (Jansen and Pooch 2001). These include Hoelscher's (1998) analysis of 16 million queries from the German search engine Fireball; Jansen et al. (2000) study of a sample day's worth of search activity from the Excite search engine; and Silverstein et al. (1999) detailed analysis of one billion queries submitted to the Alta Vista search engine over a 42-day period. These studies of transaction log data provide valuable information about search query structure and complexity, including insights about common search topics, query length, Boolean operator usage, search session length, and search results page viewing (Spink and Jansen 2004).

Notwithstanding the value of transaction log data analysis, these types of studies offer limited insights into the behavior of Web searchers beyond the search queries submitted. Hargittai's (2002, 2004) use of surveys and in-person observation of search engine usage helps alleviate these shortcomings, providing insights into how people find information online in the context of their other media use, their general Internet use patterns, and their social support networks. Broadening the analysis of user behavior beyond transaction logs allowed Hargittai (2004) to reveal the ways that factors such as age, gender, education level, and time spent online are relevant predictors of a user's Web searching skills. The work of Machill et al. (2004) and Hölscher and Strube (2000) also combined surveys, interviews, and transaction log analysis to characterize a number of information seeking behaviors of Web search engine users.

Recent scholarship has moved beyond the technical and individual focus of the user studies described above to include research into broader cultural, legal, and social implications of Web search engines. For example, cultural scholars (Hellsten et al. 2006; Wouters et al. 2004) have explored the ways in which search engines "re-write the past" due to the frequent updating of their indices and the corresponding loss of a historical record of content on the Web. Introna and Nissenbaum's (2000) seminal study, "Shaping the Web: Why the Politics of Search Engines Matter," was among the first to analyze search engines from the political perspective, noting how search engines have been heralded as "a democratizing force" that will

...give voice to diverse social, economic, and cultural groups, to members of society not frequently heard in the public sphere. It will empower the traditionally disempowered, giving them access both to typically unreachable nodes of power and to previously inaccessible troves of information. (Introna and Nissenbaum 2000: 169)

Search engines, then, act as a powerful source of access and accessibility within the Web. Introna and Nissenbaum reveal, however, that search engines "systematically exclude certain sites and certain types of sites, in favor of others, systematically giving prominence to some at the expense of others" (2000: 169).

Such a critique resembles the stance that political economists take against the contemporary mass media industry (Castells 1996; Habermas 1992; McChesney 1999), a critique that has recently been extended to Web search engines. For example, Hargittai (2004) has extended her user studies to include investigations of how financial and organizational considerations within the Web search engine industry

impact the way in which content is organized, presented, and distributed to users. And Van Couvering (2004) has engaged in extensive research on the political economy of the search engine industry in terms of its ownership, its revenues, the products it sells, its geographic spread, and the politics and regulations that govern it. Drawing comparisons to concerns over market consolidations in the mass media industry, Van Couvering fears that the market concentration and business practices of the search engine industry might limit its ability to serve “the public interest in the information society” (Van Couvering 2004: 25).

Extending from these social and cultural critiques, Web search engines have also recently been scrutinized from a moral or ethical perspective. A recent panel discussion at the Santa Clara University Markkula Center for Applied Ethics was one of the first to bring together ethicists, computer scientists, and social scientists for the express purpose of confronting some of the “unavoidable ethical questions about search engines,” including concerns of search engine bias, censorship, trust, and privacy (Norvig et al. 2006). A special issue of the *International Review of Information Ethics* on “The Ethics of Search Engines” (Nagenborg 2005) brought into focus many of the particular privacy concerns with search engines.

*Web Search: Multidisciplinary Perspectives* contributes to this rich library of research by showcasing the latest multidisciplinary theories, models, and perspectives on Web searching. Unlike many volumes on Web search engines, our book does not provide an analysis of Web searching from computer science or other Web-related technological disciplines. Rather, *Web Search: Multidisciplinary Perspectives* is focused on investigating Web search from the non-technological perspective. The editors focused on collecting papers that broaden and deepen the framework for our understanding of Web search, and invited authors from many disciplines to contribute chapters that represented emerging research directions and ideas, in an effort to build a perspective that extends beyond traditional models and research, and provide new directions for further research. In particular, the book includes papers by outstanding, yet often less established, researchers from different disciplines who challenge the established views and paradigms of Web search research. The chapter authors – as well as the editors – are drawn from the international boundaries of Web search scholarship, and this global perspective contributed greatly to the multidisciplinary depth of the volume.

## 1.2 Book Outline

*Web Search: Multidisciplinary Perspectives* is organized into five sections. Following this introductory section, Part II presents chapters that provide social, cultural and philosophical perspectives for conceptualizing Web search. Alejandro Diaz’s “Through the Google Goggles: Sociopolitical Bias in Search Engine Design” provides an opening examination based in communication and political theory on how bias in search engines – Google, in particular – might threaten the utopian and democratic ideals associated with the Web. In “Reconsidering the Rhizome: A Textual Analysis of Web Search Engines as Gatekeepers of the Internet,” Aaron Hess

performs a textual analysis of four major search engines to determine how they might resemble Deleuze and Guattari's notion of the rhizome. Rosa Mikeal Martey's contribution, "Exploring Gendered Notions: Gender, Job Hunting and Web Searches," argues that the social and cultural contexts of both the search tools and the search tasks impact how these Web-based technologies serve women in their information-seeking needs. The philosopher Lawrence Hinman provides a necessary ethical analysis of Web searching in his contribution, "Searching Ethics: The Role of Search Engines in the Construction and Distribution of Knowledge," while Michael Zimmer's chapter, "The Gaze of the Perfect Search Engine: Google as an Infrastructure of Dataveillance," focuses on the particular ethical concern with the privacy and surveillance implications Web search engine practices.

Part III includes chapters that propose political, legal, and economic perspectives for understanding Web search. The first contribution, "Search Engine Liability for Copyright Infringement" by Brian Fitzgerald, Damien O'Brien, and Anne Fitzgerald, provides a broad overview of the topic of search engine liability for copyright infringement. Eric Goldman's contribution, "Search Engine Bias and the Demise of Search Engine Utopianism," provides an additional legal analysis of Web search, using legal theory to support the position that search engine bias can be a beneficial consequence of how Web search engines increasingly customize content for individual users. In "Search Engines, Chance Exposures and Emergent Organizations," Azi Lev-On relies on political theory to reveal how search engines can provide unplanned exposures to diverse viewpoints, as well as empowering what he calls "organizational hubs of collective action." Paul Reilly continues this political analysis of Web searching by discussing the relative "visibility" of organizations on search engines in his contribution, "'Googling' Terrorists: Are Northern Irish terrorists visible on Internet Search Engines?" Finally, Elizabeth Van Couvering's chapter, "The History and Geography of the Internet Search Engine: Processes of Consolidation and Processes of Expansion," provides a detailed historical and economic analysis of Web search engines, drawing out concerns over the commercialization and consolidation of the search engine industry.

Part IV presents explorations of Web searching from the information behavior perspective. The section opens with Shirlee Knight and Amanda Spink's chapter, "Towards an Integrated Information Behavior Model of Web Search," exploring the history of information retrieval research in order to propose a "macro model" of Web-based information seeking and searching behavior. In "Web Searching for Health: Theoretical Foundations and Connections to Health Related Outcomes," Mohan Dutta and Graham Bodie utilize theories of information seeking to determine how search engines might fit within an "integrative model of health information seeking." Jenny Fry, Shefali Virkar, and Ralph Schroeder follow with "Search Engines and Expertise about Global Issues: Well-defined Landscape or Undomesticated Wilderness?", an investigation of the "winner-takes-all" effect in online information resources to help determine if search engines function as facilitators in accessing expertise or as influential gatekeepers. "Conceptual Models for Search Engines," by David Hendry and Efthimis Efthimiadis, examines the conceptual and technical understanding that people have of search engines to measure levels of "literacy" of

Web search engine design and practices. Finally, Dirk Lewandowski and Nadine Höchstötter propose and evaluate various quality measures for Web search engine performance in their contribution, “Web Searching: A Quality Measurement Perspective.”

In Part V the editors provide a concluding overview of the key trends, theories and models emerging these multidisciplinary studies, along with a range of new directions proposed in the chapters for further research.

## References

- Brin S, Page L (1998) The anatomy of a large-scale hypertextual web search engine. *WWW/Computer Networks* 30: 107–117
- Castells M (1996–1998) *The information age: economy, society and culture*, vol. i, ii, and iii. Blackwell, Cambridge, MA
- Fallows D (2005) Search engine users: Internet searchers are confident, satisfied and trusting – but they are also unaware and naïve. Pew Internet & American Life Project. Available: [http://www.pewInternet.org/pdfs/PIP\\_Searchengine\\_users.pdf](http://www.pewInternet.org/pdfs/PIP_Searchengine_users.pdf) via the Internet.
- Habermas J (1992) *The structural transformation of the public sphere : an inquiry into a category of bourgeois society*. Polity, Cambridge
- Hargittai E (2002) Beyond logs and surveys: in-depth measures of people’s web use skills. *Journal of the American Society for Information Science and Technology* 53: 1239–1244
- Hargittai E (2004a) Informed web surfing: the social context of user sophistication. *society online: the Internet in context*, Thousand Oaks: Sage Publications, Inc 257–274
- Hargittai E (2004b) The changing online landscape: from free-for-all to commercial gatekeeping. Available: <http://www.eszter.com/research/c03-onlinelandscape.html> via the Internet.
- Hellsten I et al. (2006) Multiple presents: how search engines re-write the past. *New Media and Society* 8: 901–924
- Heydon A, Najork M (1999) Mercator: a scalable, extensible web crawler. *World Wide Web* 2: 219–229
- Hoelscher C (1998) How Internet experts search for information on the web. *World Conference of the World Wide Web, Internet, and Intranet*, Orlando, FL
- Hölscher C, Strube G (2000) Web search behavior of Internet experts and newbies. *Computer Networks* 33: 337–346
- Introna L, Nissenbaum H (2000) Shaping the web: why the politics of search engines matters. *The Information Society* 16: 169–185
- Jansen BJ, Pooch U (2001) A review of web searching studies and a framework for future research. *Journal of the American Society for Information Science and Technology* 52: 235–246
- Jansen BJ, Spink A, Saracevic T (2000) Real life, real users, and real needs: a study and analysis of user queries on the web. *Information Processing and Management* 36: 207–227
- Kleinberg J, Lawrence S (2001) The structure of the web. *Science* 294: 1849–1850
- Lawrence S, Giles CL (1998) Searching the world wide web. *Science* 280: 98–100
- Lawrence S, Giles L (2000) Accessibility of information on the web. *Intelligence* 11: 32–39
- Machill M et al. (2004) Navigating the Internet. *European Journal of Communication* 19: 321–347
- McChesney R (1999) *Rich media, poor democracy: communication politics in dubious times*. University of Illinois Press, Urbana, IL
- Nagenborg M (2005) The ethics of search engines (special issue). *International Review of Information Ethics* 3.



- Norvig P et al. (2006) The ethics and politics of search engines. Panel at Santa Clara University Markkula Center for Applied Ethics. Available: <http://www.scu.edu/sts/Search-Engine-Event.cfm> via the Internet. Accessed Panel discussion
- Page L et al. (1998) The pagerank citation ranking: bringing order to the web. Technical report
- Rainie L (November 2005) Search engine use shoots up in the past year and edges towards e-mail as the primary Internet application. Pew Internet and American Life Project. Available [http://www.pewinternet.org/pdfs/PIP\\_SearchData\\_1105.pdf](http://www.pewinternet.org/pdfs/PIP_SearchData_1105.pdf) via the internet
- Silverstein et al. (1999) Analysis of a very large web search engine query log. SIGIR Forum 33: 6–12
- Spink A, Jansen BJ (2004) Web search: public searching of the web. Kluwer Academic, New York
- Van Couvering E (2004) New media? The political economy of Internet search engines. Annual Conference of the International Association of Media & Communications Researchers, Porto Alegre, Brazil, 7–14
- Wouters P et al. (2004) Internet time and the reliability of search engines. First Monday. 9(10). Available: [http://www.firstmonday.org/issues/issue9\\_10/wouters/index.html](http://www.firstmonday.org/issues/issue9_10/wouters/index.html) via the Internet.

**Part II**  
**Social, Cultural,**  
**and Philosophical Perspectives**

## 2

# Through the Google Goggles: Sociopolitical Bias in Search Engine Design

A. Diaz

**Summary** Search engines like Google are essential to navigating the Web's endless supply of news, political information, and citizen discourse. The mechanisms and conditions under which search results are selected should therefore be of considerable interest to media scholars, political theorists, and citizens alike. In this chapter, I adopt a "deliberative" ideal for search engines and examine whether Google exhibits the "same old" media biases of mainstreaming, hypercommercialism, and industry consolidation. In the end, serious objections to Google are raised: Google may favor popularity over richness; it provides advertising that competes directly with "editorial" content; it so overwhelmingly dominates the industry that users seldom get a second opinion, and this is unlikely to change. Ultimately, however, the results of this analysis may speak less about Google than about contradictions in the deliberative ideal and the so-called "inherently democratic" nature of the Web.

## 2.1 Introduction

As knowledge, commerce, and politics continue move online and to the Web in particular, search engines have quickly become the "gatekeepers" of cyberspace. What's more, a *single* search engine – Google – now handles the majority of Web queries. Google directs hundreds of millions of users towards some content and not others, towards some sources and not others. As with all media gatekeepers, if we believe in the principles of deliberative democracy – and especially if we believe that the Web is an open, "democratic" medium – then we should expect our search engines to disseminate a broad spectrum of information on any given topic.

In the first section of this chapter, I describe how a "deliberative media" ideal can be used to evaluate search engine and why, as media critics have done with prior innovations, we should examine Google's content biases, its advertising policies, and consolidation in the industry as a whole. Subsequent sections will dive into each of these areas: first, we will look at the deliberative implications of the PageRank algorithm Google uses to crawl and order Web content; next,

we will critically examine the role advertising plays in Google’s search results; lastly, we discuss the implications of a highly concentrated and commercial competitive search landscape. It is hoped that through this investigation, we might start to uncover the sociopolitics of search.

## 2.2 In Search of a Democratic Medium

The Supreme Court once observed that “the dissemination of the widest possible information from diverse and antagonistic sources is essential to the welfare of the public” (Associated Press v. United States 1945: 20). This goes to the heart of ‘deliberative democracy,’ a concept that has in recent years gained considerable currency among political scientists and media critics (Benhabib 1996; Elster 1998; Fishkin 1991; Sunstein 1997). For Benjamin Page, “In order that the public as a whole can collectively control what its government does, the public, collectively, must be well informed. Some kind of *public* deliberation is required” (Page 1996: 5). Individuals’ exposure to “diverse and antagonistic views” is central to such debate, as John Stuart Mill (1859) once argued:

He who knows only his own side of the case knows little of that. His reasons may be good ... but if he is unable to refute the reasons of the opposite side, if he does not so much as know what they are, he has no ground for preferring either opinion ... [H]e should hear the arguments ... from the persons who actually believe them, who defend them in earnest and do their very utmost for them. He must know them in their most plausible and persuasive form. (p. 35)

For Mill, it does not matter whether arguments are popular or unpopular, correct or incorrect, offensive or pleasing; what matters is that public opinion is given the *opportunity* to “be set right when it is wrong” (p. 19). This is why “streets and parks,” according to Justice Roberts, “have immemorially been held in trust for the use of the public and...have been used for purposes of assembly, communicating thoughts between citizens, and discussing public questions” (Hague et al. v. CIO et al., 1939, 515).

The deliberative model appears to capture what we usually mean by “democratic media”: forums in which every corner of society is represented fairly – spaces where the debate isn’t dominated by corporations, politicians, or privileged groups. Given that we are a nation too large and too distributed to engage in a singular, Habermasian debate (1990, 89) the media have an important role to play in ensuring that speakers have access to heterogeneous citizens. As Justice Kennedy observed, “minds are not exchanged in streets and parks as they once were. To an increasing degree, the more significant exchanges ... occur in mass and electronic media” (DAETC et al. v. FCC 1996, 132). Given the enormous reach of radio, television, and newspapers, the media could allow citizens to access a range of perspectives they might not otherwise encounter.

### 2.2.1 *The Traditional Media and a New Hope*

But as countless critics have argued, the mass media have fallen far short of these aspirations. “[M]arket forces,” writes Cooper (2003), “provide neither adequate incentives to produce the high quality media product, nor adequate incentives to distribute sufficient amounts of diverse content necessary to meet consumer and citizen needs” (p. 43). The economics of dissemination and the politics of deregulation, rather than encourage the formation of alternative outlets, have concentrated the media in fewer and fewer hands (p. 141). The scarcity of alternative channels has allowed media companies to pursue ever-greater profit margins through advertising, sponsorship, and product placement with little fear of consumer retaliation (McChesney 2000: 39–42). “The media,” according to Bagdikian (1992), “have become partners in achieving the social and economic goals of their patrons” (p. 151). The value of large audiences has tended to yield “middle-of-the road,” nonpolitical, mainstream content that creates a “buying mood” but fails to represent unpopular or diverse opinions. The result is a media landscape characterized by sameness, by a suppression of controversy, and by hypercommercialism.

But a new medium has recently emerged, and it has promised to change all this. Decentralized and distributed, the global Internet – and, in particular, the Web – allows anyone and everyone to make their views accessible, and to access anyone’s views. It is arguably more like the printing press than radio and, indeed, information on the network is not constrained by the limits of printed matter, by delivery distances, or seemingly by time, space, and matter itself. With a click of the mouse, you can read information and opinions that have not been “filtered” by profiteering corporations or corrupt governments. At the same time, underrepresented and unheard groups can cheaply bypass the “monolithic media empire” to have a voice. The Internet is many-to-many, all-to-all, and it has for many restored faith mediated deliberation. Aspirations are expressed repeatedly, and with understandable excitement:

The Web...breaks the traditional publishing model. ... [It] says instead, “You have something to say? Say it. You want to respond to something that’s been said? Say it and link to it. ... And you never have to ask anyone’s permission.” (Weinberger 2002: vii–ix)

You don’t have to be writing for an organization to have a credible voice. The Net elevates those voices. What the large media were about was distribution capacity to communicate with hundreds of thousands of people. Now the Net does that. (Barlow, qtd. in Lasica 1996)

The prospects seemed so exhilarating that some jumped to label the Internet “inherently” democratic (Gilder, qtd. in Schuler 2003: 72).

And indeed, the Web has had many tangible, positive effects for diverse, democratic discussion. Access and content on the new medium has exploded; the majority of Americans now have Internet connections in their homes (Wellman and Haythornthwaite 2002: 13). Anyone with basic computer competence can now publish a Web site viewable around the globe. Activists have grown their own “grassroots” communities to pursue particular policy objectives while bloggers – self-made ‘journalists’ who report their findings and solicit comments in a

sort of “deliberative diary” – have gained loyal followings and the attention of the mainstream media (Rodzvilla 2002). Real-world community projects have sprung up online, “evidence of an overdue renewal of interest in democracy” (Schuler 2003: 73).

And yet, over the last ten years, user traffic on the Web has gravitated around a few, large, and increasingly commercial sites. In a fascinating book, Notre Dame physicist Albert-László Barabási (2002) recounts how his team of scientists mapped the Web’s structure to reveal disturbing evidence about the supposed “egalitarianism” of the network. He found that a small number of pages – what he called “hubs” – are linked to by a great many other pages, while the vast majority of documents are linked to by few or no sites at all. Hubs are very easy to “come across” from anywhere on the Web; they are therefore more likely to be linked to, which further increases their discoverability (the so-called “rich get richer” phenomenon). Meanwhile, a typical page – one pointed to by only couple documents – remains almost impossible to find. It’s no wonder that, by 2001, over half of users’ online time was being spent at four sites; one third of the total time was spent at AOL-Time Warner properties (CNN.com 2001). On the political Web – the set of sites dealing with democratically urgent issues such the death penalty, Congress, and gun control – Hindman et al. (2003) found “strong and consistent” patterns consistent with Barabási’s research: “the number of highly visible sites is small” and “almost all prominent sites are run by long-established interest groups, by government entities, by corporations, or by traditional media outlets” (p. 26). The link structure of the Web suggests the medium exhibits the same old problems: “it is hard for all but a few ‘ordinary citizens’ to post their views prominently – and conversely, to read the views of other ordinary citizens, unless they are highlighted by a small number of prominent sites” (p. 30). Or as Barabási put it, “The hubs are the strongest argument against the utopian vision of an egalitarian cyberspace. Yes, we all have the right to put anything we wish on the Web. But will anybody notice?” (p. 58).

### **2.2.2 Search Engines as Intermediaries**

That so many accessible pages go unseen suggests that the Internet has done away with “spectrum scarcity” but not with *attention scarcity* (Kottke 2003). Sure, there are literally billions of pages (“channels”) available on the Web. But there is a rather fixed limit to how many we, as individuals, can consume. With television, radio, and the print media, we rely on the mass media to condense the available opinions and make them easily accessible through newspapers, the evening news, radio broadcasts, and so on. And the same sort of intermediation is required online.

The key “general interest intermediaries” of the Web, I argue, are the search engines. These sites are the primary means by which Internet users are directed towards particular sources of information and are among the first and most frequently accessed pages for the vast majority of users. Consider: each one of the top 5 sites

is either a portal or search engine (Burns 2007); by 2004, 84% of online Americans had used search engines, and a majority of these used them at least once a day (Fallows and Rainie 2004); search engines are the most popular way to locate medical, governmental, and religious information on the Web (Fallows 2005); fully 79% of those seeking online election information began their journeys at portals and search engines (Cornfield and Rainie 2003, p. 25).

So when Steven Levy (1995) said that “instead of a gatekeeper, users get an open invitation to the electronic world and can choose whatever they want” (p. 59), he was being less than accurate. Internet users *do* get a gatekeeper – the search engine – and they choose primarily among the sites it offers to them. As with all such intermediaries, we expect search engines to present the available information in a fair and diverse manner; we expect them, in other words, to be “democratic.” We should ask about search engines like Google the same questions scholars have asked about the traditional media: *Can underrepresented voices and diverse viewpoints be heard through the filter of search engines? What role does advertising play in the returned results? Do a few players dominate the industry?* Only by answering these questions – as we will do in turn – can one assess the true “deliberativeness” of the Web itself.

## 2.3 The Politics of PageRank

Just as the mass media have the power and responsibility to disseminate unusual and heterodox views, so too do search engines have the capability to highlight those high-quality, out-of-the-mainstream sites that would otherwise be lost in the deafening din of the Web. Automated crawling and ranking can do what we, as individuals, cannot: find, catalog, and consider millions of poorly-linked and under-represented pages – what Chris Anderson (2004) has called “the long tail” – and ultimately break through the link inequality that calls into question the egalitarian ideal of the Web.

### 2.3.1 *The Mathematics of PageRank*

So does Google actually promote those dissident and minority views so critical for a “well-functioning democracy”? Given the complex and proprietary nature of Google’s search technology its software looks at over 100 features of a page to ascertain “relevance” (Mayer 2005) – answering this question is exceedingly difficult. But we can start with what Google (2004) calls “the heart” of its software: the PageRank algorithm. PageRank estimates the “importance” of an arbitrary page by looking at how many *other* “important” pages link *to* it. Mathematically, the PageRank of your page is the weighted count of links pointing to it, with links from high-PageRank documents contributing more to your score than links from low-PageRank documents.

An analogy may be useful: an academic paper is “important” if many other papers cite it – and especially if it is referenced by other, highly cited works (“canons”).

With PageRank, then, being “important” means being “popular” or “visible.” PageRank actually turns out to be the precise probability that a “random surfer” clicking links from page to page will come across a given document. Thus the highly-referenced “hubs” Barabási worried about have the highest PageRanks; these tend to be the sites of large, famous, technology-oriented companies such as Amazon and eBay (Upstill et al. 2003). In contrast, the millions of “typical” pages – those we are already unlikely to “randomly” stumble across – have among the lowest PageRank values. Google apparently uses PageRank to guide its crawlers such that popular sites have a better chance of being indexed (Cho et al. 1998). Sites with high PageRank also tend to be more prominent among the search results (Diaz 2005: 81–85).

### 2.3.2 PageRank as a Voting Mechanism

According to Google’s public relations literature (2004), PageRank is not only consonant with democratic principles; it in fact embodies the very process of democracy itself: “Google interprets a link from page A to page B as a vote, by page A, for page B.” Princeton computer science professor and cyberactivist Ed Felten (2004) puts it more colorfully:

Google is a voting scheme ... not a mysterious Oracle of Truth. ... It’s a form of democracy – call it Googlocracy. Web authors vote by creating hyperlinks, and Google counts the votes. If we want to understand Google we need to see democracy as Google’s very nature, and not as an aberration.

But what Ed Felten and other PageRank proponents fail to recognize is the important distinction between the ideal process of “democratic” *governance* and that of “democratic” *discourse*. Sure, a political democracy generally requires that the aggregated preferences of the majority be put into practice. But this does not imply that only the majority’s views should be heard during deliberation, nor does it suggest that popular opinions should be preferred *ipso facto*. To the contrary, the validity of voting – of aggregating preferences – depends precisely on the dissemination of a broad spectrum of opinions, especially those put forth by unpopular or minority groups (Mill 1859: 16).

From the perspective of deliberative democrats, then, PageRank is highly problematic. Unpopular but nevertheless democratically critical voices face a double bind: search engines like Google are “biased against [these pages], ignoring them as they crawl the Web” (Barabási 2002: 58) and, even if the pages make it to the index, they may find themselves buried among the results. To the degree that Google adopts a PageRank bias, it *mirrors* rather than mitigates the Web’s link inequality.



Indeed, some scholars have argued that the use of PageRank actually *magnifies* the Web's skewed distribution of links, making it increasingly difficult for new sites to be discovered (Fortunato et al. 2006; Hindman et al. 2003; Pandey et al. 2005). The problem is this: a well-linked page appears prominently on search engines like Google; this page therefore enjoys greater traffic; and, as users become even more aware of the site, they link to it on their own pages, increasing the document's PageRank and visibility even further. The result is a "vicious cycle," "entrenchment bias," or "googlearchy" wherein popular pages are, over time, increasingly likely to maintain their prominence while new pages become more difficult to discover. Cho and Roy's (2004) computer simulation indicated that "it takes 66 times longer" for a new page to become popular by means of highly PageRank-biased search engines than by pure "random surfing."

### 2.3.3 *The "Common Case" and Majoritarian Interests*

PageRank therefore seems to reproduce the same sort of "antideliberative" bias typically associated with the traditional media. To recall Cooper's (2003) remarks about big media: "In the commercial model, popular, mainstream, and middle of the road ideas will almost certainly find a voice, one that is likely to be very loud. However, the unpopular, unique, and minority points of view will not" (p. 16). Similarly, "search engines wishing to achieve greatest popularity ... tend to cater to majority interests" (Introna and Nissenbaum 2000: 176). According to Google's founders, this bias was by design:

One of the design goals of PageRank was to handle the common case for queries well. ... [T]he goal of finding a site that contains a great deal of information ... is a very different task ... There is an interesting system that attempts to find sites that discuss a topic in detail ... this results in good results for queries like "flower"; the system will return good navigation pages from sites that deal with the topic of flowers in detail. Contrast that with the common case approach which might simply return a commonly used commercial site that had little information except how to buy flowers ... [W]e are concentrating only on the common case approach. (Page et al. 1999: 10–11).

PageRank, in other words, abandons the goals of actually reflecting a page's "importance" or "authoritativeness" on a given subject, and instead aims to mirror the "common" wishes of users. This, as the creators' own example illustrates, can have the problematic effect of promoting popular, commercial pages over more detailed, noncommercial sources of information.

To be sure, these problems are more or less typical of commercial search engines in general. In their groundbreaking overview of search engine bias, Introna and Nissenbaum (2000) observed that "while markets undoubtedly would force a degree of comprehensiveness and objectivity in listings, there is unlikely to be much market incentive to list sites of interest to small groups of individuals ... or, for that matter, individuals of lesser economic power" (p. 177). PageRank's

“one size fits all” approach does little for the atypical, outside-the-mainstream individuals that might actually wish to see or communicate controversial content.

### ***2.3.4 Suppression of Controversy***

Susan Gerhart (2004) makes a similar point in a unique content analysis of Web search results. Gerhart queried Google, Teoma, and AllTheWeb for information on five broad topics, each of which she knew to contain some controversial subtopic that was well documented on the Web. Gerhart then recorded, in painstaking detail, whether and how such disputed perspectives were raised within the search results. She looked, for example, at whether a search for “distance learning” would return sites that shared David Noble and other academics’ concern about “the loss of control over their intellectual products, as well as contact with students” and the tendency of these programs to act as “digital diploma mills.” Similarly, she looked at whether the results for “Einstein” mentioned the debate over whether his first wife received appropriate credit for contributions to his work.

Her findings indicate that when a controversy was frequently discussed within a topic and widely recognized as important (e.g., the effectiveness of St. John’s Wort) the disputed matters were, indeed, represented among the results. When searching for female astronauts or St. John’s Wort, for example, it was possible for a user to “definitely recognize the existence of controversy, which [a result] explains in some detail.” But for three of the topics – distance learning, Albert Einstein, and Belize – the respective disputes were to a great extent “suppressed,” such that most surfers would not “be exposed to the controversies by [a general] search...alone.” In these cases, the controversies were overrun by “organizational clout” (e.g., official Belizean tourism sites or distance learning programs) or by pages that reflected what users “wanted to see” (e.g., Einstein quotations, ‘bland’ biographies for term papers, etc.). In the end, the controversial viewpoints that perhaps matter most from a deliberative point of view – those antagonistic perspectives that haven’t garnered widespread attention – are precisely those that are left out of the search engine’s results. Gerhart concludes that

Search technology tends to present the ‘sunny side’ of a topic. This bias reflects authors’ links and searchers’ choices. A few organizations often exert strong commercial (or non-profit) influence through Web site investments and accrue high link counts through their off-Web prominence. (‘Conclusion’).

If we really believe that through “democratic media” like the Web individuals “must have the freedom to communicate radical and unpopular ideas and opinions” – and, what’s more, that citizens should be exposed to what “they don’t want to hear” – then search engines fall short of these aspirations when they fail to disseminate those dark, uncomfortable views on a given topic.

Of course, “Web search engines do not conspire to suppress controversy.” Rather, this is direct consequence of the seemingly laudable attempt to please its

users. As Gerhart suggests, “On the simplest query for a topic, a searcher expects to see the most influential organizations appear, not a bundle of dirty laundry or diatribes attacking the topic’s leaders or ideas ... Searchers use a particular engine because its biases give them the results they usually want.” The deliberative model may ask of too much from users: pushing for them to see what they *don’t* want to see because, really, it’s “good for them” (Rostbøll 2005). To this extent, it conflicts with intuitive and reasonable ideas for how search engines should work.

### 2.3.5 *Small Players (Still) Matter*

Even if PageRank does, in theory, encode an antideliberative, antidemocratic bias, a few caveats are in order. First, as Dan Bricklin (2002) has pointed out, even if popular sites do get a sizeable boost for some queries, rarely do the same corporate megasites pop up across different search topics. As a result, “small players [still] matter,” especially when we are conducting ‘typical’ searches for specialized information not easily found in the traditional media. Although it is difficult for a page to gain visibility on established topics – Microsoft, “abortion,” or “flowers” – an unprecedented number of “ordinary citizens” may still be reaching sizeable publics through the Google search engine.

Second, PageRank is only one element of Google’s ranking algorithm; consider, for example, that PageRank is completely *query-independent*, capturing the “importance” of a page irrespective of the user’s stated interest. In practice, Google takes many other factors into account when ordering search results: whether the query appears in the page’s title, what words people use to link to the page, and so forth. While it is true that PageRank predicts rank position *in the aggregate*, individual result sets exhibit at best a weak correlation (Diaz 2005: 84). For this reason, Cho and Roy’s simulation – which assumed search results were strictly ordered by PageRank – may be unrealistic. Indeed, a more recent study suggests that search engines’ query-dependent heuristics actually “smear out the traffic attraction of high-degree pages...counteracting the skewed distribution of links in the Web [by] directing some traffic toward sites that users would never visit otherwise” (Fortunato et al. 2006: 6). Clearly, there is a need for continued and systematic research into the many; sometimes counteracting biases of today’s advanced search technology.

## 2.4 Commercialism, Advertising, and “Mixed Motives”

Advertising is, by and large, how the commercialized media make money. Newspapers, magazines, radio, and television outlets provide free or inexpensive content to their readers, listeners, and viewers; in exchange, they sell advertisers access to these audiences. Advertising is, however, a mixed blessing. On the one hand, it makes it viable to disseminate information to a broad audience at a low

cost; on the other hand, there is the persistent threat that the wishes of sponsors will subtly work their way into the content itself, narrowing the range of opinions that can be profitably and widely expressed.

These competing forces come strongly into play in the arena of search engines. As the primary gatekeepers of the Web, search engines not only direct users to particular pages but can also direct consumers towards particular services and products. This presents an enormous opportunity for targeted advertising: search engines can “sell” access to highly segmented audiences while marketers can target individuals who are actively expressing interest in a topic or product. The money-making potential is enormous and, indeed, one industry report predicted as early as 2003 that “worldwide search revenue estimates of \$7B by 2007 are conservative” (Raschtchy and Avilio 2003). By 2005, advertisements on Google *alone* brought in over \$6 billion – or over 99% of the company’s yearly revenue (Google 2006).

But by selling advertising, Google and its competitors have an enormous financial incentive to direct users away from the “free,” “organic” results and towards the sites of its sponsors. These “mixed motives” are stated eloquently by none other than Google’s founders themselves, in an appendix to their 1998 Stanford research paper:

The goals of the advertising business model do not always correspond to providing quality search to users. For example, in our prototype search engine one of the top results for cellular phone is ... a study which explains [the] risk associated with conversing on a cell phone while driving ... It is clear that a search engine which was taking money for showing cellular phone ads would have difficulty justifying the page ... For this type of reason and historical experience with other search engines we ... expect that advertising funded search engines will be inherently biased towards the advertisers and away from the needs of the consumers. ... Since it is very difficult even for experts to evaluate search engines, search engine bias is particularly insidious ... [and] less blatant bias are likely to be tolerated by the market. (Brin and Page 1998: 17–18)

### ***2.4.1 A Brief History of Search Advertising***

When Google’s founders wrote those words, the predominant form of search advertising was the so-called “banner” ad. As it turns out, these ads tended not to work well in the context of search. For one, only a few banners can reasonably be placed on each page, and searchers would often click their result before the image had finished loading (Sullivan 2003a). More importantly, users quickly developed an ability to unconsciously spot and ignore banners, focusing – with “laser beam accuracy” – on what they perceived to be the actual search results (Pagendam and Schaumburg 2001). If sponsors wished to be noticed, their solicitations must look like, and appear amongst, the actual results. As the CEO of one search engine company put it, “The money is in the search results themselves, not the billboards on the site of that road. The question is how do you profit from the search results, when they have been given away for free” (Thornley qtd. in Pagendam and Schaumburg 2001).

The way many of Web search engines have gone about “profiting from their results” is by offering various kinds of “paid listings.” The most common scheme, called *paid placement*, allows sponsors to purchase search-result-like text ads that appear above, below, or alongside the “organic” results for their chosen keywords. Sometimes these paid results are marked as “sponsored” listings; other times, “it may be hard for the average user to distinguish” (Crowell 2003). Unsurprisingly, paid placement proved vastly more effective than previous methods at drawing users’ attention towards sponsors’ sites. To the degree that these “matches” walk, talk, and act like relevant results, users click them. As *Business Week* puts it, paid placements have become “the Holy Grail of Internet advertising, and no wonder” (Reinhardt 2003). These ads have caught on, in some form or another, among virtually all of Web’s most popular search engines (Google, Altavista, AOL, AskJeeves, Hotbot, Google, Lycos, MSN, and Yahoo! have similar offerings). The demand for paid listings quickly became so great that, according to *The Economist* (2004), they “lead the recovery in advertising expenditure on the Internet.”

While paid listings may be a bonanza for search companies, investors, and advertisers alike, their implications for online, egalitarian discourse are depressingly obvious:

[The] concept that Web sites should be able to buy their way to the top of search listings is being copied in one way or another by every major search and portal site. As they do, the search engines, which are still the most popular gateways to the Web, are transforming themselves from infinite electronic encyclopedias to the more prosaic, if profitable, role of universal commercial directories. (Hansell 2001)

To the extent that the commercial interests of the rich dominate the results of even noncommercial queries, the practice of selling prominence can seriously distort what the Web consists of for millions of users.

But just as market forces drive search engines to paid placement, so too do market forces push back. If, as commercial listings become more numerous, the relevancy of a search engine’s results decline, dissatisfied users may switch to a competitor, resulting in an overall decline in advertising revenues. From this angle, the amount of paid listings to include is a straightforward optimization problem. Economists Bhargava and Feng (2002) respond to it by proposing “a mathematical model for optimal design of a paid placement strategy” that would “give a search engine the best balance between revenues from content providers and revenues based on user base” (p. 122).

### 2.4.2 ‘Clearly Labeled’?

For search engine critics, however, such economic models are not very comforting. We might reasonably wonder – as Brin and Page did in 1998 – whether users will actually see what’s missing from their search results. Although Bhargava and Feng assume “that search engines cannot hide the fact that they perform paid placement” (p. 118), it appears that many Internet users remain unaware of such practices.

In 2002, a study commissioned by Consumers Union found that *fewer than one in four* Internet users had ever heard of search engines “taking fees to list some sites more prominently than others” (Princeton Survey Research Associates 2002: 17). After being told that, in fact, most search engines do exactly this, “a solid majority (80%) say it is important for search engines to tell users about their fee details, including 44% who say it is *very* important” (p. 17). At the time of the study, several search engine companies were using remarkably vague and misleading terminology to demarcate their paid listings (e.g., “Featured Sites,” “Products and Services”). So in 2001 the watchdog group Commercial Alert filed a complaint with the FTC alleging that seven search companies were engaging in “deceptive advertising” practices (Miller 2001). When FTC responded in June 2002, it did not call for immediate action against the search engines named in the complaint (Gallagher 2002), but it did agree that there was a “need for clear and conspicuous disclosures of paid placement... to advise consumers as to when they are being solicited, as opposed to being impartially informed” (Hippsey 2002).

Google has largely avoided criticism for its AdWords paid placement program and the company was noticeably absent from the Commercial Alert complaint. While other search engines were happily crowding their search results with “Featured Links,” Google insisted on drawing a line – quite literally – between “paid” and “organic” results. Algorithmically, advertising was to have no effect on the selection and ordering of the free results, and ads were “clearly marked” as “Sponsored Links.” These results initially appeared only to the right of the “organic” results, but today Google includes up to three sponsored links directly above the top result. Even though these are also labeled as “sponsored,” selected by relevance not price, and appear over a colored background (Sullivan 2002, AdwordsRep 2004), the fact remains: a considerable portion of Google’s revenue comes from moving ads to the most prominent positions *above* the “first hit.” It is unclear whether, in practice, users perceive these as ads; Google, after all, has an enormous interest in blurring that line.

In any case, disclosure alone does not solve the problems of paid listings. If we really wish to promote ideals of democratic discourse, then we should worry about any policy that allows those with money to be featured prominently among results for a given topic. This concern, it should be emphasized, is not with advertising in general. It is with a particular type of advertising competes with “organic,” relevant content; it is with advertising that supplants, rather than complements, the pages individuals might otherwise see. Despite what Brin and Page say today, paid listings, even if disclosed, are not “just like” advertising in the traditional media. Industry reporter Danny Sullivan (2003a), however, disagrees:

Think newspapers. Newspapers have both “editorial” copy, which is not supposed to be influenced by advertising, as well as ads themselves. You may read the paper primarily for the articles, but there are certainly times when you may find the advertisements useful, as well ... In “old” media ... most people can readily identify ads because they look or act so very different from “content.”

But there’s the rub. In the new media of search engines, paid listings (as opposed to banner ads) *don’t* “look or act so very different” from normal results. Search