Re: Search

Mark Graham, Ralph Schroeder and Greg Taylor
University of Oxford, UK

Abstract

Search has assumed a position of central importance in the way that people access and use online information and services. In this introduction we summarize the four articles constituting this themed section, and in so doing explore the ascendancy of search, the power it bestows upon those who control it, its role in shaping access to information, and its capacity to function as a mirror for society. We point to important outstanding questions and suggest some avenues for future work in this area.

Keywords

Bias, information seeking, power, search engines, social mirroring

Introduction

Never before have so many people engaged so intensively in practices of information search. Billions of searches are performed every day through the Internet. Searches connect us to information that helps us find the best route through a city, learn about a debilitating illness, or watch videos of cats playing the piano. We can now search for words, numbers, images, videos, sounds, places, people, products, and more besides.

Search is a process of separating the relevant from the irrelevant and the knowable from the unknowable. Search is thus suffused with power; the power to access and shape information. Digital searches mediate our interactions with an enormous networked store of knowledge, with social networks, with commercial trading partners, and with the world around us. Practices, algorithms, and rules of search govern the content, ideas, places, and commercial opportunities to which users are exposed.

Corresponding author:
Ralph Schroeder, Oxford Internet Institute, University of Oxford, 1 St Giles, Oxford OX1 3BJ, UK.
Email: ralph.schroeder@oii.ox.ac.uk

There is comparatively little research on the social implications of search (see Zimmer, 2010 for a review). The main contributions on this topic have been in two specialist areas. First, computer and information scientists have focused on rankings and information-seeking behavior, with little attention paid to the wider social aspects of this phenomenon.[1] Second, the literature in economics and marketing is burgeoning but highly technical and largely inaccessible to non-specialists.[2]

In order to begin a more inter-disciplinary discussion on the role of Internet searching in contemporary society, this themed section has brought together four papers that make important contributions to scholarship on search. We see two primary themes threaded through these papers: the social, economic, and political power associated with search engines and, as a result, search engines' ability to hold up a mirror to society. These themes, and the contributions made by the four articles, are discussed in more detail later in this introduction. It is important, though, to first explore in more detail what exactly we mean by "search."

Search!

We search in order to find and discover. Sometimes we search for the expected or anticipated (the known unknowns), and sometimes we search to discover and explore (the unknown unknowns). Before the advent of search engines, searching tended to happen in either a direct and relatively unmediated way (e.g., searching for the keys in your garden or searching for a particular paragraph in a book) or was mediated through some sort of indexing or tagging system (e.g., finding a book in a library with the help of the Library of Congress Classification or finding a street in a street atlas using an index and grid system). In both cases, there was a clear path between the searcher and the object of the search. In other words, keys were next to the shed because you left them there, or books on folklore were in the "GR" section of the library because that was where they were supposed to be. In the latter example, the mediation between the searcher and the object of the search was successful because it operated under rules apparent to both the searcher and the mediators (the library).

But the size and scope of contemporary corpuses of information makes these older practices cease to function effectively. The Internet contains a staggeringly large volume of information.[3] Enormous amounts of content about all imaginable facets of the human experience and our biophysical environment are created, stored, and indexed. In addition to its size, the very topology of the Internet makes it impossible to (effectively) perform the digital version of searching for keys—the proverbial garden is too big. The size of the Internet similarly means that it is difficult to imagine human-curated digital index cards for all online content.[4] As such, we have turned to search engines and search algorithms to order, structure, filter, rank, and make meaning out of the massive volume of information that we find ourselves swimming in. In fact, most of us are so accustomed to the answers to almost any question being at our fingertips having to retrieve information in any other way seems almost inconceivable. It is therefore easy to forget that practices of information search with the aid of search engines and algorithms are surprisingly new. It is only in the past two decades that powerful companies have come to mediate our access to information in this way.

   These search companies provide a combination of web crawlers, highly efficient indexes and servers, and sophisticated, complicated, and often highly targeted ranking algorithms. The net result for most users is typically a set of ranked results—ostensibly separating relevant from less relevant information. We see the increasing influence of search engines through the battles that are played out in order to be listed on the first page of results for search terms, the multi-billion dollar industries of search advertising and search engine optimization, and the fact that that browsers now increasingly allow for the input of search terms as well as static URLs into address bars. Search engines and practices of digital search thus play a central role in the ways in which much of the world accesses, enacts, and brings into being specific types of information. However, because of both the newness of many of these services and their opaque nature, this process of mediation remains relatively poorly understood.

   This themed section of *New Media & Society* was therefore put together in order to bring together empirical and theoretical studies of online information search and how they reshape and reflect everyday life.

Information, power, and knowledge

By shaping both what we know and how we know it, search engines—and those who design and control them—are able to wield an immense amount of social power. To justify continued possession of this power, search engine operators often attempt to point to both an algorithmic "objectivity" of their services, and a portfolio of socially virtuous practices. Google, for instance, both highlights the scientific and objective ways in which it ranks and orders information (e.g. the company's faith in engineering solutions to a range of human problems[5]) and presents an image of social responsibility summarized by the company motto: "Don't be evil."

   Despite their apparent good intentions, major search engines—and Google in particular—are not without their critics. Competitors,[6] trading partners,[7] users, academics,[8] and regulators[9] alike have expressed concern that too much power is being concentrated in the hands of search engines, and that the resulting influence has been abused in the pursuit of self-interest through practices such as own-content bias. A perennial question when considering such matters is: How can we validate the veracity of either side's claims, given that the inner algorithmic functioning of search engines is veiled in secrecy? One approach, adopted by Rieder and Sire (2014) in the first article of this themed section, is to examine the underpinning incentive structure faced by search engines—the general principle being that transgressions are more likely when they are associated with greater rewards. Employing a political microeconomic analysis, they demonstrate that search engines face a wide range of incentives to introduce bias into their results, and argue that these incentives cannot be properly understood without deconstructing the dynamic process by which those results are assembled. Rieder and Sire's paper is valuable not only as an independent examination of the issue of search engine bias, but also for its potential for facilitating a dialog between economists and media scholars.

   The theme of objectivity is also picked up in Jiang's (2014) analysis of Baidu and Google in the second article in this themed section, in which she demonstrates how

search is neither objective nor complete.[10] Jiang's conclusion, that knowledge mediated through search engines is both politically and socially constructed, is supported by an examination of how Baidu and Google fare in top results for 316 events undertaken in various parts of China, and a comparison of how accessible the links for these results are. The results differ in important ways and seeing this point documented in detail (for example, by reference to the concrete differences in versions of Chinese Wikipedia, Baidu Baike, and others) makes for a greater appreciation of how search shapes an online source which has become very widely used in everyday life.[11]

Robinson (2014), the third article in this themed section, shows how information seeking crucially depends on skills through a close examination of practices in a California high school with a population of disadvantaged youth. Rather than telling a one-sided story about digital divides, she shows that more and less successful groups coexist within an under-resourced school. The structure of these groups depends not so much on technology as on, among other things, networks of support (or lack thereof) at home, among peers, and in learning opportunities that are "given" by putting pupils into different tracks. Instead of presenting a fatalistic picture of divides in skills, she points to how these can potentially be overcome "from the ground up."

Every search relies on a vast "ecosystem" of networked information that is both created and ordered by a crowd of contributors. As such, these transactional data offer a unique pipeline into the desires of hundreds of millions of people. Sanz and Stančík (2014), in the themed section's final paper, develop the notion of ontological security (drawing on Giddens) to examine trust in relation to search results. To do this they use surveys about trust from a variety of countries together with an analysis of variations in search terms related to four major disruptive events (Wikileaks, the 2011 Japanese earthquake and tsunami, the killing of Bin Laden, and a 2011 US hurricane), drawing links between the two. They seek to explore the cultural dimension of search, arguing that this has so far been neglected, and point to rich future agendas and methods to pursue this topic further.

In sum, we see that search engines are powerful because they are both gatekeepers and intermediaries. Furthermore, the largest and most used search engines are all private entities, funded by advertising, and with little incentive to allow their users to understand how they structure knowledge. There is thus undoubtedly much still to be done to further open up and deconstruct the black boxes of search engines and search algorithms that play an increasingly important role in our everyday lives.

Future research directions

Although the four papers in this section focus in particular on the links between information and power, this area of research has barely begun to scratch the surface of the social aspects of search behavior. While we know, for example, that search is second only to email in terms of Internet usage (at least in the United States; see Purcell et al., 2012: 5) and that people go typically to a search engine rather than to a specific page when they look for information (at least in the UK; see Dutton and Blank, 2011: 22), we know little of how information seeking and search fit into the context of people's everyday lives (but see Rieh, 2004).

It is fairly easy to appreciate search engines' direct capacity to shape the visibility and success of content. However, for social scientists to gauge the importance of this gate-keeping function, much more needs to be known about what people search for. Again, little is known about this topic; in one of the few large-scale studies carried out (Waller, 2011, in relation to Australians), the findings were surprising—namely, that the vast bulk of search queries are for consumer or leisure content, and very few are carried out for content like news, politics or other content that could be regarded as knowledge-seeking. Furthermore, as information and physical places are increasingly intertwined, search mediates not just our access to information, but also the ways in which we interact with, learn about, move through, consume in, and enact physical and material places (Graham et al., 2013). In other words, search will play an ever-greater role in how we understand the material environments that we live in.

Because of their powerful role as gatekeepers, search engines also tell us much about human needs and behavior. Every search relies on a vast "ecosystem" of networked information that is both created and ordered by a crowd of contributors. These transactional data offer a unique pipeline into the interactions of billions of people, and future work will undoubtedly set out to ask what important social insights can be gleaned from them.

It is likely that an increasing amount of research will also have to focus on the unstable and shifting nature of answers retrieved from search engines. The incessant online activity of billions of people means that not only is content itself constantly changing, but so are the preferences feeding into rankings that separate the visible from the invisible. The sorting algorithms that mediate the enormous amount of content yield results that have also become less spatially, socially, and temporally fixed (e.g. Pariser, 2011).[12] In a decentered cloud of information, how do we then gaze through the opacity built into the system? How do we map and measure, study and assess the increasingly ephemeral tools that so many of us rely on for our informational needs?

Many additional questions on the intersections of search and society will undoubtedly emerge in future research, such as: What does it mean to "volunteer" information to search engines? How do we balance the utility of search engines with issues of ownership and privacy? What domains of contemporary life have been left relatively untouched by (or are consciously omitted and excluded in) digital tools and content? And again, how does the blending of tools and content thorough search engines shape our everyday lives? Thus, the boundaries of the topic of search are bound to expand, and will require media and communications scholars to work with other disciplines to examine how people experience information seeking or search more narrowly in otherwise very different contexts.

Ultimately, we hope that this special issue has helped to reinforce the need for empirical inquiry and theoretical focus on the issue of digitally mediated search. To some, the machine-mediated nature of search may seem objective and neutral. To others, the opacity of processes by which information is sorted and ranked may make it seem as if search providers are engaging in a form of algorithmic alchemy, and that it is impossible to ever fully understand the myriad complex processes through which knowledge is indexed, ranked, and retrieved. However, that is precisely why we need the type of sustained inquiry seen in the papers in this issue.

This themed section has made a start in these directions, and presents us with a moment to reflect on how the coming together of online information and proprietary algorithms have the power to mediate information and produce our understandings of the world. But there is much more to do in understanding how search captures our digital trails and mediates the nexus between information and power.

Funding

Notes

1. For example, in the computer science literature, Page et al. (1998) describe the PageRank algorithm for ranking search results. Meanwhile, Jansen and Resnick (2005) and Jansen et al. (2009) provide prototypical examples of the information science approach of examining the way that users perceive and interact with search results.
2. The economics literature has focused primarily on the efficiency of auctions for allocating sponsored search advertisements (e.g., Edelman et al., 2007; Varian, 2007), the rational search processes of users (e.g., Athey and Ellison, 2011; Chen and He, 2011), and search engines' incentives to work at improving the quality of their results (e.g., Eliaz and Spiegler, 2011; Taylor, 2013).
3. Some estimates suggest that there are more than one trillion URLs (http://googleblog.blogspot.co.uk/2008/07/we-knew-web-was-big.html/) and that Google indexes 24 petabytes of user-generated data every day (http://portal.acm.org/citation.cfm?doid=1327452.1327492/).
4. Digital catalogs such as the *Open Directory Project* (http://www.dmoz.org/) do continue to exist, but their importance has faded as it has become impracticable to keep pace with the rate at which content is being created.
5. Think, for instance, of the tag that is placed underneath every Google News page: "The selection and placement of stories on this page were determined automatically by a computer programme."
6. A number of Google's competitors, for instance, have set up the Fair Search website (http://www.fairsearch.org/) to promote accountability.
7. For example, Nexttag and Yelp have both given evidence against Google in antitrust hearings.
8. See, for example, Edelman and Lockwood (2011), and Wright (2011).
9. Both the US Senate and the European Commission deemed this issue worthy of investigation. See Manne and Wright (2010) for a legal analysis.
10. Baidu is the largest search engine in China by market share, whereas Google enjoys a majority market share in Europe, North America, and much of the rest of the world.
11. Baidu Baike is a Chinese-language, collaboratively authored online encyclopedia hosted by Baidu; Wikipedia is also a collaborative encyclopedia and is available in 285 languages.
12. For example, a Google search for "blue bicycle" is likely to yield fundamentally different results when performed by different people at different times or in different places.

References

Athey S and Ellison G (2011) Position auctions with consumer search. *Quarterly Journal of Economics* 126(3): 1213–1270.

Chen Y and He C (2011) Paid placement: advertising and search on the Internet. *Economic Journal* 121(556): F309–F328.

Dutton WH and Blank G (2011) Next generation users: the Internet in Britain. Oxford Internet Survey 2011, Oxford Internet Institute, University of Oxford. Available at: http://www.oii.ox.ac.uk/events/?id=453 (accessed 16 April 2012).

Edelman B and Lockwood B (2011) Measuring bias in organic "Web Search." Working paper. Available at: http://www.benedelman.org/searchbias/

Edelman B, Ostrovsky M and Schwarz M (2007) Internet advertising and the generalized second-price auction: selling billions of dollars worth of keywords. *American Economic Review* 97(1): 242–259.

Eliaz K and Spiegler R (2011) A simple model of search engine pricing. *Economic Journal* 121(556): F329–F339.

Graham M, Zook M and Boulton S (2013) Augmented reality in the urban environment: contested content and the duplicity of code. *Transactions of the Institute of British Geographers* 38(3): 464–479.

Jansen BJ and Resnick M (2005) Examining searcher perceptions of and interactions with sponsored results. In: *Proceedings of the workshop on sponsored search auctions*, 5–8 June, Vancouver, BC, Canada, pp. 5–8.

Jansen BJ, Zhang M and Schultz CD (2009) Brand and its effect on user perceptions of search engine performance. *Journal of the American Society for Information Science and Technology* 60(8): 1572–1595.

Manne GA and Wright JD (2010) Google and the limits of antitrust: the case against the antitrust case against Google. *Harvard Journal of Law & Public Policy* 34(1): 171–244.

Page L, Brin S, Motwani R, et al. (1998) *The pagerank citation ranking: bringing order to the web*. Stanford University Technical report no. 1999-66. Stanford University, Stanford, California, USA.

Pariser E (2011) *The Filter Bubble*. New York: Penguin Press.

Purcell K, Brenner J and Rainie L (2012) Search engine use 2012. Pew Internet and American Life Project, Pew Research Center. Available at: http://pewinternet.org/Reports/2012/Search-Engine-Use-2012.aspx (accessed 17 April 2012).

Rieh SY (2004) On the web at home: information seeking and web searching in the home environment. *Journal of the American Society for Information Science and Technology* 55: 743–753.

Taylor G (2013) Search quality and revenue cannibalisation by competing search engines. *Journal of Economics & Management Strategy* 22(3): 445–467.

Varian HR (2007) Position auctions. *International Journal of Industrial Organization* 25(6): 1163–1178.

Waller V (2011) Not just information: who searches for what on the search engine Google? *Journal of the American Society for Information Science and Technology* 62(4): 761–775.

Wright JD (2011) Defining and measuring search bias: some preliminary evidence. Working paper. 03 November, George Mason University School of Law, Arlington, Virginia, USA.

Zimmer M (2010) Web search studies: multidisciplinary perspectives on web search engines. In: Hunsinger J, Klastrup L and Allen M (eds) *International Handbook of Internet Research*. Dordrecht: Springer, pp. 507–521.

Author biographies

Mark Graham is a Geographer and a Research Fellow at the Oxford Internet Institute, where he focuses on Internet and information geographies and the intersections between ICTs and economic development. He is also the Director of Research at the Oxford Internet Institute and a Visiting Research Associate at the University of Oxford's School of Geography and the Environment.

Ralph Schroeder is a Professor and Director of the Master's degree in Social Science of the Internet at the Oxford Internet Institute. Before coming to Oxford University, he was a Professor in the School of Technology Management and Economics at Chalmers University in Gothenburg (Sweden). His recent books include *Rethinking Science, Technology and Social Change* (Stanford University Press, 2007), *Being There Together: Social Interaction in Virtual Environments* (Oxford University Press, 2010), and *An Age of Limits: Social Theory for the 21st Century* (Palgrave Macmillan, 2013).

Greg Taylor is a Research Fellow at the Oxford Internet Institute and an associate member of Nuffield College, University of Oxford. An economist by training, his research focuses on the microeconomics of markets that exist on or are mediated by the Internet. Within this area he is particularly interested in the industrial organization of the Internet search engine industry, the design of electronic marketplaces, and the effect of their existence upon advertising and consumer search behavior. He holds a PhD in economics from the University of Southampton.