

The Hyperlinked Society: Questioning Connections in the Digital Age

Joseph Turow and Lokman Tsui, Editors

Permalink: <http://hdl.handle.net/2027/spo.5680986.0001.001>
[<http://hdl.handle.net/2027/spo.5680986.0001.001>]

Published: Ann Arbor: University of Michigan Press, © 2008.

 [<http://creativecommons.org/licenses/by-nc-nd/3.0/us/>]

C=NMW;IDNO=5680986.0001.001;RGN=DIV1;VIEW=TEXT;CC=NMW;XC=1;G=DCULTURE;NODE=5680986.0001.001%3A2]

C=NMW;IDNO=5680986.0001.001;RGN=DIV1;VIEW=TEXT;CC=NMW;XC=1;G=DCULTURE;NODE=5680986.0001.001%3A4]

Page 19

Part 1: Hyperlinks and the Organization of Attention

Page 20

Page 21

Preface to Part 1

In a digital era where information is seemingly in abundance, the hyperlink organizes our attention by suggesting which ideas are worth being heard and which are not. Hyperlinks do not exist in a vacuum, however. They are created and situated in a political-social context. Despite their ubiquity, we know little about the social and political factors that drive the production of hyperlinks. The essays in this section cross a variety of disciplines to explore the forces that guide and constrain the creation of hyperlinks and the way they organize our attention.

James Webster draws from economics, sociology, and communication to develop a conceptual model he calls the “marketplace of attention.” He argues that the hyperlink can be seen as a form of currency in a marketplace where different producers of online content vie for the attention of the public. After describing the conditions under which this market operates, he focuses on the often expressed concern that linking might lead to social polarization such that people are no longer exposed to a diversity of views but instead retreat into “information enclaves.” Webster argues that the best way to understand the extent to which social

polarization will come to pass is to realize that neither the actors nor the structure of the marketplace will be all-determining. Instead, he says, what happens at the individual level will influence how the marketplace responds, and vice versa. He states that how people will be guided toward information will influence the extent to which social enclaves will come to pass. In the digital world, he contends, search engines give audiences new ways to determine what society will share as important, because their results involve amassing people's preferences.

Whereas Webster highlights the importance of understanding how the hyperlink structures the marketplace of attention, Alex Halavais takes a step back and helps us ground our understanding of the hyperlink from a historical perspective. He explains the original function of the hyperlink as a citation mechanism and shows how it evolved over the years to involve full-fledged networks. He also argues, however, that we can no

Page 22

longer afford to treat the hyperlink simply as a request for a Web document, since an entire industry has now grown up around the manipulation of links connected to search engine results. To grasp the social meaning of the hyperlink, then, requires exploring the struggles between various entities to come out on top of the rankings.

Philip Napoli looks at ways that mainstream media firms try to come up on top of the rankings. While acknowledging that some elements of the Internet are indeed challenging mainstream media, he insists that we should not lose sight of the fact that many others are undeniably becoming highly commercialized and targeting mass audiences. In his view, big media exert substantial institutional and economic power over the shape of the emerging digital environment.

Lokman Tsui's essay explores one facet of mainstream media's relation to this new world. He compares the ways prestigious newspapers and major political blogs are using the hyperlink and finds stark differences. Whereas blogs link heavily to external Web sites, some newspapers hardly link, and others link only to themselves. Considering that online versions of major newspapers are used as a primary means of directing the public's attention to what is deemed valuable information, Tsui's findings have important implications with regard to how the public learns about the world.

Eszter Hargittai is also interested in what the public learns from links, but she takes a different tack. She looks critically at the potential for abusing users via hyperlinks and at the extent to which the users themselves are likely to know that abuse is happening or be aware of this potential risk. She shows that certain users are better positioned than others to note which links are advertisements or online scams and which ones are not. Finally, Hargittai insists that we need to understand the processes that contribute to people's online literacy if we are to avoid exacerbating the current divide whereby the savvy are able to use the Internet to their advantage while the less knowledgeable remain vulnerable to misleading or sometimes even malicious content.

Seth Finkelstein sheds a somewhat different light on the problem of link manipulation. He argues that people think search results imply a Web site's authority on a topic, whereas they are in fact simply popularity measures. Using a number of case studies involving Google, he demonstrates the social dilemmas that confusing popularity with authority can cause. At a time when search engines play a pivotal role in shaping our experience online, Finkelstein's essay

reminds us that it is critical to understand the processes that create the messages we see.

Page 23

Structuring a Marketplace of Attention

James G. Webster

At the conference “The Hyperlinked Society” at the Annenberg School for Communication, Eric Picard of Microsoft asserted that with the exception of maintaining personal networks, people blogged for one of two reasons: fame or fortune. It seems to me that those motives propel most media makers, old and new. And the recipe for achieving either objective begins with attracting people’s attention. Patterns of attention, in turn, establish the boundaries within which the economic and social consequences of the new media environment are realized. This essay invites the reader to think about the hyperlinked environment as a marketplace of attention. I begin with a brief description of market conditions, outline a theoretical framework for thinking about the marketplace, and then use that framework to explore two socially consequential patterns of public attention: fragmentation and polarization. The former addresses the overall dispersion of cultural consumption. The latter addresses the tendency of people to retreat into comfortable “enclaves” of information and entertainment. Finally, I’ll suggest questions and concerns about a hyperlinked society that I believe deserve our attention.

Market Conditions

The hyperlinked environment can be thought of as a virtual marketplace in which the purveyors of content compete with one another for the attention of the public. Three realities set the conditions for the marketplace. I take these to be axiomatic.

Convergence. A popular term that means different things to different people, [1]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N101;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N101;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N101;note=ptr)]

convergence here describes the move toward fully integrated media delivery systems. While the conference focused on media that have emerged in the hyperlinked environment, like blogs, social networking sites, and other forms of user-generated content (e.g., Wikipedia,

Page 24

YouTube), all content is increasingly being distributed on the same high-speed networks. Traditional media, including newspapers, radio, television, and movies, are moving into the hyperlinked space. Consumers, in turn, function in an environment where they can move fluidly among what were once discrete media outlets. In the long term, it makes sense to think about a common media environment where all manner of content is readily available to consumers.

Abundance. The sheer volume of content is vast and increasing at an explosive rate. Exact numbers are hard to come by, in part because they change so quickly. At this writing, Technorati is tracking some sixty million blogs, MySpace has over one hundred million accounts, and the number of podcasts and video clips in circulation seem without end. A great many of their authors undoubtedly seek public attention. Once the delivery of more traditional broadcast content becomes ubiquitous, it will add perhaps one hundred million hours a year of new

programming to the mix. [2] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N102;note=ptr>] All this will be in addition to whatever repositories of movies, music, and news are available on demand. Media are, if nothing else, abundant.

Scarcity. While the supply of content seems endless, the supply of human attention to consume that content is not. There are a limited number of Internet users in the world and a limited number of waking hours. The problem of too much content and too little time is not new. In the early 1970s, Herbert Simon famously noted that “a wealth of information creates a poverty of attention, and a need to allocate that attention efficiently among the overabundance of information sources that might consume it.” [3] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N103;note=ptr>] Obviously, the problem is more acute today and getting worse.

As a result, public attention is spread thin. A relatively small handful of items will achieve widespread notice, but most will not. Those that do will have the potential to produce the fame and/or fortune that their authors desire. Richard Lanham recently wrote: “Assume that, in an information economy, the real scarce commodity will always be human attention and that attracting that attention will be the necessary precondition of social change. And the real source of wealth.” [4] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N104;note=ptr>] What is less understood is how public attention actually takes shape in the new media environment. The following sections outline a theoretical framework for structuring the marketplace of attention and the patterns of consumption that emerge as a result.

Page 25

Toward a Theory of the Marketplace

Adopting a marketplace metaphor may suggest that the operative theory is grounded exclusively in the rational choice model of neoclassical economics. I have in mind a somewhat more flexible framework borrowed from sociology, drawing especially on Giddens’s theory of structuration. [5] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N105;note=ptr>] While structuration has been adapted to explain the use of information technology within organizations, [6] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N106;note=ptr>] it hasn’t been used for a more wide-ranging consideration of media consumption. The principle components of this framework are agents, structures, and the interaction between them that is characterized either as “duality of structure” [7] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N107;note=ptr>] or “dualism.” [8] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N108;note=ptr>] These provide the conceptual tools needed to imagine a marketplace of attention.

Agents. In this context, agents are the people who consume media. Their use of media is purposeful and done at a time and place of their choosing, though in practice, it is often embedded in the routines of day-to-day life. Media consumption is rational in the sense that it

satisfies various needs and preferences. Agents know a good deal about the media environments within which they operate, reflect on how best to use those environments, and can be expected to give a reasoned account of their choices. It doesn't follow, however, that they know all there is to know about the environment or the causes and consequences of their own behaviors. In fact, they are complicit in many unintended consequences of which they are probably ignorant.

The marketplace has far too many offerings for any one person to be perfectly aware of his or her options. It is for this reason that actions can deviate from the rational, utility-maximizing viewer behavior assumed by traditional economic models of program choice. [9]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N109;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N110;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N110;note=ptr)

Rather, agents operate with “bounded rationality.” [10] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N110;note=ptr>] In part, they deal with the impossibility of knowing everything by using habits and routines.

Television viewers confine themselves to idiosyncratic “channel repertoires.” [11]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N111;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N111;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N111;note=ptr)] These are manageable subsets of ten to fifteen familiar channels. Repertoires tend to level off even as the total number of available channels skyrockets. The hyperlinked analog is bookmarks that guide users to previously helpful or interesting sites. Invoking these time-saving habits minimizes “search costs,” but they may cause people to miss content or services that might better gratify their needs and desires. In Simon's term, consumers “satisfice” rather than maximize. [12] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N112;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N112;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N112;note=ptr)] The world of hyperlinking, of course, offers users more powerful tools for finding content, which I address below as an example of duality of structure.

Structures. Structures—or, in some bodies of literature, “institutions”—cover a multitude of macrolevel constructs, including social conventions, language, legal systems, and organizations. Giddens described these as “rules and resources.” [13] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N113;note=ptr>] They can be internalized by an agent or stand apart as external constraints. Either way, they are more durable than agents to the extent that they exist before and after individual actors appear on the scene. I see structures primarily as the resources that people use to enact their media preferences. This includes the technologies that power the hyperlinked environment, as well as the organizations and producers that animate those systems with content and services. For the most part, governments and media industries provide the necessary structures. Of course, they have their own motives for doing so and attempt to manage patterns of attention toward those ends.

The case of user-generated content, so often the topic of conversation at the conference, presents an interesting wrinkle in the neat division between agents and structures. While “distributed construction” is hardly new, [14] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N114;note=ptr>] the hyperlinked environment enables it as never before. Benkler has argued that we are witnessing

the dawn of a “networked information economy,” in which decentralized peer production will shift the balance of power away from established media industries. [15]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N115;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N115;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N115;note=ptr)] An important motivation for this form of production is what he calls the “Joe Einstein” phenomenon, in which people “give away information for free in return for status, benefits to reputation,” and so on. [16] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N116;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N116;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N116;note=ptr)] Surely, contributors to this volume will recognize the syndrome. It’s hard to know what sort of equilibrium will eventually emerge between industrialized and consumer-generated production. But for the purposes of this discussion, the question seems moot.

Whether the producer seeks fame or fortune, the operative strategy is to attract attention by catering to people’s preferences and/or to direct attention by exploiting the structures of the environment. To do this, purveyors constantly monitor the marketplace, judging failures and successes and otherwise looking for opportunities to gain advantage. Suffering from their own form of bounded rationality, they can only respond to what they “see.” The sophistication of their surveillance depends largely on the size and sophistication of the institutions. Typically, the actions of agents are most salient when they are aggregated to form markets, or publics, or audiences. These are what Ettema and Whitney have referred to as “institutionally effective audiences.” [17] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N117;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N117;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N117;note=ptr)] It is this manifestation of agency that most effectively fuels the duality of structure.

Duality of Structure. There is a tendency in many quarters of academe

Page 27

to attribute social behavior almost entirely to purposeful, reasoning agents or, conversely, to macrolevel social structures. This schism is also evident in the literature on media consumption. [18] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N118;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N118;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N118;note=ptr)]

Duality of structure suggests that the two are mutually constituted; [19]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N119;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N119;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N119;note=ptr)] that is, people use structures as vehicles to exercise their agency and, in doing so, reproduce those very structures. This notion can be adapted to explain how the marketplace of attention actually takes shape. In the short term, the structure of the marketplace is relatively “hard” and may limit or direct attention. In the long term, however, it is heavily dependent on the choices of individual media consumers for its very architecture.

The hyperlinked environment is particularly well suited to accomplish this reciprocal act of creation, because of its ability to easily aggregate and make visible the behaviors of many discrete individuals. It creates institutionally effective audiences with a vengeance. Nowhere is this more evident than in the operation of search and recommendation systems, both of which are indispensable tools with which agents address their problems of bounded rationality.

Using a search system is an exercise in finding what you’re looking for. The idea of consulting a

guide to find content is nothing new, as the fortune that endowed our hosts at the Annenberg School will attest. What is new is the way in which modern search engines construct the guide. While algorithms vary, the basic strategy is to sort items in terms of their popularity. Google, for example, ranks Web sites that possess the requisite search terms according to the number and importance of their inbound links. [20] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N120;note=ptr>] Hence, the linking architecture of the Internet, which is itself the product of thousands of more or less independent decision makers, provides the principle guide for navigating hyperlinked space.

Recommendation systems alert us to things we aren't affirmatively looking for. Here again, the basic function is nothing new. Advertising is an old, if transparently self-serving, variation on this genre. The outbound links on Web sites, which are the input for search engines, constitute a decentralized network of recommendations. The tagging, bookmarking, and rating features of social networking programs serve similar functions. The most elaborate recommendation systems, based on collaborative filtering software, aggregate the behaviors of large numbers of anonymous individuals to divine what a person "like you" might prefer. Those of us who use Amazon.com to buy obscure academic books are all too familiar with the seductive power of those systems.

Search and recommendation systems, as well as many other collaborative features of the hyperlinked environment, share a number of noteworthy

Page 28

characteristics. The most elaborate systems are built by amassing people's preferences and behaviors. No one opinion leader or vested interest is able to dictate the output of these systems; hence they have a compelling air of objectivity. In effect, we trust the "wisdom of crowds." [21] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N121;note=ptr>] It is unlikely that individuals in the crowd fully understand how their actions produce a given output, if they are aware of having made any contribution at all. Yet they create, perpetuate, and/or modify structures that direct the attention of others. This duality of structure is an essential and increasingly pervasive dynamic of the marketplace. But what patterns of attention does the marketplace actually produce?

Fragmentation

Certainly, from the perspective of old media, the most consequential and widely noted feature of the new media environment is fragmentation, the tendency of audiences to be widely distributed over the many outlets or items of content competing for their attention. Its conceptual opposite is audience concentration. In the days of old media, public attention was inevitably concentrated on the few stations or newspapers available in local market areas. Since the 1970s, the increased penetration and capacities of cable and satellite systems have caused steady "erosion" in the size of broadcast television audiences. [22] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N122;note=ptr>]

The Internet added even more capacity and global reach, seemingly overnight. With huge national and international markets available, media makers could sustain themselves with niche offerings. The expansive structure of the new environment, populated by institutions

desperately seeking attention, provided the necessary conditions for fragmentation.

Setting aside the economic woes it causes old media, the trend toward ever greater fragmentation (and the consequent demise of “lowest common denominator” programming) has generally been greeted with approval. While the cultural landscape has undoubtedly changed (for reasons I develop shortly), the demise of mass appeal content is, in the words of Mark Twain, “greatly exaggerated.” In his popular book *The Long Tail*, Chris Anderson notes, “The era of one-size-fits-all is ending, and in its place is something new, a market of multitudes.” [23]

[\[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N123;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N123;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N123;note=ptr)

Consumers, empowered by “infinite choice” [24] [\[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N124;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N124;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N124;note=ptr) and

equipped with the tools of search and recommendation, can find whatever suits their preferences, no matter how obscure. For Anderson, this shift manifests itself in a migration from “hits,” which have concentrated attention on the “short head” of a distribution,

Page 29

to niches, which inhabit the increasingly long tail of consumption. Other pundits, noting how the new environment enables various forms of consumer-generated expression, have adopted a similarly celebratory tone. [25] [\[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N125;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N125;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N125;note=ptr) All of

these developments apparently lead to greater diversity of choice in the media environment and, so it would seem, promote a more perfect cultural democracy.

Even if we take these trends at face value, they are not without their worries. Elihu Katz, wistfully recalling the days when one broadcast network served the entire State of Israel, has suggested that the proliferation of new media runs the risk of denying societies a common forum with which to promote national identity and a shared sense of purpose. He has warned:

Throughout the Western world, the newspaper was the first medium of national integration. It was followed by radio. When television came, it displaced radio as the medium of national integration, and radio became the medium of segmentation. Now, following radio again, television has become a medium of segmentation, pushed by both technology and society. Unlike the moment when television assumed radio’s role as the medium of national integration, there is nothing in sight to replace television, not even media events or the Internet. [26]

[\[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N126;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N126;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N126;note=ptr)

Indeed, it’s plausible that fragmentation will make it harder for issues to reach the “threshold of public attention” necessary for agenda setting. [27] [\[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N127;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N127;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N127;note=ptr) Even

more troubling is the prospect of the common public sphere breaking into many tiny

“sphericules” [28] [\[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N128;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N128;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N128;note=ptr) that

fail to interact with one another. The possibility that people will effectively retreat into comfortable little enclaves of like-minded news and entertainment is a topic I address below as the polarization of attention.

Before accepting fragmentation as a fait accompli, however, I think it's worth considering a number of countervailing forces that pull public attention in the opposite direction. While it's fascinating to contemplate the cultural and business implications of long tails, what is even more noteworthy is the persistence of the short head in the distribution. Despite the availability of infinite choice, a relative handful of outlets continue to dominate public attention. Ironically, as we look across media that offer consumers progressively more options, audiences become more, not less, concentrated. Using various measures, researchers consistently find that the most abundant media produce the most concentrated

Page 30

markets. Radio and television, it turns out, are more egalitarian media than the World Wide Web. [29] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N129;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N129;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N129;note=ptr)]

The persistence of short heads undoubtedly has much to do with the operation of “power laws,” which accounts for the success that physicists have had modeling the architecture of the Web. [30] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N130;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N130;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N130;note=ptr)] Such models need not make assumptions about the quality of offerings to produce an expectation of short heads. But quality and social desirability do enter into the allocation of public attention.

One possibility is that the most popular items are, in fact, worthy of that attention. A number of arguments I note shortly suggest just that. Rather than moving consumption in the direction of obscure niches, many new technologies let people spend even more time with what's popular. Early indications on DVR usage, for example, suggest that people typically record top-rated TV programs. [31] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N131;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N131;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N131;note=ptr)] As panelist Jack Wakshlag noted, people use TiVos as “hit machines.” Similarly, many of the most frequently viewed clips on YouTube are the professionally produced work of networks and marketers. [32] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N132;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N132;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N132;note=ptr)] In a world of limited attention, such media use necessarily displaces watching less popular fare. Another irony of moving to an on-demand media environment, then, is that good old-fashioned linear media may have enforced a measure of exposure to things that weren't hits. Even Anderson noted how the move from CDs to iTunes has allowed listeners, “with the help of personalized recommendations,” to cherry-pick the “best individual songs” from albums and skip the “crap” in between. [33] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N133;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N133;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N133;note=ptr)] The best, it would seem, are the most recommended. And the most recommended will inevitably be the most popular.

An accounting of fragmentation is usually made by measuring the attention paid to relatively discrete outlets or items of content. Such numbers are often readily available in the form of audience ratings or paid attendance. Another phenomenon, less easily documented, may further

...the fragmentation of public attention. Some of the most popular...
quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;idno=5680986.0001.001;rgn=d...

mitigate the fragmentation of public attention. Suppose the many nodes across which attention was distributed offered essentially the same thing. There are a number of reasons why the environment might move in that direction. Several observers have noted that consumer-generated production makes liberal use of the most popular (often copyrighted) output of culture industries. [34] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N134;note=ptr>]

If new outlets are simply repurposing existing content and if petty producers are simply playing with the culture's most salient themes and products, fragmentation may be more apparent than real. Phil Napoli, another contributor to this volume, noted at the conference, "I could put ten more water faucets at different places in my home, but ultimately that water is still coming from the same reservoir."

Page 31

More specific mechanisms seem to be at work in the world of news and opinion. Recent studies in the production of online newspapers suggest that the Internet, coupled with competitive journalistic practices, actually contributes to the homogenization of news content. It appears that journalists use the online environment to continuously monitor their competition. Not wanting to be scooped and relying heavily on commonly available wire services and electronic media, newspapers increasingly replicate the same stories. [35] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N135;note=ptr>]

One can imagine a similar dynamic operating in the blogosphere. In fact, Benkler's analysis of how meritorious news and opinion percolate to the A-list blogs seems to be a related phenomenon. For him, this is the mechanism that overcomes the "Babel objection" about the democratizing effects of the Internet. [36] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N136;note=ptr>] It does suggest, however, that public attention is not as fragmented as it might at first seem.

Polarization

What is sometimes harder to see is the extent to which attention is being polarized. Unlike our view of fragmentation, which comes in the form of a snapshot showing the distribution of attendance across sources, polarization requires a consideration of which media people consume over time. Fragmentation provides evidence that public attention is, in the aggregate, spread across many more sources than was the case a decade ago. This is what Napoli has called "horizontal exposure diversity." [37] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N137;note=ptr>]

It does not follow, however, that each individual's diet of media content is also more diverse. It might be that people avail themselves of the abundance by sampling a little bit of everything. That would be evidence of a "vertical diversity" of exposure and would, by most accounts, be a socially desirable pattern. [38] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N138;note=ptr>]

Alternatively, it could be that people use the environment to binge on a few favorites. Either pattern could lie beneath the veneer of fragmentation. [39]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N139;note=ptr>]

The latter, however, has potentially chilling social implications, since it suggests that people withdraw

into various "echo chambers." [39] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N139;note=ptr>]

into various cocoons. [40] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N140;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N140;note=ptr>] Two factors will determine the outcome: the psychological predispositions of agents and the structures of the environment.

There is certainly reason to believe that rational, utility-maximizing consumers will selectively choose media materials that conform to their preferences. Traditional economic models of program choice assumed fixed preferences that were systematically related to viewer-defined program types. [41] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N141;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N141;note=ptr>] While it seems likely that preferences are, in the long term,

Page 32

cultivated by the environment, people do have relatively stable likes and dislikes. These operate along many dimensions, including (1) an appetite for specific program genres, perhaps as broadly defined as information versus entertainment; [42]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N142;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N142;note=ptr>] (2) the utility of information; [43] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N143;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N143;note=ptr>] (3) language or cultural proximity; [44] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N144;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N144;note=ptr>] (4) conservative versus liberal political ideologies; [45] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N145;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N145;note=ptr>] and (5) various manifestations of fandom. [46] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N146;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N146;note=ptr>] In short, much of what we know about the psychology of media choice suggests that people will consume relatively restricted diets to suit their tastes. With virtually every type of content available in limitless supply, it remains to be seen when or if people become satiated.

The media environment does more than simply offer an abundance of choice, however. It structures and filters what is available and, in so doing, privileges some things over others. While search and recommendation systems are apparently objective aggregations of many independent decisions, they may exacerbate the tendency of people to retreat into comfortable enclaves of like-minded speech. Cass Sunstein, among others, fears that these filtering technologies encourage people to seek out what is agreeable and to avoid anything that challenges their predispositions. Over the long haul, he writes, this is likely to promote “group polarization.” [47] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N147;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N147;note=ptr>] But even if one assumes that search engines simply do our bidding, not all structural features of the hyperlinked environment are so benign. The media institutions that, in large part, create the environment will attempt to manage and concentrate our attention with all the means at their disposal. Joseph Turow asks:

Who will create opportunities for various social groups to talk across divisions and

shape experiences when major marketing and media firms solidify social divisions

The Hyperlinked Society: Questioning Connections in the Digital Age / Josep...
 share experiences when major marketing and media firms soundly social division
 by separating people into data-driven niches with news and entertainment aimed
 primarily at reinforcing their sense of selves? [48]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N148;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N148;note=ptr)

Of course, not all writers have concluded that countless niches are a bad thing. With his characteristic enthusiasm for infinite choice, Anderson has imagined a “massively parallel culture” formed into “millions of microcultures” or “tribal eddies.” [49]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N149;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N149;note=ptr) Good or bad, it’s worth developing a better understanding of how, if at all, public attention is being polarized.

As best I can tell, the jury is still out. The clearest evidence so far is that the new marketplace allows a substantial segment of the population to avoid news and information altogether. Increasingly, we are becoming a nation of people who do or do not know about world events. While the

Page 33

old world of linear media succeeded in enforcing almost universal exposure to TV news, the new world of choice does not. Markus Prior has argued convincingly that changes in the structure of television have enabled differential patterns of news viewing, which effectively polarize the public into those with and those without political knowledge. [50]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N150;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N150;note=ptr) While a case can be made that people who avoid hard news are “rationally ignorant,” [51]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N151;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N151;note=ptr) I find Prior’s results a troubling prospect for democracy. What is less clear, though, is the extent to which consumers of news and information limit their diets largely to one ideological point of view. Iyengar and Morin’s study of online news readership [52]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N152;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N152;note=ptr) and Adamic and Glance’s analysis of the linking structures among political blogs [53]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N153;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N153;note=ptr) suggest systematic “blue/red” biases in people’s patterns of consumption across time. Conversely, Hargittai, Gallo and Kane’s study of political blogs [54]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N154;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N154;note=ptr) and a Pew study of Internet use [55] [\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N155;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N155;note=ptr)

emphasize the tendencies of people to reference and/or know opposing points of view. Prior

himself has noted that viewers of the Fox News Channel see other sources of TV news, which

munson has noted that viewers of the Fox News Channel see other sources of TV news, which suggests a vertical diversity of exposure consistent with broader findings on TV viewing. [56]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N156;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N156;note=ptr)]

Questions about the Marketplace

The shape of public attention is important because it suggests how the marketplace of ideas will operate in an age of on-demand digital media. Two questions are, to me, particularly salient. The first is whether our society's cultural center will hold in the wake of all these changes. This is and should be subject to ongoing empirical investigation. The second normative question addresses the wisdom of the data-aggregating systems that increasingly define the character of the nonlinear media environment.

Will the Center Hold? In his zeal for the long tail, Anderson has asserted that “infinite choice equals ultimate fragmentation.” [57] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N157;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N157;note=ptr)]

It's hard to imagine a common culture—let alone a vibrant democracy—whose patterns of attention are evenly spread across an infinite number of choices. Nor do I think that's likely to happen. I suspect the forces that concentrate attention, as already outlined, will save us from flying off in every conceivable direction. That said, it's clear that many more things are competing for our attention. These inevitably come at the expense of the older forms of media that once commanded center stage. They were sometimes derided as offering only the lowest common denominator, but

Page 34

by their very commercial nature, they steered a course through the heart of culture. And for all practical purposes, attendance was mandatory.

Now attendance is up to us. A veneer of fragmentation does not mandate social polarization if individual media consumers do the work necessary to “connect the dots.” They might spend time moving from the obscure to the popular or simply from niche to niche and still manage to construct a fully featured marketplace of ideas. But the very concept of a niche suggests a degree of stickiness. Every niche maker, commercial or not, wants repeat customers. Most would be happy if those customers settled exclusively on what they had to offer. Many will, undoubtedly, do what they can to make that happen. If customers are happy with their niches, they'll stay put. It's only rational.

For all those reasons, public attention is likely to be reorganized along dimensions of taste and structure. For now, we should do our best to monitor the social and cultural divisions that emerge. That will be a daunting task for two reasons. First, a complete view of how people navigate the marketplace will require following them across time and across media. The world of media research is still largely balkanized by medium. [58]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N158;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N158;note=ptr)] Today,

it's virtually impossible to know with any precision what a person reads, watches on television, hears on the radio, and consumes on the Internet. Yet all those sources are competing for attention and, in turn, shape that person's environment. As media converge on a common

distribution system, it will become easier to paint a complete picture of consumption. Second

distribution system, it will become easier to paint a complete picture of consumption. Second, assessing exposure alone will not fully answer the question. We must also have a nuanced understanding of the content that's being consumed (e.g., how links are referenced, how issues are framed) and what sense people make of those representations. Only then will we see what fault lines are forming within the culture.

Are Crowds Wise or Stupid? While hyperlinking is, on one level, about technology, the hyperlinked spaces that we use are given life by ordinary human beings. Sometimes it is the work of individual agents, but often it happens through the instant and ever-changing aggregation of choices made by others. This is true of the most powerful tools we use to navigate the environment, and it goes to the heart of what many commentators find so revolutionary about the technology. [59] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N159;note=ptr>] It is hard to imagine an arena of human activity that is more heavily dependent on the wisdom or stupidity of crowds. And it is on this point that many social commentators strike me as a bit schizophrenic.

The phrase “wisdom of crowds” was popularized by James Surowiecki. [60]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N160;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N160;note=ptr>] In his eponymous best seller, Surowiecki argued that averaging input from many ordinary, diverse, and independent decision makers

Page 35

often produces better results than the judgments of experts. A marketplace offers one example of such a mechanism. Anderson frequently repeats the “wisdom of crowds” mantra, pointing to any number of apparently successful collaborative endeavors, from Wikipedia to various forms of recommendation. [61] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N161;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N161;note=ptr>

Benkler and many others also put stock in the ability of the blogosphere to sort through and collaboratively produce the most accurate news or meritorious ideas. [62]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N162;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N162;note=ptr>

When I first read the conditions that Surowiecki suggested will unleash the wisdom of crowds, I was reminded of Blumer's classic definition of a “mass” in social theory and of its adaptation to audiences. [63] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N163;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N163;note=ptr>] A

mass audience is a heterogeneous collection of many anonymous, independent decision makers. [64] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N164;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N164;note=ptr>] The

wisdom of such crowds is routinely measured in audience ratings. Recently, the top-rated program on TV was *Dancing with the Stars*. I suspect you could find similarly reassuring gems if you checked the most viewed clips on YouTube. Anderson tries to finesse such uninspiring measures of public taste by insisting that “what matters is the rankings *within* a genre (or subgenre), not *across* genres.” [65] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N165;note=ptr)

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N165;note=ptr>

Apparently, it's only when we dig deeply into our niches that crowds become wise. To me, this reads like an invitation to cultural polarization. If we want to encourage sampling the best across genres, why is it that crowds should no longer be our guide? At what point do they become stupid? At the very least, we need to develop a more discriminating stance on the wisdom of crowds.

But, like it or not, crowds increasingly shape our world. The actions of agents are instantly aggregated and made available for all to see. These, in turn, affect the structures and offerings of the media environment. Online newspapers are discovering that it's "soft news" (e.g., stories about celebrities, sex, and animals) that attracts readers' attention. A recent piece in the *American Journalism Review* warned print journalists, "Television news veterans predict papers will face huge challenges maintaining their editorial independence while seeking to grab the attention of Web readers." [66] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N166;note=ptr>] In matters of taste, no empirical test will tell us whether the decisions of crowds are wise or not. More realistically, it will be for each of us to judge whether the results of the process offer a path to enlightenment or the road to perdition.

Notes

1. H. Jenkins, *Convergence Culture: Where Old and New Media Collide* (New York: New York University Press, 2006).[⌘] [#]

Page 36

2. P. Lyman and H. R. Varian, "How Much Information?" <http://www.sims.berkeley.edu/how-much-info-2003> [<http://www.sims.berkeley.edu/how-much-info-2003>] (accessed September 14, 2006).[⌘] [#]

3. H. A. Simon, "Designing Organizations for an Information-Rich World," in *Computers, Communications, and the Public Interest*, ed. M. Greenberger (Baltimore: Johns Hopkins University Press, 1971), 41.[⌘] [#]

4. R. Lanham, *The Economics of Attention: Style and Substance in the Age of Information* (Chicago: University of Chicago Press, 2006), 46.[⌘] [#]

5. A. Giddens, *The Constitution of Society: Outline of the Theory of Structuration* (Berkeley: University of California Press, 1984).[⌘] [#]

6. G. DeSanctis and M. S. Poole, "Capturing the Complexity in Advanced Technology Use: Adaptive Structuration Theory," *Organization Science* 5, no. 2 (1994): 121–47; W. J. Orlikowski, "The Duality of Technology: Rethinking the Concept of Technology in Organizations," *Organization Science* 3, no. 3 (1992): 398–427.[⌘] [#]

7. Giddens, *Constitution of Society*.[⌘] [#]

8. N. Mouzelis, "The Subjectivist-Objectivist Divide: Against Transcendence," *Sociology* 34, no. 4 (2000): 741–62.[⌘] [#]

9. E.g., B. M. Owen and S. S. Wildman, *Video Economics* (Cambridge, MA: Harvard University Press, 1992).[⌘] [#]

10. H. A. Simon. *Administrative Behavior: A Study of Decision-Making Processes in*

Administrative Organizations (New York: Free Press, 1997).[\[#\]](#)

11. K. A. Neuendorf, D. J. Atkin, and L. W. Jeffres, "Reconceptualizing Channel Repertoire in the Urban Cable Environment," *Journal of Broadcasting and Electronic Media* 45, no. 3 (2001): 464–82; E. J. Yuan and J. G. Webster, "Channel Repertoires: Using Peoplemeter Data in Beijing," *Journal of Broadcasting and Electronic Media* 50, no. 3 (2006): 524–36.[\[#\]](#)

12. Simon, *Administrative Behavior*.[\[#\]](#)

13. Giddens, *Constitution of Society*.[\[#\]](#)

14. P. J. Boczkowski, *Digitizing the News: Innovation in Online Newspapers* (Cambridge, MA: MIT Press, 2004).[\[#\]](#)

15. Y. Benkler, *The Wealth of Networks* (New Haven, CT: Yale University Press, 2006).[\[#\]](#)

16. Ibid., 43.[\[#\]](#)

17. J. S. Ettema and D. C. Whitney, "The Money Arrow: An Introduction to Audiemcemaking," in *Audiemcemaking: How the Media Create the Audience* (Thousand Oaks, CA: Sage, 1994), 5.[\[#\]](#)

18. J. G. Webster, "The Role of Structure in Media Choice" (forthcoming); J. G. Webster and P. F. Phalen, *The Mass Audience: Rediscovering the Dominant Model* (Mahwah, NJ: Erlbaum, 1997).[\[#\]](#)

19. Giddens, *Constitution of Society*.[\[#\]](#)

20. J. Battelle, *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture* (New York: Portfolio, 2005).[\[#\]](#)

21. J. Surowiecki, *The Wisdom of Crowds: Why the Many Are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations* (New York: Doubleday, 2004).[\[#\]](#)

22. J. G. Webster, "Beneath the Veneer of Fragmentation: Television Audience Polarization in a Multichannel World," *Journal of Communication* 55, no. 2 (2005): 366–82.[\[#\]](#)

Page 37

23. C. Anderson, *The Long Tail: Why the Future of Business Is Selling Less of More* (New York: Hyperion, 2006), 5.[\[#\]](#)

24. Ibid., 12.[\[#\]](#)

25. E.g., Benkler, *Wealth of Networks*; Jenkins, *Convergence Culture*.[\[#\]](#)

26. E. Katz, "And Deliver Us from Segmentation," *Annals of the American Academy of Political and Social Science* 546, no. 1 (1996): 33.[\[#\]](#)

27. W. R. Neuman, "The Threshold of Public Attention," *Public Opinion Quarterly* 54, no. 2 (1990): 159–76.[\[#\]](#)

28. T. Gitlin, "Public Sphere or Public Sphericules?" in *Media, Ritual, and Identity*, ed. T. Liebes and J. Curran (London: Routledge, 1998), 175–202.[\[#\]](#)

29. E.g., M. Hindman. "A Mile Wide and an Inch Deep: Measuring Media Diversity Online and

Offline,” and J. G. Webster, “Diversity of Exposure,” in *Media Diversity and Localism: Meaning and Metrics*, ed. P. Napoli (Mahwah, NJ: Erlbaum, 2006), 327–47, 309–25; J. Yim, “Audience Concentration in the Media: Cross-Media Comparisons and the Introduction of the Uncertainty Measure,” *Communication Monographs* 70, no. 2 (2003): 114–28. [#]

30. E.g., A.-L. Barabási, “The Physics of the Web,” *Physics World*, July 2001, <http://physicsweb.org/articles/world/14/7/9> [<http://physicsweb.org/articles/world/14/7/9>] (accessed July 10, 2006); B. A. Huberman, *The Laws of the Web: Patterns in the Ecology of Information* (Cambridge, MA: MIT Press, 2001). [#]

31. F. Aherns, “Fears over TiVo on Pause,” *Los Angeles Times*, August 29, 2006, <http://www.latimes.com/entertainment/news/homeentertainment/la-et-tivo29aug29,1,7984227.story?coll=la-entnews-homeent> [<http://www.latimes.com/entertainment/news/homeentertainment/la-et-tivo29aug29,1,7984227.story?coll=la-entnews-homeent>] (accessed September 27, 2006). [#]

32. W. Friedman, “CBS Scores Viewers with YouTube Alliance,” *Media Post*, November 27, 2006, <http://publications.mediapost.com/index.cfm?fuseaction=Articles.san&s=51543&Nid=25374&p=263743> [<http://publications.mediapost.com/index.cfm?fuseaction=Articles.san&s=51543&Nid=25374&p=263743>] (accessed November 27, 2006). [#]

33. Anderson, *Long Tail*, 22. [#]

34. Benkler, *Wealth of Networks*; Jenkins, *Convergence Culture*; L. Lessig, *Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity* (New York: Penguin, 2004). [#]

35. P. J. Boczkowski and M. de Santos, “When More Media Equals Less News: Patterns of Content Homogenization in Argentina’s Leading Print and Online Newspapers,” *Political Communication* 24, no. 2 (April 2007): 167–80. [#]

36. Benkler, *Wealth of Networks*, 233. [#]

37. P. M. Napoli, “Deconstructing the Diversity Principle,” *Journal of Communication* 49, no. 4 (1999): 7–34. [#]

38. Webster, “Diversity of Exposure,” 309–25. [#]

39. Webster, “Beneath the Veneer of Fragmentation,” 366–82. [#]

40. C. R. Sunstein, *Infotopia* (Oxford: Oxford University Press, 2006). [#]

41. Owen and Wildman, *Video Economics*. [#]

42. E.g., M. Prior, *Post-Broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections* (New York: Cambridge University Press, 2007); R. T. Rust, W. A. Kamakura, and M. I. Alpert, “Viewer Preference Segmentation and Viewing Choice Models for Network Television,” *Journal of Advertising* 21, no. 1 (1992): 1–18. [#]

43. E.g., M. S. Y. Chwe, *Rational Ritual: Culture, Coordination, and Common Knowledge* (Princeton, NJ: Princeton University Press, 2001); J. T. Hamilton, *All the News That’s Fit to Sell: How the Market Transforms Information into News* (Princeton, NJ: Princeton University Press, 2004). [#]

44. E.g., J. Straubhaar, "Choosing National TV: Cultural Capital, Language, and Cultural Proximity in Brazil," in *The Impact of International Television: A Paradigm Shift*, ed. M. G. Elasmr (Mahwah, NJ: Erlbaum, 2003), 77–110. [\[#\]](#)
45. E.g., S. Iyengar and R. Morin, "Red Media, Blue Media: Evidence for a Political Litmus Test in Online News Readership," *Washington Post*, May 3, 2006, <http://www.washingtonpost.com/wp-dyn/content/article/2006/05/03/AR2006050300865.html> [<http://www.washingtonpost.com/wp-dyn/content/article/2006/05/03/AR2006050300865.html>] (accessed November 30, 2006); C. R. Sunstein, *Republic.com* (Princeton, NJ: Princeton University Press, 2001). [\[#\]](#)
46. E.g., Jenkins, *Convergence Culture*. [\[#\]](#)
47. Sunstein, *Republic.com*. [\[#\]](#)
48. J. Turow, *Niche Envy: Marketing Discrimination in the Digital Age* (Cambridge, MA: MIT Press, 2006), 33. [\[#\]](#)
49. Anderson, *Long Tail*, 182. [\[#\]](#)
50. Prior, *Post-Broadcast Democracy*. [\[#\]](#)
51. Hamilton, *All the News That's Fit to Sell*. [\[#\]](#)
52. Iyengar and Morin, "Red Media, Blue Media." [\[#\]](#)
53. L. A. Adamic and N. Glance, "The Political Blogosphere and the 2004 U.S. Election: Divided They Blog," in *Proceedings of the 3rd International Workshop on Link Discovery* (New York: ACM, 2005), 36–43. [\[#\]](#)
54. E. Hargittai, J. Gallo, and M. Y. Kane, "Cross-Ideological Discussions among a Group of Conservative and Liberal Bloggers," unpublished ms. [\[#\]](#)
55. J. Horrigan, K. Garrett, and P. Resnick, *The Internet and Democratic Debate* (Washington, DC: Pew Internet and American Life Project, 2004). [\[#\]](#)
56. Webster, "Diversity of Exposure," 309–25. [\[#\]](#)
57. Anderson, *Long Tail*, 181. [\[#\]](#)
58. J. G. Webster, P. F. Phalen, and L. W. Lichty, *Ratings Analysis: The Theory and Practice of Audience Research*, 3rd ed. (Mahwah, NJ: Erlbaum, 2006). [\[#\]](#)
59. E.g., Benkler, *Wealth of Networks*; Sunstein, *Infotopia*. [\[#\]](#)
60. Surowiecki, *Wisdom of Crowds*. [\[#\]](#)
61. Anderson, *Long Tail*. [\[#\]](#)
62. Benkler, *Wealth of Networks*. [\[#\]](#)
63. H. Blumer, "The Field of Collective Behavior," in *New Outlines of the Principles of Sociology*, ed. A. M. Lee (New York: Barnes and Noble, 1946), 167–222. [\[#\]](#)
64. Webster and Phalen, *Mass Audience*. [\[#\]](#)
65. Anderson, *Long Tail*, 114. [\[#\]](#)
66. R. Shiver, "By the Numbers," *American Journalism Review*, June–July 2006.

http://www.ajr.org/article_printable.asp?id=4121 [http://www.ajr.org/article_printable.asp?id=4121] (accessed November 11, 2006).¹ [#]

Page 39

The Hyperlink as Organizing Principle

Alexander Halavais

The Hyperlink

What does a hyperlink mean? The question itself is problematical. We might be satisfied with the simpler and related question of what a hyperlink is and what a hyperlink does. But in trying to understand what the larger social effects of hyperlink networks are, it is not enough to be able to define a hyperlink, we need to understand its nature, its use, and its social effects.

This meaning is neither unitary nor stable. There are a number of ways the hyperlink could be theorized, at different levels and toward different ends. This essay will argue that with the explosion of the World Wide Web, we are beginning to see increasing awareness of hyperlink networks as meaningful, malleable, and powerful. This is in contrast to initial views of hyperlinks, which only barely glimpsed the degree to which they are able to express meaning at a conscious and intentional level. With the ability to see hyperlinks within a larger networked structure, we have already begun to understand that, en masse, they reflect deep social and cultural structures—a kind of collective unconscious. That understanding has in turn changed the ways in which hyperlinks are used and exploited. As hyperlink networks become more easily manipulated and reach farther into our social and physical lives, we will have a continuing need to understand the hyperlink as more than a way of automatically requesting documents from a Web server. As these networks increasingly represent the structures of knowledge and social interaction, they acquire the ability to influence themselves and attain some form of self-awareness.

Hyperlinks as Citations

The idea of a hyperlink—a reference that automatically brings the user to a particular point in a cited work—is deceptively simple. Those who first

Page 40

implemented hyperlinks were sometimes blind to their wider implications. There is, however, a relationship between the traditional uses of citation and the development of the hyperlink. If we wish to understand what individuals mean when they use a hyperlink, it is worthwhile to understand what they mean when they cite.

Identifying the first use of a hyperlink is difficult in part because the concept of a hyperlink is simple enough to be applicable to a wide range of technologies. If it is merely an automatic citation device and not limited to any particular medium, we might suggest a number of precursors to what we traditionally think of as a hyperlink. Indeed, it could be argued that the hyperlink is in a continual process of reinvention, or what Neville Holmes refers to as

“cumulative innovation.”^[1] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N301;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N301;note=ptr) Any claim to be the inventor of the hyperlink, as in the case of British Telecom's short-lived attempt to assert a patent on the concept, quickly becomes mired in a long history of scholarly citation and other forms of linking.

So where does the history of hyperlinking begin? If we do not stipulate that the document be digital in format, quotation and citation do certainly appear to be forms of textual links. Commentaries on religious texts, especially within Talmudic scholarship, are often seen as the earliest exemplars of citation, in part because the commentators could generally rely on standardized copies of the texts in question. Even when deviating from standard copies could have dire effects, hand-copying virtually guaranteed that fidelity would be difficult to maintain. The emergence of the printing press—and with it, standardized, distributed libraries—provided a fertile platform for the practice of citation. As Elizabeth Eisenstein has argued, this standardization allowed for new forms of cataloging and indexing, which led to increased citation, which in turn allowed for the distribution of scholarship and eventually the Enlightenment. [2]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N302;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N302;note=ptr)

But the function of citation is not as obvious as it may appear. The most obvious function is as a way of presenting others' ideas as support for the author's own argument—that is, it allows the use of “sources.” In some ways, this might be seen as a way of allowing the author, as Eisenstein suggests, to assume access to a generalized pool of authoritative texts and avoid recapitulating the development of an entire field. In this way, the reference might also serve a pedagogical role, pointing the reader to helpful resources. Because teaching and persuasion go hand in hand, the pedagogical function bleeds easily into a persuasive one, as Gilbert suggests. [3]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N303;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N303;note=ptr)

However, as Collingwood argues, the ability to interrogate previous authors—presenting their ideas in part rather than as a whole—by “dissecting

a tradition” also allows for the undermining of authoritative sources, even as it reinscribes them. [4] [\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N304;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N304;note=ptr) The motivation for linking to previous work is often to criticize, analyze, or refute that work, as well as to build on it. Polanyi argues that the development of modern scholarship, in contrast to the motto of the Royal Society (*Nullius in verba*), is not original thought alone but engaging in a distributed conversation facilitated through the use of technologies of indexing and referencing. [5] [\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N305;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N305;note=ptr)

Without effective citation, scholarly thought would have remained a relatively solitary endeavor rather than a textual conversation.

To be sure, it is a strange sort of conversation. Early citation was more likely to feel like a chat with the dead, and links to authors no longer among the living were predictably unlikely to be

reciprocated. More recent journal publication may engage in something more interactive, though

generally only linking backward in time. As a conversation, it is thus a strangely disjointed and asynchronous one. Taking a cue from Walther's "hyperpersonal" relationships, we might consider citation to be a sort of "hyperconversation," in that it occurs across contexts and across time. [6] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N306;note=ptr>]

At the very least, a reference may be a nod of thanks that acknowledges the efforts of others or a more direct demonstration of gratitude. But the social affinity may also be stronger than just a thank-you. Indeed, in an essay subtitled "Scholarly Citation Practices as Courtship Rituals," Rose explicitly emphasizes the sociable nature of citation and indicates that creating a citation is as much about entering a discourse community as it is about establishing an authoritative basis for an author's argument. [7] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N307;note=ptr>]

None of these possible motivations is obviously dominant. As critics of citation analysis have suggested, citations are created for a wide variety of reasons. Brooks identifies a set of motivations—persuasiveness, positive credit, currency, reader alert, operational information, social consensus, and negative credit—and others have created similar taxonomies. [8]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N308;note=ptr>]

Many agree that citation varies by time, by culture, and by personal style. [9]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N309;note=ptr>]

The relative dearth of empirical data and the variability of motivations and practices that characterize citation make it difficult to ascribe a simple, precise meaning to citation practices, though it seems clear that they have played an important role in the development of social knowledge. [10] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N310;note=ptr>]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N311;note=ptr>]

Vannevar Bush was aware of the importance of citation and the need for automating it, when he wrote his prescient "As We May Think" in 1945. [11] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N311;note=ptr>] While it is possible to identify precursors to hyperlinks as automated citations in the use of punch cards or tabs, the visionary potential of hyperlinking is nowhere as clear as it is in Bush's imagined personal file

system, the *memex*. Bush suggested that since the mind was organized as an associative network, a similar organizational structure would be an effective way of creating personal files and libraries as well. Moreover, the "associative trails" produced by a researcher as he or she examined the literature would provide pathways that others could follow. In this way, the scientific enterprise could be accelerated, and collaborative knowledge could be improved.

Bush's ideas helped to inspire the work of Ted Nelson and Douglas Engelbart, who provided the earliest versions of what we recognize as the familiar hyperlink. Nelson's Project Xanadu set out to create a new, broadly associative way of organizing knowledge, and in so doing, he coined the terms *hyperlink* and *hypertext*. The first working computer-based hyperlink system was

demonstrated as part of the oNLine System (NLS) by Engelbart in 1968. [12]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N312;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N312;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N312;note=ptr)] Nelson has expressed his disappointment in the limits of the hyperlink as it ended up being used: it is generally unidirectional, for example, and unable to reflect the richness of associative thought. [13] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N313;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N313;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N313;note=ptr)]

Nonetheless, the demonstration of the utility of a hyperlink led to a number of hypertext systems, culminating in the popular HyperCard application by the 1980s.

Throughout these early implementations, there was a clear conceptual relationship to previous uses of citation. As hypertext systems advanced however, differences became clear. Although hyperlinks may perform the functions of a scholarly reference, they often function in ways that references cannot, and they are often used for different reasons. For example, because electronic documents are more easily updated, it is possible to have two documents with hyperlinks pointing to each other, something that generally does not occur in printed literature. [14] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N314;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N314;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N314;note=ptr)]

Because of the instantaneous nature of hyperlinks, it was also clear that they could do much more than static references. Unlike a traditional citation, which requires an investment of time to locate and read the target document, hyperlinks allow for the instant “jump” to other texts. This immediacy allowed hyperlinks to be used to more directly structure documents, collections of documents, and—as the World Wide Web rapidly expanded—recorded media more broadly.

The emergence and popularity of the World Wide Web moved hyperlinks farther from citation, and the terminology changed to reflect that of the memex. Reading became “browsing” or “surfing,” and the user was transported along the pathways generated by large collections of hyperlinks. Those hyperlinks were created not by a single designer but by millions of authors linking to one another’s work. As database systems and programmatic interfaces became the norm, not only were a greater number

Page 43

of people able to create hyperlinks (e.g., with the increasing popularity of wikis), but the hyperlinks were often created on the fly by the servers themselves.

As a result of the increasing popularity of the hyperlink, their uses and significance have expanded and changed. Even within the world of academic linking, the role of the link has gone beyond citation to focus more on navigational issues. [15]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N315;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N315;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N315;note=ptr)]

Indeed, online scholarship tends to retain the traditional text-based citation, even on pages that are replete with hyperlinks.

Outside of the scholarly world, hyperlinks have taken on an even greater role. Clicking a hyperlink may lead to a camera changing its orientation, to a book being ordered and sent through the mail, to an e-mail in-box being reorganized, or to a closer view of a satellite image.

These potential uses were not outside of the expectations of some of the originators of the

hyperlink. In 1968, Engelbart was already integrating e-mail functions with hyperlinks, and in 1965, Nelson wrote that the “ramifications of this approach extend well beyond its original concerns, into such places as information retrieval and library science, motion pictures and the programming craft; for it is almost everywhere necessary to deal with deep structural changes in the arrangement of ideas and things.” [16] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N316;note=ptr>]

The universal nature of hyperlinking makes it a very difficult sort of artifact to understand. The question of what someone means when they create a hyperlink or when they activate one is entirely determined by the context of the hyperlink’s use. While there have been some initial attempts, at least within scholarly sections of the Web, to discover the conditions under which hyperlinks come into being, these remain at an early stage. It seems that hyperlinks clearly hold some social meaning, but beyond this broad implication, it is difficult to characterize a single hyperlink in any rigorous way. [17] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N317;note=ptr>]

Hyperlink Networks

Why should we be interested in the nature of individual hyperlinks? If hypertext is structured by hyperlinks, understanding the psychology behind those connections is valuable. With that understanding, we would be better able to comprehend what individual hyperlinks indicate. Even more important, those hyperlinks seem to provide an opportunity to understand social behavior when taken in the aggregate.

Focusing on the structural properties of hyperlinks has been particularly

Page 44

important for Web search technologies, especially for Google. By measuring which pages are most central to the network of hyperlinks on the Web at large, Google is able to rank its search results according to some indication of salience. [18] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N318;note=ptr>] Other systems have followed suit, collating links to provide a guide to the most popular sites. Technorati, for example focuses on the network of blogs, providing indications of which blogs garner the largest number of inbound hyperlinks from their peers.

In some ways, the effectiveness of these approaches is sufficient justification of their use. While using measures of network prestige is certainly not the only reason for Google’s success, there can be little doubt that it has been effective. The reason for this effectiveness does not, at the surface, appear to be surprising. Measures of prestige within social networks have a long history, and the conceptual relationship between the structure of these networks and their behavior has been well considered. [19] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N319;note=ptr>] When such measures were used, it was often based on information gathered directly from individuals regarding their behaviors or their attitudes. While there is certainly room for systematic error in such a process, there was generally a clear connection between the data collected and the inferences made.

Likewise, the measurement and tracking of citation networks in order to map a field and its

development has a long and successful history. The relationship between citation networks and hyperlink networks is clearer: citations, like hyperlinks, represent a latent, unobtrusive measure. An advantage is that respondents are not shaping answers to fit the preconceptions of the researchers. However, researchers are left with interpreting the nature of the citations themselves. In most cases, the citations are taken as a whole and considered to be links—either present or not. As suggested earlier, references may be made for a wide range of reasons, including to signal that a work is faulty or lacking in some way. Given the range of meanings that might attach to any citation, can we make sense of measures taken from the network of those citations?

The question becomes even more pressing when it comes to the interpretation of hyperlink networks. Park and Thelwall detail the increasing use of hyperlink network analysis to help understand the structure of everything from debate over public policy to e-commerce. [20]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N320;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N320;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N320;note=ptr)] While researchers have engaged the question of interpretation of hyperlinks to varying degrees, such investigations have remained tentative and generally have investigated linking within scholarly domains. This has not slowed the use of hyperlink analysis, however. In its most extreme form—studying only the hyperlinks themselves—such analysis approaches the

Page 45

purely theoretical. [21] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N321;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N321;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N321;note=ptr)] When combined with the text of the target pages, their geographical location, or other data, hyperlink analyses provide what appears to be useful information about structural relationships.

As an example, if we measure the hyperlinks stretching between cities, does this tell us something valuable? I measured links between sampled Web sites in eight world cities to determine the degree to which they interlinked. [22] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N322;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N322;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N322;note=ptr)] The resulting network of hyperlinks (fig. 1) provided interesting insights. When the first study was conducted, New York was considered a bit of a latecomer to the information revolution, with attention focused heavily on California. Instead, it was by far the most central of the global cities studied, while Tokyo remained relatively isolated, neither linking to nor linked from other global cities. Lin, Zhang, and I took a similar approach to measure the links among blogs in various U.S. cities (fig. 2). [23] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N323;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N323;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N323;note=ptr)]

Tracing the links among several hundred thousand blogs, we found clusters among cities that had similar characteristics, and we found that New York was once again at the core of the network.

While structure is revealed, what do these links really mean? Why are these blogs linked at all? Individuals linking between blogs were in some cases expressing a social connection (i.e., they knew the person in “real life”) and in others pointing to a blog that contained an interesting idea. But both within the world of blogging and more generally on the Web, links can serve a very wide range of social and technical functions. A link may represent an advertisement determined by a third party and intended to entice a customer (e.g., Google AdWords), a source of further

discussions on the same topic in other blogs, a path for connecting via e-mail or voice, a way of demonstrating ownership, a link to other sites authored by the same individual, a link to a group (blog ring, church group, collective) to demonstrate membership, or any number of other things. These myriad meanings are all tied up in the code of a hyperlink, and it may not be immediately obvious how these are related—except, of course, that they represent a pathway between pages. Are the results, which suggest affinities or proximities between cities, to be believed? What do these characteristics mean in terms of “real” social structure?

The question is a vexing one, since the reason for performing such an analysis is to reveal latent structures that are not already obvious to an observer. In the case of the observations here, the network was similar to other geographically distributed communications networks, from telephones to package deliveries. These similarities and other similarities suggest that something structural is happening and provides a way forward. I have elsewhere suggested that the process is like inferring social

Page 46

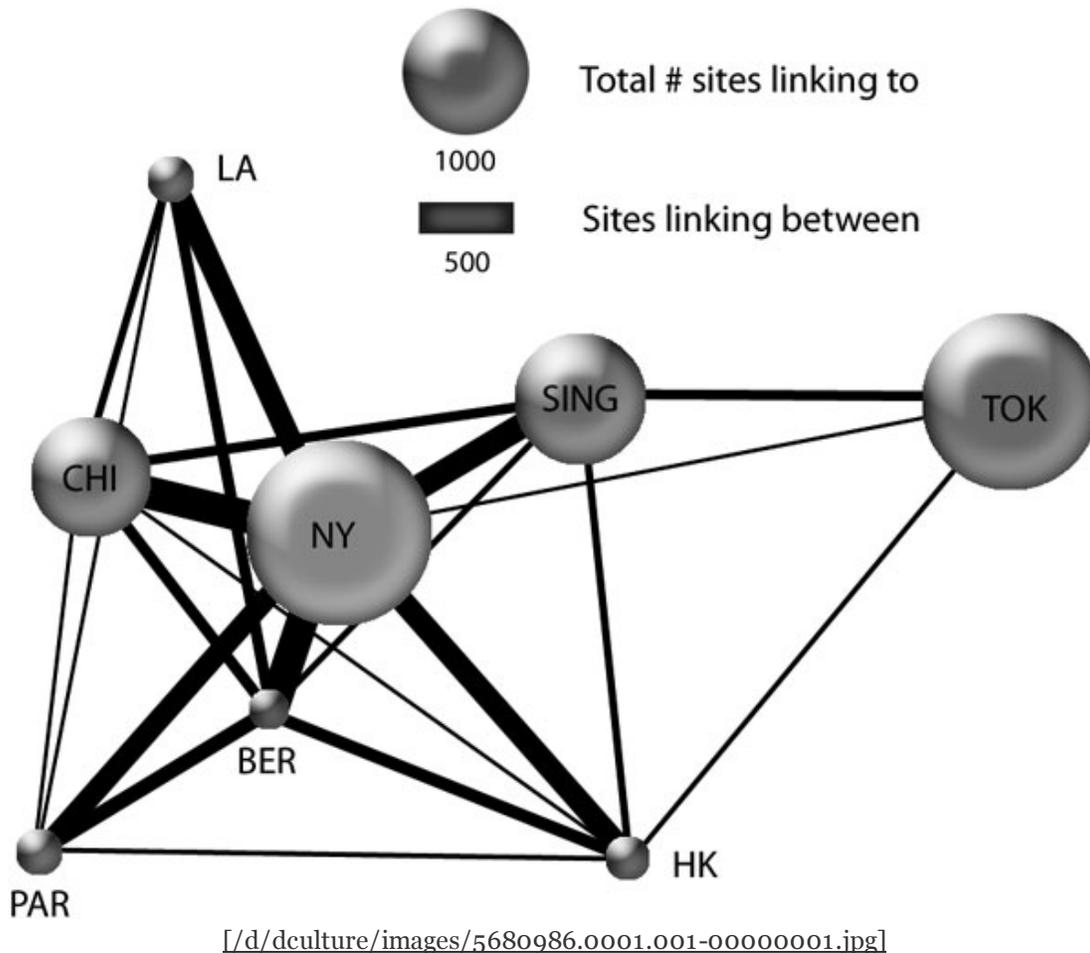


Fig. 1.

Links among eight world cities: Berlin (BER), Chicago (CHI), Hong Kong (HK), Los Angeles (LA), New York (NY), Paris (PAR), Singapore (SING), and Tokyo (TOK). (From Halavais 1999.)

relationships from road maps, or Bush’s “associative trails,” even when it may not be entirely clear why any individual has followed a given path [41] <http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;idno=5680986.0001.001;rgn=d...>

clear why any individual has followed a given path. [24] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N324;note=ptr>] I drew on a history of the telephone by Casson, who suggests that “wherever there is interdependence, there is bound to be telephony.” [25] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N325;note=ptr>] Likewise, interdependence now breeds hyperlinking. The precise mechanism by which a particular path is cut, or a hyperlink formed, is not important. What is important is that some connection must have existed in order for that connection to be cemented hypertextually; hyperlinks are not random.

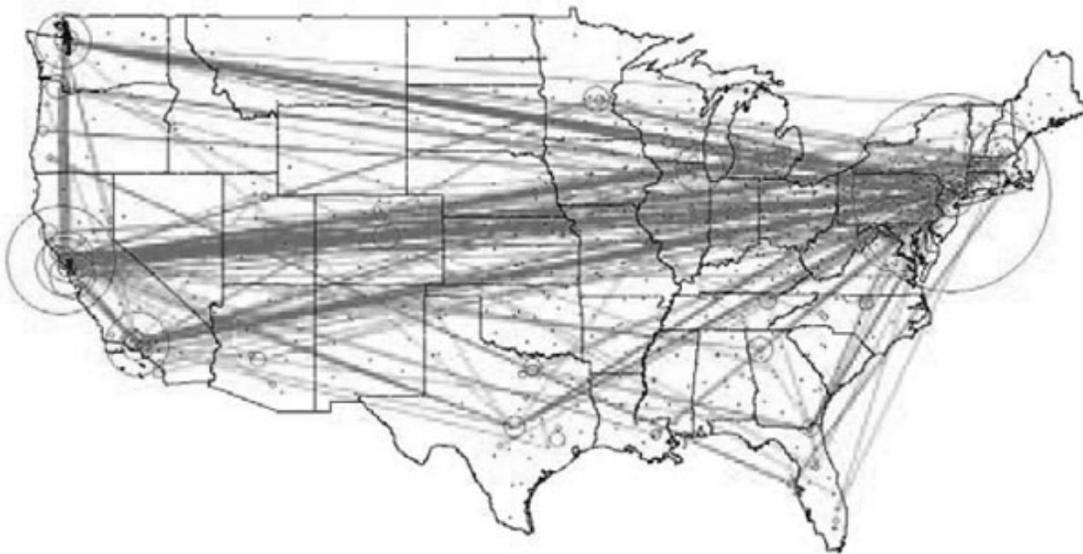
The Durkheimian tradition in sociology, like the cybernetic tradition, suggests not only that we can understand social behavior separately from the individual but that knowledge of individual motives is not required to understand “emergent properties” of society. Some have argued that despite the uncertainty surrounding individual citations, the observation of patterns at the network level can still provide valid and useful information. [26]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N326;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N326;note=ptr)]

Naturally, the ultimate aim is to integrate the microlevel and

Page 47



[</d/dculture/images/5680986.0001.001-00000002.jpg>]

Fig. 2.

Links among blogs in the United States. (From Lin, Halavais, and Zhang 2006.)

macrolevel understanding of hyperlinking behaviors, but it is not clear that the microlevel understanding must come first.

Understanding the structure of hyperlinked networks is not sufficient, nor is it intended to be. Whether geographical or not, examining the structural components of a network produces interesting questions. It insists that we ask why two documents, two organizations, or two people are linked together. It makes us wonder how certain parts of the Web come to be well

regarded and others do not. There is something satisfyingly analytical about reducing mass

impulses and hyperlinked networks to their constituent parts, and there continues to be value in doing so. But understanding the nature of the network can be an effective antecedent to understanding the nature of the individual hyperlink. It may be that the meaning of a hyperlink is best understood within the context of a hyperlinked network.

Curse of the Second Order?

For a short time, the field of cybernetics thrived in the United States. It was thought that the behaviors of complex systems were more telling than their constituent parts. Human systems are difficult to observe, in part because the observers are themselves necessarily a part of the system.

Page 48

Second-order cybernetics attempted to incorporate the observer in the observations; that is, it insisted that a degree of reflexivity is required in studying social systems. [27]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N327;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N327;note=ptr)] The trick with human systems is that they evolve and adapt not only to their environments but also to the observers who study them. When hyperlink analysis is present on the Web, especially when it is presented in real time, the same sorts of second-order effects begin to arise.

As noted, there are certain advantages to understanding the structure and dynamics of hyperlink networks. Indeed, even if we do not accept this hypothesis, a significant proportion of society has. The popular press is rife with advice about how to thrive in a newly networked society. Companies are throwing aside traditional hierarchies in favor of more agile, networked organizations. The network is becoming an organizing social principle, and in recognizing this, we are condemned to try to understand it from within.

Hyperlinks are not causing this shift alone, but the networks that are built up of hyperlinks allow for it. The hyperlink provides a basic building block through which complex, multidimensional, and easily changed documentation and communication systems may be constructed. The structural skeleton of an organization is the system by which it stores and transmits its accumulated knowledge. The move from the hierarchical, bureaucratic taxonomies of the traditional corporation—required, at least in part, by the nature of their records systems—to the more dynamic networked organization favored by terrorists and revolutionaries is the leitmotif of current business and political magazines and journals. [28]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N328;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N328;note=ptr)]

Manuel Castells remains the most identifiable proponent of the idea that networks are an organizing principle of modern society. He argues:

As a historical trend, dominant functions and processes in the information age are increasingly organized around networks. Networks constitute the new social morphology of our societies, and the diffusion of networking logic substantially modifies the operations and outcomes in processes of production, experience, power, and culture. [29] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N329;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N329;note=ptr)]

Of course, these networks do not require hyperlinks, but, notwithstanding Castells's own dismissal of hyperlinks, their ability to bind together and restructure media means that they are becoming the currency and connective tissue of the networked society. [30]

[\[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N330;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N330;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N330;note=ptr)

As a result, there is new attention being paid to network measures. Not

Page 49

only social scientists but their subjects are increasingly interested in where individuals are placed in the network. Google is built on the assumption that hyperlinks somehow transmit power or credibility. On the basis of that assumption, the search engine sends more traffic to the heavily linked sites, reinforcing that position of authority and leading to even more links. This occurs, arguably, to an even greater extent in the blogo-sphere, where some bloggers closely watch their ratings on Technorati and seek to rise in the rankings to the coveted A-list. Those who reach the most linked positions are likely to attract not only fame but fortune. [31]

[\[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N331;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N331;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N331;note=ptr)

There has always been some interest in uncovering, for example, the informal network of communication within an organization, but it has never been as easy as it is now to see who links to whom. As automatically gathered network measures become increasingly available, it is likely that behaviors will continue to change in an effort to affect these metrics. The result will be similar to what is already seen in academic circles, as tenure committees have adopted impact factors as an important way of measuring the performance of scholars. Once measures are visible, it is possible to play to the measures and to game the system.

Perhaps the most obvious example of this is a practice called "Google bombing," an attempt to associate a keyword search with a particular Web site. This manipulation is accomplished by encouraging a large number of Web authors to create a hyperlink to a Web site with anchor text containing a specific word or phrase. For example, by 2007 it was still the case that if someone queried one of the popular search engines for "failure" or "miserable failure," the biography of George W. Bush would be the first result; likewise a search for "liar" produced the biography of Tony Blair. Naturally, these phrases do not appear in the biographies of these two leaders, but because large numbers of individual links leading to the sites contained these key phrases, the Google search engine came to associate these pages with the phrases, as did several other search engines. Once it became clear that Google was susceptible to this form of manipulation, there were calls for Google to change its algorithms to reduce Google bombing. While Google initially refused to make changes, subsequent adjustments appear to have reduced the efficacy of this particular form of manipulation.

An entire industry has grown up around the manipulation of search engine results. Even those who do not consult an expert in "search engine optimization" remain interested in how best to improve their search rankings. In the blogging world, they may adjust their writing to attract a larger number of hyperlinks, just as a young academic might publish a literature

Page 50

review, rather than original work, knowing that a review is more likely to be cited. Young people come to attach as much value to the number of “friends” they have on MySpace as they do to other markers of social capital.

The effect on network measures is twofold. First, they are no longer as unobtrusive as they might once have been. As people within the network become more network conscious, their behaviors change in an attempt to affect who links to them and why. Second, at least in the case of commercial systems like Google and Technorati, the algorithms are changed to make gaming the system more difficult. This game of cat and mouse really means that those measuring the system have become a significant part of it. This may not affect social researchers as directly as it does search engines, but the conscious attempt to achieve favorable positions within the network structure means that network measures become as complicated by self-interest as are surveys and other obtrusive measures.

This is disappointing for those who would hope to be able to use hyperlinks as an unobtrusive way of mapping the structure of the collective unconscious, but it also suggests new, more “network-aware” uses of hyperlinks by those who create them. On the one hand, attempts to manipulate the network structure among bloggers are often seen in a negative light; the terms “link whoring” and “link doping” have emerged to describe behaviors directed specifically at shaping a blog’s position in the network. On the other hand, that kind of conscious awareness of the importance of deep hypertext structures by those who make use of the structures suggests a kind of maturation of collective consciousness, a striving for self-awareness, and a collaborative move toward more effective manipulation of those network structures.

The idea that the hyperlink network is becoming self-aware may sound a bit like science fiction —and with good reason. While Vannevar Bush is often credited with forming the concept of hypertext, his hypertext was intended for the scientific community, not for a broader global population. [32] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N332;note=ptr>]

H. G. Wells had recently suggested the emergence of a “world brain,” which sounded, in substance, much like Bush’s “associative trails.” [33] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N333;note=ptr>]

Wells’s book on the topic addressed the creation of an encyclopedia that could (through microfilm, as with the memex) be accessed and modified by anyone in the world. Many have seen the World Wide Web as a reflection of Wells’s vision, and some have taken this a step further and suggested that the Web is moving us toward a thinking superorganism. [34]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N334;note=ptr>]

While the particulars may differ, Wells clearly considered a global, hyperlinked encyclopedia to be a necessary part of a new form of global self-governance. [35]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N335;note=ptr>]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N335;note=ptr>]

The possibility for self-governance arises not just from an increase in global communication but from the emergence of a particular kind of communication. Deutsch suggests that self-

government, or “steering” of society, is accomplished through a special set of control channels. Not all communication is used for self-government, but those channels that allow for the organization and action of a society are particularly important and constitute the “nerves of government.” [36] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N336;note=ptr>]

Hyperlinks are, essentially, text. They differ from other content only in that they may be interpreted as a kind of control language: a code that provides for organization, coordination, and structure. Hyperlinks form the basis for this learning, adaptive, self-aware social system.

The Changing Nature of the Link

I want now to return to this essay’s initial question—what is the meaning of a hyperlink, or, as I have refigured it, what is the meaning of a hyperlink network? It seems that the answer depends in part on when the question is asked and by whom. The citation and hyperlink have a long history, and in the last decade, the role and meaning of a hyperlink seem already to have changed considerably. The future of linking is likely to be even more convoluted as hyperlinks carry more semantic value and reach even further into our everyday lives. It seems that such networks are growing more complex and adaptive and that their use is becoming more introspective. Futurists and science fiction authors have long held the idea that the network is becoming self-aware. If such self-awareness is to come to pass, hyperlink networks are likely to be central to that process.

If we are to take a long view of the hyperlink, we see it appropriating an ever-increasing role in our interactions. Its initial function mirrors that of citations and annotations, shaping text into structures that are more useful than linear arrangements. Citing the works of others, like any other form of communication, is essentially a social act, and as the norms of citation evolve, citation practices continue to have social influences. With the initial extension of the Web, hyperlinks took on an increasing role as tools for navigation, transporting attention from place to place. Especially with the rise of user-created media online, the social and navigational functions have come to predominate.

Even into the late 1990s, the Web was just one of many applications available on the Internet. Increasingly, the terms *Internet* and *World Wide Web* are used interchangeably. While not strictly correct, this conflation speaks to the degree to which the Web and hypertext have become organizing

Page 52

structures for communications online. Much of what people do on a computer these days, from e-mail to word processing, happens from within a Web browser. Hyperlinks have become the default user interface for the Internet.

The penetration of hyperlinks is likely to become even more ubiquitous as our computing devices do. The idea that clicking can only occur with a mouse on a computer monitor is already fading as touch screens on portable devices are seen more frequently and other switches become programmable. The border between hyperlinks and other forms of actuators is already dissolving to a certain degree. While the start button on some cars may not feel like a hyperlink, the functions of the stereo or navigation computer probably do. The ultimate trajectory of

hyperlinks may indeed be invisibility, the blue-underlined text merging with light switches and

voice commands to become one of a superset of links.

As the hyperlink becomes more ubiquitous, it is also layered with more meaning. The long-predicted Semantic Web, a Web that provides both content and information about how that content is related, has been slow in coming. In part, this is because it has been difficult to encourage people to create metadata, explicit descriptions of what content is and how it relates to other content on the Web. In the last few years, as the value of such data has become clearer, users have slowly started creating explicit metadata that can be read and manipulated. There is a press toward adding tags to hyperlinks to make them more meaningful. Certainly, the OpenURL has gained some ground within library circles. Google's support of the nofollow tag came, in part, as a response to spammers commenting on blogs, as well as to a general interest (also largely among bloggers) in being able to link to something without appearing to endorse it. There is an incipient effort to tag certain hyperlinks "not safe for work." [37]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N337;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N337;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N337;note=ptr)] Some of this metadata is created automatically by content management and blogging systems. Such uses of semantic markup in hyperlinks are only at their earliest stages, but the practice of what might be called "reflexive hyperlinking" is already widespread, leading to a new awareness about what it means to make a link.

While tagging and folksonomies have largely been conceptualized independently of their influence on hyperlink networks, several of the systems that allow for tagging are annotating particular pages on the Web and, by extension, the links that lead there. [38]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N338;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N338;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N338;note=ptr)] Given the long-standing difficulty in producing metadata for electronic content, this "bottom-up" approach provides a great deal more context for Web links. As noted earlier,

Page 53

much of the research power in hyperlink networks comes not from the networks themselves but from their combination of structural information with other sources of data about the pages and links. This human-generated metadata, along with other sources of metadata about content on the Web and how it is linked, will make analyses of hyperlinked networks richer and more revealing.

The increased reach of hyperlinks and the richness of information that may be associated with those links demands the continued study of hyperlinked networks and the links that make them up. There are some obvious targets for further research. While several studies have attempted to measure the motivations, gratifications, and cognitive processes surrounding the creation of hyperlinks, as well as similar contexts for choosing to follow a hyperlink, the majority of these studies have concentrated on portions of the Web with a scholarly or academic function. [39]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N339;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N339;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N339;note=ptr)

Understanding the contexts in which links are created and used represents an interesting challenge for the researcher, one that is likely to be rewarding.

This essay has suggested that the study of macrolevel hyperlink networks has been complicated

.....

by users becoming more aware of hyperlink structures. While this makes the interpretation more elaborate, it in no way obviates the interest or need to study hyperlink networks. In fact, the wider interest in such studies encourages public scholarship that provides data to the communities we study. Moreover, as the need for tools to understand the structure of hyperlink networks extends beyond the world of researchers, new tools are created that are useful in analyzing, visualizing, and making sense of these complex networks. Ideally, programs of research will allow for the macrolevel analysis to be integrated with a better understanding of why and how people contribute to the creation of these structures. Moreover, as such understanding is advanced, it is likely to alter the ways in which people create and use hyperlinks. Awareness of hyperlink networks is in some ways only an intermediary step toward networks that are able to understand and interpret themselves to an increasing degree.

In understanding what a hyperlink means, we need to look at what a hyperlink does. Over time, it has come to do more and more. At present, it stands as the basic element of organization for the Web, and as more and more of our lives are conducted through the Web, it becomes increasingly important that we understand how hyperlinked structures are formed and change.

Page 54

Notes

1. N. Holmes, "The KWIC and the Dead: A Lesson in Computing History," *Computer* 34, no. 1 (2001): 144. [#]
2. E. L. Eisenstein, *The Printing Press as an Agent of Change: Communications and Cultural Transformations in Early-Modern Europe*, vol. 1. (Cambridge: Cambridge University Press, 1979). [#]
3. G. N. Gilbert, "Referencing as Persuasion," *Social Studies of Science* 7, no. 1 (1977), 113–22. [#]
4. R. G. Collingwood, *The Idea of History*, rev. ed. (Oxford: Oxford University Press, 1994), 62. [#]
5. M. Polanyi and H. Prosch, *Meaning* (Chicago: University of Chicago Press, 1975), 184. [#]
6. J. B. Walther, "Computer-Mediated Communication: Impersonal, Interpersonal, and Hyperpersonal Interaction," *Communication Research* 23, no. 1 (1996): 3–43. [#]
7. S. Rose, "What's Love Got to Do with It? Scholarly Citation Practices as Courtship Rituals," *Language and Learning across the Disciplines* 1, no. 3 (1996): 34–48. [#]
8. T. A. Brooks, "Private Acts and Public Objects: An Investigation of Citer Motivations," *Journal of the American Society for Information Science* 36, no. 4 (1985): 223–29. [#]
9. A. Grafton, *The Footnote: A Curious History* (Cambridge, MA: Harvard University Press, 1999), 93. [#]
10. E. Garfield, "Citation Behavior: An Aid or a Hindrance to Information Retrieval?" *Current Contents* 18 (1989): 3–8. [#]
11. V. Bush, "As We May Think," *Atlantic Monthly*, July 1945,

<http://www.theatlantic.com/doc/194507/bush> [http://www.theatlantic.com/doc/194507/bush]. [#]

...

[#]

12. D. Engelbart, "The Click Heard Round the World," *Wired* 12, no. 1, <http://www.wired.com/wired/archive/12.01/mouse.html> [<http://www.wired.com/wired/archive/12.01/mouse.html>].[☛][#]
13. T. H. Nelson, *Remarks on Xanadu* (Maastricht: Internet Research 3.0, 2004).[☛][#]
14. L. Egghe, "New Informetric Aspects of the Internet: Some Reflections—Many Problems," *Journal of Information Science* 26, no. 5 (2000): 329–35.[☛][#]
15. A. Scharnhorst and M. Thelwall, "Citation and Hyperlink Networks," *Current Science* 89, no. 9 (2005): 1518–23; M. Thelwall, "What Is This Link Doing Here? Beginning a Fine-Grained Process of Identifying Reasons for Academic Hyperlink Creation," *Information Research* 8, no. 3 (2003), <http://informationr.net/ir/8-3/paper151.html> [<http://informationr.net/ir/8-3/paper151.html>].[☛][#]
16. T. H. Nelson, "Complex Information Processing: A File Structure for the Complex, the Changing, and the Indeterminate," in *Proceedings of the 1965 20th National Conference*, ed. Lewis Winner (New York: ACM, 1965), 84–100.[☛][#]
17. T. Ciszek and X. Fu, "An Annotation Paradigm: The Social Hyperlink," in *Proceedings of the American Society for Information Science and Technology* 42, no. 1 (2005), <http://www3.interscience.wiley.com/journal/112785658> [<http://www3.interscience.wiley.com/journal/112785658>]; T. Bardini, "Bridging the Gulfs: From Hypertext to Cyberspace," *Journal of Computer-Mediated Communication* 3, no. 2 (1997), <http://jcmc.indiana.edu/vol3/issue2/bardini.html> [<http://jcmc.indiana.edu/vol3/issue2/bardini.html>]; M. H. Jackson, "Assessing the Structure of Communication on the World Wide Web," *Journal of Computer-Mediated Communication* 3, no. 1 (1997), <http://jcmc.indiana.edu/vol3/issue1/jackson.html> [<http://jcmc.indiana.edu/vol3/issue1/jackson.html>].[☛][#]
18. S. Brin and L. Page, "The Anatomy of a Large-Scale Hypertextual Web Search Engine," *WWW7 / Computer Networks* 30, nos. 1–7 (1998): 107–17.[☛][#]

 Page 55

19. L. C. Freeman, "Centrality in Social Networks: Conceptual Clarification," *Social Networks* 1, no. 3 (1979): 215–39.[☛][#]
20. H. W. Park and M. Thelwall, "Hyperlink Analyses of the World Wide Web: A Review," *Journal of Computer-Mediated Communication* 8, no. 4 (2003), <http://www.jcmc.indiana.edu/vol8/issue4/park.html> [<http://www.jcmc.indiana.edu/vol8/issue4/park.html>].[☛][#]
21. A.-L. Barabási, *Linked: The New Science of Networks* (Cambridge, MA: Perseus, 2002).[☛][#]
22. A. Halavais, "Informational City Limits: Cities and the Infostructure of the WWW" (paper presented at the workshop "Cities in the Global Information Society: An International Perspective," Newcastle upon Tyne, November 1999).[☛][#]
23. J. Lin, A. Halavais, and B. Zhang, "The Blog Network in America: Blogs as Indicators of Relationships among US Cities," *Connections* 27, no. 2 (2006): 15–22.[☛][#]

24. A. Halavais, “National Borders on the World Wide Web,” *New Media and Society* 2, no. 1 (2000): 7. [1#]
25. H. N. Casson, *The History of the Telephone* (Chicago: A. C. McClurg and Company, 1910), 97. [1#]
26. B. Van der Veer Martens, “Do Citation Systems Represent Theories of Truth,” *Information Research* 6, no. 2 (2001), <http://informationr.net/ir/6-2/paper92.html> [<http://informationr.net/ir/6-2/paper92.html>]. [1#]
27. H. Von Foerster, *Understanding Understanding: Essays on Cybernetics and Cognition* (New York: Springer-Verlag, 2006), chap. 13. [1#]
28. J. Arquilla and D. F. Ronfeldt, *Networks and Netwars: The Future of Terror, Crime, and Militancy* (Santa Monica, CA: Rand Corporation, 2001). [1#]
29. M. Castells, *The Rise of the Network Society* (Oxford: Blackwell, 1996), 469. [1#]
30. M. Castells, *The Internet Galaxy: Reflections on the Internet, Business, and Society* (Oxford: Oxford University Press, 2001), 201–3. [1#]
31. C. Thompson, “Blogs to Riches: The Haves and Have-Nots of the Blogging Boom,” *New York Magazine* 20 (February 2006), <http://nymag.com/news/media/15967> [<http://nymag.com/news/media/15967>]. [1#]
32. J. M. Nyce and P. Kahn, *From Memex to Hypertext: Vannevar Bush and the Mind’s Machine* (San Diego: Academic Press, 1991), 50–51. [1#]
33. H. G. Wells, *World Brain* (London: Methuen and Company, 1938). [1#]
34. F. Heylighen and J. Bollen, “The World-Wide Web as a Super-Brain: From Metaphor to Model,” in *Cybernetics and Systems ’96*, ed. R. Trappl (Vienna: Austrian Society for Cybernetics, 1996), 917–22. [1#]
35. W. B. Rayward, “H. G. Wells’s Idea of a World Brain: A Critical Reassessment,” *Journal of the American Society for Information Science* 50, no. 7 (1999): 557–73. [1#]
36. K. Deutsch, *The Nerves of Government* (New York: Free Press, 1966). [1#]
37. P. J. Doland, “Genius Grant Please, or The NSFW HTML Attribute,” *Frosty Mug Revolution*, December 28, 2006, <http://pj.doland.org/archives/041571.php> [<http://pj.doland.org/archives/041571.php>]. [1#]
38. C. Marlow et al., “HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, to Read,” in *Proceedings of the Seventeenth Conference on Hypertext and Hypermedia* (New York: ACM, 2006), 31–40. [1#]
39. C. Hine, “Web Pages, Authors, and Audiences: The Meaning of a Mouse Click,” *Information, Communication, and Society* 4, no. 2 (2001): 182–98. [1#]

Philip M. Napoli

The role of hyperlinking in the development of the Internet warrants investigation for a number of reasons. First, along with the Internet's inherently global reach and its virtually unlimited content capacity, hyperlinking is one of the key factors that distinguishes the Internet from traditional media. Second, the dynamics of hyperlinking have evolved in a number of interesting and unexpected ways, particularly as a result of the mechanisms by which search engines choose to generate and display links. Finally, the underlying choices and dynamics of hyperlinking are, of course, central to the distribution of audience attention (and, consequently, dollars) online and can therefore exert considerable influence over how the Internet evolves as a medium.

An important component of the study of new media involves the investigation of the relationship between old and new media. Exploring how new media can either disrupt or become integrated into the existing media system offers valuable insights that can guide policy makers, industry decision makers, and scholars seeking to understand the organizational ecology of media, the evolution of media systems and media technologies, and the dynamics of media usage. Scholars from a wide range of disciplines have sought to understand the push and pull between the Internet's undeniable revolutionary potential as demonstrated by links and the various influences and constraints imposed by the existing media system that it has entered. In my own efforts to address this issue in the Internet's early stages of development, I focused on the then unclear question of the extent to which the Internet would ultimately demonstrate the characteristics of more traditional mass media and on the reasons the Internet might be likely to adopt many of the characteristics of traditional mass media rather than evolve as the entirely unique and revolutionary medium that many were hoping for and anticipating in those heady early days. [1] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N401;note=ptr>]

I dubbed the pressures compelling the Internet down more traditional media evolutionary paths the forces of "massification"—a term that referenced the then-common argument that the Internet represented the

Page 57

end (or at least the beginning of the end) of traditional mass media. Developed when the medium essentially was in its infancy, this analysis of the Internet and the predictive propositions it entailed managed to hold some water in the ensuing decade. The Internet has indeed come to serve many of the functions, feature many of the same institutions, exhibit many of the same audience behavior patterns, and provide much of the same content as many of the mass media that preceded it.

The present essay revisits some of these claims in light of the current status of the Internet and enlarges the analytical frame with an eye toward teasing out exactly how the process of linking online may or may not factor into the massification of the Internet. The first section of this essay provides an overview of the forces of massification that have traditionally influenced all new media (including the Internet); this section also considers recent developments online through this analytical lens. The next section looks specifically at the act of hyperlinking, asking whether it reinforces or undermines these forces of massification; this section draws on the growing body of literature analyzing the patterns of hyperlinking online as well as recent developments involving the process of hyperlink selection and generation. The concluding section assesses the implications of the dynamics of hyperlinking for the evolution of the Internet, considers policy

implications of the dynamics of hyperlinking for the evolution of the Internet, considers policy implications, and offers suggestions for future research.

New Media and the Forces of Massification

New media technologies do not exist in a vacuum. Rather, they enter into a diverse, complex, and dynamic mix of established and emerging media. Consequently, understanding any new medium requires an understanding of its interaction with the existing media environment, both from the standpoint of consumer adoption and usage and from the standpoint of institutional responses. [2] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N402;note=ptr>] Such an approach makes it necessary to focus not only on the interactions between old and new media but also on the key institutional and economic forces that act on any new medium as it begins to carve out its place within the established media system. Many of these forces (often the ones neglected by those providing the earliest assessments of new media technologies) in fact compel new media technologies along evolutionary lines established by traditional media. It is these that I have labeled the forces of “massification.” [3] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N403;note=ptr>] These forces fall into three broad categories: audience behavior, media economics, and institutional forces. Each of these will be reviewed briefly here.

Before examining each of these forces, however, it is important to outline

Page 58

the basic criteria that we associate with traditional mass media. Detailed discussions of this issue can be found elsewhere. [4] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N404;note=ptr>] To briefly summarize, the common characteristics of traditional mass media include a one-to-many orientation (and an associated lack of interactivity); the prominence of “institutional communicators”; a strong commercial orientation; and an associated emphasis on audience maximization and, consequently, mass appeal content. [5]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N405;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N405;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N405;note=ptr)]

Audience Behavior

Certain well-established aspects of audience behavior—across many media—can compel new media technologies to function along the lines of traditional media, particularly by encouraging audiences to maintain strong connections with one-to-many and noninteractive communicative forms, as well as connections to content with traditional mass appeal (as opposed to highly targeted and specialized niche content). There is, for instance, the well-documented tendency toward passivity in audience behavior. [6] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N406;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N406;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N406;note=ptr)] There is a limit to the extent to which audiences want their media consumption to involve substantial interactivity or substantial search activities, although this limit may (or does?) vary across media, as well as across usage categories and demographic groups.

From an audience behavior standpoint, it is also important to recognize that there is a well-documented tendency across media for audiences to prefer content with higher production

documented tendency across media for audiences to prefer content with higher production budgets and to interpret production budgets as some sort of (imperfect) manifestation of quality. [7] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N407;note=ptr>]

Of course, higher production budgets require the presumption of a satisfactory return; therefore, higher-budget content typically is geared toward having greater mass appeal. Thus the distribution of audience attention in most media contexts tends to cluster around high-budget, mass appeal content, which of course also tends to be the content produced by the traditional institutional communicators (with the resources to expend on big-budget content). [8]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N408;note=ptr>]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N408;note=ptr>]

Media Economics

The preceding discussion of audience behavior leads naturally into some of the basics of media economics. Perhaps the first key principle involves the powerful economies of scale that exist in the production of media content. Media content is defined in economic terms as a “public good.” [9] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N409;note=ptr>]

Some key characteristics of public goods are high fixed costs,

Page 59

very low variable costs, and nondepletability. It is very expensive to produce and sell the “first copy” of a public good (e.g., a television program or Web site). But to sell additional copies to additional consumers requires very little additional cost, particularly given the fact that one consumer’s consumption of the media product does not prevent another consumer from consuming the same media product (i.e., only one Web page needs to be created whether a thousand or a million people visit the site). There are enormous economies of scale to be achieved with such products, as production costs can be distributed over large audiences over long periods of time (consider the fact that *I Love Lucy* episodes are still collecting revenues).

This has a few implications for the massification of any medium. First, it creates a tremendous incentive for any new medium to function—if not primarily, at least significantly—as an ancillary distribution mechanism for content produced in older media. Second, it creates a powerful incentive for producers of content for the new medium to try to appeal to and thereby distribute production costs across as large an audience as possible. Third, the tremendous risk naturally associated with any product with very high fixed costs creates powerful incentives to employ traditional media industry strategies of risk reduction, such as derivations or recyclings of content already proven to be successful in other media or reliance on proven strategic approaches most likely to attract a large audience. [10] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N410;note=ptr>]

Institutional Forces

Finally, we come to what are termed “institutional forces,” those institutional characteristics of the media system that compel new technologies to adopt the characteristics of traditional media. First and perhaps most obvious, there is the well-documented historical pattern for existing media organizations to (somewhat belatedly, as it usually turns out) migrate into new media and, in so doing, transplant existing content (as already discussed), strategic approaches, and

and, more broadly, transparent content (as in early YouTube), strategic approaches, and business models. [11] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N411;note=ptr>]

A second significant institutional force involves the process of audience measurement. Audience attention data is a vital commodity across all ad-supported media. It has proven to be particularly important to the establishment of any new technology as a viable advertising medium. [12] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N412;note=ptr>]

Unfortunately, one unavoidable by-product of most established audience measurement methodologies is that, given the nature of sampling, the larger the size of the audience, the more accurate and reliable are the audience

Page 60

data. [13] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N413;note=ptr>]

This creates an inherent bias in the audience marketplace, favoring content providers that attract large audiences.

The Massification of the Internet

When we consider these forces in relation to the Internet, it is important to acknowledge that the Internet has undoubtedly confounded traditional notions of a mass medium. Its interactive capacity is tremendous, and it facilitates not only one-to-many but also one-to-one and many-to-many forms of communication. Institutional communicators remain tremendously prominent, but opportunities for other types of actors to achieve prominence exist to an extent that cannot be found in other media. And while substantial portions of the Internet are highly commercialized and certainly devoted to pursuing large audiences, other components of the online realm are not. In these ways, the Internet has both adopted and expanded well beyond the characteristics of traditional media. But certainly, the traditional characteristics of mass media have become integral to the institutional structure and orientation of the Internet and to how consumers use it as an information and entertainment resource. [14]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N414;note=ptr>]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N414;note=ptr>]

From an audience behavior standpoint, it is somewhat telling that the typical television viewer, in an environment of channel abundance, regularly consumes only about thirteen of the available channels—and that this is roughly the same as the number of Web sites that the typical person visits on a regular basis. [15] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N415;note=ptr>]

It is not surprising, either, that the typical Web search seldom involves looking beyond the first page of links returned by the search engine or that a user looks beyond the first three pages of links less than 10 percent of the time. [16] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N416;note=ptr>]

The search-and-retrieve dynamic, perhaps the most basic attribute of an interactive medium, is one that extracts costs from the audience member. Consequently, we see audience behavior

patterns, such as these, that illustrate important limitations in the extent to which the Internet's full potential to dramatically reconfigure the nature of audiences' interaction with their media

can be realized.

Consider also the rise to prominence of content aggregation sites such as YouTube and MySpace. While these sites have received tremendous attention for empowering individuals to serve as content producers, facilitating a many-to-many communication dynamic and thereby “deinstitutionalizing” the media (all things, it should be noted, that the Internet was already facilitating without such sites), what has been largely ignored to date is the extent to which these sites function largely to confine the vastness and complexity of the Web into a simpler and more manageable

Page 61

framework. The days of scouring the Web for individual home pages or video clips are now being replaced by individual repositories/destinations that are subject to centralized editorial control. It is as if the large-scale gatekeeper bottlenecks characteristic of old media are being re-created in an environment in which they are not technologically necessary (or, presumably, desirable). Suddenly, many of the chaotic and independent features of the Web are being voluntarily placed under the control of a single institutional communicator (i.e., News Corp. in the case of MySpace and Google in the case of YouTube). This is a kind of downsizing or consolidation of the Web itself. Such patterns are a reaction to what has been inarguably described as “an enormous oversupply of web offerings that no human being can navigate without aides that give some structure to this ever-growing universe.” [17] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N417;note=ptr>] To the extent that this kind of aggregation of Web content is proving highly desirable or even necessary to users (in the same way that Amazon and eBay have consolidated online shopping), a potentially successful business strategy for going forward would simply be to identify other broad content categories currently scattered about the Web that are in need of aggregation and then develop the appropriate aggregation, search, and display mechanisms.

Related to this phenomenon, we also see a strong tendency for online audiences to cluster around relatively few content options, in a behavioral pattern that has been well established among the traditional mass media. Audience behavior research frequently has documented a “power law” distribution of audience attention and/or dollars, with 20 percent of the available content attracting 80 percent of the audience. [18] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N418;note=ptr>]

Recent research examining the distribution of audience attention across different media has found that the concentration of audience attention around relatively few sources in the traditional media realm has been largely reproduced in the online realm. [19]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N419;note=ptr>]

Some comparative studies have found an even greater concentration of audience attention online than is found in traditional media, such as newspapers, radio, and television. [20]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N420;note=ptr>]

Equally important is the fact that this audience attention is clustering around many of the

institutional communicators that characterize the traditional media realm, as powerful media entities—ranging from News Corp. (particularly with its purchase of MySpace), to Time Warner

(which, contrary to expectations, absorbed AOL rather than vice versa), to Disney—all have established prominent positions online. [21] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N421;note=ptr>] Among the top ten “parent companies” online for the month of November 2006 were Time Warner, News Corp., the New York Times Company, and Disney (Nielsen//NetRatings, 2007). Of course, given this institutional migration and the “public good”

Page 62

characteristics of media content, it is not surprising that the Web has developed as a key mechanism for accessing and distributing “old media” products, such as recorded music, television programs, motion pictures, and magazines. The Internet has been well described as “swallow[ing] up most, if not all, of the other media in an orgy of digital convergence.” [22] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N422;note=ptr>] To the extent that this is the case, the Internet’s ability to exhibit fundamentally different characteristics from the media that preceded it seems limited.

This clustering of audiences also continues to be associated with patterns in advertiser behavior that are consistent with the massification effects of audience measurement. Established audience measurement systems naturally favor sites that attract large audiences (in the perceptions of advertisers) over sites that attract smaller audiences, even if the latter, niche sites might be attracting audiences that are more desirable (from the advertisers’ standpoint). Advertisers have shown themselves to be willing to pay a premium for accuracy in audience measurement, which can help explain why, even today, the most popular Web sites attract a share of online advertising dollars that exceeds their share of the online audience. [23]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N423;note=ptr>] This creates important economic disincentives for serving narrower, more specialized audiences online.

Hyperlinking and the Forces of Massification

As the preceding section illustrated, the technological forces compelling a new medium such as the Internet to defy the confines of traditional media are to some degree offset by a number of countervailing social and institutional forces that are clearly influencing both the structure of the online realm and the ways that consumers navigate the online space. The questions that this section seeks to answer is whether and how the practice of hyperlinking—a practice that, to a large degree, distinguishes the realm of online media—factors into the push and pull between old and new media that is at the core of the Internet’s evolutionary process. [24]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N424;note=ptr>]

Hyperlinks have been described as “the heart of the World Wide Web.” [25]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N425;note=ptr>]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N425;note=ptr>] In thinking broadly about the process of linking online, it is important not to think only in terms of

the links to text and video that can be embedded in discrete Web pages (thereby creating the distinctive “intertextuality” of the Web and Web navigation). We also need to consider the processes of link generation and display associated with the functioning of search engines (given the centrality of search engines to online

Page 63

navigation). And we need to note the processes of link generation that accompany—and are meant to assist or manipulate—consumer choices online (i.e., the recommendations about other potentially interesting content that now frequently accompany Web users’ content selections). These represent perhaps the most fundamental contexts for exploring the potential significance of linking to the process of massification online.

A potentially useful conceptual lens for examining these various contexts involves the concept of gatekeeping. Despite early proclamations to the contrary, it has become very clear by this point that the Internet has not, by any stretch of the imagination, eliminated gatekeeping or made it obsolete. Rather, the dynamics of the gatekeeping process have changed significantly, perhaps becoming a bit more covert. [26] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N426;note=ptr>]

Much gatekeeping can now be handled via technological means, though the human factor remains prominent. [27] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N427;note=ptr>]

Hyperlinking is perhaps the most significant mechanism of online gatekeeping. [28] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N428;note=ptr>]

Through their decisions about when and where to hyperlink and, most important, what to link to, content providers exert substantial editorial control. [29] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N429;note=ptr>]

As Park has noted, Web sites can very usefully be perceived as “actors,” and “through a hyperlink, an individual website plays the role of an actor who could influence other website’s trust, prestige, authority, or credibility.” [30] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N430;note=ptr>]

Hyperlinking thus serves as a primary mechanism via which an online content provider exerts control over its audience and, to use terminology drawn from traditional media (specifically television), manages “audience flow.” [31] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N431;note=ptr>]

The concept of the “walled garden” arose primarily to describe AOL’s early efforts to keep its subscribers within AOL-generated content and away from the true World Wide Web. [32] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N432;note=ptr>]

But it continues to have relevance in the context of contemporary linking activities. Research shows that online news sites overwhelmingly hyperlink only to internal Web pages and seldom link to

outside sources. [33] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N433;note=ptr>]

Other

research suggests that search engines produce results that suppress links to controversial information or news stories. [34] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N434;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N434;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N434;note=ptr)]

Recent efforts at mapping the distribution of links online (in terms of who links to who, how often, etc.) document a clear and coherent “information politics” that suggests that very deliberate editorial decisions are being made with an effort toward guiding audience attention down certain preferred paths as opposed to others. [35] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N435;note=ptr>]

When these types of traditional editorial dimensions of hyperlinking are coupled with the technical dimensions of link generation by search engines (in which the quantity of inbound links is a key driver of a link’s placement in the search results), the question frequently has arisen whether the dynamics of linking are such that the imbalances in content

Page 64

accessibility and prominence that characterize the traditional mass media world are being replicated in the online world. [36] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N436;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N436;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N436;note=ptr)]

Research suggests that this may very well be the case. [37]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N437;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N437;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N437;note=ptr)]

Koopmans and Zimmerman, for instance, find that in terms of political news coverage, the same institutional actors and information sources achieve virtually identical levels of prominence (as measured, in part, by link quantity) in both the online and print media realms. [38]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N438;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N438;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N438;note=ptr)]

The persistence of such patterns is in some ways surprising given the dramatic technological differences in how content is stored, exhibited, and accessed in online versus offline contexts. These important differences and their potentially dramatic implications are explored perhaps most extensively in Anderson’s “long tail” analysis. [39] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N439;note=ptr>]

The essence of the long tail argument is that the combination of the greatly expanded content storage capacity of a digitized space such as the Internet (versus, say, a traditional book or record store) and the enhanced, highly interactive search tools that such a space can provide (e.g., peer recommendations; site-generated recommendations; and robust, multidimensional search features) contribute to a media environment in which the traditional power law distribution of audience attention can be altered or at least can become more lucrative than was possible in the offline world. [40] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N440;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N440;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N440;note=ptr)] A

consumption dynamic in which 20 percent of the content generates 80 percent of the revenue (and in which nobody knows what that 20 percent is going to be) [41]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N441;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N441;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N441;note=ptr)] can be more profitable in an environment in which “shelf space” is much less scarce (and less

expensive) and in which the consumer's ability to effectively and satisfactorily navigate this expansive shelf space is enhanced via a wide range of search tools and linking systems.

In such an environment, the content provider can make all of the relevant content available and not have to make editorial judgments about which content to include or exclude based on (often wrong) predictions regarding consumer tastes. The content provider can also be reasonably sure that all of the content will generate at least some revenue, even if the bulk of the revenues continue to be generated by only 20 percent of the content. Under this model, the chances of success are increased because (a) the content provider never has to worry about not having any of the 20 percent of content options that prove to be enormously successful and (b) the remaining content (the long tail) can be stored and exhibited cheaply enough and located and accessed easily enough by the consumer to become a meaningful contributor to profits.

This description of the long tail model has tried to emphasize an issue that has received surprisingly little attention: the extent to which these

Page 65

radical changes in content distribution, access, and exhibition do anything to alter the well-established dynamics of how audiences distribute their attention across various content options. The long tail phenomenon (i.e., the 80/20 rule) that characterized traditional media remains a defining characteristic of the new media space, as the research already cited suggests, though other recent research suggests that some very modest shifts toward a broader allocation of audience attention can result from the migration to online distribution and exhibition. [42]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N442;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N442;note=ptr)] It seems safe to say that the online environment simply provides a potentially more profitable context in which to navigate the traditional constraints under which content providers have operated. But the fact that this dramatically changed technological environment can apparently do relatively little to alter the fundamental distribution patterns of audience attention is, in many ways, as remarkable, if not more remarkable, than the ways in which this changed technological environment can alter the economics of content distribution and exhibition. The persistence of such patterns in the distribution of audience attention may be a reflection of the fact that the exact same power law patterns can be found in the distribution of inbound and outbound links on the Web. [43] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N443;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N443;note=ptr)] Thus the ecology of hyperlinks may itself represent a set of paths that is compelling a distribution of audience attention that bears a striking resemblance to the distribution of audience attention in the traditional mass media.

Conclusion

As this essay has illustrated, even the process of hyperlinking, which is representative of the distinctive, boundary-defying, and interactive character of the Internet, in many ways complies with or is influenced by a set of forces that help compel the medium to function (from both a content producer and a content consumer's standpoint) along lines established by traditional

media. This is not to say that the innovative potential of the Internet has gone completely unrealized. But it does suggest that the evolutionary trajectory of any new medium—even one as

dramatically different as the Internet—is significantly constrained by a set of stable and influential social and institutional forces. [44] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N444;note=ptr>]

There are also some important policy implications to be drawn from the patterns reviewed in this essay. Perhaps the most important of these is to question the argument increasingly heard in policymaking circles that regulation of traditional media's ownership and market structure is no

Page 66

longer necessary because the Internet provides a robust and viable alternative to them. Clearly, the more the Web exhibits the characteristics of traditional media, the less relevant this argument becomes. [45] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N445;note=ptr>]

From a research standpoint, however, we still have much to learn about the processes of linking and how they impact the dynamics of content production, distribution, and access. As Wellman has illustrated, early Internet research focused primarily on prognostications. [46]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N446;note=ptr>] The second stage involved the basic mapping of user behavior, and only now have we entered the stage where the dynamics of Internet usage are being subject to robust empirical analysis.

However, not all aspects of Internet research are at the same evolutionary stage. While we are developing a sophisticated understanding of the dynamics of Internet usage, our understanding of the production side is not as far along. Today, we are still very much embedded in Wellman's second stage of analysis as it relates to the production and presentation of Web content. [47]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N447;note=ptr>] This “mapping” of the online space is well developed. We are developing a strong sense of the distribution of links—of who links to whom and how often. [48]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N448;note=ptr>] However, we do not yet understand very well the dynamics of the linking decision-making process. What factors determine whether or not a site is linked to another site? Why do certain sites become important nodes in Web space while others languish in relative obscurity?

Inquiries in this vein have been infrequent up to this point. [49]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N449;note=ptr>]

Moving forward, it seems important that researchers make further efforts to move beyond the consumption side of the Internet (i.e., how users navigate the online space and distribute their attention) and delve deeper into the processes surrounding the generation of content and how these content sources interact with one another (e.g., via linking). For instance, in light of the tremendous amount of attention that blogging is receiving as an alternative to traditional news

media, we need to ask to what extent the links provided by bloggers are pointing readers to traditional news media sources? Similarly, we should investigate the extent to which the content

populating sites such as YouTube is really “user-generated” content or simply content “ripped” from traditional media (e.g., TV and movie clips). Equally important, how is audience attention distributed across these different content types? Is traditional media content being consumed in proportion to its availability on such platforms? Or is it being consumed in greater or lesser proportion to its availability?

In some ways, this pattern in our understanding of the Web as a medium mirrors the evolution of the field of communications research, where the initial empirical focus was directed at the receivers of the information

Page 67

(their usage patterns, effects, etc.). Only after this line of inquiry matured did we see researchers turn their attention to the organizations involved in the production and distribution of content. However, focusing greater attention on questions such as these is essential for developing a clearer portrait of the interaction between old and new media and the extent to which a new medium is really performing new functions, instituting new communications dynamics, and providing new content.

Notes

1. P. M. Napoli, “The Internet and the Forces of ‘Massification,’” *Electronic Journal of Communication* 8, no. 2 (1998), <http://www.cios.org/www/ejc/v8n298.htm> [<http://www.cios.org/www/ejc/v8n298.htm>].[Ⓢ] [#]
2. S. Lehman-Wilzig and N. Cohen-Avigdor, “The Natural Life Cycle of New Media Evolution: Inter-Media Struggle for Survival in the Internet Age,” *New Media and Society* 6, no. 6 (2004): 707–30; P. M. Napoli, “Evolutionary Theories of Media Institutions and Their Responses to New Technologies,” in *Communication Theory: A Reader*, ed. L. Lederman (Dubuque, IA: Kendall/Hunt, 1998), 315–29.[Ⓢ] [#]
3. Napoli, “The Internet and the Forces of ‘Massification.’”[Ⓢ] [#]
4. Ibid.; W. R. Neuman, *The Future of the Mass Audience* (New York: Cambridge University Press, 1991); J. Turow, *Media Systems in Society: Understanding Industries, Strategies, and Power* (White Plains, NY: Longman, 1992); J. G. Webster and P. F. Phalen, *The Mass Audience: Rediscovering the Dominant Model* (Mahwah, NJ: LEA, 1997).[Ⓢ] [#]
5. J. Turow, “The Critical Importance of Mass Communication as a Concept,” in *Mediation, Information, and Communication: Information and Behavior*, ed. B. D. Ruben and L. Lievrouw, vol. 3 (New Brunswick, NJ: Transaction Publishers, 1990), 9–20; Napoli, “The Internet and the Forces of ‘Massification.’”[Ⓢ] [#]
6. Webster and Phalen, *Mass Audience*.[Ⓢ] [#]
7. B. M. Owen and S. S. Wildman, *Video Economics* (Cambridge, MA: Harvard University Press, 1992).[Ⓢ] [#]
8. Ibid.[Ⓢ] [#]
9. Ibid.[Ⓢ] [#]

10. Media products, in particular, have proven to be a very risky business across a wide range of

10. Media products, in particular, have proven to be a very risky business across a wide range of technologies; see P. M. Napoli, *Audience Economics: Media Institutions and the Audience Marketplace* (New York: Columbia University Press, 2003).[Ⓢ][#]

11. Napoli, "The Internet and the Forces of 'Massification'"; Napoli, "Evolutionary Theories of Media Institutions."[Ⓢ][#]

12. Napoli, *Audience Economics*.[Ⓢ][#]

13. Ibid.[Ⓢ][#]

14. L. D. Intron, "Shaping the Web: Why the Politics of Search Engines Matters," *Information Society* 16, no. 3 (2000): 169–85; J. G. Webster and S. F. Lin, "The Internet Audience: Web Use as Mass Behavior," *Journal of Broadcasting and Electronic Media* 46, no. 1 (2002): 1–12.[Ⓢ][#]

Page 68

15. D. A. Ferguson and E. M. Perse, "The World Wide Web as a Functional Alternative to Television," *Journal of Broadcasting and Electronic Media* 44, no. 2 (2000): 155–74.[Ⓢ][#]

16. iProspect, "Search Engine User Behavior Study," http://www.iprospect.com/premiumPDFs/WhitePaper_2006_SearchEngineUserBehavior.pdf [http://www.iprospect.com/premiumPDFs/WhitePaper_2006_SearchEngineUserBehavior.pdf] (accessed January 9, 2007).[Ⓢ][#]

17. R. Koopmans and A. Zimmerman, "Visibility and Communication Networks on the Internet: The Role of Search Engines and Hyperlinks" (paper presented at the CONNEX workshop "A European Public Sphere: How Much of It Do We Have and How Much Do We Need?" Amsterdam, November 2–3, 2005).[Ⓢ][#]

18. A.-L. Barabási and R. Albert, "Emergence of Scaling in Random Networks," *Science* 286, no. 5439 (1999): 509–12; Webster and Lin, "The Internet Audience"; M. Hindman, "A Mile Wide and an Inch Deep: Measuring Media Diversity Online and Offline," in *Media Diversity and Localism: Meaning and Metrics*, ed. P. M. Napoli (Mahwah, NJ: Erlbaum, 2007), 327–48.[Ⓢ][#]

19. Webster and Lin, "The Internet Audience"; Webster and Phalen, *Mass Audience*; J. Yim, "Audience Concentration in the Media: Cross-Media Comparisons and the Introduction of the Uncertainty Measure," *Communication Monographs* 70, no. 2 (2003): 114–28.[Ⓢ][#]

20. Hindman, "A Mile Wide and an Inch Deep."[Ⓢ][#]

21. L. Dahlberg, "The Corporate Colonization of Online Attention and the Marginalization of Critical Communication?" *Journal of Communication Inquiry* 29, no. 2 (2005): 160–80; Koopmans and Zimmerman, "Visibility and Communication Networks."[Ⓢ][#]

22. Lehman-Wilzig and Cohen-Avigdor, "The Natural Life Cycle of New Media Evolution," 707.[Ⓢ][#]

23. Webster and Phalen, *Mass Audience*; Napoli, *Audience Economics*; A. Klaassen, "The Short Tail: How the 'Democratized' Medium Ended Up in the Hands of the Few—at Least in Terms of Ad Dollars," *Advertising Age*, November 27, 2007, 1.[Ⓢ][#]

24. R. Cover, "Audience inter/active: Interactive Media, Narrative Control, and Reconceiving Audience History," *New Media and Society* 8, no. 1 (2006): 139–58.[Ⓢ][#]

25. J. Giuffo, "The Web: Unlock Those Links," *Columbia Journalism Review*, September–October 2002, 9. [#]
26. J. Zittrain, "A History of Online Gatekeeping," *Harvard Journal of Law and Technology* 19, no. 2 (2006): 253–98. [#]
27. Introna, "Shaping the Web." [#]
28. Ibid.; F. Menczer et al., "Googearchy or Googlocracy?" *IEEE Spectrum* 43, no. 2, <http://spectrum.ieee.org/print/2787> [<http://spectrum.ieee.org/print/2787>]. [#]
29. S. L. Gerhart, "Do Web Search Engines Suppress Controversy?" *First Monday* 9, no. 1 (2004), http://firstmonday.org/issues/issue9_1/gerhart/index.html [http://firstmonday.org/issues/issue9_1/gerhart/index.html]. [#]
30. H. W. Park, "Hyperlink Network Analysis: A New Method for the Study of Social Structure on the Web," *Connections* 25, no. 1 (2003): 53. [#]
31. Webster and Phalen, *Mass Audience*; M. McAdams and S. Berger, "Hypertext," *Journal of Electronic Publishing* 6 (2001), <http://www.press.umich.edu:80/jep/06-03/McAdams/pages/> [<http://www.press.umich.edu:80/jep/06-03/McAdams/pages/>]. [#]

Page 69

32. P. Aufderheide, "Competition and Commons: The Public Interest in and after the AOL–Time Warner Merger," *Journal of Broadcasting and Electronic Media* 46, no. 4 (2002): 515–32. [#]
33. D. V. Dimitrova, C. Connolly-Ahern, and A. Reid, "Hyperlinking as Gatekeeping: Online Newspaper Coverage of the Execution of an American Terrorist," *Journalism Studies* 4, no. 3 (2003): 401–14. [#]
34. Gerhart, "Do Web Search Engines Suppress Controversy?" [#]
35. R. Rogers, *Information Politics on the Web* (Cambridge, MA: MIT Press, 2004). [#]
36. Introna, "Shaping the Web"; Koopmans and Zimmerman, "Visibility and Communication Networks." [#]
37. Hindman, "A Mile Wide and an Inch Deep"; Yim, "Audience Concentration in the Media." [#]
38. Koopmans and Zimmerman, "Visibility and Communication Networks." [#]
39. C. Anderson, *The Long Tail: Why the Future of Business Is Selling Less of More* (New York: Hyperion, 2006). [#]
40. Ibid. [#]
41. Napoli, *Audience Economics*. [#]
42. A. Elberse and F. Oberholzer-Gee, "Superstars and Underdogs: An Examination of the Long Tail Phenomenon in Video Sales" (working paper no. 07-015, Division of Research, Harvard Business School, 2007); D. M. Pennock et al., "Winners Don't Take All: Characterizing the Competition for Links on the Web," *Proceedings of the National Academy of Sciences* 99, no. 8 (2002): 5207–11. [#]

43. L. A. Adamic and B. A. Huberman, "The Web's Hidden Order," *Communications of the ACM* 44, no. 9 (2001): 55–59. [\[#\]](#)
44. Neuman, *Future of the Mass Audience*; B. Winston, *Media Technology and Society: A History from the Telegraph to the Internet* (New York: Routledge, 1998). [\[#\]](#)
45. Gerhart, "Do Web Search Engines Suppress Controversy?" [\[#\]](#)
46. B. Wellman, "The Three Ages of Internet Studies: Ten, Five, and Zero Years Ago," *New Media and Society* 6, no. 1 (2004): 123–29. [\[#\]](#)
47. Ibid. [\[#\]](#)
48. Rogers, *Information Politics on the Web*; Park, "Hyperlink Network Analysis," 49–61. [\[#\]](#)
49. M. Tremayne, "The Web of Context: Applying Network Theory to the Use of Hyperlinks in Journalism Stories on the Web," *Journalism and Mass Communication Quarterly* 81, no. 2 (2004): 237–53. [\[#\]](#)

Page 70

The Hyperlink in Newspapers and Blogs

Lokman Tsui

Following are links to the external Web sites mentioned in this article. These sites are not part of the New York Times on the Web, and the Times has no control over their content or availability. When you have finished visiting any of these sites, you will be able to return to this page by clicking on your Web browser's "Back" button or icon until this page reappears. [\[1\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N501;note=ptr)

The hyperlink poses a dilemma for news organizations. On the one hand, links can be very useful in their ability to directly link to source material, such as public reports or official transcripts, in providing support for a news article. Considering that trust in what the people hear, see, and read has been steadily declining since the 1980s, the ability of the hyperlink to link a claim to its source can increase transparency of the news and subsequently restore some of the credibility of the mass media. [\[2\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N502;note=ptr)

On the other hand, news editors may fear to link to Web sites over which the news organization has no control, as the preceding disclaimer from the *New York Times* exemplifies. While the disclaimer itself is no longer used, it does nicely capture an anxiety regarding the clarity of boundaries in the digital space. Yet newspaper editors worried about readers' confusion may also consider that competition with blogs in the "marketplace of attention" [\[3\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N503;note=ptr)

[\[3\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N503;note=ptr) may have made concern about linking moot. Most definitions of blogs include the hyperlink as one of its characteristics, suggesting that bloggers are not at all constrained by the attributional worries that might concern newspaper editors.

These comparisons may seem logical, but it must be said that no research or writing exists on the norms that bloggers or workers at the online divisions of newspaper firms hold toward use of

the hyperlink. In fact, there are few studies of the ways hyperlinks are used by online news organizations in the coverage of news areas, such as politics. Such explorations are almost nonexistent regarding bloggers. The purpose of this

Page 71

essay is to report on a systematic comparison of the ways a sample of leading newspapers and blogs used hyperlinks. My central finding is that while the blogs link heavily to external Web sites, some major newspapers barely link at all, and others link exclusively to themselves. The strategies that explain these findings and their implications for democratic deliberation are topics deserving of further academic and public discussion.

How News Directs Attention

News has always been and is still a crucial means for organizing and directing our attention to valuable information. It distinguishes itself from other forms of public knowledge in its claim to truth. Crafting the news, journalists buttress the claim to truth by relying on the use of factual information. Facts, according to Tuchman, are “pertinent information gathered by professionally validated methods specifying the relationship between what is known and how it is known.” [4] [\[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N504;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N504;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N504;note=ptr) It is this process of sourcing, which includes fact-checking and verification, that defines news vis-à-vis other forms of public knowledge. [5] [\[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N505;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N505;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N505;note=ptr)

However, the process of sourcing has traditionally been problematic in terms of transparency. How do we know whether the journalist really did verify sources properly? Tuchman argues that the notion of objectivity is a crucial strategy journalists developed to establish a relationship of trust with the public. Well known and widely accepted, for example, is the “two source” rule. It stipulates that a journalist has to check with at least two different sources before publishing something as fact.

Professionalism, objectivity, and a code of ethics all factor in the journalist’s strategy in a bid for the public’s attention and trust. These conclusions are drawn from what are considered a set of classic newsroom ethnographies. [6] [\[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N506;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N506;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N506;note=ptr)

Obviously, notions of objectivity and professionalism continue to guide the production of news. However, considering that these ethnographies were conducted decades ago, do they still provide a comprehensive picture of how newsrooms function today? While we don’t know for sure, it is doubtful regarding online news. A crucial difference in the way online news directs our attention is through the use of the hyperlink. The hyperlink allows news providers to suggest which voices are worthy of our attention and which voices are not. The hyperlink also is able to support the facticity of news, because of its inherent ability to specify “the relationship between what is known and how it is known,” simply by providing a link to the source. With over 70 percent of the U.S.

Page 72

population having accessed online news, it becomes paramount to have a better understanding of the production of online news and the role the hyperlink plays in it. [7]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N507;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N507;note=ptr)

How Online News Directs Our Attention

People trust the New York Times and Washington Post and link to them, but there are a huge number of people who are going outside the bounds of traditional media to these new media forms to get their information and, more importantly, to participate in the discussions around news and topics. (David Sifry, blogger and CEO of Technorati) [8] [\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N508;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N508;note=ptr)

Digital network technology has drastically altered the social conditions of speech. [9]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N509;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N509;note=ptr) It has enabled the change from a situation where journalism as a practice is constrained by technology and reserved for a select few to a situation where barriers to publish are lowered to such a degree that Hartley argues that now “everyone is a journalist.” [10]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N510;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N510;note=ptr)

Jenkins similarly describes the rise of what he calls a “convergence culture,” which is blurring the lines between old and new media and is resulting in “a changed sense of community, a greater sense of participation, less dependence on official expertise, and a greater trust in collaborative problem solving.” [11] [\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N511;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N511;note=ptr)

This change in the cultural environment is perhaps best exemplified by the incredible rise in popularity of blogs. Many definitions of blogs point to the notion of a Web site with regularly updated entries, presented in reverse chronological order. Most definitions include the hyperlink as an important and even essential characteristic of what constitutes a blog. [12]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N512;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N512;note=ptr)

Herring distinguishes different genres of blogs, ranging from blogs that function as personal diaries to blogs that link to, comment on, and cover news. [13]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N513;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N513;note=ptr) While most blogs (65 percent) do not make claims to be a form of journalism, they do mention that they sometimes or often practice journalistic standards, such as including links to original sources (57 percent) and spending extra time to verify facts they want to include in their postings (56 percent). [14] [\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N514;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N514;note=ptr) Some research has framed the relationship between bloggers and journalists as adversarial. Others suggest that the question of bloggers versus journalists is over and that the two have a synergistic relationship. [15] [\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N515;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N515;note=ptr)

Lowrey, for example, suggests that a division of labor exists between the two, with bloggers

relying on the work of journalists and taking up what they fail to cover at the same time. Because

Page 73

of a relative lack of institutional constraints, bloggers can afford to be specialized and partisan; to cite nonelite sources; and, in general, to cater to a niche audience. [16]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N516;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N516;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N516;note=ptr)] At a conference panel on blogging, journalism, and credibility, Rosen stated: “One of the biggest challenges for professional journalists today is that they have to live in a shared media space. They have to get used to bloggers and others with an independent voice talking about them, fact-checking them, overlooking them, and they no longer have exclusive title to the press.” [17]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N517;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N517;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N517;note=ptr)]

Clearly, the boundaries of what constitute news are blurring, and we need to have a more inclusive understanding of online news that goes beyond what is offered by the traditional mainstream media. [18] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N518;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N518;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N518;note=ptr)] This sentiment is echoed by Jenkins, who argues that it would be “a mistake to think about either kind of media power in isolation.” [19] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N519;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N519;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N519;note=ptr)]

Phrased in terms of the imperatives of media firms, the question is this: now that news is increasingly being created and read online, how have strategies for gaining public attention and trust adjusted according to the possibilities the Internet as a new medium offers? As a fundamental characteristic of the Internet, the hyperlink stands at the center of this subject.

The Functions of the Hyperlink for Newspaper Sites and Blogs

In its most basic form, the hyperlink makes it possible to connect one Web site to another. Due to its open-ended character, the hyperlink is a simple yet powerful tool that can be employed for many uses. The meaning of the connection is not implemented in the hyperlink itself and must often be inferred from the context. [20] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N520;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N520;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N520;note=ptr)] With regard to the possible functions the hyperlink can take on in online news, we can distinguish between linking for two purposes: citation and reciprocity.

Citation

Perhaps the most classic function of the hyperlink is to use it for citation. [21]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N521;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N521;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N521;note=ptr)] In its ability to connect a claim directly to its source, the hyperlink creates transparency in “the relationship between what is known and how it is known,” something Tuchman has referred to as the defining feature of factual information. Much of the strength of the claim, however, still depends on the credibility of the source it is linked to. This might explain the reluctance to link to external Web sites, since there is no control over

either their content or availability, as the previously quoted disclaimer from the *New York Times* exemplifies. As an existing news organization with an already well-established reputation, linking to less credible, external Web sites might form a threat rather than an opportunity. It becomes paramount to distinguish between internal links, which are considered safe, and external links, which there is no control over. One way to do this is to put a firewall between internal and external links; in practice, this means clearly marking what is internal and external—for example, by adding a disclaimer and clearly positioning the links outside the news article. Another way would be to dispense with external links altogether.

Reciprocity

The second function of the hyperlink is to foster relationships of reciprocity. Blogs in particular seem to depend on a strategy of reciprocity, of exchanging links, to build up both credibility and popularity. When asked by an audience member at the conference “The Hyperlinked Society” what he could do to have his blog mentioned and linked on Jay Rosen’s popular blog PressThink, Rosen, a professor of journalism at New York University, answered that his best bet was to link to his Web site first. [22] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N522;note=ptr>]

Many search engines build on this concept of reciprocity. Measuring the relevance of a Web site through the number of incoming links is the basic idea behind PageRank, a crucial part of the success of Google as a search engine. It is also the basic idea behind Technorati, a search engine that keeps track of what is happening in the blogosphere. It measures which blogs are the most popular by their number of incoming links—by how many other Web sites link to them. The leading political blogs receive well over ten thousand incoming links from other Web sites. This includes such blogs as Michelle Malkin (10,240 incoming blog links) and group blogs, such as the Huffington Post (15,007 incoming blog links) and Daily Kos (11,475 incoming blog links). [23] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N523;note=ptr>]

Incoming links are not just valuable for blogs, however, but also may carry great value for the traditional mass media. The idea of measuring incoming links—the idea behind PageRank and Technorati—is similar to a concept Tuchman has called “the web of facticity.” It is the idea that facts can be supported and validated by other related facts, cross-referencing each other. Tuchman was certainly not referring to the World Wide Web back in 1978, but the idea of a “web” of facticity gains an added layer of meaning in the context of the hyperlink and online news: it is now possible to make the web of facticity explicit through the examination

of the use of hyperlink in news articles. In other words, a journalist is now able to write a story with factual information and directly link the fact to the source, showing the public explicitly how that journalist got to know what she or he got to know. In turn, the story can be validated by other Web sites linking to it.

In addition to considerations regarding audience understanding of the facts of a story or opinion

piece, important commercial concerns regarding reciprocal linking may guide news Web sites and blogs. All newspaper sites and many blogs carry advertising. The price of the ads goes up with the number of people who come to the site and, often, by the time they spend on the site. Newspaper sites consequently have an interest in keeping readers in their territory for as long as possible, and we might assume that external linking would work against that. Bloggers also have an interest in keeping readers, but their desire to rank highly in blog search engines so that people will visit them may lead them to follow Jay Rosen's previously noted advice and link to other bloggers.

Previous Research on Linking

Research on news production in the digital age has been sparse, with little attention being paid to the role of the hyperlink. [24] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N524;note=ptr>]

No writings examine the norms and strategies that the people who edit news or blog sites have toward links. A handful of studies do look at the presence of hyperlinks on newspaper sites. In a study published in 2002, Barnhurst concludes that online newspapers rarely make any use of hyperlinks in news articles, with more than 75 percent having no link at all. [25]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N525;note=ptr>]

Dimitrova and others found in 2003 that the destination of hyperlinks to an external Web site only happened in a stunningly low 4.1 percent of the total number of hyperlinks in newspaper articles. [26] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N526;note=ptr>]

This seems to be in line with the findings of Tremayne, who reports a steady decline in the proportion of external links over the period of 1999–2002. [27] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N527;note=ptr>]

Note, though, that these investigations were conducted during the Web's early years, and the current robust environment for the Internet might have brought changes in newspaper organizations' online procedures.

What about blogs? Contrary to what might be a general sense, by people who follow blogs, that heavy interlinking is widespread, Herring finds that about only half (51.2 percent) of the blogs that she surveyed link to other blogs, with even fewer (36.1 percent) linking to news sites. [28]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N528;note=ptr>]

However, she also notes that this number is likely to be skewed by the

high number of blogs that act as personal diaries, many of which do not link to other Web sites. One might expect the blogs that link to, comment on, and cover news to display a strategy that links heavily to other Web sites. Herring's sample, however, does not include enough of these political blogs to say anything (statistically) significant about what their typical linking pattern would be like.

For this study, I am particularly interested in link patterns of political blogs as an alternative form of online news. Political blogs are interesting because many of them are stars of the

Blogosphere, attracting the most attention. Snirky argues that blogs follow a distribution that closely resembles a power law, meaning a winner-takes-all situation, where a small minority of the total number of blogs gets the majority of attention, while there is a long tail of the remaining blogs that does not get the amount of traffic remotely near those at the top. [29]

[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N529;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N529;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N529;note=ptr)

A significant number of blogs at the high end of the power law distribution consists of these political blogs. Although the size of their audience is not quite comparable to those of the mass media, they are rapidly gaining influence.

Besides anecdotal evidence, however, there has been surprisingly little empirical research looking at link patterns of these leading political blogs. An exception is a study done by Adamic and Glance, who found in 2005 that the top forty political blogs refer to the mainstream media about once every post and referred to other blogs only one post out of ten. [30]

[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N530;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N530;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N530;note=ptr) This result, however, might not be fully generalizable, as it sampled posts during the 2004 presidential election, a time where it is particularly likely for blogs to link to coverage in the mainstream media. That the top forty political blogs linked more often to the mainstream media than to other blogs is particularly striking because these political blogs live and die by the link.

The goal of this study is to address some of the gaps in the literature on hyperlinks. It seeks to answer two sets of questions. First, is the hyperlink used at all in online news; and if so, to what Web sites do they link, in what way, and how often? Second, from the specific ways hyperlinks are used or not used, can we infer strategies regarding editorial control, the desire for high site ranking, and the interest in keeping people on the site?

Study Design and Method

I examined the online editions of four leading newspapers and five leading political blogs. The newspapers selected were the *New York Times*, the

Page 77

Washington Post, *USA Today*, and the *Los Angeles Times*. The five political blogs selected were the Huffington Post, Michelle Malkin, Daily Kos, Crooks and Liars, and Think Progress. These five blogs were listed as the five most popular political blogs by Technorati based on the number of incoming links. [31] [http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N531;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N531;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N531;note=ptr)

The study focused on the coverage of political news in two periods, together making up one full week. The first period was March 1–4, 2007. The second period was March 26–28, 2007. By focusing on two distinct periods, the hope was to limit issues of periodicity bias. Political news was chosen because this type of news provides many opportunities to link to external sites.

The news articles and blog postings were downloaded on March 4, 2007, for the first period and on March 28, 2007, for the second period. Starting from the front page of the politics section for

the newspapers and the front page of the political blogs, all articles and postings were downloaded and saved. This was critical due to the habit of newspapers to put older articles

behind (sometimes locked) archives.

Answering the research questions required a content analysis of the news articles and blog postings. Two units of analysis were used: the article and the hyperlink. This design was chosen to make the content analysis more functional by breaking down the articles into hyperlinks and to aggregate them back again once the analysis was finished. The articles were coded for the following categories: URL, date of story, author, title, and source. Another code sheet was developed to capture the characteristics of hyperlinks. Links were coded for URL, label (the underlined text that is being linked), placement (inside or outside the article body), destination (internal or external Web site), category of the destination Web site (blog, mainstream news site, governmental or other institutional site, and other), and type of content being linked to (text, video, photo, audio, or contact information—e.g., an email address). The destination Web site was coded for four categories: blog, mainstream news, government or other institution, and other.

For the purposes of this study, a blog was defined as a Web site with regularly updated entries, presented in reverse chronological order. A “mainstream news site” was a Web site of any major news organization; when in doubt, a site was coded as other. The category “government or other institution” included any Web site by any government or other major institution, such as Gallup; when in doubt, the Web site again was coded as other. Other destination Web sites who did not fit in any of the other categories were coded as other.

All collected news articles and blog postings were analyzed for content.

Page 78

Hyperlinks were coded insofar as they were deemed relevant to the news article or blog posting. The decision of what was deemed relevant was left to the coder but specifically excluded tags, trackbacks, and comments. Tags were defined as links that are used to categorize the news article, often located outside the main body of the article and internally linked. The justification for exclusion here is that they are not used for the purpose of citation or reciprocity. Trackbacks and comments were excluded to eliminate issues involved with the lack of conformity across newspapers and blogs in offering these two functionalities. To determine intercoder reliability, three news articles or blog postings for each day for each newspaper or blog were randomly selected and coded. The average intercoder reliability was established at 0.97 for the news articles and blog postings and at 0.87 for the hyperlinks, using Krippendorff’s alpha.

Findings

Do leading newspapers and political blogs link heavily? The answer seems to be yes. The total number of articles coded was 806, and the total number of links was 3,876, with a mean number of 4.8 links per article. Two newspapers, the *Los Angeles Times* and *USA Today*, were exceptions to this. *USA Today* had, despite the highest number of articles, only a little more than one link per article. The *Los Angeles Times* had even fewer links, on average only one link per three articles. The political blogs and the other two newspapers, the *Washington Post* and the *New York Times*, all linked frequently in their political news articles. Surprisingly,

TABLE 1. Frequency of Links

Source	Number of Links	Number of Articles	Mean Number of Links Per Article
<i>LA Times</i>	11	37	0.3

<i>USA Today</i>	196	180	1.1
Think Progress	424	117	3.6
Crooks and Liars	356	76	4.7
<i>Washington Post</i>	429	73	5.9
Michelle Malkin ¹	93	30	6.4
Huffington Post	1,114	163	6.8
Daily Kos	726	89	8.2
<i>New York Times</i>	427	41	10.4
Total	3,876	806	4.8

Page 79

it was not a blog but the *New York Times* that linked the most of all, with more than ten links per article on average.

Do the leading newspapers and political blogs link to external Web sites? Here is a stark difference between the newspapers and the blogs in this study. The political blogs all linked heavily and also linked heavily to external Web sites. More than a third of the links of the Huffington Post and Daily Kos and over three-quarters of the links of Think Progress and Michelle Malkin pointed to external Web sites. This is in sharp contrast with the newspapers. While both the *Washington Post* and the *New York Times* linked heavily in their news articles, they linked almost exclusively to themselves. Less than 1 percent of the links in the political news articles of the *New York Times* pointed to external Web sites, while only 3 percent of the links in the political news articles of the *Washington Post* did so.

How many links are placed outside the main body of an article? The leading newspapers all placed well over half of their links outside the main body. The leading political blogs, however, seemed to exclusively place their links within the body. The exception was the Huffington Post, which placed well over two-thirds of its links outside the main body. How many links placed outside the main body of the article also point to external Web sites? Here the picture is very clear: practically none of those in this study were linked to external Web sites.

Finally, when blogs link externally to Web sites, to what kind of Web sites do they most often link to? The newspapers were here omitted from this analysis, considering that only the blogs linked to external Web sites. In roughly one-third of the cases, the political blogs linked to other blogs, with Michelle Malkin (42.4 percent) and Crooks and Liars (47.5 percent)

TABLE 2. Destination of Links

Source	Number of Links to External Web Sites	Number of Links	Percentage of Links to External Web Sites
<i>New York Times</i>	3	427	0.7
<i>USA Today</i>	4	196	2.0
<i>Washington Post</i>	13	429	3.0
<i>LA Times</i>	3	11	27.3

Huffington Post	395	1,114	35.5
Daily Kos	273	726	37.6
Crooks and Liars	202	356	56.7
Think Progress	320	424	75.5
Michelle Malkin	151	193	78.2
Total	1,364	3,876	35.2

Page 80

TABLE 3. Placement of Links

Source	Number of Links Outside Article	Number of Links	Percentage of Links Placed Outside Article	Percentage of Links Placed Outside Article to External Web Sites
New York Times	225	427	52.7	0.0
USA Today	181	196	92.4	1.0
Washington Post	277	429	64.6	0.0
LA Times	7	11	63.6	0.0
Huffington Post	766	1,114	68.8	0.0
Daily Kos	0	726	0.0	0.0
Crooks and Liars	0	356	0.0	0.0
Michelle Malkin	9	193	4.7	0.0
Think Progress	0	424	0.0	0.0
Total	1,465	3,876	37.8	0.0

TABLE 4. Categories of Externally Linked Web Sites by Percentages

Source	Blog	News	Gov.	Other
Michelle Malkin	42.4	27.8	4.6	25.2
Crooks and Liars	47.5	16.8	6.0	29.7
Think Progress	26.9	42.5	10.3	20.3
Daily Kos	28.6	38.1	12.1	21.3
Huffington Post	28.9	21.8	7.6	41.8
Mean	34.9	29.4	8.1	27.7

in particular being fond of linking to other blogs. The blogs also link frequently to mainstream news Web sites, including those of the *Washington Post* and the *New York Times*. Crooks and Liars (16.8 percent) linked the least to mainstream news Web sites, while Think Progress (42.5) linked the most frequently to mainstream news Web sites. A moderate number of the links were directed toward governmental or other institutional Web sites. Michelle Malkin (4.6 percent) linked the least frequently to such sites, while Daily Kos (12.1 percent) linked to them the most often. Finally, a fair share of their links went to other Web sites that did not fit into one of the three categories (blogs, mainstream news sites, and governmental or other institutional sites).

Page 81

Discussion

J. D. Lasica, a media critic, blogger, and citizen media expert, has lamented the sparse use of the hyperlink by journalists.

Equally important—and still underused, in my view—is the ability to link to source materials, transcripts, public records and other original documents to buttress an article’s reporting. In this age of public mistrust of the media, such steps enhance a news organization’s credibility. In my freshman year at college my journalism professor told us that the first rule of good journalism is: Show, don’t tell. So: Don’t tell readers to trust you. Show them the goods. [32]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N532;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N532;note=ptr)

Is Lasica’s lament valid? The findings of this study show that it is a mixed bag for the leading newspapers. While the *Los Angeles Times* and *USA Today* still do not rely on the hyperlink much, the *Washington Post* and, in particular, the *New York Times* certainly do not underuse the hyperlink, as this study has shown that they link heavily in their news articles. However, Lasica is correct in his sense of the lack of use of the hyperlink by journalists to source original material, or what he refers to as “showing the goods.” Even though the *Washington Post* and the *New York Times* link heavily, they also only exclusively link to themselves.

The leading newspapers do not use the hyperlink in their political news coverage for the purpose of citation. This is particularly unfortunate given the nature of political news, which affords many opportunities to link to external sources. Blogs, by contrast, link heavily and also link heavily to external Web sites. But how do we know that they use the hyperlink for purposes of citation and not for reciprocity? In cases when the political blogs link externally to other blogs, we cannot tell for sure. A blog might link to another blog to back up a claim but might also do this in the hope that the other blog will link back. However, the political blogs also link frequently to mainstream news Web sites. In this case, we can be pretty sure that the political blogs link for the purpose of citation. As this study has shown, there is zero expectation that the mainstream news Web sites will actually consider linking back to the blogs.

So why are newspapers reluctant to link to external Web sites? Although further research on the institutional processes behind online news production is needed, the findings of this study

add support for several hypotheses that seek to explain the lack of external links. First, the study suggests little support for the hypothesis that the reluctance to link to external

Web sites results from fear of losing control because it might threaten credibility. In the past, as the *New York Times* disclaimer has shown, links to external sites were often placed outside the main body of the news article; nowadays, both the disclaimer and the links to external sites are gone. But there really is no particular reason to hesitate linking externally for fear of losing credibility, as there are many credible Web sites that could be linked to. Pointing to press releases from the White House Web site might be useful to readers, for example. The fear of losing control, then, might be because of gatekeeping purposes. Dimitrova and others have previously suggested this second hypothesis. [33] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N533;note=ptr>]

A third reason sometimes mentioned for mainstream news organizations' slowness to pick up on the potential of new technology, such as the hyperlink, points to technical or organizational inertia. The idea is that it takes time to get used to the new online environment. But inertia is clearly not the reason for news organizations' reluctance or even refusal to link to external Web sites. Tremayne has shown a clear decline in the number of external links over the years, [34] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N534;note=ptr>] and the findings of this study confirm this trend. Ironically, it seems that the more comfortable newspapers grow with the Web, the more inclined they are not to link to external Web sites and, instead, to link only to themselves.

That leaves a fourth possible suggestion for the lack of external links: newspapers' fear of losing advertising revenues by sending people out of their sites. It seems a reasonable possibility, though more research is needed to make this claim definitive. In general, more work is needed to validate the findings presented here and to determine why newspapers virtually ignore the use of links for citations. The point is not merely an academic one. In view of the importance of the *New York Times*, the *Washington Post*, *USA Today*, and the *Los Angeles Times* online as well as offline, the way these news organizations draw attention to and verify ideas ought to be a topic of concern to anyone interested in expanding the quality of democratic discourse in the digital age.

Notes

I want to extend my gratitude to Dr. Joseph Turow for his guidance and valuable feedback and for giving me the opportunity to work with him. I also want to thank Brigitte Ho and Anne-Katrin Arnold for their insightful comments and help with coding. All errors in this essay remain, of course, mine.

1. Quote from *New York Times* disclaimer retrieved from Mark Glaser, "Open Season: News Sites Add Outside Links, Free Content," *Online Journalism Review*, October 19, 2004, <http://www.ojr.org/ojr/glaser/1098225187.php> [<http://www.ojr.org/ojr/glaser/1098225187.php>] (accessed March 23, 2007). [1]

2. Pew Research Center for the People and the Press, *Online Papers Modestly Boost*

Newspaper Readership, 2006, <http://people-press.org/reports/display.php3?PageID=1069> [<http://people-press.org/reports/display.php3?PageID=1069>] (accessed March 23, 2007).[Ⓢ] [#]

3. See the essay by Webster in this book.[Ⓢ] [#]

4. G. Tuchman, *Making News* (New York: Free Press, 1978), 82.[Ⓢ] [#]

5. B. Kovach and T. Rosenstiel, *The Elements of Journalism: What Newspeople Should Know and the Public Should Expect* (New York: Crown, 2001).[Ⓢ] [#]

6. Tuchman, *Making News*; H. J. Gans, *Deciding What's News: A Study of "CBS Evening News," "NBC Nightly News," "Newsweek," and "Time"* (New York: Pantheon, 1979); M. Fishman, *Manufacturing the News* (Austin: University of Texas Press, 1980).[Ⓢ] [#]

7. For the latest information on how many people use online news, see Project for Excellence in Journalism, *The State of the News Media 2007*, http://www.stateofthenewsmedia.org/2007/narrative_online_audience.asp?cat=2&media=4 [http://www.stateofthenewsmedia.org/2007/narrative_online_audience.asp?cat=2&media=4] (accessed April 23, 2007).[Ⓢ] [#]

8. Quote by David Sifry cited in J. D. Lasica, "Transparency Begets Trust in the Ever-Expanding Blogosphere," *Online Journalism Review*, August 12, 2004, <http://ojr.org/ojr/technology/1092267863.php> [<http://ojr.org/ojr/technology/1092267863.php>] (accessed March 24, 2007).[Ⓢ] [#]

9. J. M. Balkin, "Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society," *New York University Law Review* 79, no. 1 (2006), <http://www.law.nyu.edu/journals/lawreview/issues/vol79/no1/NYU101.pdf> [<http://www.law.nyu.edu/journals/lawreview/issues/vol79/no1/NYU101.pdf>].[Ⓢ] [#]

10. J. Hartley, "Journalism as a Human Right: A Cultural Approach to Journalism," in *Journalism Research in an Era of Globalization*, ed. M. Loeffelholz and D. Weaver (London: Routledge, 2005), 39–51. Also see D. Gillmor, *We the Media: Grassroots Journalism by the People, for the People* (Sebastopol, CA: O'Reilly, 2006).[Ⓢ] [#]

11. H. Jenkins, *Convergence Culture: Where Old and New Media Collide* (New York: New York University Press, 2006), 208–9.[Ⓢ] [#]

12. R. Blood, *The Weblog Handbook: Practical Advice on Creating and Maintaining Your Blog* (Cambridge, MA: Perseus, 2002).[Ⓢ] [#]

13. S. C. Herring et al., "Conversations in the Blogosphere: An Analysis 'from the Bottom Up,'" in *Proceedings of the 38th Hawaii International Conference on System Sciences* (Los Alamitos: IEEE Press, 2005), 107–18.[Ⓢ] [#]

14. A. Lenhart and S. Fox, *Bloggers: A Portrait of the Internet's New Storytellers* (Washington, DC: Pew Internet and American Life Project, 2006), <http://www.pewinternet.org/pdfs/PIP%20Bloggers%20Report%20July%2019%202006.pdf> [<http://www.pewinternet.org/pdfs/PIP%20Bloggers%20Report%20July%2019%202006.pdf>] (accessed April 24, 2007).[Ⓢ] [#]

15. J. Rosen, "Bloggers vs. Journalism Is Over," PressThink, January 21, 2005, http://journalism.nyu.edu/pubzone/weblogs/pressthink/2005/01/21/berk_essay.html [http://journalism.nyu.edu/pubzone/weblogs/pressthink/2005/01/21/berk_essay.html] (accessed April 24, 2007). Also see J. D. Lasica, "Blogs and Journalism Need Each Other," *Nieman Reports* 57.

no. 3 (2003): 70–74. [link icon] [#]

16. W. Lowrey, "Mapping the Journalism-Blogging Relationship," *Journalism* 7, no. 4 (2006): 477–500. [link icon] [#]

17. Quote by Rosen cited after R. MacKinnon, "Blogging, Journalism, and Credibility," *Nation*, March 17, 2005, <http://www.thenation.com/doc/20050404/mackinnon> [http://www.thenation.com/doc/20050404/mackinnon] (accessed April 24, 2007). [link icon] [#]

18. J. B. Singer, "Who Are These Guys? The Online Challenge to the Notion

Page 84

of Journalistic Professionalism," *Journalism* 4, no. 2 (2003): 139–63; M. A. Deuze, "The Web and Its Journalisms: Considering the Consequences of Different Types of Newsmedia Online," *New Media and Society* 5, no. 2 (2003): 203–30. [link icon] [#]

19. Jenkins, *Convergence Culture*, 212. [link icon] [#]

20. For the sake of simplicity of argument, I am omitting developments with regard to the Semantic Web, which seek to add layers of meaning onto the hyperlink. [link icon] [#]

21. See the essay by Halavais in this book. [link icon] [#]

22. Videotaped panel discussion at the conference "The Hyperlinked Society," Annenberg School for Communication, University of Pennsylvania, June 9, 2006, available at http://appcpenn.org/HyperlinkedSociety/download/HyperLinked_Panel1_Full.wmv [http://appcpenn.org/HyperlinkedSociety/download/HyperLinked %20Panel1_Full.wmv]. [link icon] [#]

23. Retrieved from Technorati as of April 4, 2007. See <http://www.technorati.com/pop/blogs/> [http://www.technorati.com/pop/blogs/]. [link icon] [#]

24. P. J. Boczkowski, *Digitizing the News: Innovation in Online Newspapers* (Cambridge, MA: MIT Press, 2004); E. Klinenberg, "Convergence: News Production in a Digital Age," *Annals of the American Academy of Political and Social Science* 597, no. 1 (2005): 48–64. [link icon] [#]

25. K. Barnhurst, "News Geography and Monopoly," *Journalism Studies* 3, no. 4 (November 2002): 477–89, http://www.ksg.harvard.edu/presspol/research_publications/papers/working_papers/2002_2.pdf [http://www.ksg.harvard.edu/presspol/research_publications/papers/working_papers/2002_2.pdf]. [link icon] [#]

26. D. V. Dimitrova et al., "Hyperlinking as Gatekeeping: Online Newspaper Coverage of the Execution of an American Terrorist," *Journalism Studies* 4, no. 3 (2003): 401–14. [link icon] [#]

27. M. Tremayne, "News Websites as Gated Cybercommunities," *Convergence* 11, no. 3 (2005): 28–39. [link icon] [#]

28. Herring et al., "Conversations in the Blogosphere." [link icon] [#]

29. C. Shirkey. "Power Laws, Weblogs, and Inequality," in *Reformatting Politics: Information Technology and Global Civil Society*, ed. J. Anderson, J. Dean, and G. Lovink (London: Routledge, 2006), 35–42. [link icon] [#]

30. L. A. Adamic and N. Glance, "The Political Blogosphere and the 2004 U.S. Election: Divided

They Blog,” in *Proceedings of the 3rd International Workshop on Link Discovery* (New York: ACM, 2005), 36–43. [#]

31. Retrieved from Technorati as of April 4, 2007. See <http://www.technorati.com/pop/blogs/> [<http://www.technorati.com/pop/blogs/>]. [#]

32. J. D. Lasica, “How the Net is Shaping Journalism Ethics,” July 2001, <http://jdlasica.com/articles/newsethics.html> [<http://jdlasica.com/articles/newsethics.html>] (accessed March 24, 2007). [#]

33. Dimitrova et al., “Hyperlinking as Gatekeeping.” [#]

34. Tremayne, “News Websites as Gated Cybercommunities.” [#]

Page 85

The Role of Expertise in Navigating Links of Influence

Eszter Hargittai

In this essay, I focus on how the influence of links may be mediated by the skills and expertise that both content producers and viewers are able to mobilize when using the Internet. My main argument is that while lots of factors influence how links are presented on the Web and how users respond to the content that shows up on their screens, people’s Internet user abilities remain an important and understudied aspect of navigating links of influence. Both content creators and content users (readers, listeners, viewers) can benefit from a more in-depth understanding of how the Web works. Since such skills are not randomly distributed among the population, certain content providers and content users stand a better chance of benefiting from the medium than others. Relevant know-how will help producers attract attention to their materials. Savvy about the medium will assist users in sidestepping potentially misleading and malicious content.

Links’ control over what people see is less of a factor in the online behavior of savvy users than it is with those who know less about the Internet. Knowledgeable users know how to interpret various types of links and are able to approach information seeking in a myriad of ways. While some people are considerably dependent on what content is presented to them by aggregators and content providers, others can sidestep many supply-side decisions by turning to alternative ways of browsing the Web’s vast landscape. Both provider and seeker have the potential to influence which links will matter to any particular user’s experience in the course of a particular information-seeking incident or when confronted with particular content. My main argument is that the weight of how much of this relationship is influenced by the provider versus the user shifts based on the savvy of actors at both the supply and demand sides of the equation.

I start the essay by discussing why links matter and the main types of links that exist on the Web, including a brief consideration of how the

Page 86

presentation of sponsored search engine results has changed over time. In the first section, I also consider the types of manipulations that content presenters can employ in order to attract more attention than would otherwise be possible. Then I introduce the concept of user skill,

providing examples of what we know regarding people's Internet uses in order to argue that expertise is an important component of how user attention is allocated to online content and how people navigate links of influence. I end by discussing what questions remain about predictors of user savvy and the type of research that would be helpful in answering them.

Why Links Matter

From the early days of the Web, hyperlinks have allowed users to move from one page to another, finding content either with intent or through serendipity. While there are other means of getting to material on the Web, links remain an important way for users to move around online, whether within a known site or by venturing to new destinations. [1]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N601;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N601;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N601;note=ptr)] Links are important precisely because they allocate user attention. They can have both positive effects and negative ones. By driving much needed eyeballs to material, they can spread updates about important health matters, draw attention to significant political issues, encourage people to donate to a cause, or help small businesses and independent artists thrive through sales of items that would not otherwise have the chance of garnering attention were it not for the low cost of online presentation.

But links can also have negative consequences. Too much popularity can overwhelm a system and make the material at least temporarily inaccessible. More important, drawing audiences to unsubstantiated rumors can lead to harmful outcomes in people's lives. Links can compromise relationships, personal and professional. An article in the *Washington Post* reported on an incident that damaged a recent law school graduate's career advancement. [2]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N602;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N602;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N602;note=ptr)] Some negative comments left on a message board by anonymous commentators about a candidate showed up prominently when users did a search on the candidate's name. Employers are turning to the Web to gather information about applicants, so having negative comments show up high on the result list when searching on a particular name can have significant repercussions. [3] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N603;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N603;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N603;note=ptr)]

To counter such incidents, one can now turn to a whole new set of professionals to help achieve desirable rankings on search engines. Experts in search engine optimization (SEO) work with both businesses and individuals

to maximize the chances of a good position on search engine results pages. Interestingly, much of the advice given by such professionals is of the kind that a somewhat more nuanced understanding of how the Web works makes relatively simple to implement. This is one area where the importance of online skill comes into play from the perspective of content providers. Those who know more than others about how to achieve prominent exposure can respond to situations like the one just described relatively quickly and at low cost. The perceived influence of links has jump-started a new profession centered on the idea that organizations and individuals need help and are willing to pay to improve the positioning of links that pertain to

them.

Link Types and Manipulation

Links matter in a broader sense, beyond direct issues of corporate or personal reputation. To understand how, it is important to highlight the many ways in which we can categorize links from their location on a page to their source, from attached financial incentives to design principles. Technically speaking, all hyperlinks are created equal. They can be easily inserted into any page with the simple code `text or image`. At the same time, the potential of links to influence users' actions differs based on the way they are actually used. Consequently, a discussion of how a particular type of link relates to content presentation and user activity is worth consideration.

Of course, there are several ways one can arrive at a Web page without clicking on a link; these include, for example, using a bookmark or favorites listing or typing a URL in the location bar of the browser. [4] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N604;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N604;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N604;note=ptr)] A common form of moving from page to page, however, does involve clicking on a link. The simplest type of link is one that connects to additional information about a detail in some text that constitutes the main content on a page. There are also links whose main purpose is to facilitate navigation. They are not part of core content on a page. Rather, they exist solely to guide people to a destination. These links range from directory categories on large portal sites, such as Yahoo, to sidebar menus on Web sites of all sizes and complexity. These two types of links share one feature: for the most part, they are a relatively steady part of the site on which they are located. Obviously, pages can be edited easily, and links may change as a result. But these kinds of links have fairly stable positions, and producers of these sites maintain a say over their specific placement.

Page 88

In a substantively different category are links that show up on aggregator and recommender sites. These links are not based on one content producer's decisions. Rather, placement is determined by the link's popularity among users. Sites such as Digg and Reddit are examples of this presentation and organization. Any registered user can submit a link that then gets added to the pool of sites made available for users to browse. If enough site members support the link and it gains popularity relative to other submissions, it makes it onto the cover page of the site and garners increasing amounts of attention. These links are not stable the way the previous set of links are. Rather, their position and potential to be clicked changes rapidly with input from users. Thus, while visiting Reddit one minute will yield a certain link list, revisiting it a few minutes later will result in a different set of links.

Another category of links is comprised of those on search engine results pages. Here, the main purpose of the page is to redirect the user to content elsewhere. Such links depend on the proprietary algorithms used by search engine companies to rank pages. Results may be based on relevance and quality—however these two concepts are understood in a given context—but they may also be dependent on financial considerations. Search engines sometimes sell

prominent placement on their results pages. Some search engine companies, like Google and Yahoo, also have systems set up where players large and small can bid for placement on their ad link section. Those links can usually be found on a sidebar next to the unsponsored (“organic”)

search results, although they are occasionally also included within the organic listings.

Another form of sponsored links tied to search results shows up on a plethora of Web sites that have affiliations with ad placement programs offered by ad-serving companies, like Google and Yahoo. These ad links appear on sites across the Web covering numerous topics targeted at diverse communities of users. There is no standard for where they are placed. They can be embedded within the main body of text on a page or on the sidebar, depending on the preferences of the publisher of the page. It is customary for these ads to be accompanied by a note that identifies them as such, but this information is not always clearly visible.

Are such sponsored links ever effective in gaining users' attention? Evidence suggests that they are. One of the most successful Internet companies, Google Inc., has launched numerous products over the years, only very few of which have been profitable to date. One of its most important products is the AdWords program that supplies links to affiliates. Each time someone clicks on such a link, both the owner of the Web site and

Page 89

Google itself, as ad system provider, make money. Without people clicking on such links regularly, the company could not have achieved the revenue stream it has.

Whether users are clicking on these links because they are the most relevant for their needs is another matter. Layout and context of the links can, at times, be confusing or outright deceiving. Some sites display ads very clearly and mark them as such. Others are not as forthcoming about the source and reasons for the links. Take, for example, the case illustrated in figure 1. The Web site featured in this illustration focuses on photo editing. In a prominent place on its welcome page are some smaller images with links right below them. The links are ads, in this case from Yahoo's ad network. However, this is not immediately obvious. Looking at the rightmost picture, one notices an image of dishes, and the link below this picture states "San Francisco Dish." Clicking on the link, despite appearances, has nothing to do with the image of dishes displayed on the page. Rather, the link goes to an advertisement for an American Express program. The images are randomly rotated in what seems to be an effort to entice clicks despite little connection between the images and the links below them.

As suggested by the earlier examples, search engines play a special role in allocating user attention to links and thus online content, given that they are some of the most popular destinations by users. [5] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N605;note=ptr>]

Over time, there has been a considerable amount of change in how links are included and presented on search engines. John Battelle does a nice job of tracing the history of changing search engine results pages. [6] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N606;note=ptr>]

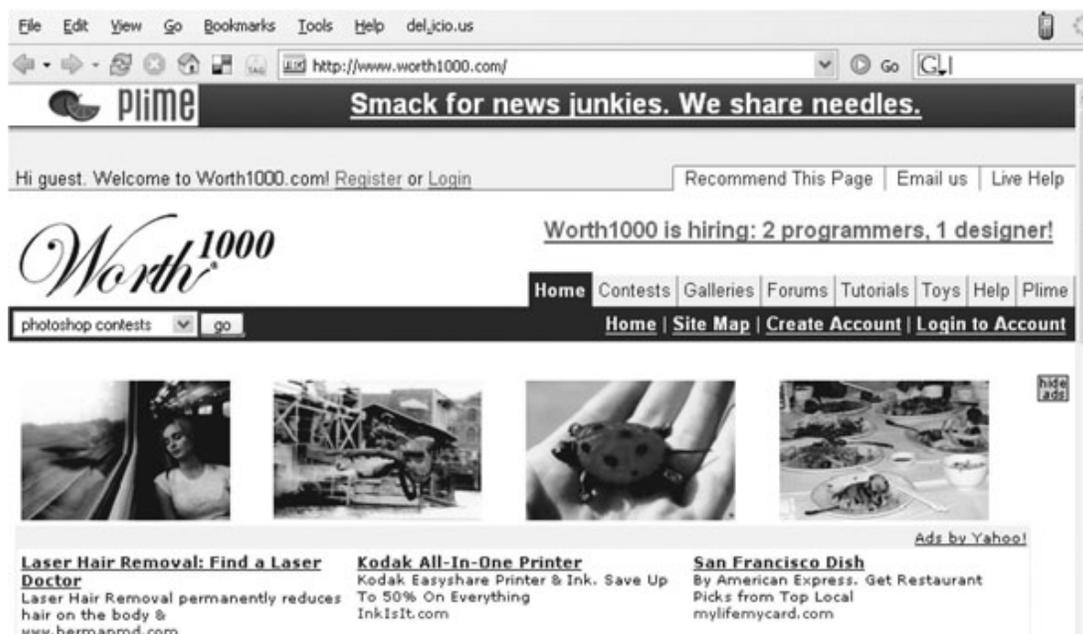
Initially, search engines just brought up sites that included at least one of the search terms entered by the user. As the Web grew, the default Boolean operator "OR" was replaced by "AND," resulting in search engines now returning results that contain all terms in a user's query. Changes also occurred in the financial domain of searching. Goto.com was the first search engine to allow payment for search positioning. These practices of the service were quite explicit. The amount of money the featured link sponsor would pay upon a click by the user was made public and listed right next to the link. Figure 2 depicts a screen shot taken on June 6, 2001, during the

online browsing actions of a forty-one-year-old woman using Goto.com for searching. [7]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N607;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N607;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N607;note=ptr)] Note the cent amounts next to the links. This example shows results to a search query for the phrase “lactose intolerance.” The top advertiser was willing to pay thirty cents per click. Then there is a sharp drop, with the following links going for seven, six, five, and four cents, respectively. This explicit manipulation of search engine results caused

Page 90



[[/d/dculture/images/5680986.0001.001-00000003.jpg](http://d/dculture/images/5680986.0001.001-00000003.jpg)]

Fig. 1.

Example of ad links presented in a confusing manner at www.worth1000.com (2007)

considerable stir in the industry. Ironically, later manifestations of sponsored links have included even less explicit mention of what the advertisers may have done to achieve their products' ranking. Despite the initial resistance by many, this practice has become commonplace across search engines.

What determines which links feature prominently on results pages? Detailed information about search engine rankings is proprietary information, so it is difficult to answer this question. [8]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N608;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N608;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N608;note=ptr)]

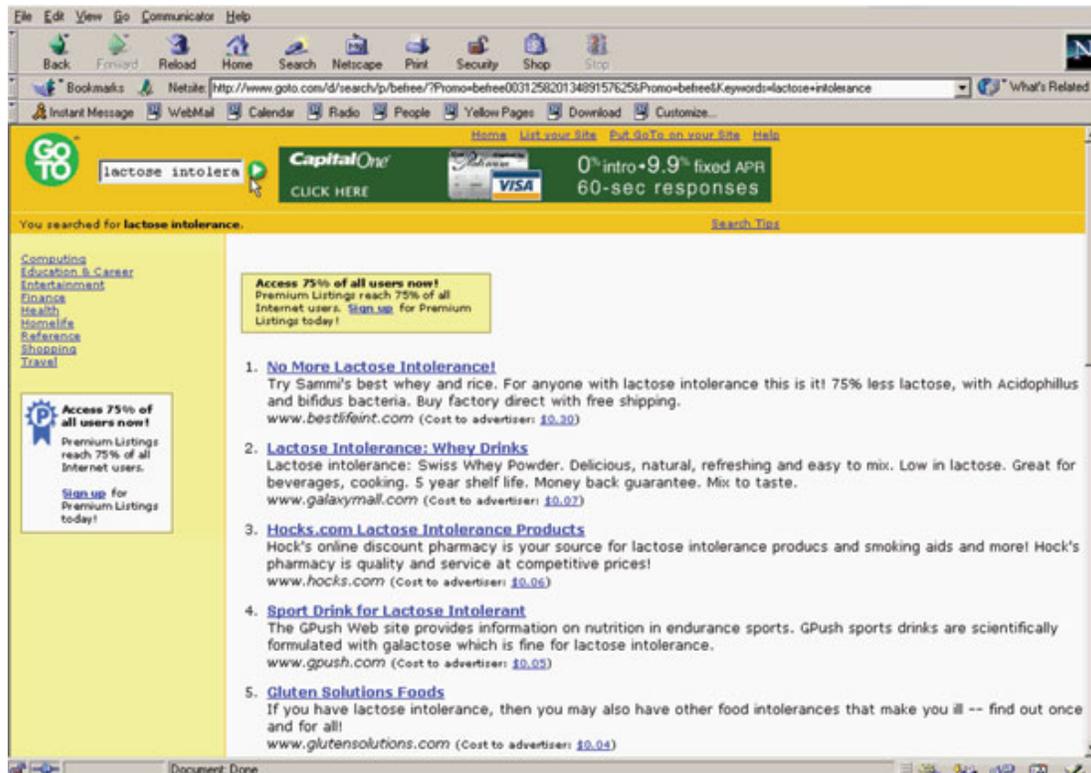
However, there are some generally understood factors that influence rankings, and this is precisely the type of know-how on which the SEO industry has been built. At the most basic level, search engines rely on programs to crawl the Web to create an index of Web site content. [9]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N609;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N609;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N609;note=ptr)] When a query is submitted to a search engine, the service returns sites that include the requested

terms and possibly considers whether the specified terms are in the title or in various tags (underlying information about the page file), possibly with attention to their position on the page. Of course, in most cases, there are numerous pages that meet these criteria. Search engines use additional information to rank results. An important factor, introduced in the late 1990s by Google founders Sergey Brin and Larry Page, concerns the reputation of the page on the Web. [10] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N610;note=ptr>]

Page 91



[/d/dculture/images/5680986.0001.001-00000004.jpg]

Fig. 2.

Screen shot of a Goto.com search engine results page (2001)

To explain the basic idea behind this reputational system, I will draw on an analogy. Imagine a classroom full of students. Each student is liked by some people, and each student, in turn, likes some other students. Let us assume that Brigid is the most popular student, because most people in the class like her. There are two students who are also liked by quite a few students: Sam and Jamie both get the affection of several classmates, although not as many as Brigid. While Brigid is friends with Sam, Brigid does not care much for Jamie, and this is widely known, since she rarely socializes with Jamie. If an outsider came into the classroom and asked a student whether she should befriend Sam or Jamie, most students would likely suggest Sam. The reason is that although Sam and Jamie are liked by the exact same number of people, Sam is also liked by the most appreciated student in class, Brigid. A vote of confidence from Brigid

plays an important role in the evaluation of the students in the context of a larger group. Now, let us replace the students in this story with Web pages, the sentiment of liking a person with a link going from one page to another. If we thus translate the story to Web pages and search engine rankings,

Page 92

the main idea is that having many links pointing to you and especially having ones from popular, established and well-regarded sites is valuable (these aspects of a site would, again, be determined based on some of the linking features of the site).

Search Engine Manipulations

Knowing that linking is important to search engine rankings, it is possible to engage in practices that may help boost a site's position on a results page. There are various ways in which content producers and distributors can influence the amount of attention their content manages to attract online. Many of these concern the manipulation of search engine rankings. The goal is to drive traffic to one's Web site, and this is often done without any regard to the needs of users who may then end up on the page.

The term "Google bombing" refers to the practice of manipulating search engine results by aggressively targeting links to a specific site with the same anchor text where the anchor text refers to the text that links to another page. Several such movements have been documented over the years. Bar-Ilan analyzed some of the most popular ones and identified their sources to be varied, ranging in motivation from personal (e.g., for people with common names wanting to be the first result in response to their names) to political (e.g., links to a page denying the existence of the "Arabian Gulf" despite the use of that name by some for the "Persian Gulf"), humorous (e.g., a search for "French political victories" yielding a link to a spoof search engine page on "French military defeats"), or financial. [11] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw:g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N611;note=ptr>] Users achieve surprisingly high rankings for specific sites in these cases by organizing a movement of people linking to a specified page using a particular term as the anchor text. If the Google bomb is successful, future searches on the anchor text will yield the page that was being targeted by this effort.

While many Google bombs have a larger social or political purpose, some are much less controversial and simply target the popularization of a private individual's ranking on the search engine. For example, freelance journalist and photographer David Gallagher decided in 2002 that he wanted his site to have the top spot in the results listings in response to a search on his name. [12] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw:g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N612;note=ptr>] This was not a trivial goal, given that many people share his name, including a Hollywood actor. Nonetheless, in a few months, he achieved his goal and remained in the top spot for three years, occupying the second position as of this writing. [13] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw:g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N613;note=ptr>]

Page 93

Mobilizing many people to help out with a Google bomb requires a convincing story to motivate

participants. Political or humorous motives seem to work well. Commercial ones from which only a handful of people or entities benefit are less likely to gain wide popularity; in such a case, boosting a site's rankings is left to the actions of just a few people. This is where sites like splogs come in. Splogs, or "spam blogs," are Web sites that include nothing but links with one of two purposes: either they are filled with revenue-generating links, or they feature links to a site with the same goal as the links just described in the Google bombing scenario. The sole purpose of these sites is to come up high on search engine results and then make money by getting people to click on revenue-generating links.

Search engines have been vulnerable to such practices. Google often lists splogs prominently on its results pages, including in the top ten results. For example, at the time of this writing, a search on the words "origami tulip" yields a link to <http://www.origamitulip.com> in the top ten results on Google but not on any of the other top three engines. Curiously, however, there is no material on this Web site that directly addresses folding paper into tulip shapes. Instead, the page is completely made up of links that point off-site. This is precisely the type of site that has no original content (at the time of this writing) and simply contains links pointing elsewhere.

Staying ahead of such empty and confusing content is a cat-and-mouse game between spammers and search engines. However, while search engines catch up with the imaginative, ever-evolving approaches of spammers, users are caught in the middle, having to deal with the resulting confusion. One approach used by spammers is setting up for-profit sites that mimic government sites but use the suffix ".com" rather than ".gov" in URLs, as in "whitehouse.com" instead of "whitehouse.gov." Many users do not understand the distinction between different top-level domain names (here ".com" versus ".gov") and thus are vulnerable to clicking on the wrong link when faced with several seemingly interchangeable options. Analyzing the methods by which users find tax forms, I found that many are derailed and confused by profit-making ventures that claim to assist with tax forms but, in the end, do not include relevant information. [14] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N614;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N614;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N614;note=ptr)

Whether splogs and other such sites continue to mislead users is a question of how well search engines and other aggregators can stay ahead of such malicious practices, in addition to what extent users understand such practices. A paper looking at the source of spam redirection content found that just a few sites are responsible for a large portion of spam content. [15]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N615;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N615;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N615;note=ptr)

Ironically, the Google-owned free blog-hosting site Blogspot appears to be one of the most spam-infested sites, hosting thousands of splogs. In a related realm, people (or, often likely, automated robots or programs) leave strategic comments on blogs to drive traffic and rankings to their sites. When a user leaves a comment on a blog, the username is often linked to a site specified by the user. In this case, the spammer includes a link to the site that is being promoted. Many of the splogs previously mentioned gain popularity precisely through this

practice. Once a splog is set up, the next step is to create links to it by leaving comments on legitimate blogs with good search engine rankings, so as to boost the splog's reputation.

Whether vying for people's attention as the provider of information or looking for the most relevant material to meet one's needs as a user, links are at the forefront of how user attention is allocated to content on the Web. Consequently, exploring how users interpret and approach them is crucial for a better understanding of how attention is allocated online, why some content gets audiences while other content does not, and why some people are better than others at finding content of interest to them. This is an area that has only begun to be investigated. My research and studies by others suggest that users differ with respect to their know-how about the Internet, the sources of various links, and the motivations behind their placements. To get a feel for the nature and importance of what people do and do not know about hyperlinking, it is useful to explore the topic through three categories: general user savvy, users' understanding of search engine rankings, and users' understanding of links in e-mails.

General User Savvy

Based on data I have gathered over the years, it is clear that people differ considerably in their understanding of various Internet-related terms and activities, and these abilities are not randomly distributed across the population. Here, I will draw on various studies to illustrate these differences. Based on surveys administered to hundreds of mostly first-year college students at a diverse urban public research university in the winters of 2006 and 2007, I found that even members of the wired generation are not necessarily savvy about terms that are important for informed Internet

Page 95

use and understanding links in particular. [16] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N616;note=ptr>] While most students exhibit a relatively high level of familiarity with mainstream terms, such as *spam* and *bookmark*, know-how is much lower when it comes to terms relating to more recent Web developments, such as *widget* and *malware*. Moreover, this knowledge is not randomly distributed. Students who scored higher on their college entrance exam (measured by their reported American College Testing score) and students whose parents have higher educational levels reported a higher level of familiarity with both mainstream and more advanced Internet-related terms. [17] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N617;note=ptr>]

Surveying such a highly connected population is especially relevant since students represent the wired generation and thus make it possible to control for exposure to and experience with the medium. The fact that some people are not necessarily knowledgeable about Internet-related terms and activities despite high levels of connectivity and frequent usage suggests that mere exposure to and use of the medium does not result in savvy users. As per the findings already cited, students' socioeconomic background is related to their online know-how. This suggests that those in more privileged positions are more likely to understand their online actions well and thus are less likely to be derailed by confusing content presentation.

Knowing how to interpret URLs is an important part of user abilities. Understanding how a user can tell whether a site is secure is an essential part of staying secure when submitting certain types of information to sites, such as financially sensitive data. In a questionnaire administered to hundreds of undergraduate students in the winter of 2007, I gathered information about a

related know-how. First, it is important to note that this is truly the wired generation. On average, respondents in this study had been online for over six years, and the majority (88 percent) reported using the Internet more than once a day. When asked to rate on a five-point scale, ranging from “strongly disagree” to “strongly agree,” how confident they feel about “knowing the difference between http and https”—the latter of which signals to users that they are on a secure site—only 18 percent agreed with the statement. Over half (57 percent) disagreed (over a quarter of the full sample disagreed strongly) suggesting that many young adults even among the wired generation are not fully aware of how to be really safe in their online actions, since it is not clear that they could tell when they are on a secure site. While the relationship is not large, there is a statistically significant positive correlation between parents’ education and reported level of know-how concerning “https,” and there is a similar relationship with college entrance exam scores.

Page 96

Understanding Search Engine Rankings

Regarding the special case of understanding how search engines make decisions about what content to display, some surveys have collected data on users’ understanding of the practice of sponsored versus paid search results. Findings from these studies suggest that people are not particularly savvy about the behind-the-scenes aspects of search engines. For example, when asked in one study whether they were aware of the distinction between paid and unpaid results, the majority of adults interviewed (62 percent) indicated that they were not. [18]

[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N618;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N618;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N618;note=ptr) These findings were mirrored by another study, asking similar questions, where 56 percent of adult respondents did not know the difference between the two types of results. [19]

[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N619;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N619;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N619;note=ptr)

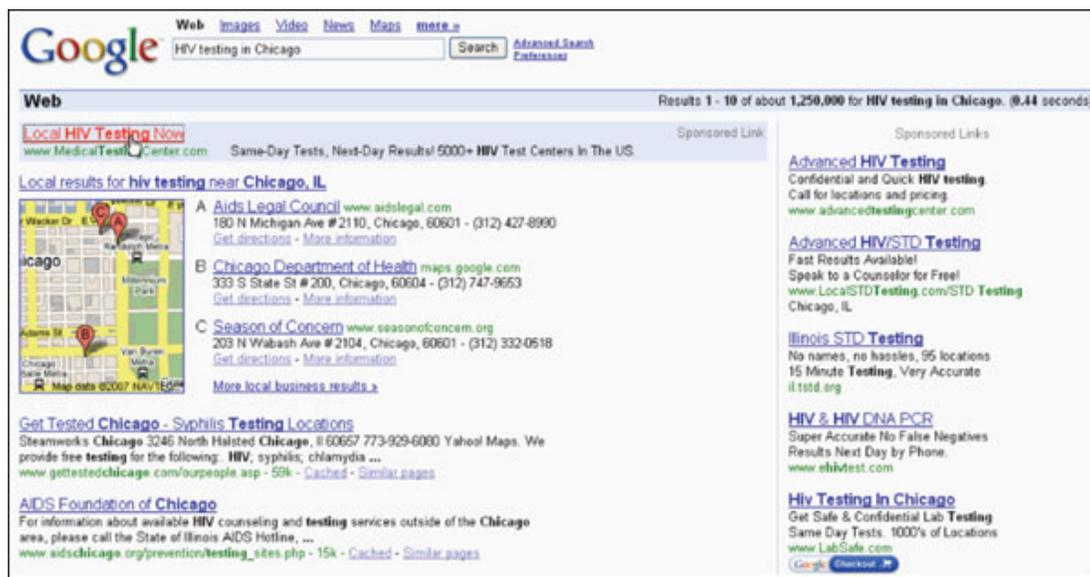
Moreover, findings suggested that this know-how is not randomly distributed among users, as men and younger adults claimed to be more informed about this aspect of search engines than women and older users. Howard and Massanari also found that more experienced users were considerably more confident in their ability to tell apart paid and unpaid content on search engines. [20] [http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N620;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N620;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N620;note=ptr)

How do members of the wired generation respond to similar questions? I asked about related issues in a study I conducted in the winter of 2006 on a group of 150 undergraduate students at a private research university. These students had been, on average, Internet users for over seven years, and 98 percent of them claimed going online several times a day, signifying that the Internet is very much a part of their everyday lives. Among them, over 37 percent claimed never having heard about the fact that search engines are “paid to list some sites more prominently than others in their search results.” Following up, all of the students in the sample were asked, on a four-point scale, how important they think it is that search engines tell users about this practice “in the search results or on an easy-to-find page on the site.” Less than a quarter (24 percent) found this to be “very important,” with an additional 46 percent considering this practice “important.” Over 24 percent, however, thought this was “not too important,” and a remaining 5 percent found it to be “not at all important.”

There are limitations to what we can learn through surveys, so using other methodologies to address these questions can be helpful as well. Follow-up observations can help shed some light on the extent to which students understand links. Drawing on data from a study conducted in 2007, figure 3 shows the action of a first-year female college student at an urban public research university in response to a search query looking for

Page 97



[/d/dculture/images/5680986.0001.001-00000005.jpg]

Fig. 3.

Screen shot of a study participant's selection of a sponsored link result (2007)

HIV testing options in the city of Chicago. The respondent entered "HIV testing in Chicago" into the search box at Google.com and was presented with a list of results, including a highlighted link explicitly designated as sponsored and numerous ad links on the right side of the screen. She clicked on the sponsored link at the top of the page, right below the query box. This page did not yield the desired information.

When asked, later, to explain her choice here, the respondent stated: "I know that the ones that are in here [points to sponsored link section], they're the most relevant to what I'm looking for." There was no mention of sponsorship in her response. Later, in an effort to see whether she would say more about this, she was asked to recount how she learns what she knows about search engines. She stated that it comes "from using it frequently for school and for when you have to do homework." This response was fairly generic and suggests that her assumptions have received no external validation by other sources (whether people from her social networks or other resources). In the end, there is no basis for her assertion that the highlighted link is the most relevant result. It may be on occasion, but it is not always. Certainly, in this case it was not, as it led to a confusing site that did not include information on what she was seeking. Overall, it seems that this user does not have a good grasp of how search engines make decisions about what results to display. This user seems to

put quite a bit of trust in Google's rankings, regardless of outcome, a finding that has been shown to be true for other student users of this service as well. [21]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N621;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N621;note=ptr)]

Understanding Links in E-mails

When we think about links, we tend to think about clickable words or images on Web sites. Links in e-mail messages are increasingly common as well and pose a set of their own unique challenges. It can be convenient to receive a link in an e-mail message, but it can also be dangerous. The medium of e-mail is especially vulnerable to exploitation, because some people assume that seeing the name of a trusted source in the "From" line of the message automatically means that it contains legitimate content.

The term *phishing* refers to the practice of directing a user to a Web site other than one that the link and surrounding message context would seem to suggest, with the goal of extracting sensitive information from the user. For example, many users receive messages claiming to be from a bank (e.g., Chase) or an online commerce-related Web site (e.g., eBay or PayPal). [22]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N622;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N622;note=ptr)] These

messages ask users to follow the provided link and then the instructions on the Web site to which the link leads. The instructions often ask users to enter their username and password into a form secretly monitored by the malicious originators of the message. Once users have shared their login data, they may be exposed to fraudulent activity by the scammers.

Given technological advances, it is relatively easy to configure an e-mail message so it seems to be sent from a source other than the actual sender, resulting in what seems like a legitimate note to the recipient. However, once the user clicks on the included link, it may well lead to a malicious Web site. How many users are aware of these malevolent practices? In my surveys of a diverse group of undergraduate students, I asked respondents to indicate their level of understanding about the term *phishing*. (This question was part of a longer survey item asking about a myriad of terms, an item validated in earlier work as a good measure of people's actual online skills.) [23] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N623;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N623;note=ptr)] In

both 2006 and 2007, the reported level of understanding was extremely low: 1.6 and 1.7, respectively, on a scale of 1–5. Placing the term *phishing* in the context of other terms is also revealing. From among over twenty-five terms presented to the student sample in both years, *phishing* was one of the least understood. The survey included other terms, from the widely understood (e.g., *spam* and *bookmark*) to the less recognized (e.g., *tagging* and *tabbed browsing*) and the largely cryptic (e.g., *torrent* and *widget*). Nonetheless, all of these were

claimed to be better understood by students than the term *phishing*. As with other types of Internet know-how, understanding phishing exhibits a statistically significant positive relationship with a student's score on a college entrance exam.

My findings are mirrored by data collected on people's understanding of Internet-related terms by the Pew Internet and American Life Project. [24] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N624;note=ptr>] That organization's survey of a national sample of adult Internet users found that 15 percent had never heard of the term *phishing* and that 55 percent were "not really sure" what it meant (that survey only allowed three answer options, so the results of these studies—mine and Pew's—are not directly comparable). Of course, it may be that people understand the malicious practice and simply do not know the term that is used to describe it. It is possible to test this using a more nuanced method.

To examine the extent to which people are cautious about messages they receive, I have been presenting some college student study participants with hypothetical e-mail scenarios. Respondents are asked to read supposed e-mail messages and indicate how they would respond to them. Answer options include anything from reporting the message to IT support as fraudulent to following the instructions outlined within and forwarding the note to friends or family. There is also the option of choosing "other" and explaining what one might do, such as click on the link and check where it leads. Respondents are requested to check all of the actions in which they would engage upon receipt of the e-mail.

There are three messages in the study, one of which is made to look just like the e-mails students on this campus receive from the university through its official announcement list, including the appropriate sender and subject line conventions. The e-mail instructs recipients to log into a site and type in their username and password. The specified site address looks like a page on the university's Web site (i.e., it begins <http://www.university.edu/admin/> ...). The way this experiment is set up, the message is not clickable, so it is not possible for students to verify to what Web page the link actually leads. They are asked to indicate what they would do if they received this e-mail in their mailbox, by marking off all possible actions. Interestingly, very few suggest that they would contact technical support or verify where the link leads, and based on twenty-six cases, no one mentioned checking the address of the destination Web site. Over half of the students indicated that they would follow the instructions in the message and would click on the link and do what the destination page instructed, although a few did add that they would concurrently contact the IT department for more information.

Even when links are labeled as sponsored, users do not realize that

Page 100

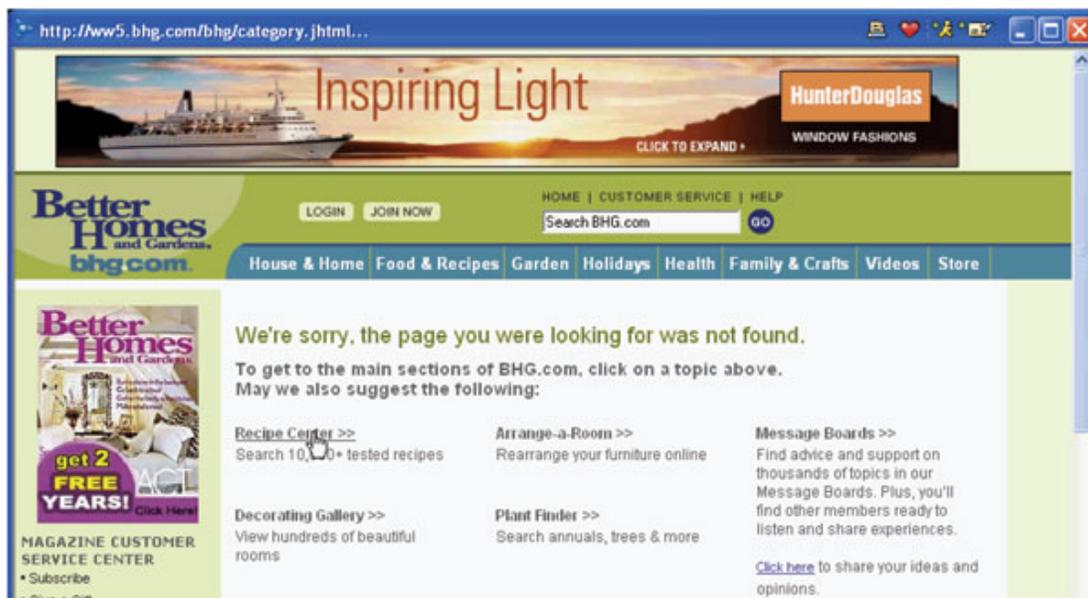
they may not be the most relevant (of course, on occasion, they may be). Take the case of a thirty-seven-year-old woman who had been using the Internet for eleven years, was frequently online, and participated in a study conducted on average adult users in the spring of 2006 in a suburban town. While searching for information on lactose intolerance, she clicked on a sponsored result that showed up at the top of the search engine results page. This link led her to a site that did not include the information of interest to her. She then returned to the original results pages and proceeded to click on another result (this time the top result under the heading "Web Results" on the AOL search results page). She was directed to a page with the necessary information.

As a next step, she was asked to look for recipes that are acceptable for lactose intolerant people. She clicked on a link that was listed on the bottom of the previous page she had been viewing.

This link was located under the heading “Sponsored Links.” The link led to a page with the following statement in the midst of lots of graphics: “We’re sorry, the page you were looking for was not found” (fig. 4). Below this statement were several links whose sponsorship was obvious to the trained eye but much less so to this particular user. She clicked on one of them and proceeded off-site to a page that no longer had anything to do with her original intent of finding a recipe that is suitable for lactose intolerant people. Based on her comments about the resulting page, however, it was clear that she did not realize this. She seemed to assume she was still on the original site at which she had started out her exploration. She was therefore confident that the recipe she had found was acceptable for lactose intolerant people, when, in reality, it was not. This is an example of the limited extent to which people understand where links lead them and of how they can be sent from one site to a completely different one, often due to strategically placed sponsored links that do not address the user’s intent and may be interpreted as something other than what they really are.

Discussion

Relying on data collected using various methods, the empirical evidence presented in this chapter suggests that many users are not particularly familiar with the behind-the-scenes issues of Web content organization and presentation, issues related to how they may be navigating links of influence. Internet users differ considerably regarding their online savvy and their understanding of link navigation in particular. This know-how



[/d/dculture/images/5680986.0001.001-00000006.jpg]

Fig. 4.

Strategically placed sponsored links (2006)

is not randomly distributed: on the contrary, socioeconomic status variables exhibit a statistically significant relationship with online savvy. Take, for example, the young woman who expressed considerable confidence in the relevance of a sponsored link on a search results page. She is a first-generation college student with parents who have no more than a high school education. This relationship between parental education and Internet skill seems to be

education. This relationship between parental education and internet skill seems to be consistent across several studies.

Despite some statistically significant relationships between user attributes and skill measures, it is safe to say that not enough work has been done in this domain for us to understand in depth what processes contribute to people's online abilities. We know, from earlier work and findings discussed in this piece, that information-seeking abilities and spelling mistakes are related to socioeconomic status, [25] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N625;note=ptr>] but we know much less about link savvy in particular. We need better measures of this concept, especially survey items that can be administered to larger numbers of users for statistical analyses and generalizable results. Also, we need to go past individual user attributes to explore the role of users' social surroundings in their online behavior.

Links play a crucial role in how attention is allocated to material online, in what content becomes popular, and in what information is seen only by a few people. Links help users meet everyday needs ranging from

Page 102

the trivial to the profound. Given that people vary in their abilities to understand the sources of different links and their relevance, and given that these skills are not randomly distributed, some users are better positioned to use the medium efficiently and to their benefit, while others are more likely to be misguided and possibly even to fall into malicious traps. Links are important, but their potential influence on users is mediated by the level of expertise people bring to their online pursuits. Since those in more privileged positions seem to exhibit higher-level savvy, the Internet may be contributing to social inequalities rather than alleviating them, despite the many opportunities it makes available, theoretically, to everyone.

Notes

1. Here, it is worth noting that I use the terms *audience*, *consumer*, *user*, *reader*, and *viewer* interchangeably. The level of agency associated with these terms may differ, but this not being the central focus of the chapter, I do not discriminate among them on that basis. ♣ [#]
2. E. Nakashima, "Harsh Words Die Hard on the Web," *Washington Post*, March 2007. ♣ [#]
3. K. Shea and J. Wesley, "How Social Networking Sites Affect Employers, Students, and Career Services," *NACE Journal* 66, no. 4 (2006): 26–32, <http://www.naceweb.org> [<http://www.naceweb.org>]. ♣ [#]
4. E. Hargittai, "Classifying and Coding Online Actions," *Social Science Computer Review* 22, no. 2 (2004): 210–27. ♣ [#]
5. D. Fallows, *Search Engine Users* (Washington, DC: Pew Internet and American Life Project, 2005). ♣ [#]
6. J. Battelle, *The Search* (New York: Penguin, 2005). ♣ [#]
7. E. Hargittai, "How Wide a Web? Inequalities in Accessing Information Online," PhD diss., Department of Sociology, Princeton University, 2003. ♣ [#]
8. Battelle, *The Search*. ♣ [#]

9. Ibid., 20–22.  [\[#\]](#)
10. S. Brin and L. Page, “The Anatomy of a Large-Scale Hypertextual Web Search Engine,” *WWW7 / Computer Networks* 30 nos. 1–7 (1998): 107–17.  [\[#\]](#)
11. J. Bar-Ilan, “Google Bombing from a Time Perspective,” *Journal of Computer-Mediated Communication* 12, no. 3 (2007), article 8, <http://www.jcmc.indiana.edu/vol12/issue3/bar-ilan.html> [<http://www.jcmc.indiana.edu/vol12/issue3/bar-ilan.html>].  [\[#\]](#)
12. D. F. Gallagher, “Top of the Heap,” *Business 2.0* (2002), <http://www.business2.com/articles/mag10,1640,41488,00.html> [<http://www.business2.com/articles/mag10,1640,41488,00.html>].  [\[#\]](#)
13. Bar-Ilan, “Google Bombing.”  [\[#\]](#)
14. E. Hargittai, “Serving Citizen’s Needs: Minimizing Online Hurdles to Accessing Government Information,” *IT and Society* 3, no. 3 (2003): 27–41.  [\[#\]](#)
15. Y. M. Wang, M. Ma, Y. Niu, and H. Chen, “Spam Double-Funnel: Connecting Web Spammers with Advertisers,” paper presented at Sixteenth International World Wide Web Conference, Banff, Canada, May 8–12, 2007.  [\[#\]](#)

Page 103

16. E. Hargittai, “A Framework for Studying Differences in People’s Digital Media Uses,” in *Cyberworld Unlimited? Digital Inequality and New Spaces of Informal Education for Young People*, ed. N. Kutscher and H.-U. Otto. (forthcoming).  [\[#\]](#)
17. Ibid.  [\[#\]](#)
18. Fallows, *Search Engine Users*.  [\[#\]](#)
19. iCrossing, How America Searches, 2005, <http://www.icrossing.com> [<http://www.icrossing.com>].  [\[#\]](#)
20. P. N. Howard and A. Massanari, “Learning to Search and Searching to Learn: Income, Education, and Experience Online,” *Journal of Computer-Mediated Communication* 12, no. 3 (2007). <http://www.jmcm.indiana.edu/vol12/issue3/howard.html> [<http://www.jmcm.indiana.edu/vol12/issue3/howard.html>].  [\[#\]](#)
21. B. Pan, H. Hembrooke, T. Joachims, L. Lorigo, G. Gay, and L. Granka, “In Google We Trust: Users’ Decisions on Rank, Position, and Relevancy,” *Journal of Computer-Mediated Communication* 12, no. 3 (2007). <http://www.jcmc.indiana.edu/vol12/issue3/pan.html> [<http://www.jcmc.indiana.edu/vol12/issue3/pan.html>].  [\[#\]](#)
22. M. Huffman, “‘Phishing’ Scam Takes New Tack,” *Consumer Affairs*, March 2007, http://www.consumeraffairs.com/news04/2007/03/phishing_tactic.html [http://www.consumeraffairs.com/news04/2007/03/phishing_tactic.html].  [\[#\]](#)
23. Ibid.  [\[#\]](#)
24. L. Rainie, *Public Awareness of Internet Terms* (Washington, DC: Pew Internet and American Life Project, 2005).  [\[#\]](#)
25. E. Hargittai, “Hurdles to Information Seeking: Spelling and Typographical Mistakes during

Users' Online Behavior," *Journal of the Association for Information Systems* 7, no. 1 (2006), <http://ais.aisnet.org/articles/default.asp?vol=7&art=1> [<http://ais.aisnet.org/articles/default.asp?vol=7&art=1>]; E. Hargittai, "Second-Level Digital Divide: Differences in People's Online Skills," *First Monday* 7, no. 4 (2002): 1–18. [\[#\]](#)

Page 104

Google, Links, and Popularity versus Authority

Seth Finkelstein

Suppose one wished to search through the data available on the Internet to find some information. Often, a user searches for Web pages associated with some particular keywords. However, the number of Web pages available is enormous. Whether millions or billions, the number of items that could potentially be read vastly exceeds any human capacity to examine them. This fundamental mathematical fact creates an opportunity for a solution by the use of automated assistance: that is, a search engine.

A search engine typically contains an index of some portion of all available existing pages and a means of returning an ordered subset of the available pages in response to a user query. Given that users likely wish to examine as few results as possible, the ordering of the results in response to the user query has become a subject of intense interest. The number of pages that merely contain the desired keywords could still be many thousands, but the user may start to lose patience when examining more than a few results.

Thus, primitive implementations of returning all pages that contain certain keywords, in an order based, perhaps, on the age of the page or on when the page was placed in the search engines database, work poorly in terms of returning results that are significant to the user. A major advance in quality of results was the PageRank algorithm of the Google system.

Academic citation literature has been applied to the web, largely by counting citations or backlinks to a given page. This gives some approximation of a page's importance or quality. PageRank extends this idea by not counting links from all pages equally, and by normalizing by the number of links on a page. This innovation proved to be extremely successful. By taking into account the link structure among a network of pages, and employing a measurement based on the results, the structure of links was used in part to impose a structure of relevancy. However, this practice of using links as a metric for meaning has proved to have many complicated social effects. [\[1\] \[http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N701;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N701;note=ptr)

Page 105

In sociological terms, it was insightful of the Google creators to realize that a popular answer would be a popular answer; that is, if someone were to search for the term *widget*, a popular answer, for the purpose of seeming to fit the needs of the searcher, would be to look for a popular page in some sense. For example, a frequently referenced (linked) page likely had some appealing or attractive aspect to many people. So when that page was returned to a searcher as

appealing or attractive aspect to many people. So when that page was returned to a searcher as a result, it would then likely have a similarly appealing or attractive aspect to that searcher.

A very naive initial concept of the functioning of PageRank in a search would include the following steps:

1. Select all pages containing the target term.
2. Order this subset by the size of their PageRank.
3. Return the top results of this ordered subset.

Some reflection would quickly show this model to be untenable. For example, the page that happened to possess the highest PageRank would then appear as the first result for a search on every word it contained, the page with the second highest PageRank would dominate for another set of words, and so on. Obviously, these results might not be very meaningful responses for the search words. Additional criteria for ranking pages for search terms must therefore be introduced, to prevent a small number of pages from dominating the results. Such criteria can include looking for large numbers of the search term; use of the search terms in emphasized or special contexts; or, crucially, hyperlinks from other pages that use the search term in the anchor text of the hyperlink.

The anchor text criterion is particularly powerful. If many people or a few prominent people refer (link) to a page with the desired term, that page is likely to be a good result to return for the desired term. So a somewhat more refined search algorithm would include the following steps:

1. Select all pages containing the target term or that have the target term in the anchor text of links to the page.
2. Calculate the number of links to the page containing the target term and the number of times the term appears on the page, as well as the PageRank.
3. Order the results by a weighted combination of the preceding factors.

As the algorithm becomes more and more elaborate, the addition of an increasing number of factors can create many unintended consequences.

Page 106

As the various ranking aspects interact with each other, several small factors can combine to be equivalent to a large amount of another factor; or, inversely, a very high scoring on one particular basis may overwhelm negligible amounts of every other score. Crucially, all such quantitative criteria do not convey any sense of quality, as to whether the page might be considered good or bad from a perspective based on truth or merit (in an academic sense). While syntactical analysis of page elements (determining how many keywords are present, where they are, and whether they have any special attributes) is easy, semantic analysis (determining what the elements mean) is hard. There can be a confusion of quantitative with qualitative value, or popularity with authority.

Both the nature of the page-ranking activity and its uses underscore the importance of seeing search results as a value-laden process with serious social implications. The following pages will elaborate this idea by exploring three propositions. First, searching is not a democratic activity. Second, searching inherently raises the question of whether, when searching, we want to see society as we are or as we should be. Third, the current norms of searching, based on popularity

The Hyperlinked Society: Questioning Connections in the Digital Age / Josep...
 society as we are or as we should be. Third, the current norms of searching, based on popularity, are not an appropriate model for civil society.

PageRank and Democracy

It's common to think about the technical examination of a network structure in terms of a political system imposing social structure. The analysis of relevancy in terms of popularity lends itself to an easy analogy of voting and democracy. But an analysis of the fundamental driver of Google's approach, the PageRank, reveals the problems with this analogy.

PageRank relies on the uniquely democratic nature of the web by using its vast link structure as an indicator of an individual page's value. In essence, Google interprets a link from page A to page B as a vote, by page A, for page B. But, Google looks at more than the sheer volume of votes, or links a page receives; it also analyzes the page that casts the vote. Votes cast by pages that are themselves "important" weigh more heavily and help to make other pages "important." [2]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw:g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N702;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw:g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N702;note=ptr)

Someone might simplistically think that a democratic practice implies that one link is one vote and might then mentally equivalence that idea to a concept of everyone having equal power. But the ranking algorithms are

Page 107

rarely simple direct democracy. They're akin to "shareholder democracy" as practiced in corporations: that is, each person doesn't have a single vote; rather, individual voting power varies by orders of magnitude (for corporations, this depends on how many shares are owned by the shareholder). The votes are more like weighted contributions from blocks or interest groups, not equal individual contributions. One link is not one vote, but it has influence proportional to the relative power (in terms of popularity) of the voter. Because blocks of common interests, or social factions, can affect the results of a search to a degree depending on their relative weight in the network, the results of the algorithmic calculation by a search engine come to reflect political struggles in society.

A Proxy for Societal Importance

The outcome to these political struggles via searching can be quite real. Being highly ranked is the end result of a complex algorithm that is often taken as a proxy for societal importance. Inversely, being lowly ranked can doom a source to marginalization. One response to this concern may be that searching is necessary because of the "information overload" in contemporary society.

While information overload may be a modern cliché, there has always been too much information, ever since the days of cavemen grunting around the campfire, when more occurred at a tribal council than could be effectively retold to an absent hunter. The need to summarize events—to present important (according to some definition) information in a short, accessible form—is hardly new. Many issues surrounding search engines can in fact be framed as instances of long-standing journalistic problems. The universe of available information needs to undergo a

strong standing journalistic problems. The universe of available information needs to undergo a winnowing process that can be described as selection, sorting, and spinning, according to the following model:

1. Selection: Which items are important?
2. Sorting: In what order should the items be presented?
3. Spinning: How should one view the items in context?

Compare this description to the colloquial summary of journalism as determining the who, what, when, where, why, and how of an event. For both journalism and search engines, crucial decisions are made as to what results to present to the end user, from among an overwhelming set of

Page 108

possibilities. And both have a concept of objectivity in theory but also inescapable problems with values that enter the decision-making process.

Consider the following passage, where a journalist outlines his education in the algorithm used for determining newsworthiness of traffic accidents, in the era before civil rights reforms took hold.

The unwritten guidelines for reporting fatal automobile accidents were more complicated, the rough rule of thumb being: No n[-]gg[-]rs after 11 p.m. on weekdays, 9 p.m. on Saturdays (as the Sunday paper went to press early). Fatal highway accidents were reported without regard to the color of the deceased until these home edition deadlines: To get a late story in the final editions required making changes, and by tradition only white traffic deaths were considered worth submitting. The exception to this rule was in the area of quantity: If two black persons died in a late evening auto crash, that event had a fair chance of making the news columns. Three dead was considered a safe number by everyone, except those reporters who were known to be viciously anti-Negro. Most of us, of course, considered ourselves neutral or objective in that regard. Yet none of us questioned the professional proposition that the loss of a white life had more news value than the loss of a black life. [3] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N703;note=ptr>]

The journalist describes determining newsworthiness by weighing various factors, such as number, time, and (crucially) social influence. All these factors might be calculated in a “neutral or objective” manner (by asking what time an accident was, how many people are dead, and what their color was). But by taking into account the relative weight (like PageRank) of race, social judgments are incorporated in the results of “news value.”

Censorship and Search Links

Sometimes authorities don't want links to be made or, at least, to be visible. Perhaps contrary to a naive impression, there are specific cases where the results of a search are affected by government prohibition; that is, search results that might otherwise be shown are deliberately excluded. The suppression may be local to a country or global to all Google results. [1]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N704;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N704;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N704;note=ptr)

Search engines do not simply present a raw dump of a database query to the user's screen. The retrieval of the data is just one step. There is much postprocessing afterward, in terms of presentation and customization.

When Google "removes" material, often it is still in the Google index

Page 109

itself. But the postprocessing has removed it from any results shown to the user. This system can be applied, for quality reasons, to remove sites that "spam" the search engine. And that is, by volume, certainly the overwhelming application of the mechanism. But it can also be directed against sites that have been prohibited for government-based reasons.

One very simplistic model of links in the world is that all nodes are ideally visible to all other nodes. But search engines act as sources or portals for a set of links. So suppressing sites in search results will be an ongoing battle.

As We Are or As We Should Be?

Some of the debate over search results echoes ancient descriptive versus prescriptive philosophical conflicts. Should the world be presented as it is (at least as created through the particular search algorithm) or as it should be? The two case studies that follow highlight how Google's approach to the world raises this issue to sometimes emotional heights.

Case Study—Chester's Guide to Molesting Google

What if the terms sometimes used to find an innocuous site are also linked to a site that seems to be associated with child predators? Such a situation led to "moral panic" and a newspaper censorship campaign to have a site removed from both its host and the Google search index. The uproar turned out to originate from a single page of text of "sick humor."

An article headlined "Sick Website Taken Down" in the U.K. *Chester Chronicle* reported: "People power and The Chronicle have won the fight to get a sickening paedophile site—in the name of Chester—removed from the web." [5] [[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N705;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N705;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N705;note=ptr)

Almost every fact in that article was wrong: the targeted site was not a pedophile site; the Google search index is not the Web. But the confusions of the people involved in this campaign (which ranged up to U.K. members of Parliament) are revealing. The article related:

Councillors and readers were disgusted earlier this month when we told how a disturbing site could be accessed after innocently typing "Chester Guide" into the popular search engine run by Google.

This week, the US firm agreed to remove the site, entitled "Chester's guide to picking up little girls," after receiving complaints from our readers.

The move also comes after Cheshire Constabulary's paedophile unit alerted the Internet Watch Foundation... .

However, they urged objectors to bombard Google and the Internet service provider Marhost.com with complaints.

A driver of the controversy was apparently that the same words that would naturally be used to find material about the town of Chester were also featured on a page of extremely tasteless material. Thus some sort of association or connection was implied. Of course, extreme bad taste is not illegal. Contrary to the inflammatory description, all that was being returned was a page of very low humor. Bizarre tastelessness makes the rounds of the Internet every day and even has a genre of books devoted to it (e.g., *Truly Tasteless Jokes*). But contrast these statements from the same article:

Google's international public relations manager, Debbie Frost, said ... :

"When an illegal site is discovered, search engines like Google will remove such sites from their indices in order to abide by the law."

"After our investigation, we have determined that the site in question is illegal and therefore it will be removed from our index."

... John Price, leader of Chester City Council, was furious when we informed him of the site's existence.

This week, he said: "It's great news the site has been removed. Good riddance to bad rubbish. However, we must now be vigilant and make sure it does not come back."

Chester MP Christine Russell was also outraged and immediately agreed to demand a change in the law to make such sickening sites illegal.

Crucially, no judicial process seems to have been applied in Google's determination. There was certainly no judicial avenue of appeal, no public evidence record to examine. One might argue that there was little value to the page that was removed from the index, but the implications of such a removal can be troubling.

Case Study—Jew Watch

While Chester's problem of a popular link that yielded unfortunate search results may sound unique, it is not. One of the most well-known examples of complex issues of unintended consequences and social dilemmas

is the high ranking of an anti-Semitic Web site, Jew Watch, for Google searches on the keyword "Jew." The Web site describes itself as "keeping a close watch on Jewish communities, organizations, monopoly, banking, and media control worldwide." The front page contains such categories as "Jewish-Zionist-Soviet Anti-American Spies," "Jewish Communist Rulers & Killers." and "Jewish Terrorists." It is unarguably a site devoted to anti-Semitic "hate speech."

However, such material, though repulsive, is completely protected under the U.S. Constitution's First Amendment, though other countries may consider it illegal.

For a long time, this objectionable site was the first result in a Google search for the keyword "Jew." As reported by ZDNet:

The dispute began ... when Steven Weinstock, a New York real estate investor and former yeshiva student, did a Google search on "Jew." ... Weinstock has launched an online petition, asking Google to remove the site from its index. [6]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N706;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N706;note=ptr)

After the controversy had been in the news for some time, Google posted an explanation of the search result.

A site's ranking in Google's search results is automatically determined by computer algorithms using thousands of factors to calculate a page's relevance to a given query. Sometimes subtleties of language cause anomalies to appear that cannot be predicted. A search for "Jew" brings up one such unexpected result. [7]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N707;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N707;note=ptr)

The explanation was in part aimed at defusing charges that Google was anti-Semitic and had deliberately placed a hate site in a high search ranking. Such a charge is completely unfounded. But the problem is more closely outlined by the Anti-Defamation League's analysis: "The longevity of ownership, the way articles are posted to it, the links to and from the site, and the structure of the site itself all increase the ranking of 'Jewwatch' within the Google formula." [8]

[\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N708;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N708;note=ptr) While Google did not in any way promote the hate site, there is more to the ranking than "subtleties of language." The Google system was, in effect, used by the site to promote itself.

Another site, Remove Jew Watch (www.removejewwatch.com), was set up to launch a petition to "get Google.com to remove Jewwatch.com from their search engine." Other people tried to have different sites rank higher for the keyword "Jew." But Jonathan Bernstein, regional director

of the Anti-Defamation League, noted that "one can stumble across plenty of Holocaust denial Web sites by simply typing 'Holocaust' into Google." He added: "Some responsibility for this needs to rest on our own shoulders and not just a company like Google. We have to prepare our kids for things they come across on the Internet. This is part of the nature of an Internet world. The disadvantage is we see more of it and our kids see more of it. The advantage is, we see more of it, so we're able to respond to it... I'm not sure what people would want to see happen. You couldn't really ask Google not to list it." [9] [\http://quod.lib.umich.edu/cgi/t/text/text-idx?

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N709;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-id?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N709;note=ptr)

It might be noted, however, that Google will place sites on certain blacklists if they are illegal. A search for the keyword “Jew” in some country-specific Google versions (in Germany and France) shows Jew Watch removed from Google. [10] [<http://quod.lib.umich.edu/cgi/t/text/text-id?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N710;note=ptr>] And in at least one situation (the “Chester’s guide” case mentioned previously), Google has blacklisted a site that was not illegal. But that way lies madness, and Google has sound reasons to duck the issues as much as it can. The problem will not disappear, and there will be constant pressure from various groups.

Ironically, all the controversy probably raised the rank and relevance of the Jew Watch site within Google’s algorithms, at least temporarily. Most important, people who made hyperlinks to the site for the purposes of reference added to the number of links to the site on the Web, which could have contributed to raising its search ranking. For a while, the site lost its service provider and, since it was not available, dropped in ranking; but then it rose back up (around April 22, 2004). Eventually, the Wikipedia entry for the word “Jew” took over the top position for a search on that word, and attention to this case subsided. But as hate groups realize the power that comes from prominent placement in searches, the topic will certainly be revisited. As an ironic aside, during the height of the controversy, one neo-Nazi was apparently jealous of all the attention received by a like-minded rival, so he tried to generate a campaign to ban his own site, presumably so publicity and anticensorship sentiment would give that site similar prominence. The campaign failed, but it illustrates the extremes of convoluted political maneuvering that can be found in the topic.

To some extent, the high position of the Jew Watch site in search results for the keyword “Jew” can represent a kind of plurality dominance over diluted opposition. If one were to ask what the most prominent associations with the word “Jew” are, anti-Semitism would sadly have to be significant. And it would by no means need to have anywhere near a majority share to be returned as a first result. If, hypothetically, anti-Semitism

Page 113

were the association 19 percent of the time and there were nine other slightly different positive associations that each had 9 percent of the remaining time, being the greatest single identifiable block could give it a ranking of “most popular” in some algorithmic sense. This is the popularity versus authority conflict all over again. A site that has a plurality of weighted link votes need not be accurate or even inoffensive to the population outside that group.

Moreover, if a goal is to return relevant results, anti-Semites also use search engines, and a hate site counts as a correct result to them. In a sense, Google argued that it was performing a descriptive function in reflecting relative prominence for a search term, against the tangle that would develop if it was prescriptive in its results. But a contrary point of view is that an algorithm that gives high ratings to hate sites is by definition flawed in some way and should not be justified merely by the fact of being an algorithm. At least, if the choice is made that a dominant plurality result is correct, even if it is sometimes offensive, it should be recognized that there are significant social implications of such a choice.

Intentionally or unintentionally, the Jew Watch site had done search engine optimization for the

keyword “Jew.” In extreme forms, an optimization activity turns into “Google spamming,” where search engine spammers try to get irrelevant pages to rank highly in order to obtain profit from ad clicks. The activity can reach a point of doing significant damage to search results, and it has generated some drastic countermeasures, where harsh antis spam actions cause problems with legitimate sites. But significant self-promotion can be done short of spamming, and search engine optimization is merely puffery, not fraud.

A different form of linking to game Google is a practice known as “Google bombing” (defined at Wordspy.com as “setting up a large number of Web pages with links that point to a specific Web site so that the site will appear near the top of a Google search when users enter the link text”). Technically, this manipulates Google search results by hyping the ranking factor associated with the words used to link to a site—for example, using the phrase “miserable failure” to link to a biography of President George W. Bush or connecting the phrase “out-of-touch executives” to Google corporate information. From a Web site’s standpoint, Google bombing is the mirror image of search engine optimization, where a site seeks to rank highly for desired keywords.

Search engine optimization for political ends is a largely unresearched area. Google bombing is now a crude process, done for laughs. In the future, it might well involve much more serious political dirty tricks. Indeed,

Page 114

political campaigning is at heart a process of manipulating information, and as search engines become more important as sources of information, we can expect more and varied creative attempts at their manipulation.

PageRank Selling and Commodification of Social Relations

The factors that Google uses to rank pages have long been a target for financial ends. Once any sort of value is created by a link, there’s an immediate thought that a market can be created to monetize that value. While many people think of linking as a purely social relationship, it’s quite possible to have such expressions of social interconnection be subverted for commercial purposes.

But search engines cannot simply let the market decide the value of a link. That would eventually produce pages of results that are nothing but advertisements, which would then drive users away from the search engine. Those would not be popular results—advertisements tend to be unpopular (even if they are sometimes effective in generating business). Moreover, paying for links on the Web usually competes with the search engine’s own paid advertising program. So a search company has an incentive to disallow outright sales of links, while marketers have an incentive to attempt to buy as much influence as possible.

A crude way to do such buying of links would be to seek out high-ranking pages and offer payment for placement. But such pages are relatively easy to monitor, and internal ranking penalties can be applied if a site owner is found to be participating in such practices. More sophisticated schemes are being refined by companies that offer independent Web writers (bloggers) small amounts of money to write about products on the writer’s own Web site. These arrangements are commonly discussed in terms of traditional journalistic ethics regarding sponsorship or disclosure. The idea is that if the writer discloses that the article is a paid placement, the reader can then apply the appropriate adjustment to the credibility of the content.

However, such a traditional framework misses an important aspect of the exchange. In the case of PageRank selling, the sources of the ranking will not be evident. It won't matter what the writer says about the product or what the reader thinks in terms of trusting the article, as the ranking algorithm will see only the link itself. If the accumulated purchasing

Page 115

of links eventually results in a high ranking, that process will be virtually invisible to the searcher. In a way, this is a disintermediation of the elite influencers—commodifying their social capital—and a reintermediation of that influencing process with an agency specializing in the task. Instead of courting a relatively few A-list writers who are highly valued for their ability to have their choice of topics echoed by many other writers, the lesser writers can be purchased directly (and perhaps more simply and cheaply).

Even for prominent writers who would decline an explicit pay-for-placement deal, the many ways linkage can be purchased (literally or metaphorically) leads to controversy over proper behavior. For example, one company, FON, set off a round of discussion by having many advisory board members who were also widely read Web writers. [11]

[[http://quod.lib.umich.edu/cgi/t/text/text-idx?](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N711;note=ptr)

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N711;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N711;note=ptr)] But the tiny company also got publicity from another source: influential commentators on the Internet who write blogs—including some who may be compensated in the future for advising FON about its business. Though an appropriate journalistic disclosure was made almost everywhere in this case, the aspect in which the social was intermingling with the commercial remained unsettling. A focus on disclaimers often assumes a certain background in separation and avoidance of conflict of interest and is insufficient when those strictures are no longer in place. While blurring the lines between business and friendship is not at all a new problem, the shifting systems of attention sorting and seeking are now bringing these issues to notice in new contexts.

To put it simply, there's an old joke that runs as follows:

Billionaire to woman: Would you have sex with me for a million dollars?

Woman: Well ... yes.

Billionaire: Would you have sex with me for ten dollars?

Woman: What kind of a girl do you think I am?

Billionaire: We've already determined that. Now we're just arguing over the price.

Two factors make up the humor in this joke: commerce itself and amount. The obvious aspect of the joke is that there are two categories of interactions, commercial and social, between which there is not supposed to be any overlap, regardless of the dollar amount at stake. A less-often-remarked aspect is that there is indeed a "class" division between high-priced commercial and low-priced commercial.

Future controversies may present a real-life version of that joke that might go roughly as follows:

Company to blogger: Would you write about me for advisory board membership?

Blogger: Well ... yes.

Company: Would you write about me for ten dollars?

Blogger: What kind of a flack do you think I am?

Company: We've already determined that. Now we're just arguing over the price.

Is a few dollars the same as an advisory board membership? No—there's a class division, in that an advisory board membership is high-class and expensive, while a few dollars is tawdry and cheap. But there's also a problem when executive “escorts” criticize street prostitutes.

The Nofollow Attribute

There's a public relations saying (attributed to many people) that goes, “I don't care what the newspapers say about me as long as they spell my name right.” The concept is that any mention, positive or negative, is helpful in terms of recognition. Links have a somewhat similar phenomena, where any link, even originating from a page making negative statements about the site, can help build the site's search ranking. This is a particular pernicious issue in the case of hate sites (as discussed earlier), as any publicity for the sites tends to generate more links to the sites even if the publicity is negative. A link, by itself, cannot distinguish fame from infamy.

One attempt to address this dilemma has been the introduction of a special attribute, nofollow, to try to distinguish the purely referential aspect of a link from any implied popularity or importance of the site that has been referenced. [\[12\] \[http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N712;note=ptr\]](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N712;note=ptr) If you're a blogger (or a blog reader), you're painfully familiar with people who try to raise their own Web sites' search engine rankings by submitting linked blog comments like “Visit my discount pharmaceuticals site.” This is called comment spam. We researchers don't like it either, and we've been testing a new tag that blocks it. From now on, when Google sees the attribute (rel=nofollow) on hyperlinks, those links won't get any credit when we rank Web sites in our search results. This isn't a negative vote for the site where the comment

was posted; it's just a way to make sure that spammers get no benefit from abusing public areas like blog comments, trackbacks, and referrer lists.

The results of this attribute have been mixed. It certainly has prevented many blog owners who have open comment areas from inadvertently adding to spam pollution. But even if some link spammers have been discouraged, more than enough remain undeterred so that the problem of spammers is still overwhelming. While many blogs have automatically implemented the nofollow attribute on all links in their public areas, a large number of spammers will apparently spam

anyway—finding it more efficient to be indiscriminate, perhaps, or in hopes of benefiting somehow in any case.

Businesses That Mine Data for Popularity: Not a Model for Civil Society

From a political standpoint, one might hope that the use of the nofollow attribute regarding hate sites would lower their rankings as people who mention them unfavorably discourage linking. But the use of this attribute in linking requires both knowledge of its existence and some sophisticated knowledge of how to code a link (as opposed to using a simple interface). So while this way of separating meanings is helpful overall, it is complicated enough to carry out so that the problem is not substantially addressed in practice.

Moving from the specifics of the nofollow attribute to the more general impact of links on people's consciousness, it should be clear by now that Google-like approaches to searching, which base rankings on the popularity of links, tend not to question the society's basic hierarchy. One initial simplistic way of thinking about link networks is to somehow lump all nodes together, as if there were no other structure for determining who received links. But since many links are made by people, all the prejudices and biases that affect who someone networks with personally or professionally can affect who they network with in terms of hypertext linkage. One writer described this (often gender-based) cliquishness in the following manner:

Point of fact, if you follow the thread of this discussion, you would see something like Dave linking to Cory who then links to Scoble who links to Dave who links to Tim who links to Steve who then links to Dave who links to Doc who follows through with a link to

Page 118

Dan, and so on. If you throw in the fact that the Google Guys are, well, guys, then we start to see a pattern here: men have a real thing for the hypertext link... .

[Later] When we women ask the power-linkers why they don't link to us more, what we're talking about is communication, and wanting a fair shot of being heard; but what the guys hear is a woman asking for a little link love. Hey lady, do you have what it takes? More important, are you willing to give what it takes?

Groupies and blogging babes, only, need apply. [13]

<http://quod.lib.umich.edu/cgi/t/text/text-idx?>

[c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N713;note=ptr](http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N713;note=ptr)

Recall that popularity can be confused with authority and that a link from a popular site carries more weight to a search engine. The self-reinforcing nature of references within a small group can then be a very powerful tool for excluding those outside the inner circle. Instead of democracy, there's effectively oligarchy.

The best way, by far, to get a link from an A-List blogger is to provide a link to the A-List blogger. As the blogosphere has become more rigidly hierarchical, not by design but as a natural consequence of hyperlinking patterns, filtering algorithms, aggregation engines, and subscription and syndication technologies, not to mention

human nature, it has turned into a grand system of patronage operated—with the best of intentions, mind you—by a tiny, self-perpetuating elite. A blog-peasant, one of the Great Unread, comes to the wall of the castle to offer a tribute to a royal, and the royal drops a couple of coins of attention into the peasant's little purse. The peasant is happy, and the royal's hold over his position in the castle is a little bit stronger. [14] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N714;note=ptr>]

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N714;note=ptr>

In fact, rather than subvert hierarchy, it's much more likely that hyperlinks (and associated popularity algorithms) reflect existing hierarchies. [15] [<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N715;note=ptr>]

This is true for a very deep reason—if an information-searching system continually returned results that were disturbing or upsetting, there would be strong pressure to regard that system as incorrect and change it or to defect to a different provider. As can be seen in some of the discussions earlier in this essay, even isolated anomalous results can draw angry reactions. Subversive results would not be acceptable.

The positive results from data-mining links for popularity are certainly impressive but have also inspired flights of punditry that project a type of divinity or mystification into the technology. *New York Times*

Page 119

columnist Thomas Friedman wrote an op-ed column entitled “Is Google God?” where he quoted a Wi-Fi company vice president as saying:

If I can operate Google, I can find anything. And with wireless, it means I will be able to find anything, anywhere, anytime. Which is why I say that Google, combined with Wi-Fi, is a little bit like God. God is wireless, God is everywhere and God sees and knows everything. Throughout history, people connected to God without wires. Now, for many questions in the world, you ask Google, and increasingly, you can do it without wires, too. [16]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N716;note=ptr>]

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N716;note=ptr>

However, in contrast to the utopianism, there is much research to show that the mundane world is very much the same as it ever was. Hindman and his colleagues note: “It is clear that in some ways the Web functions quite similarly to traditional media. Yes, almost anyone can put up a political Web site. But our research suggests that this is usually the online equivalent of hosting a talk show on public access television at 3:30 in the morning.” [17]

[<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N717;note=ptr>]

<http://quod.lib.umich.edu/cgi/t/text/text-idx?c=nmw;cc=nmw;g=dculture;xc=1;rgn=div1;view=trgt;idno=5680986.0001.001;id=N717;note=ptr>

Link popularity is itself no solution to problems in governance. Determining what opinions are popular is usually one of the least complicated political tasks. But what if the results are hateful or are manufactured by an organized lobbying campaign? How much weight should be given to

or are manufactured by an organized lobbying campaign? How much weight should be given to strong minority views in opposition to the majority? These questions, which determine the character of a society, are not answered by merely listing the popular opinions and options. Moreover, some of the lessons learned from such businesses are arguably exactly the wrong lessons needed for a pluralistic democracy, where you cannot simply ban the minority that isn't profitable. Unfortunately and maybe self-provingly, that is not a popular position.

Notes

1. S. Brin and L. Page, "The Anatomy of a Large-Scale Hypertextual Web Search Engine," *WWW7 / Computer Networks* 30, nos. 1–7 (1998): 107–17.  [#]
 2. Google Inc., "Our Search: Google Technology," 2007, <http://www.google.com/intl/en/technology/> [<http://www.google.com/intl/en/technology/>].  [#]
 3. W. Hinckle, *If You Have a Lemon, Make Lemonade* (New York: Putnam, 1974).  [#]
 4. B. Edelman and J. Zittrain, "Localized Google Search Result Exclusions," <http://cyber.law.harvard.edu/filtering/google/> [<http://cyber.law.harvard.edu/filtering/google/>]; Seth Finkelstein, "Google Censorship—How It Works," <http://sethf.com/anticensorware/general/google-censorship.php> [<http://sethf.com/anticensorware/general/google-censorship.php>];
-
- Page 120
- BBC News, "Google Censors Itself for China," <http://news.bbc.co.uk/1/hi/technology/4645596.stm> [<http://news.bbc.co.uk/1/hi/technology/4645596.stm>].  [#]
 5. "Sick Website Taken Down," *Chester Chronicle*, February 21, 2003, <http://iccheshireonline.icnetwork.co.uk/0100news/chesterchronicle/page.cfm?objectid=12663897&method=full&siteid=50020> [<http://iccheshireonline.icnetwork.co.uk/0100news/chesterchronicle/page.cfm?objectid=12663897&method=full&siteid=50020>].  [#]
 6. D. Becker, "Google Caught in Anti-Semitism Flap," 2004, <http://zdnet.com.com/2100-1104-5186012.html> [<http://zdnet.com.com/2100-1104-5186012.html>].  [#]
 7. Google Inc., "An Explanation of Our Search Results," 2004, <http://www.google.com/explanation.html> [<http://www.google.com/explanation.html>].  [#]
 8. Anti-Defamation League, "Google Search Ranking of Hate Sites Not Intentional," 2004, http://www.adl.org/rumors/google_search_rumors.asp [http://www.adl.org/rumors/google_search_rumors.asp].  [#]
 9. J. Eskenazi, "No. 1 Google Result for 'Jew' Is Fanatical Hate Site—for Now," 2004, http://www.jewishsf.com/content/2-0/module/displaystory/story_id/21783/format/html/displaystory.html [http://www.jewishsf.com/content/2-0/module/displaystory/story_id/21783/format/html/displaystory.html].  [#]
 10. B. Edelman and J. Zittrain, "Localized Google Search Result Exclusions," 2002, <http://cyber.law.harvard.edu/filtering/google/> [<http://cyber.law.harvard.edu/filtering/google/>].  [#]

[#]

11. R. Buckman, "Blog Buzz on High-Tech Start-ups Causes Some Static," 2006, http://online.wsj.com/public/article/SB113945389770169170-oDZ4wOffelheiC5fe4GISe73UwQ_20070209.html
[http://online.wsj.com/public/article/SB113945389770169170-oDZ4wOffelheiC5fe4GISe73UwQ_20070209.html]. [#]
12. Google Inc., "Preventing Comment Spam," 2005, <http://googleblog.blogspot.com/2005/01/preventing-comment-spam.html>
[<http://googleblog.blogspot.com/2005/01/preventing-comment-spam.html>]. [#]
13. S. Powers, "Guys Don't Link," 2005, <http://weblog.burningbird.net/2005/03/07/guys-dont-link/>
[<http://weblog.burningbird.net/2005/03/07/guys-dont-link/>]. [#]
14. N. Carr, "The Great Unread," 2006, http://www.routhtype.com/archives/2006/08/the_great_unrea.php
[http://www.routhtype.com/archives/2006/08/the_great_unrea.php]. [#]
15. J. Garfunkel, "The New Gatekeepers," 2005, <http://civilities.net/TheNewGatekeepers>
[<http://civilities.net/TheNewGatekeepers>]. [#]
16. T. Friedman, "Is Google God?" 2003, <http://www.cnn.com/2003/US/06/29/nyt.friedman/>
[<http://www.cnn.com/2003/US/06/29/nyt.friedman/>]. [#]
17. M. Hindman, K. Tsioutsoulis, and J. Johnson, "'Googlearchy': How a Few Heavily-Linked Sites Dominate Politics on the Web," 2003, <http://www.cs.princeton.edu/~kt/mpsa03.pdf>
[<http://www.cs.princeton.edu/~kt/mpsa03.pdf>]. [#]

C=Nmw;IDNO=5680986 80986.0001.001%3A2]

C=Nmw;IDNO=5680986.0001.001;RGN=DIV1;VIEW=TEXT;CC=Nmw;XC=1;G=DCULTURE;NODE=5680986.0001.001%3A4]