ARTICLE

Diogenes† Wanders the Superhighway: A Proposal for Authentication of Publicly Disseminated Documents on the Internet

*Kelly Kunsch**

---

† Diogenes was a Greek philosopher circa 320 BC. Tradition ascribes to him the famous search for an honest man conducted in broad daylight with a lighted lantern. 4 THE NEW ENCYCLOPAEDIA BRITANNICA 107 (1990).

* Reference Librarian, Seattle University School of Law. J.D., M.L.S. University of Washington. Mr. Kunsch lectures at the law school on legal research (including computer-assisted legal research and searching the Internet). He has logged thousands of hours on the Internet assisting faculty, students, and attorneys find both legal and nonlegal information.

AUTHOR'S NOTE: After the writing but prior to the printing of this article, new protocols for domain name registration were created. Under the new system, the National Science Foundation no longer oversees the assignation of domain names. Concurrently, entities other than Network Solutions are allowed to register domain names. In addition, several new domains (such as ".firm" and ".store") have been created. None of this should affect the arguments proposing a verification system for domain name registration set forth in this article. *See* Jon Swartz & Julia Angwin, *Web to Get Wider—New Address Options*, S.F. Chron., May 2, 1997, at A1.

> A popular Government, without popular information, or the means of acquiring it, is but a Prologue to a Farce or a Tragedy; or, perhaps both.[1]
>
> James Madison

> [L]et your reason serve
> To make the truth appear where it seems hid,
> And hide the false seems true.[2]
>
> William Shakespeare

I. INTRODUCTION

With the swiftness of a winter night, the Internet appears to have draped itself across our entire world.[3] In that time, it has taken on

---

1. 9 WRITINGS OF JAMES MADISON 103 (Gaillard Hunt ed., 1910).

2. WILLIAM SHAKESPEARE, MEASURE FOR MEASURE act 5, sc. 1, lns. 65-67 *in* I THE ANNOTATED SHAKESPEARE (A.L. Rowse ed., 1978).

3. In 1981, fewer than 300 computers were linked to the Internet. In 1989, fewer than 90,000 were linked. ACLU v. Reno, 929 F. Supp. 824, 831 (E.D. Pa. 1996). By 1993, over 1,000,000 computers were linked. *Id.* In 1996, over 9,400,000 host computers were estimated to be linked to the Internet. *Id.*

The *ACLU* case, which found unconstitutional parts of the Communications Decency Act, Pub. L. No. 104-104, § 502, 110 Stat. 133, 133-36 (1996), sets forth findings of fact that describe the Internet, its uses, and its technology. *Id.* at 830-49. Many of the findings were stipulated to by the parties. The case is an excellent source for understanding the basics of the Internet.

many capacities. To name a few of its most prominent uses, it is a vehicle for personal communication, a vehicle for publication, an advertising medium, an entertainment medium, and a research source.[4] Presently, the Internet performs some of these tasks better than others. Those functions that possess the most potential for commercial enrichment are constantly being improved to make them not merely viable, but preferred alternatives. By contrast, other (less monetarily rewarding) functions have received less consideration.

On the national level, there are proposals to make the Internet the primary, and even the exclusive, means of disseminating certain government information.[5] Concurrently, corporations and other private organizations may adopt a similar approach for their reports and other documents. Intertwined with these official and quasi-official documents are innumerable others created by individuals around the world.[6] With so many documents, there is potential for dissemination of false, biased, and even fraudulent information.[7] This is the source of the authentication problem.

Due to the plethora of documents and authors (or publishers, as the case may be), how is it possible for anyone to determine the reliability of the data in a particular document? One particular area that seems prone to abuse with potentially fatal consequences is health

---

The Internet's history (including its gestation period) spans four decades. However, its present structure and particularly its pervasiveness is a phenomenon of the 1990s. *See generally* Robert H'obbes' Zakon, *Hobbes' Internet Timeline v2.5* (last modified Aug. 15, 1996) <http://info.isoc.org/guest/zakon/Internet/History/HIT.html>; Bruce Sterling, *Internet*, THE MAGAZINE OF FANTASY AND SCIENCE FICTION, Feb. 1993, at 105-06; *An Internet Time Line*, INSIDE THE INTERNET, Sept. 1996, at 12.

4. *See* ACLU v. Reno, 929 F. Supp. at 842-43.

5. According to the Government Printing Office's three year strategic plan, fifty percent of government information will be made available through electronic means by the end of Fiscal Year 1998. FEDERAL DEPOSITORY LIBRARY PROGRAM: INFORMATION DISSEMINATION AND ACCESS STRATEGIC PLAN, FY 1996-FY 2001, *reprinted in* U.S. GOVERNMENT PRINTING OFFICE, REPORT TO THE CONGRESS: STUDY TO IDENTIFY MEASURES NECESSARY FOR A SUCCESSFUL TRANSITION TO A MORE ELECTRONIC FEDERAL DEPOSITORY LIBRARY PROGRAM E-23 (1996) [hereinafter REPORT]. Although some of these documents may be published in other formats as well, such redundant versions will be reduced over time. Decisions will be based on factors such as the usability, intended audience, time sensitivity, and costs. *Id.* at E-6. A core list of publications has been identified which must remain in paper format regardless of their availability in other formats. *Id.* at E-17.

6. *See* ACLU v. Reno, 929 F. Supp. at 837 ("The Web . . . contains a variety of documents prepared with quite varying degrees of care, from the hastily typed idea, to the professionally executed corporate profile.").

7. WALT CRAWFORD & MICHAEL GORMAN, FUTURE LIBRARIES: DREAMS, MADNESS AND REALITY 79 (1995).

care information.[8] Before the issue is resolved, Pierre Salinger will not be the only victim that has fallen prey to misinformation posted to the Internet (a "suppost"?).[9]

In addition to the basic authentication problem, this author believes there is the problem with authenticating historical digital documents that have been stored in a site other than the original one. As information ages, it is often stored (or archived) at a different site than it was stored when originally published. This will be just as true of digitized documents as it is of printed manuscripts. However, when digitized information moves, it has the potential to lose its imprint of authority. Those who seek historical information in a digital archive must also be able to authenticate documents they receive.

To a limited extent, some Internet users have identified the above problems and are creating solutions catered to their specific needs. Both the government[10] and some commercial users[11] are in the process of creating modes of authentication for certain of their transactions. The Internal Revenue Service and private financial institutions, for example, are in the process of implementing security and authenticative measures to allow individuals to files taxes and transfer money using the Internet as their medium.[12] These modes,

---

8. *See, e.g.*, Marilynn Larkin, *Health Information Online*, FDA CONSUMER, June 1996, at 21. In Larkin's article, an FDA compliance officer describes a web site that purported to have a cure for a very serious disease. Visitors to the site were advised to stop taking their prescription medication and buy the product sold at the site. *Id.* at 22.

> The article also points out the difficulty in determining reliability of medical sites:
> Since anyone with a Web page can create links to any other site on the Internet—and the owner of the site that is "linked to" has no say over who links to it—then a person offering suspect medical advice could conceivably try to make his or her advice appear legitimate by, say, creating a link to FDA's Web site. What's more, health information produced by FDA or other government agencies is not copyrighted; therefore, someone can quote FDA information at a site and be perfectly within his or her rights. By citing a source such as FDA, experienced marketers using careful wording can make it appear as though FDA endorses their products.

Id. at 23.

9. Salinger believed the authenticity of an Internet document that claimed TWA Flight 800 had been shot down by a missile fired from a Navy ship on a training exercise. He even went so far as to assert to the press the veracity of the statements in the document, for which he was widely ridiculed. *See* Mike Royko, *One Good Story from the Internet Deserves Another*, CHICAGO TRIB., Nov. 13, 1996, § 1, at 3.

10. *See* Gary H. Anthes, *Feds to Secure 'Net Access*, COMPUTERWORLD, May 27, 1996, at 69.

11. *See* John Fontana, *VeriSign Aims to Secure E-Commerce*, COMMUNICATIONS WEEK, July 29, 1996, at 4 (announcing joint ventures between digital authentication service provider VeriSign and Visa). The article also says MasterCard was forming a similar alliance with an authentication service provider named CyberTrust.

12. *See* Kevin Power, *IRS SSA to Let Public Try Digital Signatures*, GOVERNMENT COMPUTER NEWS, Nov. 13, 1995, at 1.

however, seem limited to transactions between particular parties, as opposed to being applicable to documents intended for public dissemination. There remains a need for an authentication system that encompasses all components of the Internet—government, corporations, political organizations, and even individuals.

In resolving this problem, the challenges for any solution are not merely the technological impediments.[13] Novel governmental and legal issues created by the Internet's multinational and international nature are inevitable.[14] Also entangled in this web are the denizens of the Internet and their preferred laissez faire method of operation.[15] For them, it is important to emphasize that a system for authentication is not an attempt to monitor content. Rather, it is an attempt to empower Internet users to assess the value of the information on it.

The purpose of this article is to extend the discussion beyond the mere identification of needs to actual proposals of solutions. From that point, interested parties can suggest alternative solutions until (hopefully) the optimal solutions are implemented.

---

13. The first barrier is to overcome viewing the Internet as an alternative version of a different medium. Perhaps the most important finding of fact made in *ACLU v. Reno* is as follows: "The Internet is . . . a unique and wholly new medium of worldwide human communication." ACLU v. Reno, 929 F. Supp. at 844.

This author's opinion is that the Internet's technical problems are as follows:

(1) Keeping information secure from tampering while at the same time allowing wide-scale access. If sites like the Pentagon are subject to tampering, less secure sites will likely become prey to scam artists. *See Pentagon Closes Its Web Sites to Repair Damage by Hacker*, LOS ANGELES TIMES, Dec. 31, 1996, at A12.

(2) Allowing for the migration of information from one site to another. This is substantially different than reshelving a book in a different building. When digitized information moves, its accession point may also change, making it akin to a book changing its title every time it is moved to a new building.

(3) Creating an attitude of cooperation among the innumerable commercial enterprises on the Internet so that a solution can be arrived at in a logical and, perhaps, linear manner.

14. In the author's opinion, the legal problems raised by the Internet include the following:

(1) Whether Congress should attempt to regulate the Internet (or specifically, the domain name registration process).

(2) Even if Congress chooses to regulate the Internet, sites in other countries will remain outside of that mandated regulatory system. *See* David Post, *The New Electronic Federalism*, AMERICAN LAWYER, Oct. 1996, at 93-94. How should those be regulated (if at all)? One possibility would be an international convention on the registration of domain names. But even so, countries could refuse to sign.

(3) Whether the government should compel certain industries or businesses to archive nongovernmental information, or leave that decision to the private sector.

15. Although recognizing the necessity for some governmental regulation of the Internet, the Internet community generally favors a self-regulating environment. *See* Cynthia Flash, *Task Force Trying to Develop Internet Guidelines*, THE TACOMA NEWS TRIB., May 26, 1996, at A1. *See generally Internet Law and Policy Forum* (visited Mar. 17, 1997) <http://www.ilpf.org>; *The Electronic Frontier Foundation* (visited Mar. 17, 1997) <http://www.eff.org>.

Before proceeding, however, certain terms used in this article should be defined.

II. DEFINITIONS

Document: The word "document" is used throughout this article. What "document" means in this context is a data file. When the data in the file is accessed, it may, in fact, turn out to be a writing, picture, motion picture, sound, combination of these, or possibly some other method of communication or sensation.[16]

Publicly Disseminated: The phrase "publicly disseminated" is used instead of "public" because the latter is often interpreted as pertaining to a governmental entity.[17] Many of the documents on which people rely in researching information and offering exhibits in court, for example, are not produced by the government. Thus, publicly disseminated documents would include corporation reports, press releases by organizations like Amnesty International, and even announcements by individuals that are meant to be read by members of the public at large.[18]

Authentication: Authentication, here, means proving that a document is, in fact, what it appears to be or purports to be.[19] Such a definition includes, but extends beyond, the legal definition that is concerned with a document's admissibility into evidence.[20] With more people using the Internet to do research, the reliability of the information on it becomes important.

There are three important components to authenticating a document on the Internet.

Origin: The document must have been written or published by the person or entity that claims authorship.[21]

---

16. *See* ACLU v. Reno, 929 F. Supp. at 836. The opinion uses the word "document" while recognizing that the display may be something other than a traditional document. *Id.*

17. *See* BLACK'S LAW DICTIONARY 1227 (6th ed. 1990).

18. *See* JOE MOREHEAD & MARY FETZER, INTRODUCTION TO UNITED STATES GOVERNMENT INFORMATION SOURCES (4th ed. 1992). The definition of *public document* included in 44 U.S.C. § 1901 deleted the words "reproduced wholly or partially at government expense" and substituted the words "reproduced for official use of a government entity." *Id.* at 13. "This was intended to clarify the status of scientific or scholarly works produced under government grants, which are not public documents unless reproduced for official use of a government agency." *Id.* The definition of *public document* was later changed to "*government publication.*" *Id.* at 14.

19. *See* I THE OXFORD ENGLISH DICTIONARY 796 (J.A. Simpson & E.S.C. Weiner, eds., 2d ed. 1989). Definition 6 reads as follows: "Really proceeding from its reputed source or author; of undisputed origin, genuine. (Opposed to *counterfeit, forged, apocryphal.*)" *Id.*

20. *See, e.g.*, BLACK'S LAW DICTIONARY at 132.

21. *See* CRAWFORD & GORMAN, *supra* note 7, at 78.

Integrity: In addition to merely verifying that a document came from the entity that created it, a user must also be assured that the document has not in any way been altered. Unlike its print counterpart, information in electronic formats may be suspect due to the ease with which it may be altered.[22]

Currency: Finally, documents on the Internet are often not conspicuously dated. Thus, if a document is subject to updating or revising, the user needs to make certain the document reflects the desired time period.[23]

III. AUTHENTICATION IN TRADITIONAL PUBLIC DISSEMINATION OF INFORMATION

From the time of Gutenberg's invention of the movable type press until the last half century, large scale dissemination of information was accomplished primarily through the print medium.[24] More recently, other media such as radio, television, and videotape were utilized.[25] The information disseminated through all these media was, in a sense, unconsciously authenticated. Factors such as the format, process, and expense of producing the information filtered out most of the unreliable information.[26] First, the author's name was usually prominently noted. The author's reputation gave evidence of the reliability of a document. In addition, an author could be held liable

---

22. Maynard Brichford & William Maher, *Archival Issues in Network Electronic Publications*, LIBRARY TRENDS, Mar. 22, 1995, at 701, 704. One should note, however, that print copy is not without its potential for tampering or falsifying; even judicial opinions have been falsified. *See* Catt v. Ark., 691 S.W.2d 120 (Ark. 1985). This case, reported April 1, was a figment of the imagination of Arkansas Supreme Court Justice George Rose Smith. Carrie Rengers, *Delaware Court Finally Catches Judge's Joke*, ARKANSAS DEMOCRAT-GAZETTE, Apr. 9, 1996, at 8E.

23. *See* CRAWFORD & GORMAN, *supra* note 7, at 78.

> A user following up a citation needs to know that the article he or she is reading is the article as it was when it was cited. At the very least, if it is not exactly that article but the author's current version, or one that has been changed by another person, the fact that the article has been changed (and ideally the changes themselves) should be clearly indicated.

Id. The printed text, by contrast, represented the words of an author in a definitive or "final" form. DAVID CROWLEY & PAUL HEYER, COMMUNICATION IN HISTORY: TECHNOLOGY, CULTURE, SOCIETY 91-143 (2d ed. 1995).

24. *See generally* CROWLEY & HEYER, *supra* note 23, at 91-143. "In the later Middle Ages print helped democratize the reading public. It lessened the control over literacy exercised by church scribes." *Id.* at 307.

25. *See generally* CROWLEY & HEYER, *supra* note 23, at 225-358.

26. *See* CRAWFORD & GORMAN, *supra* note 7, at 25-26 (discussing the filtering and gatekeeping function of traditional publication).

for certain statements (e.g., fraud and defamation).[27] Likewise, a publisher evaluated the information for potential liability.[28] This provided a second check on a document's truth and authenticity. Finally, the expense of printing or otherwise producing copies and distributing them assured that they had enough value to offset the costs incurred in their dissemination.[29] With the Internet, by contrast, the publisher has virtually disappeared from the equation and costs of large-scale dissemination have fallen dramatically.[30] Only the author is left in the public dissemination process and that person has the option of remaining anonymous.[31] The only form of identification available to a reader is often the URL[32] (Internet address) of the website.

A. *Public Dissemination of Governmental Information*

Within the larger schemata of information publication, the United States government has taken a special role in assuring that its citizens are guaranteed the information necessary to understand and utilize its services. Following is a brief history and description of the federal government's public dissemination system.

---

27. For defamation, see W. PAGE KEETON ET AL., PROSSER AND KEETON ON THE LAW OF TORTS §§ 111-116A (5th ed. 1984). For other causes of action, see Steve Reitenour, *Liability for Injuries Caused by Printed Media*, 14 J. OF PROD. LIAB. 71 (1992). The author lists six different theories by which to find a defendant liable: (1) negligence; (2) breach of warranty (express and implied); (3) strict liability; (4) misrepresentation; (5) malpractice; and (6) incitement, imitation, or invitation. *Id.* at 72-73.

28. *See* W. PAGE KEETON ET AL., *supra* note 27, § 113, at 810.

> Those who manufacture books by way of printing and selling them, and those who print and sell newspapers, magazines, journals, and the like, are subject to liability as primary publishers because they have the opportunity to know the content of the material being published and should therefore be subject to the same liability rules as are the author and originator of the written material.

Id.

29. "[M]arket itself is for most literary and informative writing a simple and true test of merit: if a book cannot be published viably perhaps it is simply not good enough and should not be published at all. Vox populi—vox Dei." JOHN P. DESSAUER, BOOK PUBLISHING: A BASIC INTRODUCTION 37 (1989).

"A printed book in today's economy of writing must . . . speak to an economically viable or culturally important group of readers." CROWLEY & HEYER, *supra* note 23, at 337-38.

30. ACLU v. Reno, 929 F. Supp. at 843 ("[T]he Internet provides an easy and inexpensive way for a speaker to reach a large audience, potentially of millions. The start-up and operating costs entailed by communication on the Internet are significantly lower than those associated with use of other forms of mass communication . . . .)."

31. Anonymity is important to many Internet users. *Id.* at 849.

32. "URL" stands for "Uniform Resource Locator." It is the official technical description of the document's Internet location. HARLEY HAHN & RICK STOUT, THE INTERNET COMPLETE REFERENCE 507 (1994).

Since its founding, the United States has adopted a policy of making its most important publications readily accessible to its citizens.[33] "[T]he Constitution," said Justice Brennan, "protects the right to receive information and ideas."[34] The means for carrying out this function is the federal government's Depository Library Program.[35] The program sets up a system in which certain libraries are deemed depository libraries.[36] Through these libraries, all government publications are made available to the public except those which have no public interest or educational value or are protected for national security.[37]

Over the last decade, several legislative and administrative initiatives, including the Paperwork Reduction Act of 1995, the Government Printing Office Electronic Information Access Enhancement Act of 1993, and the 1994 revision of OMB Circular A-130, have attempted to address and advance the shift in government dissemination methods from paper to electronic media.[38] In August 1995, the United States Government Printing Office (GPO), at the direction of Congress,[39] initiated a cooperative study to identify measures necessary for a successful transition to a more electronic Federal Depository Library Program. The final report was submitted to Congress in June 1996. In the report, the GPO identified numerous issues including both authentication and archiving,[40] but they have yet to suggest any potential solutions to the problem.

---

33. 44 U.S.C. § 1902 (1988). Publications excluded from the act are "those determined by their issuing components to be required for official use only or for strictly administrative or operational purposes which have no public interest or educational value and publications classified for reasons of national security." *Id.* A "government publication" is defined as "informational matter which is published as an individual document at Government expense, or as required by law." 44 U.S.C. § 1901 (1988).

34. Bd. of Educ. v. Pico, 457 U.S. 853, 867 (1982). Brennan claimed that the right was "an inherent corollary of the rights of free speech and press that are explicitly guaranteed by the Constitution." *Id.*

35. *See* 44 U.S.C. §§ 1901-1916 (1988). For a discussion of the history and procedures of the Depository Library System, see JOE MOREHEAD & MARY FETZER, INTRODUCTION TO UNITED STATES GOVERNMENT INFORMATION SOURCES 47-77 (4th ed. 1992).

36. 44 U.S.C. § 1907 (1988).

37. 44 U.S.C. § 1902 (1988).

38. REPORT, *supra* note 5, at 1.

39. *See* S. REP. NO. 114, 104th Cong., 1st Sess. at 48-49 (1995).

40. REPORT, *supra* note 5, at 4-5.

B. Public Dissemination of Nongovernmental Information

Government entities are not the sole parties interested in public distribution of information. Businesses want to report their quarterly earnings, research institutions want to announce the results of their studies, and public interest or other associations want people to know and understand their beliefs on important issues. In addition, every entity selling a product wants the buying public to know the value of its product.

Traditionally, public dissemination of this information has been done using the print media.[41] For information that did not hold long-term value, this was done with a press release or paid advertisement.[42] If greater permanency of the information record was sought, the creator of the document would have numerous copies of the report or study printed. The entity would then either sell copies of the document or give them away to the desired audience.

The foregoing discussion describes the framework upon which the Internet has been overlaid. The following discussion will show that this overlaying of a new medium is not a perfect fit. Ultimately, the misalignment of media is the impetus for the authentication problem that is the focus of this article.

IV. USE OF THE INTERNET FOR RESEARCH AND THE AUTHENTICATION PROBLEM

A comparison of Internet research with other on-line research will aid a fuller appreciation of the unique authentication problem in using the Internet for research.[43]

---

41. The "culture of consumption" arrived with the market-industrial society. *See generally* CROWLEY & HEYER, *supra* note 23, at 218-222. Business and industry awakened to the need for a greatly intensified selling effort in order to move the goods cascading off their assembly lines. *Id.* During this period, advertising came to constitute the largest share of print media revenues. By 1920, it accounted for about two-thirds of all newspaper and magazine income. *Id.*

42. In such cases, the newspaper or magazine that published an article based on the press release or printed the advertisement could act as the authenticating agent. Verification of the source in such instances was generally informal.

43. As a point of information, this author uses and is aware of three major methods of searching for information on the web. The first is by typing a known address into the browser software to go directly to that site. This method can only be used if a user knows the precise address where the document is located. The second method for finding information on the Internet is by searching through the hierarchical indexes that are available from companies like Yahoo. In this way, a user slowly narrows down the subject matter until the precise document can be pinpointed. The third method of searching for information on the web is by using one of the numerous search engines to search for key words. This allows a user to find the exact document with minimal browsing. Even if the exact document is not found with the search, the

The initial movement to on-line publication has been a reprinting of print sources rather than original dissemination in an on-line format.[44] Thus, there has existed at least one, and usually many, original documents to compare the on-line version with for verification of its accuracy. The lack of such verifiable copies is at the center of the problem.

For lawyers, a sharp contrast can be made between the Internet and on-line services like LEXIS, Westlaw, or the various CD-ROM services.[45] When lawyers use these latter products, they rely without thinking on the accuracy of the documents in them. Although they are not searching and retrieving original source documents, they rely on the accuracy of the vendors (*e.g.*, West Publishing and Reed-Elsevier) in constructing their databases.[46] They trust that the on-line versions accurately display the original documents: legislative enactments, judicial opinions, and the like. Even if lawyers might be skeptical of the truth of what they read in an on-line newspaper or law review, they trust that the services are accurate in reporting who wrote the article and where it was published and that the on-line version accurately represents the print version. They can then assess the authors and publishers for possible biases.

The Internet, by contrast, has no content control. No overseeing entity checks for truth of content or accuracy in the posting of documents. The major burden is moved from publishers and vendors of on-line service (such as LEXIS and Westlaw) to the user. And the only way the user can begin to determine the authoritativeness of a document on the web is by knowing where it is disseminated from. And that can only be done through a cumbersome process that may include substantial guesswork.

---

links that are retrieved may lead the user to the document. This method seems particularly open to retrieving questionable information because it merely operates by matching words. There is absolutely no check like that provided by a topical arrangement, which allows a user to check other documents against the one in question.

44. According to Marshall McLuhan, a new medium first tries to incorporate the form and content of previous media. CROWLEY & HEYER, *supra* note 23, at 309. McLuhan refers to the "bias and blindness induced in any society by its pre-existent technology." MARSHALL MCLUHAN, UNDERSTANDING MEDIA: THE EXTENSION OF MAN 304 (1964). In a similar vein, other commentators state that new technologies usually complement and change older ones rather than displace them. CRAWFORD & GORMAN, *supra* note 7, at 48.

45. *ACLU v. Reno* makes the distinction between the Internet and what it calls "closed databases" (like Westlaw and LEXIS). 929 F. Supp. at 838.

46. Even though the digitization of print resources creates potential for mistranslation, a user is to some degree protected by the addition of a potentially liable party (the on-line service provider). For possible grounds of liability, see generally CRAWFORD & GORMAN, *supra* note 7.

The system as it now exists might be analogized to snipping bylines and publication information from all newspaper and magazine articles and placing them in large tumblers by what their headlines contain. For example, an article about Elvis Presley might come from the *New York Times*, *Christian Science Monitor*, *Easyriders* or *The Weekly World News*. Unfortunately, the user would not know from which publication the article came. If searchers wanted to know something about a subject, they would reach into the tumbler with articles on the subject and read. In such a system, a parent might be surprised to discover what a child "learned" about Elvis Presley. So it is with research on the Internet.

Presently, few people seem concerned with the lack of reliable information on the Internet. Still, at least one national legal journal has identified the problem. "Most of what you will find . . . is misleading or plain wrong. . . . With rare exceptions, don't rely on the Net for authoritative research."[47] We should not, however, stop at acknowledging this phenomenon—particularly as more publishing migrates to this medium. Rather than saying don't use the Internet or use it with caution, we need to remedy its problems. And the remedy does not lie in correcting inaccurate information, but instead in identifying the correct information.

V. SOLUTIONS

A. The Private Communications Model

Although some public users of the Internet have identified the authentication problem, at this time the proposed solutions have been made with private communications in mind. Because the concerns of those transmitting these documents have some overlap with those of persons disseminating to the general public, the progress on this front is worth considering. Persons communicating privately are most likely concerned with whether a document is from whom it purports to be from and has not been tampered with or altered in any way. However, in addition, the private sector would be interested in the confidentiality of certain documents and a procedure for making representations and promises in these documents legally enforceable. These latter concerns are not shared by those publishing and researching publicly disseminated documents.

47. Lewis R. Clayton, *Ten Tips On Using The Internet Creatively*, AMERICAN LAWYER, Dec. 1995, at 34, 34 (Supp. 1995).

The dominant system of authentication appears to be that of dual key encryption. A discussion of that system follows.

1. Dual Key Encryption

In 1996, Washington became the second state[48] to pass an Electronic Authentication Act,[49] effective January 1, 1998.[50] The impetus for the act was the increasing amount of commerce, particularly international commerce, that is transacted electronically.[51]

The system that the Legislature created is based on a process called "dual key encryption"[52] or "public key encryption."[53] In such a system, a person (in this case, the sender) purchases a "key pair."[54] A key pair consists of two keys: a private key and a public key. The sender would keep the private key and give a copy of the public key to the person with whom the sender wanted to authenticate communication. Thus, both the sender and receiver of an electronic message possess a key. Each key can read a message or signature that has been encrypted by the other.[55] For most communications, the keys will be used to create a digital signature that verifies the source of the message (and thus its contents).[56] The public key tells the recipient of the electronic communication two things. First, it indicates whether the signature at the bottom of the message is the actual signor. Second, it shows whether the communication has been tampered with. To the extent that another person on the Internet may have intercepted the sending of the public key, the system does not protect confidentiality. In other words, it does not tell the recipient whether anyone else has read the communication, only that the communication has not been changed.

---

48. Utah was the first state to pass such an act. *See* Utah Digital Signature Act, 1995 Utah Laws ch. 61 (codified at UTAH CODE ANN. § 46-3 (Michie Supp. 1996)).

49. 1996 Wash. Laws ch. 250.

50. *Id.* at § 602, 1209.

51. *See* HOUSE BILL REPORT, ESB 6423, 55th Leg., Reg. Sess. (Wash. 1996).

52. FINAL BILL REPORT, ESB 6423, 55th Leg., Reg. Sess. (Wash. 1996).

53. Robert T. Haslam & Thomas P. Maliska, *Encryption Ensures Privacy of Online Expression*, NAT'L L.J., Feb. 12, 1996, at C13.

54. A key is a sequence of bits that is used with a complex mathematical function (or algorithm) to encrypt or decrypt a message. Lyle T. Millham, Note, *Recent Legislative Developments in Utah Law: Digital Signatures Act*, 1995 UTAH L. REV. 1167, 1171 n.22 (1995). Key pairs will be available from certification authorities, the United States Post Office, computer stores, and other retail and wholesale outlets. *Id.*

55. Thus, at times, the original sender may also be a receiver and the original receiver may be a sender. Both types of keys can read the encrypted messages as well as encrypt messages.

56. The signatures may also be used in place of actual signatures to create enforceable contracts. *See* 1996 Wash. Laws ch. 250, §§ 401-406, 1206.

The Washington system revolves around an intermediary called a "certification authority."[57] The purpose for having a certification authority is to insure that the sender (or holder of the private key) is truly the person she claims to be.[58] The certification authority could be considered a notary for electronic communications. The certification authority issues certificates that contain the subscriber's public key.[59]

A similar system is being utilized by the federal government for future uses in such transactions as income tax filing.[60]

The difficulty with using this system for authenticating publicly disseminated information is in distributing public keys. Every publisher would use a private key to prove authentication of their document. In turn, every recipient who wanted to authenticate a document posted on the Internet would have to locate the corresponding public key and then use it to verify the "signature" on the document made with a private key. Such a system would be too cumbersome for the general populace. Furthermore, the cost of issuing keys and using certification authorities would be prohibitive.

For this reason, the dual key encryption method of authentication is inappropriate for authentication of publicly disseminated documents. Even if the federal government adopted a policy of placing authenticating signatures on its documents, they could not effectively mandate that private organizations do the same. Although parts of the private model could be utilized in a parallel public model,[61] there remains the need for such an alternative form of verification.

---

57. 1996 Wash. Laws ch. 250, § 201, 1195. (Enunciates the qualifications and requirements for receiving a license for being a certification authority.)

58. "[T]he strength of cryptographic mechanisms relates to protecting the confidentiality of the private keys and the integrity of the public keys." Bruce W. McConnell & Edward J. Appel, Co-Chairs, Interagency Working Group on Cryptography Policy, draft paper, *Enabling Privacy, Commerce, Security and Public Safety in the Global Information Infrastructure* (last modified May 17, 1996) <http://www.epic.org/crypto/key_escrow/white_paper.html>.

59. The certification authority issues a "certificate" which is a computer based record including the name of the subscriber and the subscriber's public key. 1996 Wash. Laws ch. 250, § 103(3), 1191.

60. The Paperless Federal Transactions for Citizens project will be tested and gradually expanded beginning in the Summer of 1996. *See* Gary H. Anthes, *Feds to Secure 'Net Access*, COMPUTERWORLD, May 27, 1996, at 69, 72. Information about the General Services Administration's Federal Security Infrastructure Program can be found on the Internet at <http://www.gsa.gov/fsi>. *Id.*

61. The certification authority, for example, would be an ideal intermediary in the domain name registration process described *infra*, section V.B.2.

B. The Proposed Solution: Synopsis

The authentication solution proposed in this Article is to utilize the existing system of creating and registering domain names to allow for verification of site ownership. By tightening the registration procedure and modifying browser software to easily access registration information, users could then determine the source of Internet documents. This proposal will be treated in more depth following an account of the existing domain system.

The proposed solution attempts to utilize preexisting components and procedures of the Internet. The intent is to resolve the problem with a minimal amount of regulation and imposition on the users, which include service providers as well as end users. The proposal does, however, require some additional formalities in the domain name registration process.

The discussion must begin with an overview of the domain system and the current protocol for domain name registration.

1. The Domain System

The computers that utilize the Internet are classified and identified by what is called the domain system.[62] By utilizing this system, a method for authenticating documents could be created. Every computer on the Internet has an "address."[63] This address is actually a combination of names that are called "domains." The domain system breaks the gigantic worldwide Internet into manageable pieces.[64] A complete address name is made up of a varying number of subdomains. These subdomains are divided by dots or periods.

The way to understand a particular domain name is to look at the subdomains from right to left.[65] The rightmost subdomain is the most general; it is called the top level domain.[66] Reading to the left,

---

62. *See Registering Your Own Domain Name*, INSIDE THE INTERNET, Sept. 1996, at 1, 1.

63. Computers actually use numeric IP (Internet Protocol) addresses to find each other, but people find words and abbreviations more convenient. So InterNIC assigns each computer or resource a domain name that corresponds to its IP address. Each IP address and domain name must be unique. *Registering Your Own Domain Name*, INSIDE THE INTERNET, Sept. 1996, at 1, 1.

64. ED KROL, THE WHOLE INTERNET: USER'S GUIDE & CATALOG 34 (2d ed. 1994).

65. HAHN & STOUT, *supra* note 32, at 49.

66. *Id.* There may, in fact, be characters to the right of the primary domain name. These characters are not a part of the domain name. They are separated from the domain name by a slash (/) and identify separate pages within the domain. Such pages are often those of users who the domain owner has allowed to set up a web site on the owner's server. *Id.*

the subdomains become more specific.[67] The pieces of a domain name tell you who is responsible for maintaining the name. They may not, however, tell you anything about who maintains the computer corresponding to that Internet Protocol (IP) address[68] or even where that machine is located.[69]

There are seven top level organizational domains.[70] In addition to organizational domains, there are geographical domains for foreign countries. These two-letter domain names were adopted in response to the Internet's international growth.[71] The United States has a geographical domain but it is not generally used in this country.[72] Outside the U.S., geographical names are used almost exclusively.[73]

2. Issuance and Registration of Domain Names

In the United States, domain names are registered by Network Systems Incorporated, a Virginia-based company that controls addressing on the Internet under contract with the National Science Foundation and its umbrella organization, InterNIC.[74] In addition to registration services, InterNIC also provides a directory service.[75] This service, called WHOIS, allows a user to enter a domain name into a searching box.[76] The service will then identify the owner (if any) of the domain name.[77]

---

67. *Id.*

68. The IP address is a unique series of numbers that identifies each computer on the Internet. The domain name is an alias for that series of numbers. It is simply easier for persons to identify and remember. *See* Carol L. Sehlein, *Getting a Home Page for Small Law Firms*, NEW JERSEY LAWYER, Dec. 2, 1996, at 39.

69. KROL, *supra* note 64, at 33.

70. HAHN & STOUT, *supra* note 32, at 54-55. The seven domain names are: com (for commercial organizations, roughly equivalent to businesses); edu (for educational organizations); gov (for governmental entities and departments); mil (for military entities); org (general organizations, often nonprofits); net (for network resources); and int (for international organizations). The last of the domain names (int) was added after the rest. *Id.*

71. *Id.* at 54.

72. *Id.* at 55.

73. *Id.*

74. The National Science Foundation created InterNIC, an umbrella organization for other companies. While Network Systems was given exclusive control over distribution of domain names, American Telephone and Telegraph (AT&T) was given exclusive control over directory services to those domain names.

75. *Welcome to the InterNIC* (visited Mar. 17, 1997) <http://www.internic.net>. The Directory & Database Services are provided by AT&T. By contrast, Network Solutions provides Registration Services. *Id.*

76. *Registering Your Own Domain Name*, *supra* note 62, at 1, 3.

77. The InterNIC Directory and Database Services actually provide a unified access point to the two official Internet WHOIS servers for persons and organizations. *InterNIC Whois* (visited Mar. 17, 1997) <http://www.internic.net/wp/whois.html>. The two servers can be

Generally, business corporations and other organizations have tried to register domain names that are recognizable as connected to them. Thus, McDonald's would register the domain name <mcdonalds.com>, and the American Civil Liberties Union would own <aclu.org>. In fact, the earliest skirmishes in the domain name registration field involved trademark claims against persons registering domain names that include words that are another's trademark.

Despite the above trend, there is no requirement that the domain name have any relation to the person or entity registering the name.[78] Without an easy domain verification device, this creates a potential for mistake and fraud. For example, a computer test developer registered the domain name "dole96"[79] that featured a picture of Bob Dole and a page labeled "An Official World Wide Web Internet Site."[80] Because the page parodied Dole and his campaign, it may have been protected by the First Amendment. Still, it demonstrates the potential unverified domain names have to deceive.

Several disputes have already arisen with regard to domain names.[81] These have been fought under the auspices of trademark law.[82] Companies have claimed that a name registered by another infringed on the trademark rights of that company.[83] In response to the problems with trademark infringement, Network Solutions now requires domain name applicants to warrant that their use of a domain name will not "interfere with or infringe the right of any third party in

---

searched separately or together. *Id.* The servers are as follows: (1) <rs.internic.net>, which is operated by InterNIC Registration Services and provides civilian Internet organization and person registration information; and (2) <nic.ddn.mil>, which is operated by the Defense Information Systems Agency and provides organization and person registration information for the military and defense community. *Id.*

78. *See Registering Your Own Domain Name, supra* note 62, at 1, 2.

79. *Bob Dole for President: The Ripe Man for the Job* (visited June 14, 1996) <http://www.dole96.org>. The site contained satirical links, such as one to "Bob Dole's Courageous Stand . . . Against War (except when it's only sort of a war, like the Gulf Not-A-War, which he was for, even though he thinks congress should have instigated it rather than Bush.)." *Id.* Other links included "Assorted links to other Fruit & Vegetable enthusiasts" and "Links to disrespectful Weenies." *Id.* The 1996 Dole presidential campaign seemed to take a beating from domain name registrants. Another site, <http://www.dole-kemp.com>, contained a statement that said: "Pssst . . . The past is over. Click below to make the right choice for the future." Clicking on the arrow led to the official Clinton-Gore re-election web page. *Notes from the Campaign Trail*, UPI, Oct. 8, 1996, *available in* LEXIS, News Library, UPI File.

80. *See* Guy Alvarez, *New Legal Issues On the Net*, AMERICAN LAWYER, Dec. 1995, at 28, 31 (Supp. 1995).

81. *See* David Post, *A Domain By Any Other Name*, AMERICAN LAWYER, May 1996, at 117-18.

82. *See, e.g.*, Panavision Int'l v. Toeppen, 945 F. Supp. 1296 (C.D. Cal. 1996).

83. *Id.* at 1300.

any jurisdiction with respect to trademark, service mark, trade name, company name or any other intellectual property right."[84] An applicant must also agree to indemnify Network Solutions in the event of any such third-party claims.[85]

Even with the safeguards provided by trademark law and Networks Solutions' corresponding registration warranties, the system of registration is pregnant with the potential for fraud.[86] An applicant still is not required to prove representative capacity. Consider the following hypothetical.[87]

Sheepco is a small publicly traded company. A large percentage of its products are sold to a single customer, ZZZ, Incorporated.

Sly Shorter is a large-scale investor who decides to make a killing on the market by short-selling Sheepco (betting the stock price will drop).[88]

After arranging the deal, Shorter registers the domain name <3Z.com> with InterNIC. In the application, Shorter lists the domain

---

84. Network Solutions, Inc., *NSI Domain Dispute Resolution Policy Statement* (visited June 10, 1997) <http://www.shore.net/dns/internic-domain-1.html>.

85. *Id.*

86. In this author's opinion, there are numerous reasons why trademark protection alone is insufficient to resolve the authentication problem. Some of the major reasons follow.

In broadest terms, the trademark system is designed to protect the owner of the trademark (in a commercial setting, the seller) rather than the computer user (in a commercial setting, the buyer). When it comes to authentication, trademark law is the wrong tool for the job. Rather than giving the common person a means of assuring the accuracy of information obtained on the Internet, trademark law protects those who disseminate information from having others appear to disseminate it in their name.

Another problem with trademark protection is that many domain names are not and cannot be protected by trademark (*e.g.*, generic names such as <badbreath.com>, or abbreviations such as <b&n.com>). For those sites, trademark law adds no level of identification or authentication.

Finally, relying exclusively on trademark protection (or any other post-registration remedy) will promote litigation. Injunctions of trademark infringement or actions for fraud must be done through the court. A more stringent and verifiable registration system will make confusion of Internet sites less likely, thus reducing the need to litigate to enjoin others from using particular domain names or creating misleading web sites. Considering the difficulties inherent in Internet litigation (jurisdiction, service of process, forum, and choice of law, to name a few), any system that relies on litigation is poorly suited to resolve Internet problems.

87. This hypothetical is not too distant from what is occurring to some degree on the Internet every day. *See* Joseph Nocera, *Investing in a Fool's Paradise*, FORTUNE, Apr. 15, 1996, at 86 (discussing attempted manipulations of stock price of Iomega stock by anonymous participants of a listserv dedicated to discussing such matters about the company as its products, its prospects, and its stock price).

88. A "short sale" is defined in rule 3b-3 under the Exchange Act. 17 C.F.R. § 240.3b-3 (1996). Generally, it is a sale by a person who does not own the security. *Id.* The investor effects delivery of the securities sold usually by borrowing stock. *Id.* In turn, the investor hopes to profit if the securities decrease in value by covering the short position (that is, buying the securities) at a lower price than the original sale.

owner's name as "ZZZ, Inc."[89] No verification is necessary. Shorter then creates a web site for ZZZ Incorporated in the domain by downloading documents (such as Annual Reports, SEC filings and press releases) from ZZZ's true web site. Finally, he adds his own original document: a press release stating that ZZZ predicts increased profitability in the next quarter due to reduced expenditures as a result of a back inventory of supplies from Sheepco.

After creating the phony web site, Shorter signs onto several listservs[90] dealing with investing.[91] He posts messages that Sheepco stock will likely plummet with the decrease in demand from its major client. The posting points to the <3Z.com> web site as proof of the claims made.

Any reader of the posting wanting to verify the information would find that, in fact, the domain owner of the site was listed as ZZZ, Inc. The only way to refute the information is to contact ZZZ Incorporated directly and inquire. As a result, the stock drops and Shorter completes his short-sale and removes the documents from the website. When investors later discover the falsity of the press release, they will call for government intervention. If the Internet community does not address such existing and potential problems, the government will undoubtedly intervene.

C. *The Proposed Solution: Verification By Registered Domain Name*

Essentially, there are two major steps that must be taken to support and implement the proposed solution of authentication. The first step is to create a verified registry of domain owners. While InterNIC currently keeps a registry of domain owners, there is no verification process for identifying owners as who they purport to be. By setting up criteria for identification, InterNIC, or some other entity given the task, would provide a safeguard against fraud in domain ownership. Once they become more commonplace, certification

89. Shorter is required to provide an administrative or technical contact. This could be a fictitious person, the actual name of a ZZZ executive, or a real person unrelated to ZZZ.

90. A "listserv" is an on-line bulletin board or discussion group. A subscriber to a listserv sends a message to the listserv, which the listserv system in turn routs directly to all other subscribers of the listserv. Thomas D. Brooks, Comment, *Catching Jellyfish in the Internet: The Public-Figure Doctrine and Defamation on Computer Bulletin Boards*, 21 RUTGERS COMPUTER & TECH. L.J. 461, 467 n.53 (1995).

91. There are an ever increasing number of such listservs on the Internet. *See* Deborah Lohse, *Stock Regulators Are Worried Dangers Lurk for Investors in On-Line Chat Sites*, WALL ST. J., Sept. 12, 1996, at C1. The article gives examples of substantial stock price run-ups attributable to false information posted on such listservs. *Id.*

authorities would be a logical choice as gatekeepers for verifying persons registering domain names with InterNIC.[92]

The second step in the process is to create a quick link from a web site to the registry so that a user can identify the actual owner of the site. This could either be done by automatic display (as can be done with URLs)[93] or by creating a verification command that allowed a user to double click on an icon that would then run the domain name search on the registry.

To avoid objections that such a registration system infringed on the right to privacy, it would be a voluntary system. If a domain owner does not provide evidence of identity, that domain name and its owner will not be placed on the registry. Thus, any information residing at the web site will not be verifiable and a user is consequently put on notice not to rely on it. The presumption here is that site owners who want others to cite or rely on information published at the site will identify themselves.

This would create an overlapping double domain realm.[94] The first realm would be web sites with verified owners. It would be transparent and blend into the second realm that would include the entire Internet, including all domains (ownership verified or not). In the future, a researcher might be able to limit a search to only sites that could be verified. Because the second tier would be entirely voluntary, the larger Internet should be undisturbed by any regulation.

Ideally, every domain level would be identified in the verification process.[95] Thus, a user would know exactly what person or agency's computer posted the document. However, requiring each domain owner to register every sublevel domain name would be cumbersome. The Internet community would also probably consider such regulation excessive. In addition, the numbers of domain names on the registration lists would be staggering, possibly making the lists unusable. Because users already must register primary domain names, establishing a verification process at that level would be the least burdensome

---

92. *See supra* notes 57-59 and accompanying text.

93. *See supra* note 32 and accompanying text.

94. The concept of double realms has already been realized with the creation of the so-called "Internet II." *See* David S. Hilzenrath, *Internet II Will Put Colleges Back on the Fast Track*, THE WASHINGTON POST, Oct. 7, 1996, at F18. This is a network that will be created by agreement among a group of 34 research universities. *Id.* The new computer network will augment the Internet and allow those universities to have the higher speed capacity necessary to support many of the leading-edge applications that they require. *Id.*

95. Under current practice, the domain name that is registered with InterNIC includes only the first and second levels. *Registering Your Own Domain Name*, *supra* note 62, at 1.

on the current system and still provide a fundamental degree of site authentication.

To be sure, this registration system does not solve the entire authentication problem. It only resolves the origin issue. However, making registered domain owners legally responsible for their sites leaves it to market or liability factors to insure integrity of the site.[96] This includes not only the duty to secure the site,[97] but also that of discovering tampering on the site. Failure to take reasonable steps in performing either of these tasks could give rise to liability on the part of the domain owner. To avoid these consequences, site owners have the options of not registering as a verified site or of disclaiming liability for information on the site.[98] Finally, the domain owner would also be responsible for posting the date of documents. Ideally, this problem will disappear when future versions of software automatically post such information when a site is updated.

1. Redundancy and Mirror Sites

Because of their reliability, authoritative sites can become popular sites. When this happens, incoming traffic to the site increases and access to the information can be slow or even unavailable. A common solution to this problem is to create a "mirror site." A mirror site is a second server site that includes the same information as the original site. This gives users a second alternative to access the information. In a perfect world, it reduces the amount of traffic to the original site by half. In addition, mirror sites provide checks for the integrity of the original site.

Under the system proposed here, mirror sites, unless the domain name is owned by the original author or publisher, will lose their

96. This solves the legal problem of a document's integrity. It leaves to technology the problem of finding an easy way for a domain owner to prevent and discover tampering.

97. The degree of security provided by particular software varies greatly. *See* Gene Steinberg, *False Security*, MACWORLD, Nov. 1995, at 118. For increased protection, a domain owner may want to utilize a dedicated security program on top of that provided in a particular software program. *Id.*

98. The incentive for site owners to register as a verified site derives from the value of being considered reliable. As the Internet becomes more overloaded with information, knowledgeable users will focus their searches exclusively on sites that have indicia of trustworthiness.

authentication.[99] Along with that, they may lose some of their value. Consequently, they will not alleviate traffic to the original site.

One solution is to have a notice at the top of such sites stating that it is a mirror site to the official source. A link would then be provided to the domain of the originator of the document. That link would go to a statement of verification (authenticated by being from the originator's domain) that the mirror site is, in fact, what it purports to be. Such a link would bypass the main entrance and commonly used links of the original site, thereby avoiding bottlenecks. In addition, it would only require one search in the original domain for verification. The user could then use the alternative and less congested mirror site.

VI. ARCHIVING

A related issue in the authentication of Internet documents is that of archiving. As information on the Internet ages, those who posted it may remove it or transfer it to a less prominent location. The appropriate mechanism for digital archiving is already undergoing debate. Despite there being no system for authentication, it is still important that those making archiving decisions take it into account.

A solution to digital archiving is beyond the scope of this Article. However, a description of the needs and some of the possible resolutions to the authentication problem follows. Web owners cannot be expected to retain every document published on their sites or eventually databases (like library collections and Fibber McGee's closet[100]) will become cluttered with old and unused information. To prevent this, database vendors and owners of websites will have to perform a task librarians call "weeding."[101] Traditionally, this process resulted in discarding unused and outdated materials. Other

---

99. A similar problem exists for entities that publish documents that they do not create. For example, the opinions of many of the federal courts are disseminated by law schools. *See, e.g., United States Court of Appeals for the Ninth Circuit* (last modified Jan. 5, 1997) <http://www.law.vill.edu/fed-ct/ca09.html>. A verification check telling a user that Villanova University owned the domain on which the opinions were published would presumably not confer authoritativeness on those opinions, because Villanova is not the entity from which the opinions would naturally emanate.

100. *Fibber McGee and Molly* was an NBC radio program which aired from 1935 to 1956. Every time Fibber McGee opened the hall closet, he was buried in a deluge of clutter. *See* JON D. SWARTZ & ROBERT C. REINEHR, HANDBOOK OF OLD-TIME RADIO: A COMPREHENSIVE GUIDE TO GOLDEN AGE RADIO LISTENING AND COLLECTING 324 (1993).

101. "Weeding" is defined as the withdrawal of materials from the collection to improve the effectiveness of the collection. ARTHUR CURLEY & DOROTHY BRODERICK, BUILDING LIBRARY COLLECTIONS 308 (6th ed. 1985). *See also* STANLEY J. SLOTE, WEEDING LIBRARY COLLECTIONS 3 (3d ed. 1989).

lesser-used works were retained, but stored in a different (usually remote) location from the main collection of materials.[102] With the latter, the cataloging record of the document with its remote location noted allowed a researcher continued access to it.

In a similar manner, documents at a website might be discarded or relocated. One critical factor in on-line weeding is that, unlike published print works, an on-line document may be the only existing copy available to the public. Other public "copies" are merely links to the sole original. Unless the copy is retained and there is a referring link made when the document is relocated, access to the document from other websites will be lost. It is also critical that such a document retain its authentication when it is relocated and archived.

A. The Need for Archives

Publicly disseminated information may be as important for a secondary purpose as it is for its primary purpose. This is because, as information ages, the user clientele shifts from subscribers to researchers.[103] More specifically, initially a document may be disseminated to create an immediate awareness of its content. In time, that information will become either obsolete (outdated and untrue) or stale (common knowledge). At such a time, its primary purpose is no longer of consequence. However, there remains a secondary function[104] for such documents. They can be used retrospectively for historical research to indicate what the purpose or sense of documents was at the time of their dissemination. The Task Force on Archiving of Digital Information states:

> The pursuit of knowledge is a process in which the emergence of new knowledge builds on and reconstructs the old. Knowledge cannot advance without consistent and reliable access to information sources, past and present. It is the archival function in the system of knowledge creation and use that serves to identify and retain important sources of information and to ensure continuing access to them.[105]

---

102. *See* SLOTE, *supra* note 101, at 27. The tradeoff for lack of accessibility is a reduction in expense.

103. Brichford & Maher, *supra* note 22, at 703.

104. This function may not be a "purpose" in the sense that those disseminating the information may, in fact, not want the information used for this secondary purpose. For example, a corporation would not want certain press releases to be later used against it by shareholders in a subsequent derivative suit.

105. Task Force On Archiving Of Digital Information, report, *Preserving Digital Information*, Information Objects in the Digital Landscape (visited Mar. 17, 1997) <http://www.rlg.org/ArchTF> [hereinafter *PDI*].

Although electronic publishing has an undisputed advantage in providing rapid and broad distribution of information, it will not be able to fulfill key substantive and "political" roles of scholarly publishing unless it provides assurances for ongoing accessibility.[106] To ensure the retention, preservation, and utilization of such information, archives need to be established.[107]

B. Government and Private Activities In Archiving

The federal government has already begun the process of solving the problems in archiving digitized documents. In 1993, the Government Printing Office Electronic Information Access Enhancement Act[108] required the Superintendent of Documents to operate an electronic storage facility for federal electronic information. In addition, the privately funded Commission on Preservation and Access and Research Libraries Group created a Task Force on Digital Archiving. The charge of the Task Force was to frame the key problems, define critical issues, consider alternatives, and make recommendations relating to digital archiving.[109]

The Task Force puts forth the principle that responsibility for archiving rests initially with the creator or owner of the information.[110]

The Task Force recommends a distributed, rather than a centralized, structure of archiving.

> A distributed structure, built on a foundation of electronic networks, places archival responsibility with those who presumably care most about and have the greatest understanding of the value of particular digital information objects. Moreover, such a structure locates the economic and cultural incentives where they are most likely to prompt those preserving digital information to respond with the greatest agility to the changing digital landscape and to the shifting tides of technology.[111]

Such a distributive network would include corporations, federations, and consortia and may range over regional and national boundaries. If this network becomes reality, important government information is less likely to be lost than is information from political organizations,

---

106. Brichford & Maher, *supra* note 22, at 709-10.
107. *Id.* at 701.
108. Pub. L. No. 103-40, 107 Stat. 112 (1993).
109. *PDI*, *supra* note 105, at Appendix 1.
110. *Id.* at Archival Roles and Responsibilities—General Principles.
111. *Id.* at Archival Roles and Responsibilities.

corporations, and other private entities that will not be required to create a system for archiving.

According to the Task Force, the greatest fear about the life of information in the digital future is that owners or custodians who cannot bear the expense and difficulty will, through a simple failure to act, destroy the objects without regard to future use.[112] To compensate for this, the Task Force recommends an aggressive rescue function that acts as a fail-safe mechanism that allows one agency, acting in the long-term public interest of protecting the cultural record, to override another's neglect or active interest in abandoning or destroying parts of that record.[113]

C. *Issues In Archiving of Digitized Documents*

Traditionally, archiving was concerned with longevity of the physical media. Today, it is better understood as a matter of ensuring the future availability and intelligibility of the informational content of documents.[114] What matters most to the user is the survival of the information itself and the access points provided by the system, rather than the specific hardware, or even software, on which information is stored.[115] Internet documents provide additional difficulties because ensuring the longevity of a physical medium and the means to read it will not necessarily preserve the complex nature of a document whose informational linkages (hypertext links) are an essential component of the document itself.[116]

---

112. *Id.* at The Challenge of Archiving Digital Information—Legal and Organizational Issues.

113. *Id.* at Archival Roles and Responsibilities—General Principles. The Task Force recognized several factors that might cause custodians to act in this way. They include budgetary constraints, reorganization of priorities or focus, change of business, the need to go out of existence, or expiration of copyright. *Id.* In addition, there might be some occasions where there is no natural institutional home for the document. *Id.*

114. Brichford & Maher, *supra* note 22, at 704.

115. *Id.* at 705. Ironically, the article points out that the usable life of physical media (for example, disks and tape) is now greater than the life cycle of most software and hardware used to access the media. *Id.* at 706. The 1960 Census is a prime example. The records were stored on tapes that could only be read with a UNIVAC type II-A tape drive. Those drives became obsolete in the 1970s. Ultimately, the Census Bureau was able to copy nearly all the data judged to have long-term value onto industry-standard tapes. *See PDI*, *supra* note 105, at Introduction—The Limits of Digital Technology.

116. Brichford & Maher, *supra* note 22, at 707. According to the Task Force on Archiving of Digital Information,

> If the integrity of these objects is seen as residing in the network of linkages among them, rather than in the individual objects, or nodes, on the network, then the archival challenge would be to preserve both the objects and the linkages, a task that would today be exceedingly complex. At present, there appears to be no good archiving

A major dilemma in archiving digitized materials is making decisions within a context of conflicting principles. These principles are on the one hand, security (maintaining integrity of the document's contents) and, on the other hand, accessibility. Accessibility is necessary because users need more than merely the ability to retrieve previously identified documents. They also need the ability to search for unidentified information—through a text search of documents or even through serendipity. After all, that is one strength of digitized documents: the ability to search and find information in them without knowing ahead of time the document in which the information resides.[117] Allowing free accessibility, however, risks security.[118] Digital information is more easily contaminated or otherwise altered than hard copy.[119] One possible solution is a dual archive system. One archive would be publicly accessible; the other, an inaccessible archive (like the Sèvres model described below), would exist to preserve the untampered original document. Such a solution, however, is more difficult and expensive to maintain than other solutions.

Another critical component of digital archiving is what the Task Force calls "migration."[120] Digital information usually requires a separate mechanism to read or translate the data. The problem is that these reading mechanisms continue to change with technology. Often, documents created for an earlier reading mechanism are not upgraded to allow a later mechanism to read them. For example, a 3 1/2 inch disk drive cannot read a document on a 5 1/4 inch disk. Long-term retention of electronic publications is problematic because of the lack of archival standards of permanence for digital storage media.[121] As the operating environments (formats for archiving) change, it becomes

---

solution; a possible stop-gap measure would be to treat the network in terms of its component parts and to take periodic snapshots of the individual WWW objects. *PDI*, *supra* note 105, at Information Objects in the Digital Landscape—The Integrity of Digital Information—Context.

117. *PDI*, *supra* note 105, at Introduction—The Fragility of Cultural Memory in a Digital Age ("In full text documents, a reader can retrieve needed information by searching for words, combinations of words, phrases or ideas.").

118. *See* William Dutcher, *Locking the Corporate Vault: Achieving a Balance Between Security and Accessibility Can be Tricky*, PC WEEK, Mar. 11, 1996, at N1.

119. *PDI*, *supra* note 105, at Information Objects in the Digital Landscape—The Integrity of Digital Information—Fixity.

120. "Migration" means both: (1) the periodic refreshing or transfer of Government information products from one medium to another to minimize loss of information due to physical deterioration of storage media and (2) the reformatting of information to avoid technological obsolescence due to software or platform dependence. REPORT, *supra* note 5, at E-vi.

121. *See* Brichford & Maher, *supra* note 22, at 711. Long time computer users may recall the Bernoulli cartridges that were used less than a decade ago for storage of larger data files.

necessary to migrate the contents of previously archived documents.[122] Although acknowledging that additional research is needed to test the technical feasibility and financial costs of various approaches to the problem, the report discusses several strategies that can be employed in addressing it.[123] In addition, the Task Force suggests that tax incentives and accounting rules might create an incentive for investment in the long-term capital stock of digital archiving.[124] Otherwise, the report suggests, solutions to cost questions are likely to be found in relation to specific bodies of digital materials and the communities that are interested in them.

D. Options for Archiving

Obviously, there are numerous archiving issues that are unresolved. For one, private entities may not want to create their own archiving facility. Still, they will need to maintain some method of archival. As the digital environment emerges and its requirements become clearer, traditional institutions may need to change in various structurally significant ways and new kinds of institutions and institutional structures may emerge to perform all or parts of key archival functions for digital information.[125] In all likelihood, companies will be created that put information into archival form and then store it for other companies. Another option would be for the government to maintain an archive for private digitized publications, not unlike what the Library of Congress currently does for works sent in for copyright protection. At one extreme, publishers could even be legally bound to place a copy of their published digital works in a standard archival format with a certified digital archive.[126]

With respect to archival format, there are several possible options. The major ones are as follows:

1. Transfering and storing documents in hard copy.
2. Storing digitized documents
   a. with a firewall separating them from access. Allow access only by checkout.
   b. in accessible storage facility.
3. Storing two sets of digitized data: one behind firewall (for security) and one accessible.

---

122. *PDI*, *supra* note 105, at Archival Roles and Responsibilities—Migration Strategies.
123. *Id.*
124. *Id.* at Archival Roles and Responsibilities—Managing Costs and Finances—Financing.
125. *Id.* at The Challenge of Archiving Digital Information—Conceptual Framework.
126. *Id.* at Archival Roles and Responsibilities—General Principles.

4. Combining the previous options.

Each of these methods has strengths and weaknesses. They are briefly discussed here.

1. Storing documents in hard copy: The Task Force refers to this as a more stable media.[127] However, it potentially comes at the expense of great losses in the form or structure of digital information.[128] These "flattened" documents will not retain their hypertext links, which may be essential to the integrity of the document. Because of this, the strategy is not feasible for preserving complex data objects from complex systems.[129] Furthermore, such storage eliminates on-line accessibility. The advantage is that it is more difficult to corrupt. I call this the "Sèvres model" after the city in France where the one true unit of measure of a kilogram is located.[130]

2. Storing digitized documents: Instead of paper copy, documents could be stored on tape or disk (like an individual would store a floppy disk copy of a document). Such a copy would be more secure if it was not accessible from a remote site. Hackers would not be able to easily infiltrate such documents. However, storage of digitized documents creates the potential problem in extinction of the reading mechanism. Thus, such a policy requires a plan for migration of information.

a. Storing digitized data behind firewall: Although the concept of a firewall is not homogeneous, conventional wisdom holds that anything that provides a barrier between one network segment and another and allows only authorized access is a firewall.[131] One

---

127. *Id.* at Archival Roles and Responsibilities—Migration Strategies—Change Media.

128. *Id.*

129. *Id.*

130. In 1875, an international conference convened in Paris to establish an International Bureau of Weights and Measures. 8 THE NEW ENCYCLOPAEDIA BRITANNICA 73 (15th ed. 1990). The Treaty of the Metre signed there provided for a permanent laboratory in Sèvres, where international standards were to be kept. *Id.* Although definitions for the metre have been redefined by natural constants (so as to be calibrated anywhere), the kilogram is still defined as the mass of the international prototype at Sèvres. *Id.*

131. Julie Bort, *Firewalls Hold Down The Fort*, SOFTWARE MAGAZINE, Oct. 1995, at 130. Each firewall must be placed on its own server so that all traffic travels through the firewall. Firewalls employ two means to thwart hackers: packet filters and gateways. *Id.* Packet filters route or drop packets based on the packet's address. Thus, if a user tried to access the protected network from an unauthorized computer, the packet filter would drop the packet, preventing the user from entry to the network. *Id.* There are different kinds of gateways. Gateways funnel network traffic through a single point. At that point, the user must establish authorization to access the network for a specified purpose (usually by entering a particular login and/or password). *Id.* An application gateway limits access to specific applications. Thus, for example, a law student might be permitted to use the law school's network to access the library's collection but the gateway would prevent the student from accessing records in the financial aid or

distinct advantage of firewalls is security. The major drawback is in accessibility. If a user knows a document exists, a copy of it can be made and distributed. However, such storage makes it difficult, if not impossible, to discover the existence of an unknown document and virtually eliminates finding the information in the document.

b. Storing digitized data in accessible facility: This approach allows users to access information in document. However, document's integrity is subject to attack by hackers who can easily access the document.

3. Storing digitized data in two facilities: This is the best of both options two and three. The major drawback is the cost. In this report, the Task Force recognized that there may be multiple levels of storage in a digital archive.[132] Such levels would be based on expected use and needed performance in retrieval (the Task Force did not mention security but that should be added). The Task Force used the terms "off-line" for little used material, "on-line" for high demand objects where retrieval time is at a premium, and "near-line" storage as an intermediate solution.[133] The obvious drawback to this option is that the costs will be higher for such storage.

4. Combination of the above policies: The final approach to digital document archiving is to utilize more than one of the above policies. The option selected for any particular document would depend on the document's characteristics and importance. For instance, documents of greater importance and documents where there is only one original should be stored in both a firewall protected facility as well as one that was generally accessible by users. Effective storage management also requires providing for redundant copies of archived documents as an insurance against loss.[134]

There is one critical note with respect to authentication. Under the current system, archiving of documents on the Internet will often cause a changing of that document's Internet address. As previously stated, it is important that when digital documents are archived, there remains some method to authenticate the information in them. The Task Force on Archiving of Digital Information recommends a self-

---

registrar's office.

132. *PDI*, *supra* note 105, at Archival Roles and Responsibilities—The Operating Environment of Digital Archives—Storage.

133. *Id.*

134. *Id.*

referential mode for identifying documents so that they could be found despite changes in Internet addresses.[135]

VII. LEGAL CONSEQUENCES OF THE PROPOSED SOLUTION

If an authoritative system for authentication of Internet sites is created, that system's impact on existing law will vary depending on each adopting country's domestic incorporation of the system. There are several areas where this proposal would probably impact the law in the United States. It will affect not only what is admissible in court proceedings; the presumptions it makes could create theories of liability analogous to that in other areas of law. A brief discussion of the potential impact on key areas of law in the United States follows.

A. Evidence

The proposed solution is compatible with the existing rules of evidence. Those rules create presumptions based on the location of documents.[136] Evidentiary rules state that public records or reports are authenticated by evidence that they are from the public office where items of that nature are kept.[137] Similarly, data compilations (coupled in the rule with ancient documents) can be authenticated with evidence that it "was in a place where it, if authentic, would likely be."[138]

Unless there is evidence that a document has been altered or is otherwise lacking in integrity, a user (and likewise the court) should presume that the digitized document from a verified site is authentic. That shifts the burden of disproving authenticity to the domain owner who, by having greater access to the maintenance of the web site,

---

135. *See id.* at Information Objects in the Digital Landscape—The Integrity of Digital Information—Reference ("In order to provide a consistent means of reference for digital objects, systems of citation, description and classification will need to dispense more than name and location information.").

136. *See* FED. R. EVID. 901(b)(7), (8).

137. FED. R. EVID. 901(b)(7).

> Rule 901. Requirement of Authentication or Identification . . .
>
> (b) Illustrations. By way of illustration only, and not by way of limitation, the following are examples of authentication or identification conforming with the requirements of this rule: . . .
>
> (7) Public records or reports. Evidence that a writing authorized by law to be recorded or filed and in fact recorded or filed in a public office, or a purported public record, report, statement, or data compilation, in any form, is from the public office where items of this nature are kept.

Id.

138. FED. R. EVID. 901(b)(8)(B).

controls the evidence on the issue. Such a policy leaves security of web sites to their owners.

There may be additional considerations in determining the integrity of digital documents that have been archived.[139] This is because the data in the document may have been translated into one or more different formats to reflect changes in the technological environment.[140]

B. *Liability of Domain Owners*

In terms of liability, domain owners under such a system would be made responsible for information within their verifiable domain.[141] That would leave it to institutions to create and enforce policies for personal pages tied to the institution's official domain. Within such a system, persons allowing access to the Internet through their domain should establish technology guidelines for all such users.[142] One option for domain owners who allow subscribers, employees, or students to utilize their server would be to utilize a second server. The one server (and domain name) would be for documents published by the domain owner, the other for its users.[143] The owner's domain name would be verified for the first domain so that documents on it could be authenticated. The second (user) domain name could either be registered as a "subscriber" or "user" domain (thus documents

---

139. The Task Force on Archiving of Digital Information found that the features that determine information integrity and deserve special attention for archival purposes include content, fixity, reference, provenance, and context. *See PDI*, *supra* note 105, at Information Objects in the Digital Landscape—The Integrity of Digital Information.

140. *See supra* note 114 and accompanying text.

141. This is, in fact, happening already on several fronts. *See Users and Providers Learned Cyberspace Talk is not Always Carefree Schmoozing*, NAT'L L.J., Dec. 25, 1995-Jan. 1, 1996, at C13. The article discusses a New York Supreme Court ruling that Internet provider Prodigy was a publisher and could therefore be sued for libel. It also discusses the settlement of a case involving another provider, CompuServe, for copyright infringement. *See also* Religious Tech. Ctr. v. Netcom On-Line Communication Services, 907 F. Supp. 1361 (N.D. Cal. 1995) (copyright infringement case against a computer bulletin board service). In denying the service provider's motion for summary judgment, the court said: "Netcom is not free from liability just because it did not directly infringe plaintiffs' works; it may still be liable as a contributory infringer." *Id.* at 1373.

There are stronger legal arguments for finding liability of verifiable domain owners because the owner, to some degree, gives a warranty—that the information is emanating from the owner's site. Furthermore, the domain owner has the option of avoiding such liability by not verifying the site.

142. Clayton, *supra* note 47, at 34, 35.

143. Otherwise, companies like America On-Line or educational institutions that allow students to utilize their domain name in their home page or other Internet address will be in danger of having a multitude of subscribers or users being mistaken as representatives of the company or school.

would not be presumed as those of the domain owner) or the domain name would be left unverified. Such unregistered or "user" denominated domains would have a higher threshold of liability for domain owners. In this way, domain owners could more easily limit any liability for libelous statements, invasion of privacy actions, or other theories of liability[144] based on personal home pages utilizing their domain name.

The rules of agency, such as the doctrine of apparent authority,[145] would establish the effect of a document on a person accessing it and the purported originator. Prudent domain owners may require employees or others who have personal pages connected through the institution's domain to include disclaimers on documents that might be construed to implicate the domain owners.[146]

By placing legal responsibility with domain name owners and web site operators, the government stays out of a policing business far beyond its capacity to perform. The Internet also maintains its cherished independence from governmental intrusion.

1. Disclaiming Links

Another likely consequence of the proposed solution is proliferation of disclaimers on documents posted on the Internet. Internet documents are often made up not only of text but also of hypertext links to other documents. These other documents may be produced by the same author or publisher and, thus, may be in the same domain. Links may also, however, be to documents outside of the domain. A document linked in such a way would need to be separately and individually identified by its domain and author or publisher. This is already happening to some extent. The following is from the Web page of the United Nations:

---

144. Such theories include negligence, breach of warranty, and strict liability. *See* Steve Reitenour, *supra* note 27, at 71, 72-73.

145. Apparent authority is the power to affect the legal relations of another person by transactions with third persons, professedly as agent for the other arising from and in accordance with the other's manifestations to such third persons. RESTATEMENT (SECOND) OF AGENCY § 8 (1958). Allowing an employee (or other person) to access the Internet through the company's domain is similar to putting them in the company uniform. Apparent authority exists only to the extent that it is reasonable for a third person dealing with the agent to believe that the agent is authorized. *Id.* at cmt. c.

146. A single institutional disclaimer of liability on each personal home page might not be sufficient because a user might bypass the initial page and access an original document within the personal page and, correspondingly, within the domain. Therefore, a disclaimer on all original documents is advisable. Links to other domains, however, would not possess a presumption of institutional authenticity if they were located outside of the domain.

> Disclaimer of Warranty; Limitation of Liability. YOU EXPRESSLY ACKNOWLEDGE AND AGREE THAT ANY AND ALL USE OF UNITED NATIONS ON-LINE IS AT YOUR SOLE RISK. . . .
>
> YOU FURTHER EXPRESSLY ACKNOWLEDGE AND AGREE THAT INFORMATION, TEXT, GRAPHICS, AND HYPERLINKS PROVIDED TO YOU THROUGH UNITED NATIONS ON-LINE AND LOCATED ON OTHER SITES THROUGHOUT THE COMBINED GLOBAL ELECTRONIC NETWORKS KNOWN AS THE INTERNET AND THE WORLD-WIDE-WEB ARE PROVIDED SOLELY AS A RESOURCE AND A CONVENIENCE TO YOU. SUCH HYPERLINKS TO OTHER SITES ARE NOT AN ENDORSEMENT BY THE UNITED NATIONS OF THOSE SITES. THE UNITED NATIONS MAKES NO WARRANTY, EITHER EXPRESS OR IMPLIED, AS TO THE ACCURACY, RELIABILITY, OR CONTENT OF SUCH INFORMATION, TEXT, GRAPHICS AND HYPERLINKS.[147]

The validity of disclaimers of information and links will probably need to be examined on a case by case basis.[148] Still, prudent domain owners and end users would be well advised to use such disclaimers if there is a possibility that somebody might detrimentally rely on the documents posted on their web site. In particular, corporations may want to disclaim warranties in their on-line sites. Otherwise, potential claimants in derivative suits might conduct "discovery."[149]

VIII. NECESSARY STEPS IN IMPLEMENTING THE SYSTEM

There are several types of actors in the Internet world. If they want to keep that world in its virtually unregulated state, they must standardize and police themselves. In the effort to create the previously proposed authentication system and combat fraud and frustration

---

147. *United Nations Publications Usage Agreement* (visited Mar. 17, 1997) <http://www.un.org/Depts/Treaty/agree.html>.

148. Blodwen Tarter, *Information Liability: New Interpretations for the Electronic Age*, 11 COMPUTER/L.J. 481, 545 (1992) ("The United States lacks comprehensive legislative or extensive case law dealing with information liability (or information technology at all) and thus must resort to case-by-case precedent setting."). *See generally* Anita Cava & Don Wiesner, *Rationalizing a Decade of Judicial Responses to Exculpatory Clauses*, 28 SANTA CLARA L. REV. 611 (1988). The authors state: "When examining exculpatory clauses, courts police: (1) the technical formation of the contract, i.e., the offer and acceptance; (2) the status or position of the parties to the bargain; and (3) public policy concerns." *Id.* at 612.

149. *See* Jean Marie R. Pechette, *Electronic Records Are Discoverable in Litigation*, NAT'L L.J., June 27, 1994, at C8.

with respect to information on the Internet, each type of actor has a part to play.

The existing *registration system*—or a separate entity for domain name registration should require proof of identity from person's registering domain names. In addition, it should require individuals registering domain names in the name of organizations and other corporate entities to provide proof of representative capacity. Otherwise, even with a verification mechanism, the system will be fraught with fraud. At the very least, it should keep a separate registry of verified domain owners and make the registry easily available for searching via the Internet itself.

Creators of web *browser software* should take steps to build verification capabilities into their software. Ideally, a browser would identify the domain owner as well as the address when it linked to the web site. In the alternative, browsers could create a verify command that would link to Network Solution's list of domain names.

Courts should announce policies concerning authenticity of web sites. To promote immediate uniformity, such policies would best be announced by court rule (amending Federal Rule of Evidence 901, for example), but could be done by decision. The rule would be that a document posted at the site owned by the entity that created it is presumed to be authentic. The presumption would be rebuttable. The rule would be an extension of Federal Rule of Evidence 901(b)(7), which applies to "Public Records or Reports" (limited to documents from a "public office").[150] Some nongovernmental documents are covered in a similar manner under Federal Rule of Evidence 901(b)(8) (Ancient Documents or Data Compilation), but those must be at least 20 years old.[151]

Domain owners must become aware of the responsibilities of ownership. They are the publishers of the twenty-first century. As such, they may be held liable for any number of occurrences. Owners should become more conscious of the potential for problems with their subdomain users. A prudent domain owner will establish policies and guidelines for web pages of those using their server—not limited to providing disclaimers of liability. Part of the responsibility of self-regulation is that domain owners must become responsible for the security of their web site. This responsibility is shared with Internet software providers. In addition, domain owners will be responsible for decisionmaking with respect to archiving (discussed previously).

---

150. *See* FED. R. EVID. 901(b)(7).
151. *See* FED. R. EVID. 901(b)(8).

Software makers should identify the problems in authentication and create programs to solve those problems. Security from tampering is one such need. A related need is a kind of "flagging" system that informs a user that a site has been tampered with. That way, even if the information cannot be made entirely secure from hackers, a user will be on notice if the document has been altered. In addition, software for creating and updating web pages should automatically post the date of changes on the updated pages.

Once the system is in place, *makers of search engines* should enable their products to search domain owners. If, for instance, a user wanted a document published by the Federalist Society, a search could be run that would look for that document only on servers that were registered to the Federalist Society.[152]

Government should scrutinize the Internet community to make sure that the standards and controls created protect the citizens. If the verified registry's costs will otherwise deter registration or create user fees, government should subsidize that registry. Government should also promote education of intelligent Internet use. Finally, recognizing its limitations in the multigovernment world of cyberlaw, government should encourage and facilitate self-regulation on the Internet. Government can also provide the impetus for an international convention that would allow similar authentication for foreign Internet sites.

These are the tasks that the interested parties to the Internet must perform in the near future. The practical problem is how to get them each to perform, and furthermore, to do so in conjunction with each other, without a single regulating or supervising source. The first reason to do so is to create a better product for the consumer (which ultimately increases revenue). The second reason is that if they do not resolve the problem themselves, a regulating or supervising source may ultimately make them.

IX. CONCLUSION

As the Internet becomes a more commonly used medium for research and commerce, the need for a method to establish authenticity of the information on it becomes imperative. The solution proposed in this Article attempts to join existing components and technologies

152. Legal researchers will recognize this type of search as being similar to the "segment" or "field" search available on LEXIS and Westlaw, respectively. Search engine makers would be well advised to consider whether other components of the search engines used by these legal research vendors could be utilized in Internet searching.

of the Internet with the desire to minimize governmental regulation of it.

As the Task Force on Archiving Digital Information recognized, the digital world of information technology is in its embryonic state.[153] Even before the recommendations of this Article could possibly be enacted, the context in which they were created will have changed to some degree. Because of the numerous variables in any resolution to the authentication problem, as well as the mutability of those variables, this solution may ultimately miss the mark. In addition, the interplay of so many parties, functions, and consequences will inevitably create other issues not addressed by this Article. Even so, putting forward a proposal provides a framework and ideas from which a better or more complete solution may ultimately come to light.

153. *PDI*, *supra* note 105, at The Challenge of Archiving Digital Information—Conceptual Framework.