DISCUSSION FORUM

Disinformation and Misinformation through the Internet: Findings of an Exploratory Study

Peter Hernon*

The quantity of information available through the Internet and the evolving national information infrastructure is voluminous. For safety nets, information providers, and users of the information superhighway, an important consideration is how to navigate the sheer quantity of available information to find that which is most relevant, accessible, and of high quality. Accessibility refers to ease of access; relevancy presumably covers appropriateness to specific information needs; and quality includes components such as accuracy, timeliness, and understandability of content. It is not enough that information is readily available; before relying on any data or information, it may be important to ascertain, for example, the veracity of the content.

Complicating the verification process, there may be different versions of an electronic record. Which one is the "official" or "more official" version and merits archiving and retention, and by whom? How can agencies maintain an audit trail documenting the development, implementation, and enforcement of policy? Such questions have been raised for almost a decade, but they have not been completely resolved to the satisfaction of everyone. Furthermore, the explosion of networked information makes their resolution more critical today than ever before.

Should the contents of a government publication, electronically produced article, or news account be accepted at face value as they appear on a listserv or elsewhere on the Internet? Perhaps a better phrased question is "How often and in what circumstances can the integrity of a document—in either print or electronic form—be taken for granted?"[1] Those concerned about the erosion of privacy in the electronic information

* *Direct all correspondence to: Peter Hernon, Simmons College, Graduate School of Library and Information Science, 300 The Fenway, Boston, Massachusetts 02115; <phernon@vmsvax.simmons.edu>.*

Government Information Quarterly, Volume 12, Number 2, pages 133-139.

ISSN: 0740-624X.

age warn that there should be no automatic acceptance of content, and that data overlays through the use of geographic information systems (GISs), in some instances, erode privacy rights. As they remind us, an extensive array of publicly available information profiles individual households and reveals a lot about our lives.[2] Furthermore, it is easy to alter information content and to pass a forgery as genuine.

Paul Wallich advises us to expect that the evolving information superhighway contains "electronic bandits and other hazards,"[3] while William J. Mitchell shows that digital technology can manipulate images, thereby subverting "the certainty of photographic evidence."[4] Tom Raum notes that text itself is easily altered. According to him, the Clinton White House "deleted the word 'lie' four times in issuing a sanitized electronic version of a press release attacking a critic of President Clinton's health plan."[5] Dennis W. Viehland, in relating how the Executive Information Special Interest Group (EISSIG) "was mistakenly identified as an address for the White House's newly established electronic mail office ... [demonstrates] how misinformation spreads in cyberspace."[6] Inaccurate information might result from either a deliberate attempt to deceive or mislead (*disinformation*), or an honest mistake (*misinformation*). Either way, incorrect information gets out. Clearly, "authenticating and verifying the integrity of a document is ... [more than] simply obtaining [and using] a copy of the document."[7]

How much of a problem does inaccuracy over the Internet present at this time? Or:

> aside from a few pranks and malicious acts (most commonly people sending electronic mail with false origin addresses) it is unclear whether the fears that people seem to harbor about the deceptive and mutable nature of the electronic environment are justified by real occurrences of problems.[8]

Such questions are impossible to answer with any degree of certitude. There are increased opportunities for disinformation and misinformation to occur, and for students, faculty, and others to unknowingly reference them. For example, they might reference secondary sources which contain mistakes of substance without checking the wording and intent of the original text; at some point, mistaken information might gain general acceptance.

For what information and under what circumstances do people verify the authenticity of the information and data they use? How trusting are they (and should they be) of information content received via the Internet? Should individuals be any more or less trusting of electronic format information? Is liability changed by information being in electronic formats? The purpose of this essay is to raise such questions through the conduct of an experiment related to disinformation:

> It seems clear that if network based information distribution is to become a widely accepted context for the sorts of archival materials that libraries currently acquire and provide access to in print these concerns must be addressed.[9]

THE EXPERIMENT

The investigator took section 6(t) of the definitions from the electronic version of Circular A-130, "Management of Federal Information Resources,"[10] which the Office of Management and Budget (OMB) proposed in September 1993. He substituted *customer* for *user*:

> The term "user" means an organizational or programmatic entity that receives information processing services from an information processing services organization (IPSO). A user may be either internal or external to the organization responsible for providing information resources services, but normally does not report either to the manager or director of the IPSO or to the same immediate supervisor.

By substituting the word *customer*, the investigator directly linked the proposed revision to President Clinton's executive order on "setting customer service standards."[11]

Using a case-study approach for this exploratory study, the investigator showed the disinformation to 16 individuals (two members of OMB, six other policymakers, and eight librarians) during the winter and spring of 1994. Each of them was interviewed separately and is knowledgeable about the government information policies of the Clinton administration. The investigator did alert those interviewed to the fact that he created disinformation and that he did not find it on the Internet.

Assumptions

In conducting this research, the investigator assumed that the more dramatic the disinformation, the more likely that *knowledgeable* individuals will verify the content. Others might be more accepting of what they see on the screen or on printout. Another assumption is that the more referencing an item receives, without the explicit label of disinformation or misinformation, the more likely that information will be credible and accepted. As Lynda Gorov, a reporter, has written, "in a rush to break news, local television stations [have] aired . . . inaccuracies. And their errors were compounded when other news agencies picked up the stories and disseminated them nationwide."[12] It was outside the scope of this small-scale study to explore these assumptions or to interact with the *customers* of an agency.

Research Questions

- How trusting should we be of information available through the Internet?
- What is the reaction of those interviewed to the creation of disinformation? Do they regard disinformation and misinformation on the Internet as presenting a substantive issue?
- Are they aware of specific instances of disinformation and misinformation, intentional or other, on the Internet?
- If there are differences among sources, what do they regard as the official or authentic version, assuming one exists?

Findings

Accuracy of Information Available through the Internet

The consensus of those interviewed was that information available on the Internet is no better or worse than information distributed in other ways. Still, it is important that agencies provide a digital signature to certify the accuracy of the information which they disseminate. For this reason, the Copyright Office of the Library of Congress is investigating digital signatures. A problem, as one of those interviewed noted, will be to apply such signatures to all publications, information, and datasets.

Reaction to the Creation of Disinformation and Misinformation

Clearly, for those not well informed about the law and government information policy, the falsified information used for this experiment might seem believable. Those interviewed were concerned that the public might use disinformation and misinformation and that the government needs to develop digital signatures. Yet, "how can we protect against every possible misuse of information," one policymaker queried. When agency officials do not read the original text of a public law, but rather rely on policy instruments such as A-130, it is important that they have access to authentic versions of these instruments. Nonetheless, they take circulars and convert them into internal policy, which they enforce. Therefore, the experiment has more implications for information users, other than those in Federal agencies.

One policymaker and one librarian noted that "we can put quotation marks around anything and change meaning." For example, as they pointed out, it would be possible to define *informational matter* (44 *USC* 1901) as referring to content and format. By so doing, it would be possible to further blur the distinction which Circular A-130 attempts to make among a government publication, an information product, and an electronic information product. Who would use this misinformation or disinformation, and for what purposes? The misuse might be exposed, but the person doing the misuse might only be guilty of taking something publicly available, through a listserv or electronic journal or newsletter, without checking the original source.

Awareness of Instances of Disinformation and Misinformation on the Internet

It merits mention that, in its draft circulars, OMB may insert quotation marks even if it is paraphrasing content.[13] For example, 44 *USC* 1901 defines a government publication as "informational matter which is published as an individual document at government expense, or as required by law." OMB has dropped the "al" from "informational" and the word *matter* from drafts of Circular A-130 because it considers them to be superfluous. As long as OMB does not believe it has altered the meaning of public law, quotations within circulars might not correspond exactly to the actual wording of a statute.

The significance of this finding is disputed. One policymaker noted that the term *informational matter* might open the door for software distribution and, if so, OMB and the agencies might not favor it. The other policymakers interviewed appreciated knowing what convention OMB used in writing circulars but believed that no significant issues were involved. As one of them noted, "how important is this one matter in comparison to other issues requiring examination and oversight?" Furthermore, "does this distinction make a difference?" On the other hand, the librarians interviewed were "shocked" to learn about the practice and wanted to alert their clientele to be more selective in quoting public laws from a circular. Ultimately, the issue seems to be "To what extent do I quote from secondary sources, even key policy instruments, as opposed to the original one?"

Those interviewed shared examples of inaccuracies between drafts of A-130 and the circular itself, such as those relating to definitions of records management and information resources management. In one instance, OMB's General Counsel apparently reviewed the wording, interpreted it as changing legislative intent, and required that the wording be changed. The librarians interviewed offered examples of incorrect e-mail and remote access addresses for document receipt, and they claimed that electronic versions of draft documents, such as those from OMB, might be harder to find and download than the agency claimed. In such cases, documentation might be misleading rather than intentionally inaccurate.

Sources of Official or Authentic Versions

Those government officials interviewed noted that if there are any inconsistencies between the Internet version and the version in either the *Federal Register* or the *Congressional Record*, the latter is more "official." The contents of documents which agencies released, they noted, however, should be identical to those appearing in electronic form.

Two of those interviewed suggested that the Government Printing Office (GPO) might become the official repository for the electronic version of titles such as the *Congressional Record* and the *United States Code*. However, as was noted, Congress will always have one or two formal, official copies but it might choose to produce public copies only in electronic format. The National Archives and Records Administration would continue to be the repository of the official versions, for example, the official parchment version.

TOPICS MERITING FURTHER EXPLORATION

This investigator feels uncomfortable with recommending that disinformation and misinformation be placed on the Internet, in order to monitor patterns of referencing and use, in part because methods for retraction remain imprecise and incomplete. He would dislike for disinformation and misinformation presented through an experiment such as this one to be widely accepted by an uncritical public, one not comparing secondary to primary sources.

Still, it is possible to take something from the online version of the *Congressional Record* or *Federal Register* and to pursue the above-mentioned research questions, in greater detail, with individuals who are, at some point, informed about the nature of the experiment. An experiment might include falsified census data graphically displayed through a GIS and extend to information "customers," including undergraduate and graduate students, contractors, and the business sector. A presumption worthy of testing is that the more uncritical references that disinformation and misinformation receive, the more likely that an untruth becomes accepted as truth. Besides, what is the role of government regarding meeting the information needs of its customers? Despite the fifth interagency conference on public access to government information, which considered agency "customers,"[14] the term "customers" has received insufficient attention in the policy literature. To what extent does each and every agency meet the needs of its customers in an effective and efficient manner, and what is the extent of customer satisfaction and service quality? How can these be measured through performance measures that examine outcomes not outputs?

OMB's Circular A-130 and its drafts have defined government publication, information products, and electronic information products, but do not adequately differentiate among these terms. Moreover, OMB has not offered a clear and consistent vision of the information life cycle.[15] What, if any, are the implications of these oversights and failures from a policy and customer perspective? The opportunities for the conduct of an experiment are many, but those doing such research must adhere to the highest ethical standards.

CONCLUSION

In the case of government information, the source of disinformation and misinformation might be individuals within the agency or White House, or the agency itself, an intermediary (presumably the Government Printing Office, National Technical Service, and so forth), a user of that information, or a hacker or "electronic pirate." The more authoritative the user appears, presumably the more reliable is the information used. However, even knowledgeable individuals might be duped.

The question underlying the experiment is, "Does the vulnerability of information available through the Internet or the evolving information superhighway reduce their credibility as *safe highways*?" This is an important question which is beyond the scope of this essay to address. However, if the highways prove unsafe, what are the implications, especially as the Federal government attempts to disgorge more information into locator systems and the superhighway? Clearly, we need to support efforts to develop digital signatures and other authentication techniques.

ACKNOWLEDGMENT

The author wishes to thank the Emily Hollowell Fund at Simmons College for providing the funding necessary to conduct the research upon which this study is based.

NOTES AND REFERENCES

1. As noted in *Accessibility and Integrity of Networked Information Collections*, background paper (Washington, D.C.: Office of Technology Assessment, 1993):

> In the print environment a great deal is taken for granted about the integrity of documents. If an article appears in a journal it is extremely rare that the authorship that the journal lists for the article is called into question; when this happens it is framed either in the context of scientific and scholarly misconduct such as plagiarism and/or results in a lawsuit (p. 61).

> Perceptions and concerns in the world of networked information are quite different. It is very easy for someone to distribute information over someone's name, and hard to trace the person who does it in most cases. It is very easy to replace an electronic dataset with an updated copy, and, since there is no automatic system which distributes multiple copies of the original version to different sites ... the replacement can have wide-reaching effects. The processes of authorship, which often involve a series of drafts that are circulated to various people, produce different versions which in an electronic environment can easily go into broad circulation; if each draft is not carefully labeled and dated it is difficult to tell which draft one is looking at, or whether one has the 'final' version of a work. Because of ease with which material can be taken from one document and inserted into another which can then be circulated to a large number of people quickly, there are concerns about quotation from obsolete or draft ("unpublished") versions of a work (p. 62).

[2] See Peter Hernon, "Privacy Protection and the Increasing Vulnerability of the Public," *Government Information Quarterly*, 11 (1994): 241-244. See also the responses to this discussion forum (pp. 245-260).

[3] Paul Wallich, "Wire Pirates," *Scientific American*, 270 (March 1994), p. 91.

[4] William J. Mitchell, "When Is Seeing Believing?," *Scientific American*, 270 (February 1994), p. 68. According to *Accessibility and Integrity of Networked Information Collections*, "a public key cryptosystem can be used to attach a signature to a digital object in such a way that the contents can be associated with a given individual or organization at a given time" (p. 63). However, "while the basic technology exists to solve the problems in question (at least as long as one is satisfied with literal bit-by-bit equivalence of two digital objects as a definition of having the 'same' document, which is often really overly restrictive, since it prevents any reformatting, character code conversion or other activities that might be needed to successfully move the document from one machine to another, even if these do not change the 'content' of the document in any way) the operational problems of implementing these technologies on a large scale in environments such as the Internet are far from solved" (p. 63). Problems might relate to the need for standards, patent issues, export restrictions related to cryptographic technology, and so forth.

[5] Tom Raum, "Memo Tests Honesty as Best Policy," *Middlesex News* [Framingham, Massachusetts] (February 11, 1994), p. 4A. See also William P. Cheshire, "Hoaxing Along the Infobahn," *The Arizona Republic* (October 30, 1994), p. E1. He discusses H.R. 5904, "Sniper Prevention and Firearms Collection Act," supposedly introduced by Representatives Karen English and Coppersmith. "The 'bill' is a hoax, one of countless bogus documents circulating among the patrons of the nation's computer networks."

[6] Dennis W. Viehland, "Dear Mr. President: A Story of Misinformation Distributed in Cyberspace," *Internet Research*, 3 (Fall 1993), p. 57.

[7] *Accessibility and Integrity of Networked Information Collections*, pp. 65-66.

[8] *Accessibility and Integrity of Networked Information Collections*, p. 62.

[9] Ibid., pp. 62-63.

[10] Office of Management and Budget, "Circular A-130, Management of Federal Information Resources" [ftp: nis.nsf.net. as omb/omb.a130.rev3]. Also *Federal Register* 58 (September 10, 1993): 47790-47797.

[11] President William J. Clinton, Executive Order 12862, "Setting Customer Service Standards" (September 11, 1993).

[12] Her comments relate to the accuracy of some of the reports on the murder case against O.J. Simpson. See Lynda Gorov, "The Scoop on Simpson: Beware News," *The Boston Globe* (September 26, 1994), p. 3.

[13] OMB might take "poetic license" when it finds it necessary to make language less archaic and to shorten a definition. For instance, the agency might do so in referring to the text of the Brooks Act or the Paperwork Reduction Act, and in discussing information technology. OMB does not cite a statute if it intentionally deviates from official language. Is there a practical difference between the law and the interpretation? If OMB feels it erred in a draft circular, it may make the correction in the circular.

[14] Neil J. Stillman and Norman Oslik, "Working with the Public to Ensure Public Access to Federal Information in an Electronic Age: Proceedings of the Fifth Solomons Interagency Conference on Public Access, June 27-28, 1994," *Government Information Quarterly*, 12 (1995), pp. 163-198.

[15] See Peter Hernon, "Information Life Cycle: Its Place in the Management of U.S. Government Information Resources," *Government Information Quarterly*, 11 (1994), pp. 143-170.