Reliance and Reliability: The Problem of Information on the Internet

Robert Wachbroit

Internet use has grown astonishingly in the past few years. With little effort one can learn what legislation is pending in Congress, the status of medical research in various universities and institutes, and the biographies of both the famous and the obscure. The Internet also offers countless sources of news – from traditional sources, such as the New York Times and CNN, to more partisan but sometimes more probing advocacy groups. For information junkies, there is little that can beat the Internet. The good news is that everything is on the Internet. The bad news is that everything is on the Internet.

In all likelihood, future use of the Internet will only increase and, as it does, some of us will rely on the Internet for much – perhaps all – of our information. Not all of this information will be reliable. But even if much of it is, heavy reliance on the Internet for information may present a worry. This essay examines the problems of information reliability and reliance on the Internet and looks at some of the difficulties raised by recent attempts to address these problems.

The Problem of Reliability

As many people have noted, anyone with access to the Internet can be a publisher. If you can write it down, you can put it up on the Web for millions to read. As a result, the Internet is an enormous source of information and an enormous source of misinformation. One can easily come across rumors, gossip, ideological rantings, paranoid accusations, lunatic ravings, outright lies, and wishful thinking – all just a mouse-click away from the more reliable information on the Internet.

Is this so significant? After all, rumors, gossip, etc., have always been with us. We shouldn’t be at all surprised that misinformation has a future in every advance of telecommunications technology. Soon after the invention of the telephone, the word ‘phony’ entered the American language. The word, referring to a fake or a counterfeit, apparently arose from the increasingly common experience, by the beginning of the twentieth century, that the friendly voice on the telephone was too often the voice of a swindler. Today we have grown appropriately cautious about phone calls from strangers, realizing that the benefits of the technology allow for some abuses.

Nevertheless, the consequences of misinformation on the Internet can be significant. The Internet is much like a broadcast medium, spreading messages rapidly, widely, and effortlessly. But unlike most broadcast media, messages can be posted anonymously or with a pseudonym so that checking the source is difficult. At a single click, more people can be misled than is possible with the telephone. Consider some examples:

A teenager was charged with securities fraud, after apparently buying thinly traded stocks and then, using a number of pseudonyms, posting numerous messages on the Internet touting the stock. Those who believed that enthusiasm for the stock was growing, bid up its price, which soon dropped once they realized that no such enthusiasm existed for the stock. But by that time, the teenager had sold his holdings at a profit.

The Internet is an increasingly important resource for those seeking information about medical conditions. But individuals searching for information on cancer can easily discover Dr. Ryke Geerd Hamer’s Web site (www.geocities.com/HotSprings/3374/entdeck.htm), which claims to explain what cancer really is. Not only does the information on this site fall into the dangerous
category of quack medicine—among other advice, the doctor urges people to abandon “official medical treatments”—but the site also fails to inform the reader that Dr. Hamer was arrested for illegally practicing medicine.

Not all the misinformation on the Web is expensive or dangerous, however. Some is just utter nonsense, as is the case of the Web site for the Oklahoma Association of Wine Producers (www.members.aol.com/okawp/). Those who are aware that Oklahoma has no wine industry might be amused by the site, but one can also imagine someone planning a tour of Oklahoma’s wineries after reading the “information” on the site. (The authors of the site intend it as an educational tool regarding misinformation on the Internet, but one has to examine the site carefully to see the disclaimer.)

Intermediaries and Gatekeepers

Several writers see the problem of reliability as arising from the demise of “information intermediaries.” In traditional information outlets—major newspapers and book publishers, most notably—several layers of intermediaries exist between the writer and the reader. These intermediaries ensure the correctness of what is written. The reader knows that, for example, the reputation of the New York Times is behind each of its stories and that the editors, production staff, managers, and legal counsel work hard to ensure that the reputation of the paper is well deserved. The reader can thus trust what he reads in the New York Times because of the information intermediaries it employs. But once these layers of intermediaries are removed, a mechanism for ensuring reliability is dismantled. As Andrew Shapiro notes, “Where once there were reporters, writers, editors, fact-checkers, production staff, publishers, libel lawyers, and large media owners, now a worldwide dispatch may be the result of a quirky thought and a bit of tapping at a keyboard in one’s bedroom.”

This explanation of the problem suggests an obvious response. If the reliability of these traditional sources of information is acceptable, then reliable Web sites could consist of these sources having an online presence. And indeed, many of them do. Nearly every major newspaper has a Web site where many of its articles are posted, and these sites are as reliable as their print counterparts. This practice is not confined to newspapers—many commercial, governmental, and non-profit entities have also created an online presence.

One might respond, of course, that such efforts in no way eliminate or control the presence of unreliable Web sites. But that is not the issue. The concern is not the presence of unreliable Web sites—we will always have them, just as we have always had unreliable sources of information prior to the development of the Internet. The concern is identifying the reliable sources—distinguishing trustworthy information from fake. Having respectable sources maintain an online presence seems to ease this concern.

Unfortunately, this response doesn’t go far enough. It undervalues the potential of the Internet, treating it merely as a new medium that happens to be inexpensive and easy. But what excites so many people about the Internet is not so much that it is cheap and convenient but rather that it is decentralized and open. Anyone can post information without facing censorship and without seeking the approval of some information gatekeeper. Some might say that the more controversial the information, the more important is the availability of the Internet. The Internet can be a powerful democratizing force, especially in repressive societies where information is tightly controlled. Even in more open societies, it is essential for democracy that the public can get information outside of the mainstream. Identifying reliable Web sites simply on the basis of a reliable offline presence excludes the important new sources of information that the Web can provide.

Rather than intermediaries ensuring accuracy, one might propose some sort of certification procedure. Reliable Web sites could display some seal which attests to their adherence to a set of standards regarding the quality of the information presented. For example, a Web site that describes the treatment options for a particular form of cancer might reassure readers of the trustworthiness of its information by displaying a seal of approval from, for instance, the National Institutes of Health or the American Medical Association.

Despite its initial promise, a certification procedure amounts to the re-introduction of information intermediaries—perhaps not at the level of introducing people who are fact-checkers and editors, but at a meta-level of experts who certify the practices and authority of a Web site. While this might enable some new sites to emerge as information sources, trust in the site’s accuracy would actually reside in the site’s certifiers. Further, a certification process would prove unhelpful when information is deemed controversial. In fact, one can easily imagine someone unfairly stigmatizing some sites as unreliable simply because they did not have the time, resources, or awareness to submit themselves to certification. Certification raises obvious questions, finally, regarding who is responsible for this certification, how it is done, and the nature of the certification authority or trust.
More important, any effort at certification can address only one aspect of the problem of information on the Internet. After all, the problem is not really so much about unreliable information on the Internet as about people being misled by what is on the Internet. Being misled requires two elements: unreliable information and credulity. Information has no significance unless it is believed. If people easily believe nearly everything they see on the Web, then the element of gullibility must be added to the problem about the quality of the information on the Web. In this regard, it is worth noting that according to one report, about half of Internet users “believe that most or all online information is reliable and accurate.”

This suggests that we need to consider people’s behavior as consumers of information, or as knowers. If everyone were more careful about what they glean from the Internet, if they were more responsible searchers of information, then perhaps the concerns about unreliable Web sites might be of much less significance. What is it to be a responsible searcher on the Internet?

The Responsibility of Information Seekers

Believing something just because it is on the Web is irresponsible, but maintaining a thorough skepticism regarding everything on the Web is hardly better. One must develop appropriate habits for assessment and use of information on the Internet. A considerable literature has arisen on how to do just that, including some quite practical advice on how to evaluate Web sites. One can even find Web sites devoted to explaining how to evaluate Web sites. Much of the advice is simple, common sense: check the source, look for independent confirmation of the information, etc. The overall lesson is that a responsible information seeker is not passive.

The importance of active involvement in evaluating information becomes especially clear when one examines the behavior of information seekers using the Internet. While many people go to certain trusted Web sites to get information, the most common way people get information on the Internet is by using a search engine. In fact, some sources claim that more than 80% of users get their information from search engines.

By typing a few keywords or even a question, a search engine will proceed to identify those Web sites that more or less match the keywords or question. Although the precise workings of many search engines are closely guarded trade secrets, search engines generally operate in one of two ways: they either examine a proprietary database of (selected) Web sites, which they periodically update. Alternatively, search engines “crawl” through the World Wide Web, using various algorithms to identify Web sites that meet the search criterion. Using either type of search method, the result commonly yields tens, hundreds, even thousands of identified sites. Such large results are usually of little significance, since they are ranked by order of relevance and few people ever examine search results beyond the first dozen or so listings.

One might argue that the availability of search engines helps people be responsible information seekers on the Internet. For example, anyone can easily determine whether different Web sites report the same information; one can even determine whether certain Web sites have been subject to praise or blame for the information posted on the site. In contrast, few people have the time, resources, or inclination to see how the
The New Progressive Era: Toward a Fair and Deliberative Democracy

Peter Levine

A century ago, Americans embarked on a period of civic renewal and political reform. Today, amid deep dissatisfaction with our major institutions, there are signs that a new movement may revive the spirit of the original Progressive Era. Peter Levine draws inspiration from the great Progressive leader Robert M. LaFollette, Sr., and his circle, which included John Dewey, Jane Addams, and Louis Brandeis. He discusses the shortcomings of this group as well as their successes, but he argues that their ideal of a fair and deliberative democracy is right for our time. Bringing their Progressive philosophy to bear on contemporary concerns, Levine advocates campaign finance reform, an entirely different approach to regulation, new styles of journalism and civic education, and fundamental changes in the tax system. Combining philosophical arguments, historical background, empirical data, and concrete proposals, The New Progressive Era offers today's most comprehensive plan for civic renewal and political reform.

"Peter Levine's new book represents an important new voice in our national deliberations about how to revitalize American democracy. It is a thorough, thoughtful account of the contemporary relevance of the ideas and innovations of the Progressive Era and a persuasive case for a new progressive agenda in American politics."

—Robert D. Putnam, Kennedy School of Government, Harvard University

"Some books you read and put aside. Others you send to friends; this is one of those books. Peter Levine brings a rich historical and philosophical perspective to an immediate and practical question: What is going to be the effect of all the effort that has gone into civic renewal in the last decade? This book speaks to everyone from journalists to foundation executives to teachers to members of civic organizations—all citizens. Don't miss reading it."

—David Mathews, President, Kettering Foundation

272 pages; $65.00 (cloth)
$19.95 (paper)

Rowman and Littlefield Publishers, Inc.
http://www.rowmanlittlefield.com
Tel.: 800.462.6420; FAX: 800.338.4550
For examination copies, please call 800.273.5720
It is important to keep in mind that a search engine is not a "truth engine." No matter how good the search engine is in selecting the right sites, it cannot attest to the reliability of the information on that site. Wittgenstein in *On Certainty* presents the image of someone trying to check a story in a newspaper by buying other copies of the same newspaper and reading the story again. The absurdity of this effort is plain. It would hardly be less absurd if the individual buys different newspapers, all of which use the same source of information, such as a wire-service report or a press release. But this is more or less what an Internet information seeker could wind up doing. Because information can spread rapidly and globally on the Internet, even the information from a single, unreliable source can appear quickly on many different Web sites, and checking different sites may afford no better test of reliability than buying many copies of the same newspaper, or newspapers that rely on the same wire-service report. In short, using a search engine to check the reliability of information on a Web site has its flaws.

The World Wide Web is still young and, as it evolves and as search-engine technology is refined, one can expect many of these limitations to diminish although not entirely disappear. Better search-engine technology will likely increase public reliance on the Internet as the prime source of information. However, heavy reliance on search engines—even if searches become more relevant and rankings increase in trustworthiness—raises other difficulties.

The Problem of Reliance

Many observers have noted that Internet use can be customized to suit a variety of needs and desires. Not only can I choose an online newspaper, but I can also modify it in dramatic ways. I can specify that news items within my chosen topics appear on the "front page." News on topics I have little or no interest can be shifted to the end or even filtered out entirely. As one writer remarks, instead of the *New York Times*, the *Washington Post*, or CNN, I can have "The Daily Me." The significance of this customization, carried to an extreme, is two-fold: first, by allowing one person's knowledge of what is news to be so different from another's, the potential for common experiences essential to a sense of community may be drastically reduced, if not undermined entirely. And second, customized news services can foster a polarization, if not an extremism, in opinions, which, in turn, can threaten democratic deliberation. Let me expand upon these two features.

The commentators who have written on the threat to community and democracy posed by the customization of Web experiences—particularly, Cass Sunstein and Andrew Shapiro—generally direct their attention to filters. They argue that if we get our information entirely from the Internet and if we impose perfect filters—so that we exclude from our sight topics we are not interested in—then we run the risk of communities fragmenting along narrow lines of interest. Since it is important to be informed not only of what you want to know but also of what you may not (at first) want to know, news services perfectly filtered, in effect, consumer preferences can have a devastating effect on the survival of democratic communities.

Communities consist of people with a variety of interests, concerns, and viewpoints and any effort to reduce that diversity is often seen as repressive and undemocratic. Nevertheless, many writers have argued that democratic communities require the existence of public spaces or forums, accessible by all. Without a public space, the potential for community members to encounter other members with different viewpoints is diminished or eliminated. It is then a short step for one to lose all sense of the many and diverse interests that inform the community's interests.

This loss of community, in turn, can lead to polarization since, if all one hears are echoes of one's own concerns, then one can easily come to believe, from their repetition, that these concerns are the most important—perhaps the only—concerns. When one becomes oblivious to the interests and viewpoints of others, one becomes incapable of deliberation, since when a community deliberates about a course of action, the various interests and viewpoints of its members must be respectfully acknowledged in order for the deliberation to proceed in a fair and democratic manner. If news is customized to exclude all other concerns but one's own, it becomes all the more difficult to recognize these concerns when one must join together in public deliberation.

One response to the problem of fragmentation and polarization is the suggestion that, as more people spend time on the Internet and learn from it what is going on in the world, public spaces must be created on the Internet itself. If the more traditional public spaces—such as parks, public squares, even sidewalks—become less significant, as people occupy more time in cyberspace, then, so the argument goes, we need to create an "e-commons."
Filters and Search Engines

Many of the same problems regarding filters are apparent when one thinks about the increasing reliance on search engines. This becomes clear by realizing that a search engine is the mirror image of a filter. Whereas filters try to exclude what you do not want, search engines try to include only what you do want. At the limit, where filters and search engines are perfect and make no mistake, the theoretical differences between them begin to disappear. Moreover, in some ways, worries about search engines appear more urgent. The main use of filters on the Internet is to eliminate material unsuitable for children; filters for adult use are not common. In contrast, as mentioned earlier, the use of search engines is widespread.

To be sure, hardly anyone at present relies exclusively on search engines for their information, and for a variety of reasons, including the novelty and imperfections of the technology. But suppose, for the sake of argument, that we are at the limit—search engines can handle extremely sophisticated search criteria, successfully selecting only wanted items. Suppose also that a significant number of people get all of their information from search engines. In that case, it is unlikely that these individuals would search for information on topics in which they have no interest or find uncomfortable. As a consequence, biases regarding the urgency of one’s own concerns are thus reinforced, and blindness towards the concerns of others encouraged.

It might be helpful to think about the fragmentation problem in a more general way. Geography fragments the human population, and, to a rough extent, many traditional communities arise from this fragmentation. Since traditional communities are local, many of the virtues found in them arise from the habits that members of a community must develop in order to cooperate with one another to benefit themselves and the community.

The Internet, however, can induce fragmentation along lines that cut across geography, undermining traditional communities. One type of fragmentation encouraged by the Internet is “associational,” which results when individuals form associations with others who are very much like them, without regard to geography and perhaps at the expense of associations constrained by geography. Associations of like-minded people—custom communities—do not have to foster habits of tolerance or democracy. Another type of fragmentation encouraged by the Internet is “informational,” which, as I have argued, is the consequence of an exclusive reliance on search engines—we become better informed about topics of our choosing, but at the expense of our being informed about topics that matter to others.

Political scientist Robert Putnam suggests that associational fragmentation might be the greater threat to traditional communities. He may be right, but the influences of informational and associational fragmentation are likely to be mutually reinforcing and their bad effects compounded. Further, informational fragmentation can also exacerbate problems within any subgroup that relies on specialized information. Specialization in the sciences, for instance, is efficient for research and training, since an intellectual division of labor allows scientists to focus their energies on narrowly defined problem sets.

But divisions and specializations in the sciences often reflect more the various histories of particular professions and scholars than they do any objective divisions in nature. Consequently, specialization can inhibit progress and creativity because a narrow focus at times is tantamount to tunnel vision. Too often scientists become divorced from the stimulation of interaction with scientists and scholars from neighboring fields. The Internet aids specialization by putting the scientist in touch with like-minded scientists, regardless of geography. But this new interaction comes at the expense of scientists interacting with the broader community of scholars, and the result is ever more fragmented science.

At this point, we should reconsider the earlier supposition that search engines serve as the exclusive source of information for most people. That supposition might seem simply unrealistic since there is no evidence that search engines will come to displace all other sources of information. Furthermore, customization of information sources seems to occur to some extent already in traditional sources. Many traditional sources—e.g., newspapers, radio and television stations—have, or appear to have, biases or ideologies that shape how information is presented; many people choose which source to read, listen to, or watch because they are comfortable with its bias. People with right-wing ideologies usually do not read left-wing magazines, and vice versa. Thus, we seem to have an informational fragmentation even without the Internet. Consequently, the worry that the use of search engines specifically—and the use of the Internet more generally—will lead to informational fragmentation seems overblown: it is based on an unrealistic premise regarding Internet use and it ignores the informational fragmentation that we already encounter prior to the Internet.

One can begin to respond to both points by noting that fragmentation comes in degrees. We do already
have some informational fragmentation due, in part, to consumer choices among traditional information sources. And some degree of fragmentation is no doubt beneficial, as is some degree of specialization in the sciences. But informational fragmentation does not have to be complete in order for it to worry us. Nor does the existence of some informational fragmentation mean that we need not be concerned about that fragmentation increasing. We need to be aware of the potential costs of our growing use of search engines, including the loss of less targeted sources of information. If we (choose to) become increasingly less informed about topics that we have no interest in, then significant "stress fractures" in associations and community are a likely result.

This essay does not pursue the question of how much fragmentation is too much—and perhaps that question allows for little empirical precision. The conclusion one can draw is that we should now cultivate some awareness of the dangers of informational fragmentation rather than take a wait-and-see attitude until some critical threshold has been crossed. We also need to explore suggestions for developing search engines that effectively respond to these concerns. For example, search engines might include information that is not picked up by the intended search but also is not confused with it—such as banner news. A variation on this suggestion would be search engines embedded in general information Web sites: calling up a search engine involves accessing the front page of an online newspaper. Such suggestions are not without problems. What is the incentive for search engine owners to agree to such inclusions? Wouldn't such inclusions simply underscore the worry about bias? Wouldn't some of the questions about certification be applicable here as well?

The question of the reliability of information on the Internet points to both a narrow and a wide issue. The narrow issue concerns whether the information on Internet sites is reliable; the wide issue concerns the impact of a heavy reliance on Internet search engines as sources of information. While more attention must be given to the narrow issue, if we are to realize the benefits of reliable information on the Internet, we must address the wide issue as well.