

2009

## Best Evidence and the Wayback Machine: Toward a Workable Authentication Standard for Archived Internet Evidence

Deborah R. Eltgroth

---

### Recommended Citation

Deborah R. Eltgroth, *Best Evidence and the Wayback Machine: Toward a Workable Authentication Standard for Archived Internet Evidence*, 78 Fordham L. Rev. 181 (2009).

Available at: <http://ir.lawnet.fordham.edu/flr/vol78/iss1/5>

This Article is brought to you for free and open access by FLASH: The Fordham Law Archive of Scholarship and History. It has been accepted for inclusion in Fordham Law Review by an authorized editor of FLASH: The Fordham Law Archive of Scholarship and History. For more information, please contact [tmelnick@law.fordham.edu](mailto:tmelnick@law.fordham.edu).

# **BEST EVIDENCE AND THE WAYBACK MACHINE: TOWARD A WORKABLE AUTHENTICATION STANDARD FOR ARCHIVED INTERNET EVIDENCE**

*Deborah R. Eltgroth\**

*This Note addresses the use of archived Internet content obtained via the Wayback Machine, a service provided by the Internet Archive that accesses the largest online digital collection of archived Web pages in the world. Given the dynamic nature of the World Wide Web, Internet content is constantly changed, amended, and removed. As a result, interim versions of Web pages have limited life spans. The Internet Archive indexes and stores Web pages to allow researchers to access discarded or since-altered versions. In the legal profession, archived Web pages have become an increasingly helpful form of proof. Intellectual property enforcers have recognized the value of the Internet Archive as a tool for tracking down infringers, but evidence from the Internet Archive has rarely been admitted at trial. This Note surveys the handful of judicial opinions and orders that comment on the admission of Internet Archive evidence and explores the conflict underlying these approaches. As an alternative to the courses they have taken, this Note urges courts to treat the introduction of archived Web pages as implicating a best evidence issue in addition to an authentication question. Under this approach, courts would decide using evidence sufficient to the purpose, but not necessarily admissible at trial, whether the archived page qualifies as a “duplicate” of a page that once appeared on the Web. Beyond that, courts would apply authentication standards already developed to decide whether a reasonable jury could find, based only on admissible evidence, whether proffered evidence accurately represents the page stored on the Internet Archive server and, if necessary, whether the original page accurately represented material placed on the originating site by the site’s owner or operator. With this additional step, reliable evidence from the Wayback Machine can become as easily admitted as any other Internet-derived proof.*

---

\* J.D. Candidate, 2010, Fordham University School of Law; B.A., 2006, UCLA. I would like to thank my advisor, Professor James Kainen, for his tremendously valuable help and insight. Thank you also to my family and friends for their constant encouragement, patience, and support.

## TABLE OF CONTENTS

INTRODUCTION.....	182
I. THE AUTHENTICATION OF WEB CONTENT AND THE INTERNET	
ARCHIVE'S WAYBACK MACHINE .....	184
A. <i>Authentication and the Federal Rules</i> .....	186
1. Laying the Foundation: Federal Rule of Evidence 901 ....	187
2. The Best Evidence Rule and Admission of Duplicates .....	193
3. Role of Judge and Jury.....	196
B. <i>Anatomy of an Archived Web Page</i> .....	197
1. Origins: Where and What Is a Web Page .....	198
2. The Archival Process: Crawling the Web.....	198
3. Taking the Trip Back in Time: Finding Outdated Web Content with the Wayback Machine .....	201
II. THE EMERGING AUTHENTICATION STANDARDS FOR EVIDENCE	
FROM THE WAYBACK MACHINE .....	202
A. <i>Trusting the Intermediary To Establish the Rule 901         Foundation</i> .....	203
1. <i>Telewizja Polska USA, Inc. v. Echostar Satellite Corp.</i> .....	203
2. <i>St. Luke's Cataract &amp; Laser Institute, P.A. v. Sanderson,             M.D., LLC</i> .....	204
B. <i>Skeptics of Evidence from the Wayback Machine</i> .....	207
III. ARCHIVED WEB PAGES SHOULD BE EVALUATED AS AN ISSUE OF BEST EVIDENCE.....	210
A. <i>Channeling Xerox: Copies on the Web</i> .....	210
B. <i>Step One: Whether the Archived Web Page Qualifies as a         Duplicate</i> .....	212
C. <i>Step Two: Whether the Proffered Printout Shows What         Actually Appeared On-screen at the Internet Archive         Website</i> .....	213
CONCLUSION .....	214

## INTRODUCTION

The Web contains unfathomably more information than did the Alexandria library. If our culture ends up unable to retrieve and use that information, then all that knowledge will, in effect, have gone up in smoke.<sup>1</sup>

In April 2003, a novel form of evidence made a debut during the course of a scintillating trademark infringement suit brought by Playboy

---

1. Rick Weiss, *On the Web, Research Work Proves Ephemeral*, WASH. POST, Nov. 24, 2003, at A8.

Enterprises over use of the name “Sex Court.”<sup>2</sup> Dispositive of the issues at trial was priority of use; Mario Cavalluzzo, an Internet entrepreneur, claimed he had lawful rights to use the “Sex Court” term because he had operated a website at the *sexcourt.com* domain address prior to the air date of the Playboy show by the same name.<sup>3</sup> Playboy claimed that it had appropriated the mark through continuous bona fide use before Cavalluzzo had registered his domain name<sup>4</sup> and that, even assuming he had timely registered the domain, he failed to use the site on the continuous basis required to acquire trademark protection.<sup>5</sup> Playboy asserted that after leaving it inactive for some time, Cavalluzzo gradually began to emulate Playboy’s *Sex Court* show on the website.<sup>6</sup> Consequently, the start date of the defendant’s website in relation to the marketing and airing of *Sex Court* became crucial to Playboy’s case.

During the defense attorney’s cross-examination, Cindy Johnson, an Internet research manager employed by Playboy, explained that, although she could not testify to visiting the defendants’ website as far back as 1998, she had used a program called the Wayback Machine to obtain a snapshot of how the page had appeared then.<sup>7</sup> “A ‘Wayback Machine,’” now-Chief Judge Raymond J. Dearie quipped, “is this back to the future?”<sup>8</sup> The attorney moved on to a different line of questions without revisiting the issue; however, the cross question had opened the door for the plaintiffs. On redirect examination, counsel for Playboy introduced, over the defendants’ continuing objection, the printouts that Johnson had obtained via the Wayback Machine.<sup>9</sup> Their appearance suggested that the disputed content had appeared on Cavalluzzo’s website well after Playboy had appropriated the mark.<sup>10</sup> The parties settled the case before it went to the jury.<sup>11</sup>

This Note addresses the authentication of printouts obtained via the Internet Archive’s Wayback Machine within the framework of the Federal Rules of Evidence.<sup>12</sup> Using this framework, this Note seeks to reconcile the conflicting approaches that courts have taken to authenticate this novel form of evidence by recommending a standard where, as a preliminary

---

2. See Anthony M. DeStefano, *An Un.com-mom Revelation in Porn-Name Case*, *NEWSDAY* (Queens ed.), Apr. 10, 2003, at A23.

3. Counterclaimants’ Trial Memorandum of Law at 1–4, *Playboy Enters., Inc. v. On Line Entm’t, Inc.*, No. 00-Civ-6618 (E.D.N.Y. Mar. 31, 2003).

4. Plaintiffs’ Trial Memorandum of Law at 3, *Playboy*, No. 00-Civ-6618 (E.D.N.Y. Mar. 24, 2003).

5. *Id.* at 6.

6. *Id.* at 7.

7. Transcript of Record at 374–75, *Playboy*, No. 00-Civ-6618 (E.D.N.Y. Apr. 9, 2003).

8. *Id.* at 375.

9. *Id.* at 402.

10. See *id.* at 403–04.

11. See David Kesmodel, *Lawyers’ Delight: Old Web Material Doesn’t Disappear*, *WALL ST. J.*, July 27, 2005, at A1.

12. This Note analyzes the authentication issues that arise when proponents introduce printouts of archived Web pages; the issue of hearsay is beyond the scope of this Note.

matter to the Rule 901 authentication inquiry, courts would treat archived Web pages as copies subject to the standards of Rules 1001(4), 1002, and 1003.

Part I provides an overview of the case law and Federal Rules of Evidence relevant to this authentication inquiry. Part II discusses the approaches courts have taken thus far to authenticate Internet Archive evidence and explains why the approaches fundamentally conflict. Part III argues that as an alternative to the approaches they have taken, courts should treat the introduction of archived Web pages as implicating best evidence and decide this issue as an additional step to the Rule 901 authentication inquiry.

#### I. THE AUTHENTICATION OF WEB CONTENT AND THE INTERNET ARCHIVE'S WAYBACK MACHINE

Archived Internet evidence poses a novel scenario for the traditional evidence framework. Therefore, in order to conceptualize this form of evidence within the constructs of the Federal Rules of Evidence, the following sections provide a brief overview of the relevant Rules then describe the parties and processes involved in the compilation of the Internet Archive.

Part I.A summarizes the Federal Rules that inform the authentication inquiry and surveys how courts have applied the rules to authenticate contemporaneous Web pages introduced into evidence. Part I.B provides background on the Internet Archive, the Wayback Machine, and the process of digital archiving. In doing so, these sections should provide a picture of what, exactly, proponents seek to introduce when they present a printout from the Wayback Machine.

The Internet is hardly static. At any moment, websites are added, removed, and altered. Because of the transitive nature of the World Wide Web, nothing guarantees that a user will be able to access content in an identical form in the future.<sup>13</sup> The fleeting nature of the Internet poses a problem as the public and practitioners alike increasingly rely on the Web for information from governments, organizations, and companies.<sup>14</sup> For example, footnotes in academic journals that cite to Internet content run the risk of becoming obsolete fairly quickly, at times even before an article goes to press.<sup>15</sup> Attorneys face a similar problem if content on the Web

---

13. See Lawrence Lessig, *Innovating Copyright*, 20 CARDOZO ARTS & ENT. L.J. 611, 615 (2002) ("This is the scariest feature of the Internet, the part George Orwell would have understood best: The sense in which Net has no real history. A page can be changed without anybody noticing. It gets updated but no marks are left. . . . Remember the editors in 1984 constantly rewriting the past? Those editors are the Internet. For at any moment, we have no way of knowing what went before.").

14. See Richard Marcus, *Only Yesterday: Reflections on Rulemaking Responses to E-discovery*, 73 FORDHAM L. REV. 1, 10–11 (2004) (predicting that civil cases will become increasingly reliant on electronically stored materials).

15. See Weiss, *supra* note 1; see, e.g., David W. Opderbeck, *The Penguin's Paradox: The Political Economy of International Intellectual Property and the Paradox of Open*

stands to play a pivotal role in proving their case, as it did in the Playboy Sex Court trial. While Web pages are readily available contemporaneously, proof becomes difficult to come by once an opponent removes the content from a page. Enter the Internet Archive's Wayback Machine.

The Internet Archive is a nonprofit corporation based in San Francisco, California. Brewster Kahl and John Gage founded the organization in 1996 as an online library that would provide access to historical collections in digital format and "to prevent the Internet—a new medium with major historical significance—and other 'born-digital' materials from disappearing into the past."<sup>16</sup> When it debuted the Wayback Machine in 2001, the Internet Archive contained one hundred terabytes of data, roughly five times the content of the Library of Congress.<sup>17</sup> By 2005, the collection had grown to contain more than a petabyte, making it the largest data archive in the world.<sup>18</sup> In May, 2007 California officially recognized the Internet Archive as a library, thereby permitting the archive to benefit from several federal grant programs.<sup>19</sup>

The Wayback Machine<sup>20</sup> is a service available through the Internet Archive that allows parties to visit digitally archived Web pages.<sup>21</sup> Users can type in a URL<sup>22</sup> and select a date range, permitting them to browse

---

*Intellectual Property Models*, 18 STAN. L. & POL'Y REV. 101, 147 n.218 (2007) (citing to an archived copy of a Web page available at the Internet Archive because the source's Web page had been modified prior to the article's publication).

16. Internet Archive, About the Internet Archive, <http://www.archive.org/about/about.php> (last visited Aug. 12, 2009).

17. Beryl A. Howell, *Proving Web History: How To Use the Internet Archive*, J. INTERNET LAW 3, 5 (2006); Richard Koman, *How the Wayback Machine Works*, XML.com, Jan. 21, 2002, <http://www.xml.com/pub/a/ws/2002/01/18/brewster.html>.

18. Howell, *supra* note 17, at 5.

19. Adrian McCoy, *The Internet Gives Birth to an 'Official' Online Library*, PITTSBURGH POST-GAZETTE, June 24, 2007, at F1. The Internet Archive is not alone in the mission to catalogue the content of the Internet. The Library of Congress preserves culturally significant websites by way of a pilot project called Web Capture. See The Library of Congress, Web Capture, <http://www.loc.gov/webcapture/> (last visited Aug. 12, 2009). The Library of Congress and the Internet Archive are both members of the International Internet Preservation Consortium, whose other members include the national libraries of Australia, Canada, Denmark, Finland, France, Iceland, Italy, Norway, Sweden, and the United Kingdom. See International Internet Preservation Consortium, <http://netpreserve.org/about/index.php> (last visited Aug. 12, 2009); see also WebCite, <http://www.webcitation.org> (last visited Aug. 12, 2009) (archiving for academic purposes); Hanzo Archives, <http://www.hanzoarchives.com> (last visited Aug. 12, 2009) (providing archiving services to companies for compliance and e-discovery purposes).

20. The Wayback Machine name is a tribute to the *Rocky and Bullwinkle* cartoon in which the characters Peabody and Sherman used the "WABAC Machine" to travel back in time to famous events in history. Internet Archive, Frequently Asked Questions, <http://www.archive.org/about/faqs.php> (last visited Aug. 12, 2009) [hereinafter Internet Archive Frequently Asked Questions].

21. See Internet Archive, Wayback Machine, <http://www.archive.org> (last visited Aug. 12, 2009).

22. A URL (Universal Resource Locator) is an Internet address (for example, <http://www.archive.org>) that will take a user directly to a Web page when typed into the browser address window. RON WHITE, *HOW COMPUTERS WORK* 313 (9th ed. 2008). The address "consists of a communications protocol followed by the name or address of a

through older versions of the given site that were posted during the designated period.<sup>23</sup>

Thus far, many Internet researchers have utilized the Internet Archive's Wayback Machine regularly to monitor the Internet and protect intellectual property.<sup>24</sup> The Wayback Machine stands to play an increasingly important role at trial, particularly in intellectual property disputes.<sup>25</sup> However, some courts have excluded evidence obtained via the Wayback Machine, citing both authentication and hearsay concerns.<sup>26</sup> As a result, parties who rely on the content of a Web page, which can be changed or removed quickly, currently face an unclear evidentiary standard for the admission of what may serve as the only available record of since-vanished content. Once a party removes content from a website, the current case law fails to provide a clear standard for how to authenticate archived Web pages offered in lieu of the missing original.<sup>27</sup>

### A. Authentication and the Federal Rules

Before considering the authentication hurdles posed by the digital archival process, an overview of the Federal Rules of Evidence that inform

---

computer on the network and that often includes additional locating information." MERRIAM-WEBSTER'S COLLEGIATE DICTIONARY 1377 (11th ed. 2003).

23. Internet Archive Frequently Asked Questions, *supra* note 20. Part I.C provides an overview of the configuration and locale of digitally archived Web pages stored by the Internet Archive and available via the Wayback Machine.

24. Kesmodel, *supra* note 11. Kesmodel reports that use of the Wayback Machine has become so prevalent in the legal profession that "[a]t some law firms, litigators now ask researchers, 'can you do a Wayback on that?'" *Id.* See generally Matthew Fagan, Note, "Can You Do a Wayback on That?" *The Legal Community's Use of Cached Web Pages in and out of Trial*, 13 B.U. J. SCI. & TECH. L. 46 (2007).

25. See, e.g., Syncsort, Inc. v. Innovative Routines Int'l, Inc., No. 04-3623, 2008 U.S. Dist. LEXIS 35364, at \*13 (D.N.J. Apr. 30, 2008) (using Internet Archive evidence to prove that public had access to disputed information on the Internet and, therefore, the information did not amount to a trade secret); Allen v. The Ghoulish Gallery, No. 06-CV-371, 2007 U.S. Dist. LEXIS 86224, at \*8-10 (S.D. Cal. Nov. 20, 2007) (using Internet Archive evidence to establish validity of copyright claim); Dell, Inc. v. Innervision Web Solutions, Claim No. FA0503000445601 (National Arbitration Forum May 23, 2005), <http://www.adrforum.com/domains/decisions/445601.htm> (using Internet Archive evidence to conclude that the defendant had unlawfully diverted Web traffic, where the defendant had since changed its website to include only lawful uses); Natalie J. Spears & S. Roberts Carter III, *This Brand Is My Brand: Litigating Product Image*, 31 LITIG. 33, 38 (2005) (describing use of Internet Archive to prove past illegal uses of metatags).

26. See, e.g., Chamilia LLC v. Pandora Jewelry LLC, 85 U.S.P.Q.2d 1169 (S.D.N.Y. 2007); Novak v. Tucows Inc., No. 06-CV-1909, 2007 U.S. Dist. LEXIS 21269 (E.D.N.Y. Mar. 26, 2007) (order granting motion to bar evidence), *aff'd*, No. 07-2211-cv, 2009 U.S. App. LEXIS 9786 (2d Cir. May 6, 2009); see also Paris Glove of Can., Ltd. v. SBC/Sporto Corp., 84 U.S.P.Q.2d 1856, 1858-59 (T.T.A.B. 2007) (concluding the Internet Archive serves no purpose in authenticating contemporaneous Web pages because the Internet Archive itself is not self-authenticating).

27. See James H. Johnson Jr. & Deidre A. Francis, *Reviewing the Past Year's Top 10 Trademark Trends and an Analysis of Their Impact on 2008*, INTELL. PROP. & TECH. L.J., Feb. 2008, at 1, 5. For example, "trademark owners are left unsure of how best to prove infringement of their marks in cyberspace once a putative infringer removes the mark from its Web site." *Id.*

the authentication inquiry proves helpful.<sup>28</sup> The next section provides a background on the foundation requirement of Rule 901. As a starting point to this Note's analysis of archived evidence, this section then examines how Rule 901 has been applied to authenticate contemporaneous websites. Part I.A.2 explains the best evidence rule and the conditions under which duplicates may be admitted in lieu of originals. Part I.B then describes the process of digital archiving and the makeup of archived Web pages to provide a framework in which to appreciate the authentication and best evidence issues posed by the introduction of archived evidence.

### 1. Laying the Foundation: Federal Rule of Evidence 901

Prior to admission at trial, proponents seeking admission of evidence must authenticate that evidence.<sup>29</sup> Rule 901(a) sets out the foundation requirement for authentication of evidence.<sup>30</sup> Rule 901 provides that, as a condition precedent to admissibility, the proponent must lay a foundation sufficient to support a finding that the proffered evidence is what it purports to be.<sup>31</sup> The authenticity of evidence is ultimately an issue for the fact finder to decide;<sup>32</sup> however, because evidence lacks relevance unless it is what a proponent claims, the judge must first determine pursuant to Rules 104(a) and 901 that a factual basis exists to support a jury finding that the evidence is authentic.<sup>33</sup> For example, a proponent proffering a letter written by X must provide enough proof to establish a foundation on which a reasonable juror could conclude that the letter actually was written by X. Upon this showing, the court should admit the letter, but the jury will decide ultimately whether the evidence is credible and what weight to afford the contents of the letter.

Rule 901(b) sets forth some examples of means by which a proponent may satisfy the foundation requirement.<sup>34</sup> Rule 901(b)(1) provides that

---

28. A debate ensues over whether circumstances posed by the World Wide Web transcend traditional legal principles. Compare Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, 113 HARV. L. REV. 501 (1999), with Frank H. Easterbrook, *Cyberspace and the Law of the Horse*, 1996 U. CHI. LEGAL F. 207; see also David G. Post, *Against "Against Cyberanarchy,"* 17 BERKELEY TECH. L.J. 1365 (2002); Jack L. Goldsmith, *Against Cyberanarchy*, 65 U. CHI. L. REV. 1199 (1998). This Note proceeds on the assumption that the Federal Rules of Evidence provide sufficient framework in which to authenticate evidence from the Internet Archive. Cf. Easterbrook, *supra*.

29. 5 STEPHEN A. SALTZBURG ET AL., FEDERAL RULES OF EVIDENCE MANUAL § 901.02 (9th ed. 2006) ("Courts will not assume that evidence is what the proponent claims simply because on the face of the evidence it is possible that it might be."). The exception to this rule occurs where evidence is deemed self-authenticating. See FED. R. EVID. 902.

30. See FED. R. EVID. 901(a).

31. *Id.*

32. See 5 SALTZBURG ET AL., *supra* note 29, § 901.02; see, e.g., *United States v. Jackson*, 208 F.3d 633, 637 (7th Cir. 2000) ("Sorting truth from fiction, of course, is for the jury.").

33. See FED. R. EVID. 104(b); 5 SALTZBURG ET AL., *supra* note 29, § 901.02 ("As a question of conditional relevance, the admissibility standard is the same as that provided by Rule 104(b): Has the proponent offered a foundation from which the jury could reasonably find that the evidence is what the proponent says it is.").

34. See FED. R. EVID. 901(b).



testimony by a witness with knowledge that evidence is what it is claimed to be will generally satisfy the foundation requirement.<sup>35</sup> Rule 901(b)(9) provides “[e]vidence describing a process or system used to produce a result and showing that the process or system produces an accurate result” will also satisfy the foundation requirement.<sup>36</sup> This process or system category has played a major role in the authentication of evidence obtained through new types of technology and electronic evidence. For example, this rule has ushered in evidence in the forms of X-rays, films, and audio recordings.<sup>37</sup> Courts may also authenticate computer output pursuant to Rule 901(b)(9).<sup>38</sup> The foundation evidence used to authenticate computer output must describe the process involved and why it is accurate; for example, a proponent may need to demonstrate that the type of software or hardware employed is accepted as generally reliable.<sup>39</sup>

The history of courts’ approaches to the authentication of photographs illustrates the interplay of the provisions of Rule 901. In early cases where parties sought to admit photographs, courts treated the novel form of evidence with a high degree of suspicion. To establish the foundation requirement, courts often “place[d] upon the offering party the burden of producing the negative as well as the photograph itself, and of proving that neither retouching or other manual or chemical intervention was reflected in the proffered print.”<sup>40</sup> However, as the technology became more widespread and the admission of photographs more commonplace, courts relaxed the burden. In 1968, the U.S. Court of Appeals for the Sixth Circuit summed up in *United States v. Hobbs*<sup>41</sup> the rationale for relaxing the burden on proponents of photographic evidence:

Concerning any photographic operation only the most scholarly expert could testify as to the manner in which the original image is transmitted through the lens of the camera to the emulsion on the film or plate, the development of the latent image, the printing by a contact or projection process, and concerning the chemical procedures involved. Even where an occasional qualified witness may be available to testify as to such

---

35. FED. R. EVID. 901(b)(1).

36. FED. R. EVID. 901(b)(9).

37. GLEN WEISSENBERGER & JAMES J. DUANE, WEISSENBERGER’S FEDERAL EVIDENCE § 901.39 (5th ed. 2006). When assessing evidence pursuant to Rule 901(b)(9), “[j]udges must be careful to differentiate . . . between authentication of a process generally and a showing that a particular machine works as intended.” 5 SALTZBURG ET AL., *supra* note 29, § 901.02[11].

38. 31 CHARLES ALAN WRIGHT & VICTOR JAMES GOLD, FEDERAL PRACTICE AND PROCEDURE: EVIDENCE § 7114 (2000); *see, e.g.*, *State v. Cook*, 777 N.E.2d 882, 886–87 (Ohio Ct. App. 2002) (holding “mirror image” of a computer’s hard drive created by way of connecting the hard drive through a cable to a replica hard drive had been authenticated properly pursuant to Rule 901(b)(9)).

39. 31 WRIGHT & GOLD, *supra* note 38, § 7114; *e.g.*, *Bray v. Bi-State Dev. Corp.*, 949 S.W.2d 93, 97 (Mo. Ct. App. 1997) (“[T]he relevant technical or scientific community’s use of or reliance on such software has been held sufficient to establish the accuracy of the software.”).

40. *United States v. Hobbs*, 403 F.2d 977, 978 (6th Cir. 1968).

41. 403 F.2d 977.

details such testimony would obviously be irrelevant and immaterial. What is material is what the rankst box camera amateur knows, namely that he “gets” what he sees. We thus come full circle to the judicial test . . . as being whether the proffered photograph is an accurate representation of the scene depicted.<sup>42</sup>

Today, a party seeking to admit photographic images can generally establish the foundational predicate with testimony by a witness verifying that the photographs are a fair and accurate representation of what the witness saw.<sup>43</sup> Because the examples set forth in Rule 901(b) are “[b]y way of illustration only, and not by way of limitation,” a proponent alternatively can fulfill the foundation requirement through other means.<sup>44</sup> As the U.S. Court of Appeals for the District of Columbia Circuit has explained, “the contents of photographic evidence to be admitted into evidence need not be merely illustrative, but can be admitted as evidence independent of the testimony of any witness as to the events depicted, upon a foundation sufficient to meet the requirements of Federal Rule of Evidence 901(a).”<sup>45</sup> In similar fashion, the increasing use of surveillance cameras has led some courts to authenticate video recordings and images obtained from these recordings based on circumstantial evidence without testimony by witnesses to the recorded events.<sup>46</sup>

Norms surrounding the introduction of Internet evidence in court have evolved rather recently. While use of the Internet has grown increasingly prevalent, early cases applying Rule 901 to authenticate Web pages imposed an extremely high bar to establish the foundational predicate—in a fashion reminiscent of early cases that examined photographic evidence. In an often mentioned U.S. District Court for the Southern District of Texas

---

42. *Id.* at 978–79, *cited with approval in* Mikus v. United States, 433 F.2d 719, 725 (2d Cir. 1970).

43. See FED. R. EVID. 901(b)(1); see, e.g., Isaacs v. State, 386 S.E.2d 316, 329 (Ga. 1989); Huffman v. State, 746 S.W.2d 212, 222 (Tex. Crim. App. 1988); People v. Bowley, 382 P.2d 591, 596 (Cal. 1963); see also Mikus, 433 F.2d at 725 (admitting videotape recording upon testimony by witness that it was a fair and accurate representation of events in question because requiring expert to testify about the functioning of a video camera would be an unnecessary roadblock to admission).

44. FED. R. EVID. 901.

45. United States v. Rembert, 863 F.2d 1023, 1026–28 (D.C. Cir. 1988); see United States v. Stearns, 550 F.2d 1167, 1171 (9th Cir. 1977) (“Even if direct testimony as to foundation matters is absent . . . the contents of a photograph itself, together with such other circumstantial or indirect evidence as bears upon the issue, may serve to explain and authenticate a photograph sufficiently to justify its admission into evidence.”); see also Bowley, 382 P.2d at 594–95 (concluding that, in addition to illustrative purposes, photographs “may also be used as probative evidence of what they depict”).

46. See, e.g., Rembert, 863 F.2d at 1026–28 (upholding admission of time-stamped photographic evidence obtained from surveillance videotape); United States v. Taylor, 530 F.2d 639, 641–42 (5th Cir. 1976) (finding foundation properly established where government witnesses “testified as to the manner in which the film was installed in the camera, how the camera was activated, the fact that the film was removed immediately after the robbery, the chain of its possession, and the fact that it was properly developed and contact prints made from it,” even though the bank employees did not personally witness the recorded events and, therefore, could not attest to the accuracy of the tape).

opinion, a judge referred to the Internet as “voodoo information” and held “[t]here is no way Plaintiff can overcome the presumption that the information he discovered on the Internet is inherently untrustworthy.”<sup>47</sup> Courts often cite *St. Clair v. Johnny’s Oyster & Shrimp Inc.*<sup>48</sup> for the proposition that Internet evidence should be treated with a high degree of skepticism.<sup>49</sup>

However, to an increasing extent courts recognize that the authentication issues that arise from content obtained from the Internet can be resolved with existing evidentiary principles.<sup>50</sup> Gregory Joseph, the former Chair of the American Bar Association Section of Litigation and a former member of the Advisory Committee on Evidence Rules, advises,

Detecting modifications of electronic evidence can be very difficult, if not impossible. That does not mean, however, that nothing is admissible because everything is subject to distortion. The same is true of many kinds of evidence, from testimony to photographs to digital images, but that does not render everything inadmissible. It merely accentuates the need for the judge to focus on all relevant circumstances in assessing admissibility under Fed. R. Evid. 104(a)—and to leave the rest to the jury, under Rule 104(b).<sup>51</sup>

In line with this view, courts increasingly have found occasion to admit Web content pursuant to the Federal Rules.<sup>52</sup>

---

47. *St. Clair v. Johnny’s Oyster & Shrimp, Inc.*, 76 F. Supp. 2d 773, 774–75 (S.D. Tex. 1999).

48. 76 F. Supp. 2d 773.

49. See, e.g., *Tolliver v. Fed. Republic of Nig.*, 265 F. Supp. 2d 873, 876 (W.D. Mich. 2003) (citing *St. Clair* in striking news postings printed from the Internet); *Wady v. Provident Life & Accident Ins. Co. of Am.*, 216 F. Supp. 2d 1060, 1064–65 (C.D. Cal. 2002); *Barbour v. Head*, 178 F. Supp. 2d 758, 760 n.3 (S.D. Tex. 2001) (noting agreement with *St. Clair* that information obtained from the Internet is “inherently untrustworthy” (quoting *St. Clair*, 76 F. Supp. 2d at 774)).

50. See 5 JACK B. WEINSTEIN & MARGARET A. BERGER, WEINSTEIN’S FEDERAL EVIDENCE § 900.06[3] (Joseph M. McLaughlin ed., Matthew Bender 2d ed. 1997) (“[R]outine computer operations will not raise serious authentication challenges, primarily because the accuracy of basic computer operations has become widely accepted, dispensing the need to make an elaborate showing of accuracy. . . . So long as the requirements of Rule 803(6) are satisfied, the trustworthiness of individual computers is not separately scrutinized under Rule 901, unless trustworthiness is specifically challenged. Moreover, to the extent that the reliability of the evidence is challenged, the objection goes to its probative weight, and not to its admissibility. For example, objections raising inaccuracies in the operation of the software program or from incorrect data inputting usually go the weight of the evidence and not its admissibility.”). But see GEORGE L. PAUL, FOUNDATIONS OF DIGITAL EVIDENCE 29 (2008) (submitting that “the hallowed evidentiary concept of ‘original’ no longer has meaning in the digital realm”).

51. Gregory P. Joseph, *Internet and Email Evidence*, in TRIAL EVIDENCE IN THE FEDERAL COURTS: PROBLEMS AND SOLUTIONS 559, 564–65 (ALI-ABA Course of Study, Feb. 28–29, 2008), as reprinted in 5 SALTZBURG ET AL., *supra* note 29, pt. 4, at 20.

52. See, e.g., *Westlands Water Dist. v. United States*, 153 F. Supp. 2d 1133, 1143 n.9 (E.D. Cal. 2001) (noting that Internet citations are an appropriate source for background information); *In re F.P.*, 878 A.2d 91, 95 (Pa. Super. Ct. 2005) (rejecting an opponent’s challenge to the admission of emails into evidence and concluding “the same uncertainties [about evidence obtained from the Internet] exist with traditional written documents. A

Practically speaking, when parties want to introduce the content of a Web page at trial, they must either print out a Web page to demonstrate how it appears on-screen, or have a witness testify to the appearance of the page. In either case, the authentication inquiry must consider three subsidiary questions: “(1) What was actually on the Web site? (2) Does the exhibit or testimony accurately reflect it? (3) If so, is it attributable to the owner of the site?”<sup>53</sup>

Where the evidence provided is in the form of a printout, the proponent can answer the first two questions with testimony by any witness who has typed in the URL corresponding to the website, viewed what was posted to the website, and, consequently, can testify to the accuracy of the printout with respect to the content that appeared on the screen.<sup>54</sup> Requiring a witness to testify that a proffered printout is a fair and accurate representation of a Web page is akin to the authentication of photographic evidence.<sup>55</sup> For example, in *Ford v. State*,<sup>56</sup> the Georgia Court of Appeals permitted a participant in a chat room conversation to authenticate a transcript, reasoning,

[W]e find this situation analogous to the admission of a videotape, which “is admissible where the operator of the machine which produced it, or one who personally witnessed the events recorded[,] testifies that the videotape accurately portrayed what the witness saw take place at the time the events occurred.” Here, Keller personally witnessed the real-time chat recorded in Transcript B as it was taking place, and he testified that the transcript accurately represented the on-line conversation. Under these circumstances, Keller’s testimony was tantamount to that of a witness to an event and was sufficient to authenticate the transcript.<sup>57</sup>

---

signature can be forged; a letter can be typed on another’s typewriter; distinct letterhead stationary can be copied or stolen. We believe that e-mail messages and similar forms of electronic communication can be properly authenticated within the existing framework . . .”).

53. Joseph, *supra* note 51, at 562, *cited with approval in* Lorraine v. Markel Am. Ins. Co., 241 F.R.D. 534, 555 (D. Md. 2007); *see also* 1 JAY E. GRENIG & WILLIAM C. GLEISNER, III, EDISCOVERY & DIGITAL EVIDENCE § 14:9 (2005) (explaining that to authenticate a printout from a Web page it must be shown that the printout accurately reflects the same text or images appearing on the computer monitor on the date the printout was made).

54. *See, e.g.,* Perfect 10, Inc. v. Cybernet Ventures, Inc., 213 F. Supp. 2d 1146, 1154 (C.D. Cal. 2002) (admitting printouts of several Web pages based on the declaration of the party who printed them out verifying that the printouts were true and correct copies of the Web pages as they appeared on the screen).

55. Joseph, *supra* note 51, at 562–63 (The showing required to establish the likeness between the page as viewed and the page as printed “is no different than that required to authenticate a photograph, other replica or demonstrative exhibit.”); *see supra* notes 43–46 and accompanying text.

56. 617 S.E.2d 262 (Ga. Ct. App. 2005), *cert. denied*, No. S05C1921, 2005 Ga. LEXIS 789 (Ga. Nov. 7, 2005).

57. *Id.* at 265–66 (quoting *Hudson v. State*, 538 S.E.2d 751 (Ga. 2000)) (alteration in original); *see also* *Adams v. State*, 117 P.3d 1210, 1218 (Wyo. 2005) (“The State’s witness testified that the chat log exhibits were exact copies of the communication between the parties contained in the computer and thus, they were either appropriate computer ‘originals’ or duplicates which were properly authenticated. Whether they accurately reflected the

Further, like photographic evidence, courts have concluded in some cases that circumstantial indicia of authenticity, such as the date and Web address on a Web page printout, can substantiate a reasonable conclusion of authenticity under Rule 901.<sup>58</sup>

Whether content can be attributed rightly to the owner of a website largely depends on individual factual circumstances. In some cases, courts presume that material on a URL and date stamped printout may be attributed to the owner of the site.<sup>59</sup> In a similar vein, federal courts consider records from government websites self-authenticating pursuant to Rule 902(5).<sup>60</sup> Absent such a presumption, courts may require proof of accuracy by the owner, webmaster, or content provider of a Web page.<sup>61</sup> For example, courts are less likely to attribute the contents of a website to

---

contents of the instant messages sent between the parties was an issue for the jury to decide . . .”).

58. See *Perfect 10*, 213 F. Supp. 2d at 1154; see also *U.S. EEOC v. E.I. DuPont de Nemours & Co.*, No. 03-1605, 2004 WL 2347559, at \*2 (E.D. La. Oct. 18, 2004) (order admitting printout that showed date and URL address).

59. See *Joseph*, *supra* note 51, at 564 (“While it is reasonable to indulge a presumption that the contents of a website are fairly attributable to the site’s owner, that does not apply to chat room evidence.”); 1 GRENIG & GLEISNER, *supra* note 53, § 14:9 (“If the court has concerns, the proponent could be required to obtain certifications of the accuracy of the content of the copied Web pages from the owner of the Web page. In the alternative, courts can assume that most Web pages appear the same as they did when posted by the owner.”); see, e.g., *Page v. Lexington Co. Sch. Dist. One*, No. 06-249-CMC, 2007 U.S. Dist. LEXIS 3886, at \*21 n.12 (D.S.C. Jan. 17, 2007) (presuming that proffered screen printouts were accurate depictions of the Web page in question); *Ampex Corp. v. Cargle*, 27 Cal. Rptr. 3d 863, 867 n.2 (App. Dep’t Super. Ct. 2005) (finding printouts from a corporate website self-authenticating under California law).

60. See, e.g., *Estate of Gonzales v. Hickman*, No. 05-660, 2007 WL 3237727, at \*2 n.3 (C.D. Cal. May 30, 2007) (order admitting report issued by the Inspector General of the State of California posted on the Office of the Inspector General’s website); *Lorraine v. Markel Am. Ins. Co.*, 241 F.R.D. 534, 551 (D. Md. 2007) (“Given the frequency with which official publications from government agencies are relevant to litigation and the increasing tendency for such agencies to have their own websites, Rule 902(5) provides a very useful method of authenticating these publications. When combined with the public records exception to the hearsay rule, Rule 803(8), these official publications posted on government agency websites should be admitted into evidence easily.”); *United States ex rel. Parikh v. Premier Blue Cross*, No. C01-0476P, 2006 WL 2841998, at \*4 (W.D. Wash. Sept. 29, 2006) (order denying motion to dismiss). The rationale for including government websites in the self-authenticating category is largely premised on chain of custody considerations. See 5 SALTZBURG ET AL., *supra* note 29, § 901.02 [11].

61. 1 GRENIG & GLEISNER, *supra* note 53, § 14:9; *Joseph*, *supra* note 51, at 566; see, e.g., *In re Homestore.com, Inc. Sec. Litig.*, 347 F. Supp. 2d 769, 782 (C.D. Cal. 2004) (“Printouts from a web site do not bear the indicia of reliability demanded for other self-authenticating documents under Fed. R. Evid. 902. To be authenticated, some statement or affidavit from someone with knowledge is required; for example, Homestore’s web master or someone else with personal knowledge . . . .”); *Wady v. Provident Life & Accident Ins. Co. of Am.*, 216 F. Supp. 2d 1060, 1064 (C.D. Cal. 2002) (striking documents obtained from defendant’s website because plaintiff lacked personal knowledge to authenticate the documents as statements made by the defendant).

the owner when a site permits users to post their own content, as is the case in Internet chat rooms.<sup>62</sup>

## 2. The Best Evidence Rule and Admission of Duplicates

Because they amount to writings, Web page printouts are subject to the best evidence rule, which provides that when a party seeks “[t]o prove the content of a writing, recording, or photograph, the original writing, recording, or photograph is required” unless the circumstances justify use of a duplicate.<sup>63</sup> The best evidence rule “is principally aimed, not at securing a writing at all hazards and in every instance, but at securing the best *obtainable* evidence of its contents.”<sup>64</sup>

The Federal Rules of Evidence expressly provide that “[i]f data are stored in a computer or similar device, any printout or other output readable by sight, shown to reflect the data accurately, is an ‘original.’”<sup>65</sup> Therefore, printouts from contemporaneous websites do not violate the best evidence rule to the extent they are offered to depict the image displayed on the computer screen.<sup>66</sup> However, the printouts are considered originals only with respect to the computer data they represent; they are not originals of sources of that data.<sup>67</sup> Consequently, writings, photographs, or recordings that have been scanned or entered into a computer retain their original status for the purposes of the best evidence rule.<sup>68</sup> On the other hand, “the display may qualify as a duplicate o[f] the photograph or document under Rule 1001(4).”<sup>69</sup>

The best evidence rule does not apply unless a party seeks to prove the contents of a writing, recording, or photograph.<sup>70</sup> So for example, the Rule

---

62. *E.g.*, *United States v. Jackson*, 208 F.3d 633, 638 (7th Cir. 2000) (“Jackson needed to show that the web postings in which the white supremacist groups took responsibility for the racist mailings actually were posted by the groups, as opposed to being slipped onto the groups’ web sites by Jackson herself, who was a skilled computer user.”); *see also* SHIRA A. SCHEINDLIN, DANIEL J. CAPRA & THE SEDONA CONFERENCE, *ELECTRONIC DISCOVERY AND DIGITAL EVIDENCE: CASES AND MATERIALS* 539–40 (2009).

63. FED. R. EVID. 1002.

64. 2 KENNETH S. BROUN ET AL., *MCCORMICK ON EVIDENCE* § 237 (John W. Strong ed., 5th ed. 1999); *see United States v. Bakhtiar*, 994 F.2d 970, 977 (2d Cir. 1993) (upholding admission of duplicates of cancelled checks and explaining best evidence rule is concerned with “fairness of introducing the duplicates as a means of establishing the contents of the writing”).

65. FED. R. EVID. 1001(3).

66. *See* 31 WRIGHT & GOLD, *supra* note 38, § 7166, at 338–39.

67. *Id.* at § 7166, at 339.

68. *Id.*

69. *Id.*

70. 5 SALTZBURG ET AL., *supra* note 29, § 1002.02; *see R & R Assocs. v. Visual Scene, Inc.*, 726 F.2d 36, 38 (1st Cir. 1984) (“Rule 1002 applies not when a piece of evidence sought to be introduced has been somewhere recorded in writing but when it is that written record itself that the party seeks to prove.”); *United States v. Rusmisel*, 716 F.2d 301, 314 n.16 (5th Cir. 1983) (finding witness testimony in lieu of a court record did not violate best evidence rule where party sought to prove the content of the closing argument—not the contents of a writing); *United States v. Rose*, 590 F.2d 232, 237 (7th Cir. 1978) (finding testimony about a tape-recorded telephone conversation did not violate best evidence rule).

would not bar the introduction of photographs used to illustrate a crime scene because the contents of the photographs themselves would not be in dispute but, rather, the scene they depict.<sup>71</sup> Instead, the best evidence rule applies where the contents of the writing or recording form the basis for the action, as is the case in suits for copyright infringement, fraud, libel, obscenity, and invasion of privacy.<sup>72</sup>

The Federal Rules of Evidence account for the fact that originals may not always be available and that modern technology affords access to reliable copies. In this regard, Rule 1003 creates an exception to requiring an original by providing for the admission of duplicates in lieu of originals under some circumstances. Rule 1003 states that “[a] duplicate is admissible to the same extent as an original unless (1) a genuine question is raised as to the authenticity of the original or (2) in the circumstances it would be unfair to admit the duplicate in lieu of the original.”<sup>73</sup> Underlying Rule 1003 is the recognition that as copying technology has become widespread, parties at times inevitably will rely on duplicates of original documents. The effect is that, while duplicates remain subject to the best evidence rule, Rule 1003 treats them as presumptively admissible.<sup>74</sup>

---

71. See *United States v. Fagan*, 821 F.2d 1002, 1008 n.1 (5th Cir. 1987) (finding no violation of best evidence rule where court admitted testimony about a conversation even though a tape was available because the testimony was offered to prove the content of the conversation and not the content of the tape); cf. *Ford v. State*, 617 S.E.2d 262, 265–66 (Ga. Ct. App. 2005), *cert. denied*, No. S05C1921, 2005 Ga. LEXIS 789 (Ga. Nov. 7, 2005) (likening chat room transcript to a video recording of a conversation). The advisory committee’s note to Rule 1002 contemplates that the use of photographs seldom implicates the best evidence rule because in most cases witnesses have personal knowledge of the depicted scene and use photographs only to illustrate their testimony. The committee opines that “[c]ases in which an offer is made of the testimony of a witness as to what he saw in a photograph or motion picture, without producing the same, are most unusual,” and generally limited to the copyright, defamation, and invasion of privacy realms. FED. R. EVID. 1002 advisory committee’s note. However, one could contemplate that this would be precisely the scenario in which the Internet Archive would be most useful.

72. 31 WRIGHT & GOLD, *supra* note 38, § 7184, at 386–88; see also 5 SALTZBURG ET AL., *supra* note 29, § 1002.02.

73. FED. R. EVID. 1003.

74. 5 SALTZBURG ET AL., *supra* note 29, § 1003.02; see also SCHEINDLIN ET AL., *supra* note 62, at 540–44 (“[T]he Best Evidence Rule is likely to be a problem only if the proponent is unprepared.”). The best evidence rule is not without criticism, especially as applied to digital evidence. In his recent book, George L. Paul takes the position that originality as contemplated by the Federal Rules of Evidence proves largely irrelevant in the digital world:

[T]he concept of originality under the Federal Rules has been reduced to triviality regarding digital information. A large hole in the law has appeared in one of its main assumptions, that with regard to written records there was evidence that could be traced closer to an original or farther away from it—and that there was a requirement to go closer to the original if that was available.

Accordingly, the current system of foundations allows litigants to place into evidence almost anything they want so long as they can get a witness with some nexus to testify that a document is what it is claimed to be. They can employ a sort of legerdemain. If we are to be intellectually honest, there is almost no preliminary burden of proving digital information is authentic.

A duplicate is defined by the Federal Rules as “a counterpart produced by the same impression as the original, or from the same matrix, or by means of photography, including enlargements and miniatures, or by mechanical or electronic re-recording, or by chemical reproduction, or by other equivalent techniques which accurately reproduces the original.”<sup>75</sup> As some commentators have suggested, “[t]he requirement of accuracy does not always mean that the reproduction must be perfect. If the defects in reproduction do not materially alter the contents, the item may still be considered a duplicate.”<sup>76</sup> If a copy does not qualify as a duplicate under Rule 1003, it may still qualify for admission as secondary evidence of a writing’s contents pursuant to and subject to the conditions of Rule 1004.<sup>77</sup>

Whether or not a proffered document constitutes a duplicate under 1001(4) is a matter left to the discretion of the court.<sup>78</sup> While the advisory committee notes that, in general, original status should be self-evident, the committee advises that some circumstances call for more particularized definitions. For example, “[w]hile strictly speaking the original of a photograph might be thought to be only the negative, practicality and common usage require that any print from the negative be regarded as an original.”<sup>79</sup> Wright and Gold’s *Federal Practice and Procedure* suggests that a “‘duplicate’ [under Rule 1001(4)] is a copy produced by some technique that reduces the risk of error by avoiding dependence on human accuracy and sincerity.”<sup>80</sup> This suggestion squares with the advisory committee notes, which make clear that manual copies should not be

---

In the future, in certain defined areas we may want to devise a regulatory scheme, or even a larger system of evidence, that requires legitimate proof of authenticity before digital records are thrown into the ring of evidence. The bottom line is that today, any trivial showing is likely to serve as a foundation for getting digital evidence admitted.

PAUL, *supra* note 50, at 49.

75. FED. R. EVID. 1001(4).

76. 31 WRIGHT & GOLD, *supra* note 38, § 7167, at 345.

77. See FED. R. EVID. 1004; KENNETH S. BROUN ET AL., *supra* note 64, § 237 (“[I]f as a practical matter the document cannot be produced because it has been lost or destroyed, the production of the original is excused and other evidence of its contents becomes admissible.”); see, e.g., *In re Macmillan, Inc.*, 186 B.R. 35, 48 (Bankr. S.D.N.Y. 1995) (“Depending on the circumstances present in a given case, evidence may qualify as a ‘duplicate original’ under Rule 1003, thus meeting the ‘original writing’ rule’s requirements head on, or it may qualify as secondary evidence of the writing’s contents under Rule 1004.”).

78. See 5 SALTZBURG ET AL., *supra* note 29, § 1003.02[2]. The authors suggest that “the Trial Judge should err on the side of admitting the duplicate, leaving it to the jury to consider the possibility that the original might not be genuine or that the duplicate might be inaccurate or incomplete.” *Id.*; see, e.g., *State v. Cook*, 777 N.E.2d 882, 887 (Ohio Ct. App. 2002) (finding no abuse of discretion where trial court admitted a duplicate of the contents of defendant’s hard drive containing child pornography); *Broderick v. State*, 35 S.W.3d 67, 79 (Tex. App. 2000) (affirming trial court’s admission of duplicate of defendant’s hard drive in lieu of the original upon testimony by expert that the copy of the hard drive duplicated exactly the original).

79. FED. R. EVID. 1001(3) advisory committee’s note.

80. 31 WRIGHT & GOLD, *supra* note 38, § 7167, at 341.



included within the definition,<sup>81</sup> although, in some cases, an incomplete copy will still qualify as a duplicate for purposes of the best evidence rule.<sup>82</sup> The committee's notes to Rule 1001(4) further advise that a document may constitute both a duplicate and an original depending on the circumstances.<sup>83</sup>

### 3. Role of Judge and Jury

Courts generally agree that the authentication of evidence is an issue of conditional relevance and, consequently, Rule 104(b) governs the Rule 901 inquiry.<sup>84</sup> When making a determination pursuant to Rule 104(b), a court may only consider evidence that is admissible at trial.<sup>85</sup> In contrast, Rule 1003 falls within the purview of Rule 104(a), which delegates preliminary questions of admissibility to the court.<sup>86</sup> Consequently, as part of the inquiry into duplicate status and best evidence, the judge may consider inadmissible evidence, including hearsay, to inform the decision whether to admit a duplicate in lieu of an original. In this respect, the judge acts as fact finder when deciding whether evidence fulfills the technical requirements of Rule 1001(4) to qualify as a duplicate and conform to the best evidence rule.<sup>87</sup> The questions for the judge include:

- (1) whether a given item of evidence is an "original";
- (2) whether a given item of evidence qualifies as a duplicate and is thus presumptively admissible;
- (3) whether a genuine question is raised as to the authenticity of the original for purposes of Rule 1003;

---

81. FED. R. EVID. 1001(4) advisory committee's note.

82. See, e.g., *United States v. Hall*, 342 F.2d 849, 852-53 (4th Cir. 1965) (no abuse of discretion in allowing copy of a tape to be played to jury where a quarter of the recording was inaudible); *Fed. Deposit Ins. Corp. v. Rodenberg*, 571 F. Supp. 455, 458 (D. Md. 1983) (missing margins of one document and illegible portions of another did not render photocopies inaccurate and, therefore, they were admissible as duplicates pursuant to Rule 1001(4)); cf. *United States v. Schanerman*, 150 F.2d 941, 944 (3d Cir. 1945) ("There would be no more valid reason for exclusion of the mechanically recorded conversations than there would be for excluding competent conversations, overheard in part, by human witnesses."). But see FED. R. EVID. 1003 advisory committee's note ("Other reasons for requiring the original may be present when only a part of the original is reproduced and the remainder is needed for cross-examination or may disclose matters qualifying the part offered or otherwise useful to the opposing party." (citing *United States v. Alexander*, 326 F.2d 736 (4th Cir. 1964))).

83. FED. R. EVID. 1001(4) advisory committee's note.

84. 5 WEINSTEIN & BERGER, *supra* note 50, § 900.06[1][c] ("[U]nder Rule 104(b) a judge only makes a preliminary determination that there is sufficient evidence to support a jury finding that the evidence is relevant.").

85. *Id.*; FED. R. EVID. 901 advisory committee's note; see, e.g., *Ricketts v. City of Hartford*, 74 F.3d 1397, 1409-10 (2d Cir. 1996).

86. See FED. R. EVID. 104(a); 1 SALTZBURG ET AL., *supra* note 29, § 104.02[1].

87. See 1 SALTZBURG ET AL., *supra* note 29, § 104.02[1].

- (4) whether it would be unfair to admit a duplicate in lieu of an original as provided for in Rule 1003;
- (5) whether an original is lost or destroyed, and whether a diligent search has been conducted for the original;
- (6) whether the proponent lost or destroyed evidence in bad faith;
- (7) whether an original can be obtained by any available judicial process;
- (8) whether an adverse party has possession or control over the original and, if so, whether proper notice was given to that party;
- (9) whether evidence goes to a collateral matter or to a controlling issue.<sup>88</sup>

For example, while photocopy machines are a generally accepted means to produce a duplicate document, human error or mechanical malfunctions may at times interfere with the photocopying process. As a result, the judge must still determine whether a given photocopy in fact “accurately reproduces the original” pursuant to Rule 1001(4).<sup>89</sup> “[E]ven if an item qualifies as a duplicate under Rule 1001(4), a court may still exclude it under Rule 1003 . . . .”<sup>90</sup> The *Federal Rules of Evidence Manual* advises that when considering whether to admit a duplicate, judges should consider “the quality of the duplicate, the possibility that the duplicate does not fully reproduce all aspects of the original, the specificity and sincerity of the challenge, the importance of the evidence to the case, and the burdens of producing the original.”<sup>91</sup> In addition, the opportunity or lack thereof for discovery prior to trial may inform the decision.<sup>92</sup>

### B. *Anatomy of an Archived Web Page*

This section conceptualizes archived Web pages in three steps: first, the publication of an “original” Web page; second, the process by which this page gets archived; and third, the archived version as it exists at a separate, independent location apart from the original on the World Wide Web. Part I.B.1 provides a background on the mechanics of contemporaneous Web pages. Part I.B.2 summarizes the process by which Web “crawlers” make digital copies of these Web pages. Part I.B.3 describes the end product (the archived Web page) that the Wayback Machine makes accessible.

---

88. 5 SALTZBURG ET AL., *supra* note 29, § 1008.02[1] (footnote omitted). “All of the above questions are concerned with the policies behind the Best Evidence Rule. None would be of much interest to the jury, and would only serve to distract and confuse. Hence, they are properly left to the Trial Judge.” *Id.*

89. 31 WRIGHT & GOLD, *supra* note 38, § 7167, at 344.

90. *Id.* at 345–46.

91. 5 SALTZBURG ET AL., *supra* note 29, § 1003.02[2].

92. *See id.* (“An additional factor that Trial Judges may wish to consider in deciding whether a genuine question as to authenticity is raised or whether admission of a duplicate would be unfair is the opportunity for discovery before trial.”).

### 1. Origins: Where and What Is a Web Page

To a certain degree, to differentiate between an archived and contemporaneous Web page is misleading because, at the structural level, the two are identical. Therefore, to understand the makeup of an archived Web page, a description of its predecessor, the contemporaneous Web page, proves helpful.

In its basic file form, a contemporaneous Web page consists of a series of text and brackets; the pictures and formatting that ultimately appear on a computer user's screen are completely absent. This format is called hypertext markup language (HTML).<sup>93</sup> HTML is the programming language used to create documents for display over the Internet. The HTML codes used in a text document dictate the way in which words in the file will appear on-screen.<sup>94</sup> In addition to the text that will appear on-screen, the codes may include the URLs of other types of files such as sound, videos, and graphics. These files are stored elsewhere on the same server as the HTML file or in another, remote location.<sup>95</sup>

When a user requests to view a Web page, the server sends the HTML document over the Internet to the user's Internet Protocol (IP) address.<sup>96</sup> The server simultaneously instructs the sites containing the sound and other media files included in the HTML coding to send those files to the user's computer.<sup>97</sup> As the different parts of the Web page arrive at the user's computer, the computer will store them in a cache, a temporary storage for a collection of duplicate data.<sup>98</sup> The user's browser will then use the different elements that are stored temporarily in the cache to reassemble the Web page on-screen. To accomplish this task, the browser follows the dictates of the HTML codes in the main document, which instruct the browser from which storage place to retrieve, then where to place on-screen, the various components of the page, including text, graphics, and video.<sup>99</sup> Consequently, when viewed by an individual user, a Web page is in fact a collaboration between two computers, unlike a physical writing.

### 2. The Archival Process: Crawling the Web

The second step to the genesis of an archived Web page consists of digital archiving. Archiving on the Web occurs by way of a process called "crawling." Web crawlers systematically visit websites to extract and store

---

93. WHITE, *supra* note 22, at 370.

94. *Id.*

95. *Id.*

96. IP addresses serve as identifiers for specific computers or devices on a network. Users are able to view the Web pages they request because signals and messages get routed to the unique IP addresses that correspond to their computers. *Id.* at 312.

97. *Id.* at 370.

98. *Id.* at 371. Cached content on computer hard drives has been used in some cases as proof of possession of child pornography and other cybercrimes. *See, e.g., United States v. Tucker*, 150 F. Supp. 2d 1263, 1265 (D. Utah 2001).

99. WHITE, *supra* note 22, at 371.

their content.<sup>100</sup> Crawlers are automated software programs that can make thousands of requests at any given moment. Google is a prime example of the use of crawling technology.<sup>101</sup> Google uses this process to index millions of Web pages on the Internet, and, in turn, to facilitate user searching.<sup>102</sup> Before a person sits down at a computer to visit a website, Google's crawling software program, called "Googlebot," downloads the pages, extracts and indexes the text in the HTML files, then stores the content on Google's index servers.<sup>103</sup> Google uses the content stored on its servers in an algorithm that matches search queries with the appropriate websites based on the text in those HTML files.<sup>104</sup> Computer users see the result of this process when they browse through the results of their search queries and click through a link to their desired Web page.

In addition to indexing, crawlers can be used to duplicate Web pages.<sup>105</sup> In that case, rather than extracting key words from the HTML copy obtained by the crawl, the service displays in full the contents of the HTML copies as they are stored on the servers. Google makes such duplicate Web pages available to the public via its "cached" links.<sup>106</sup> The cached versions are the "snapshots" of the contemporaneous Web pages taken at the moment the Googlebot last crawled.<sup>107</sup> To view the version that is stored on Google's server, users may click on the "Cached" link that appears alongside their search results.<sup>108</sup> When users click on this link, they will view the most recent snapshot that was stored on the Google server instead of the contemporaneous version stored on the owner's server.<sup>109</sup> By permitting access to the cached version of Web pages, Google allows users to view Web page content when the server hosting the site goes down or when the site itself is inaccessible.<sup>110</sup>

The process by which the Internet Archive obtains its collection of archived Web pages resembles the process employed by Google. Unlike Google, however, which maintains cached files temporarily, the Internet Archive endeavors to maintain access to archived content indefinitely. Googlebot constantly replaces the cached pages stored on Google's servers

---

100. *Id.* at 374.

101. See Expert Report of Dr. John R. Levine para. 14, *Field v. Google*, 412 F. Supp. 2d 1106 (D. Nev. 2006) (No. CV-S-04-0413).

102. *Id.*

103. WHITE, *supra* note 22, at 374.

104. *Id.* at 375.

105. See, e.g., *Ticketmaster Corp. v. Tickets.com, Inc.*, No. 99-CV-7654, 2003 WL 21406289 (C.D. Cal. Mar. 7, 2003) (order granting defendant's summary judgment motion as to copyright and trespass to chattels claim) ("[T]he spider picks up all of the electronic symbols which, if it had been put on a monitor with the right software, would duplicate the TM [Ticketmaster] web page.").

106. Expert Report of Dr. John R. Levine, *supra* note 101, para. 15; see GoogleGuide, *Cached Pages*, [http://www.googleguide.com/cached\\_pages.html](http://www.googleguide.com/cached_pages.html) (last visited Aug. 12, 2009).

107. GoogleGuide, *Cached Pages*, *supra* note 106.

108. Expert Report of Dr. John R. Levine, *supra* note 101, para. 15.

109. GoogleGuide, *Cached Pages*, *supra* note 106.

110. Expert Report of Dr. John R. Levine, *supra* note 101, paras. 18–19.

once the Googlebot crawls the most recent version. Because of their temporary nature, Google's cached pages are not a particularly helpful means by which to create a record or perform extensive hindsight research. In contrast, the Internet Archive maintains its collection with precisely these goals in mind.

The Internet Archive obtains its collection of Web pages through a collaboration with Alexa Internet, Inc. (Alexa).<sup>111</sup> Alexa employs a crawler similar to Google.<sup>112</sup> Alexa's crawler methodically browses the World Wide Web, creating copies of all the pages visited and capturing about 1.6 terabytes of information per day.<sup>113</sup> Each of Alexa's snapshots of the World Wide Web, which take about two months to complete, contains about 4.5 billion pages from 16 million websites.<sup>114</sup> Alexa does not capture every page on the Web; instead, the crawler prioritizes Web pages based on the frequency with which users request them in searches.<sup>115</sup> Alexa then uses this information for commercial purposes, such as search engine indexing, but also "donates" the contents of these Web crawls to the Internet Archive.<sup>116</sup>

Once these copies have been made and donated to the Internet Archive, the Internet Archive stores and preserves this data.<sup>117</sup> The following section describes this process and conceptualizes the means by which these archived pages may then be accessed by computer users.

---

111. Internet Archive Frequently Asked Questions, *supra* note 20.

112. The legality of this archiving process has itself been questioned and upheld. *See* Healthcare Advocates, Inc. v. Harding, Early, Follmer & Frailey, 497 F. Supp. 2d 627 (E.D. Pa. 2007); Field v. Google, 412 F. Supp. 2d 1106 (D. Nev. 2006); Tom Zeller, Jr., *Keeper of Expired Web Pages Is Sued Because Archive Was Used in Another Suit*, N.Y. TIMES, July 13, 2005, at C9. For a discussion of whether copyright and contract laws proscribe Internet archiving, see James Grimmelman, *The Structure of Search Engine Law*, 93 IOWA L. REV. 1, 28–29 (2007); Rebecca Bolin, *Locking Down the Library: How Copyright, Contract, and Cybertrespass Block Internet Archiving*, 29 HASTINGS COMM. & ENT. L.J. 1 (2006). In its current form, website owners do not opt in to the archive, but rather may opt out by installing a robots.txt file or a "no-archive" metatag on their page, which will block future snapshots and render unavailable any previously archived copies. Internet Archive Frequently Asked Questions, *supra* note 20; *see* Field, 412 F. Supp. 2d at 1112 n.4 (describing use of "no-archive" metatag).

113. Howell, *supra* note 17, at 4.

114. Alexa Internet, Inc., Technology: How and Why We Crawl the Web, <http://www.alexa.com/site/company/technology> (last visited Aug. 12, 2009).

115. Howell, *supra* note 17, at 4.

116. *Id.*

117. Internet Archive Frequently Asked Questions, *supra* note 20. A great deal of the Internet Archive's work as an organization entails innovation in storage and preservation of massive quantities of data. For example, the Internet Archive is in the process of developing smaller machines on which to store massive quantities of content. For a photo, see Internet Archive, Petabox, <http://www.archive.org/web/petabox.php> (last visited Aug. 12, 2009).

### 3. Taking the Trip Back in Time: Finding Outdated Web Content with the Wayback Machine

The following section explains how digital copies exist on the Internet Archive's servers and describes the output that results from a search with the Wayback Machine. Part II goes on to summarize the instances in which courts have applied the Federal Rules of Evidence to authenticate this form of evidence.

Like cached pages available via Google,<sup>118</sup> archived pages exist independently of their original source. Once archived, the digital copies are stored on Internet Archive servers as opposed to their original host servers.<sup>119</sup> As a result, these copies remain unchanged even once webmasters update the contemporaneous versions stored on their own servers.

In addition to the text that appears on Web pages, the Wayback Machine includes many image files in the archive.<sup>120</sup> When users view an archived page, rather than retrieving the images from the original server that stored the images,<sup>121</sup> the Wayback Machine draws them from the Internet Archive server to which they were copied.<sup>122</sup> Likewise, links on archived websites do not generally link to current Web pages.<sup>123</sup> Instead, links will take users to the closest (by date) archived version of the linked page available that is also stored on the Internet Archive server.<sup>124</sup> However, if an image or linked page was not archived, then the image or link will direct to the original source, which may have since been removed or changed.<sup>125</sup>

Like contemporaneous websites, archived websites are accessible by way of a unique URL address. When a party prints out an archived Web page via the Wayback Machine, the printout will show the URL in the footer (in the same fashion as a URL is shown on printouts of contemporaneous Web pages).

The Internet Archive assigns a URL on its site to the archived files in the format `http://web.archive.org/Web/[Year in yyyy][Month in mm][Day in dd][Time code in hh:mm:ss]/[Archived URL]`. Thus, the Internet Archive URL

`http://web.archive.org/web/19970126045828/http://www.archive.org/` would be the URL for the record of the Internet Archive homepage HTML file (`http://www.archive.org/`) archived on January 26, 1997 at 4:58 a.m. and 28 seconds . . . . The date assigned by the Internet Archive

---

118. See *supra* notes 105–10 and accompanying text.

119. Internet Archive Frequently Asked Questions, *supra* note 20.

120. Greg R. Notess, *The Wayback Machine: The Web's Archive*, ONLINE, Mar.–Apr. 2002, at 59, <http://www.onlinemag.net/mar02/OnTheNet.htm>. Recall that image files and links consist of separate files embedded within the HTML code of the website. See *supra* Part I.B.1.

121. See *supra* Part I.B.1.

122. Notess, *supra* note 120.

123. *Id.*

124. *Id.*

125. *Id.*

applies to the HTML file but not to image files linked therein. Thus images that appear on the printed page may not have been archived on the same date as the HTML file. Likewise, if a website is designed with "frames," the date assigned by the Internet Archive applies to the frameset as a whole, and not the individual pages within each frame.<sup>126</sup>

The above demonstrates the similarities and the differences between contemporaneous and archived Web pages. With respect to the structure, archived and contemporaneous Web pages are identical to the extent they both consist of HTML code that is read and configured by Internet browsers. On the other hand, content is copied from a contemporaneous website to the archive server by way of an automated process controlled by a third party. The copy, in some instances, will be incomplete.<sup>127</sup> The following sections highlight the treatment that archived Web pages have received as courts grapple with these newly emerging issues.

## II. THE EMERGING AUTHENTICATION STANDARDS FOR EVIDENCE FROM THE WAYBACK MACHINE

Part I.A laid out the Federal Rules that govern authentication of Internet evidence and surveyed courts' application of the rules to authenticate contemporaneous Web page printouts. Part I.B conceptualized the process and output of digital archiving. Part II examines how courts have applied the Federal Rules to authenticate printouts from the Wayback Machine and traces the conflict underlying the approaches that courts have used.

In admitting into evidence a printout from the Wayback Machine in 2004, the U.S. District Court for the Northern District of Illinois concluded that an affidavit from an employee of the Internet Archive satisfied the Rule 901 threshold to authenticate the printout as a prior version of the opponent's website.<sup>128</sup> The U.S. District Court for the Middle District of Florida has followed a similar approach.<sup>129</sup> However, courts in the Second Circuit take a conflicting view. The courts in the Second Circuit concluded that testimony by an intermediary with personal knowledge of the archival process proved insufficient to authenticate the archived Web pages; the courts instead focused on the extremities of the transaction, namely, the original Web page and the digital copy as it appeared on the Internet

---

126. Plaintiff's Fifteenth Motion in Limine, exhibit A, *Telewizja Polska USA, Inc. v. Echostar Satellite Corp.*, No. 02-CV-3293 (N.D. Ill. Jul. 30, 2004) [hereinafter *Davis Affidavit*].

127. Internet Archive Frequently Asked Questions, *supra* note 20. For example, "when a dynamic page renders standard html, the archive works beautifully. When a dynamic page contains forms, JavaScript, or other elements that require interaction with the originating host, the archive will not contain the original site's functionality." *Id.*

128. *Telewizja Polska USA, Inc. v. Echostar Satellite Corp.*, No. 02-3293, 2004 WL 2367740, at \*6 (N.D. Ill. Oct. 15, 2004) (order denying motion in limine to bar Internet Archive evidence).

129. *See St. Luke's Cataract & Laser Inst., P.A. v. Sanderson, M.D., LLC*, No. 06-CV-223, 2006 WL 1320242 (M.D. Fla. May 12, 2006) (order denying admission of printouts from the Internet Archive).

Archive's servers.<sup>130</sup> Under this view, only a party with personal knowledge of the content of the original website, such as a webmaster or content provider, would be sufficient to authenticate the Web page's digital counterpart stored in the Internet Archive. The following sections examine these competing opinions. Part II.A discusses the approach advanced by the proponent in *Telewizja Polska USA, Inc. v. Echostar Satellite Corp.*<sup>131</sup> that was adopted by the magistrate judge in the District Court for the Northern District of Illinois. Part II.B discusses *Novak v. Tucows*<sup>132</sup> and the current view in the Second Circuit.

### A. *Trusting the Intermediary To Establish the Rule 901 Foundation*

#### 1. *Telewizja Polska USA, Inc. v. Echostar Satellite Corp.*

A case in the District Court for the Northern District of Illinois garnered attention in Internet law circles in 2004 when a magistrate judge admitted into evidence a printout from the Wayback Machine.<sup>133</sup> In *Telewizja*, a Polish television channel alleged that Echostar improperly had used *Telewizja Polska*'s trademark to promote its Dish Network satellite TV service after the contractual marketing rights to do so had expired.<sup>134</sup> In response, Echostar sought to introduce printouts of *Telewizja Polska*'s website obtained using the Wayback Machine to prove that *Telewizja Polska* similarly had touted its connection to the Dish Network after their agreement had expired.<sup>135</sup> *Polska* moved to exclude the evidence on hearsay and authentication grounds.<sup>136</sup>

Echostar argued that the foundational predicate required by Rule 901(a) had been satisfied by the affidavit of the Administrative Director of the Internet Archive, Molly Davis.<sup>137</sup> In the affidavit, Davis attested,

The Internet Archive receives data from third parties who compile the data by using software programs known as crawlers that surf the Web and automatically store copies of Web site files at certain points in time as they existed at that point in time. This data is donated to the Internet Archive, which preserves and provides access to it.<sup>138</sup>

---

130. See *Chamilia LLC v. Pandora Jewelry LLC*, 85 U.S.P.Q.2d 1169 (S.D.N.Y. 2007); *Novak v. Tucows Inc.*, No. 06-CV-1909, 2007 U.S. Dist. LEXIS 21269 (E.D.N.Y. Mar. 26, 2007) (order granting motion to bar evidence), *aff'd*, No. 07-2211-cv, 2009 U.S. App. LEXIS 9786 (2d Cir. May 6, 2009); *St. Luke's*, 2006 WL 1320242 (order denying admission of printouts from the Internet Archive).

131. 2004 WL 2367740.

132. 2007 U.S. Dist. LEXIS 21269.

133. See Stanford Law School Center for Internet and Society, Lauren Gelman's Blog, <http://cyberlaw.stanford.edu/packets002728.shtml> (Nov. 17, 2004, 12:09 PST).

134. *Telewizja*, 2004 WL 2367740, at \*5.

135. *Id.*

136. *Id.*

137. See Defendant's Response and Opposition to Plaintiff's Fifteenth Motion in Limine at 3, *Telewizja*, No. 02-CV-3293 (N.D. Ill. Sept. 20, 2004).

138. Davis Affidavit, *supra* note 126, at 1.



Davis went on to state that the proffered exhibits were “true and accurate copies of printouts of the Internet Archive’s records of the HTML files archived from the URLs and the dates specified in the footer of the printout.”<sup>139</sup>

After concluding that the website in question fell outside the ambit of the hearsay rule,<sup>140</sup> the court considered the issue of authentication. The court concluded that the threshold showing required by Rule 901 had been satisfied by the Davis Affidavit, reasoning,

Federal Rule of Evidence 901 “requires only a prima facie showing of genuineness and leaves it to the jury to decide the true authenticity and probative value of the evidence.” Admittedly, the Internet Archive does not fit neatly into any of the non-exhaustive examples listed in Rule 901; the Internet Archive is a relatively new source for archiving websites. Nevertheless, Plaintiff has presented no evidence that the Internet Archive is unreliable or biased. And Plaintiff has neither denied that the exhibit represents the contents of its website on the dates in question, nor come forward with its own evidence challenging the veracity of the exhibit. Under these circumstances, the Court is of the opinion that [the affidavit from the representative of the Internet Archive Company] is sufficient to satisfy Rule 901’s threshold requirement for admissibility.<sup>141</sup>

## 2. *St. Luke’s Cataract & Laser Institute, P.A. v. Sanderson, M.D., LLC*

Since *Telewizja*, a number of opinions and orders have endorsed the same approach. A year and a half after *Telewizja*, the U.S. District Court for the Middle District of Florida addressed the issue of authenticating archived Internet evidence in *St. Luke’s Cataract & Laser Institute P.A. v. Sanderson, M.D., LLC*.<sup>142</sup> In *St. Luke’s*, the plaintiff eye care center sued two former employees: a physician and a webmaster. St. Luke’s alleged that upon their departure, the former employees stole the copyrighted contents of an Internet website and two domain names that St. Luke’s had funded and developed, then used the domain names and website content to promote the physician’s new competing practice.<sup>143</sup> St. Luke’s alleged

---

139. *Id.*

140. *Telewizja*, 2004 WL 2367740, at \*5. The court reasoned that the images and text were introduced to show the images and text found on the website and, therefore, were not statements at all and outside the scope of the hearsay rule. The court further noted that even if they were statements introduced for the truth, “the contents of [the plaintiff]’s website may be considered an admission of a party-opponent, and are [therefore] not barred by the hearsay rule.” *Id.*

141. *Id.* at \*6 (quoting *United States v. Harvey*, 117 F.3d 1044, 1049 (7th Cir. 1997)).

142. No. 06-CV-223, 2006 WL 1320242 (M.D. Fla. May 12, 2006) (order denying admission of evidence from the Internet Archive). This case went on to trial and a verdict, and the parties subsequently appealed. See *St. Luke’s Cataract & Laser Inst., P.A. v. Sanderson, M.D., LLC*, No. 08-11848, 2009 WL 1955609 (11th Cir. July 9, 2009) (affirming in part and reversing in part judgment entered following jury trial).

143. Joint Pretrial Statement at 2, *St. Luke’s*, No. 8:06-CV-223 (M.D. Fla. Apr. 17, 2007), 2007 WL 5117992. The disputed website had initially been located at the [www.laserspecialist.com](http://www.laserspecialist.com) domain address. *Id.*

copyright infringement, cyberpiracy, trademark infringement, conversion, unfair competition, and violation of the Florida Unfair Trade and Deceptive Practices Act.<sup>144</sup>

To prove these allegations, St. Luke's moved to admit Wayback Machine printouts purporting to show the appearance of the pages at various times between 2000 and 2004.<sup>145</sup> To authenticate the printouts, the plaintiff attached a certified copy of the Davis Affidavit used in *Telewizja*, along with declarations from Benjamin Fertic and Bradley Houser, two employees of St. Luke's who had performed the Wayback Machine searches.<sup>146</sup>

In his declaration, Fertic stated that the attached printout from the Wayback Machine was "a true and correct cached image of the relevant pages of the LaserSpecialist.com Site" on the dates in question.<sup>147</sup> Houser stated that the various Wayback Machine printouts he included were "true and correct" copies of the defendants' website as it appeared on dates between 2000 and 2004.<sup>148</sup> Houser also stated that St. Luke's did not discover the infringing activity until 2006, when it attempted to update the

---

144. *Id.*

145. *St. Luke's*, 2006 WL 1320242, at \*1; Plaintiff's Motion for Admission of Evidence and Memorandum of Law in Support Thereof at 1-2, *St. Luke's*, No. 8:06-CV-223 (M.D. Fla. May 12, 2006). St. Luke's sought to admit the Wayback Machine printouts in lieu of originals because they no longer had copies of the Web page on their servers. *Id.* at 1. St. Luke's had used Internet Archive printouts previously in January 2006 when filing for copyright protection of the 2003 version of the website. *St. Luke's*, No. 08-11848, 2009 WL 1955609, at \*5 (11th Cir. July 9, 2009); see 17 U.S.C. § 408(b) (2006) (requiring a copyright registrant to deposit two copies of the best version of a published work with the Copyright Office). Notably, the Board of Appeal of the European Patent Office has weighed in on the admission of evidence obtained from the Internet Archive used to establish the related question of the state of the art upon an application for patent protection. See Decision of European Patent Office, Boards of Appeal 3.2.04, Case T-1134/06 (Jan. 16, 2007), available at <http://legal.european-patent-office.org/dg3/pdf/t061134eu1.pdf>. In so doing, the Board of Appeal examined closely the process of Internet archiving and the limited European case law on the subject, concluding,

Where a disclosure has been retrieved from a resource such as the Internet Archive, further evidence concerning the history of the disclosure, whether and how it has been modified since the date it originally appeared on a web site will be necessary. This could be in the form of an authoritative statement from the archivist. Alternatively, an appropriate statement as to the content, either from the owner or author of the archived web site which included the disclosure may suffice.

*Id.*

146. *St. Luke's*, 2006 WL 1320242, at \*1.

147. Declaration of Benjamin C. Fertic in Support of Plaintiff's Motion for Preliminary Injunction and Expedited Discovery at 2, *St. Luke's*, No. 8:06-CV-223 (M.D. Fla. Feb. 17, 2006).

148. Declaration of J. Bradley Houser in Support of Plaintiff's Motion for Preliminary Injunction and Expedited Discovery at 3, 7, *St. Luke's*, No. 8:06-CV-223 (M.D. Fla. Feb. 17, 2006); see also *id.* exhibits B, G, H.

website.<sup>149</sup> It was at that point that St. Luke's first reviewed the archived records.<sup>150</sup>

The defense opposed the admission of the printouts on the grounds that they were not properly authenticated. To this end, the defense classified the Wayback Machine printouts as highly technical evidence and, as such, within the purview of Federal Rule of Evidence 702.<sup>151</sup> They argued that the declarations of Fertic and Houser provided insufficient foundation because the technical nature of the printouts required an explanation by an expert, as opposed to fact witnesses, as to how the Web pages were collected by the Internet Archive. The defense further argued that the declarations should be excluded pursuant to the judge's gatekeeping function set out in *Daubert v. Merrell Dow Pharmaceutical, Inc.*<sup>152</sup> because Fertic and Houser were not qualified to explain the process by which the software copied and stored the Web page files.<sup>153</sup>

The court did not adopt the defense argument but, nevertheless, denied admission of the evidence. Ruling on the motion, the court noted that in Florida "Web-sites are not self-authenticating," and, therefore, to authenticate a printout from a website the proffering party would need to provide a "statement or affidavit from someone with knowledge [of the website]," such as a webmaster or owner.<sup>154</sup> It then concluded that the affidavits of Fertic and Houser were insufficient to satisfy the Rule 901 requirement because neither had "*personal knowledge* of the contents of the Internet Archive website."<sup>155</sup> The court found that the Davis Affidavit also failed to establish the foundational predicate because the affidavit was from prior litigation. However, then-Magistrate Judge Mary S. Scriven aligned with the holding in *Telewizja* and proposed that "an affidavit by Ms. Davis, or some other representative of Internet Archive with personal knowledge of its contents, verifying that the printouts Plaintiff seeks to admit are true

---

149. *Id.* at 7.

150. *Id.* St. Luke's predicament illustrates the potential value of archived evidence because parties most likely will search archives only once contemporaneous sites are no longer available. *See supra* note 71 and accompanying text (discussing the distinction between illustrative evidence and evidence offered for its content and the ensuing best evidence implications).

151. *See* FED. R. EVID. 702 (permitting testimony of experts with specialized technical knowledge only "if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case").

152. 509 U.S. 579 (1993).

153. Defendants' Memorandum Opposing Admission of Evidence from Archive.org and the Wayback Machine at 1-2, 6, *St. Luke's*, No. 8:06-CV-223 (M.D. Fla. May 5, 2006); *see Daubert*, 509 U.S. at 589 ("[T]he trial judge must ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable.").

154. *St. Luke's*, 2006 WL 1320242, at \*2 (quoting *In re Homestore.com, Inc. Sec. Litig.*, 347 F. Supp. 2d 769, 782 (C.D. Cal. 2004)) (alteration in original) (internal quotation marks omitted).

155. *Id.*

and accurate copies of Internet Archive's records" would be sufficient to satisfy the Rule 901 foundation requirement.<sup>156</sup>

### B. *Skeptics of Evidence from the Wayback Machine*

*Telewizja* appeared to open the door to the use of printouts from the Internet Archive.<sup>157</sup> The Internet Archive now maintains a page on its website specifically addressed to the legal community.<sup>158</sup> In apparent conformity with the outcomes in *Telewizja* and *St. Luke's*, this page provides information on the Wayback Machine and the process by which lawyers can obtain affidavits for trial. The sample affidavit available on this website resembles closely the affidavit that was used in *Telewizja* and discussed in Part II.A.1.<sup>159</sup>

However, two district courts in the Second Circuit have espoused a different standard to authenticate pages obtained with the Wayback Machine and the U.S. Court of Appeals for the Second Circuit has affirmed this approach. These courts focus not on the authenticity of a proffered printout vis-à-vis the Internet Archive but, rather, vis-à-vis the original publisher of the Web page. Judge Joseph F. Bianco of the U.S. District Court for the Eastern District of New York first broached the issue in *Novak v. Tucows, Inc.*<sup>160</sup>

In *Novak*, the plaintiff was a purveyor of pet supplies and livestock. He sold his goods via the trademarked website *petswarehouse.com*, which was, at one time, the fourth most-visited pet supply website in the United States.<sup>161</sup> In February 2003 an individual named John Benn obtained a default judgment against Novak in Alabama (Novak was a New York resident).<sup>162</sup> Novak then opted to transfer his domain name from a Maryland-based registrar to Nitin, a registration company based in New York. However, unbeknownst to Novak, his domain name instead was transferred erroneously to Tucows, a Canadian registration company.<sup>163</sup>

---

156. *Id.* In so holding, then-Magistrate Judge Mary S. Scriven aligned with the ruling in *Telewizja* by concluding that the personal knowledge at issue in the Rule 901 inquiry belonged to the party responsible for maintaining the digital archive. *Accord Attig v. DRG, Inc.*, No. Civ.A.04-CV-3740, 2005 WL 730681, at \*2 n.1 (E.D. Pa. Mar. 30, 2005) (unreported order in which parties to a copyright dispute involving online content agreed that the copies of the websites obtained via the Internet Archive were admissible evidence in light of *Telewizja*). This approach treats the archived Web pages as attributable to the webmasters of the Internet Archive, rather than as digital copies ultimately attributable to the owners of the originating website. Part III will criticize the suitability of this approach.

157. See Stanford Law School Center for Internet and Society, Lauren Gelman's Blog, *supra* note 133.

158. See Internet Archive, Legal, Frequently Asked Questions, <http://www.archive.org/legal/faq.php> (last visited Aug. 12, 2009).

159. See Internet Archive, Standard Affidavit, <http://www.archive.org/legal/affidavit.php> (last visited Aug. 12, 2009).

160. No. 06-CV-1909, 2007 U.S. Dist. LEXIS 21269 (E.D.N.Y. Mar. 26, 2007).

161. *Id.* at \*1.

162. *Id.* at \*5.

163. *Id.* at \*4.

Benn applied for a writ of execution against the domain name petswarehouse.com to enforce the default judgment.<sup>164</sup> The ensuing writ of execution required the domain name host to turn over the domain name to the sheriff's department to be auctioned.<sup>165</sup> Upon receiving notice of this writ from the Alabama court, the Canadian registrar suspended access to the server and turned over ownership of the domain name.<sup>166</sup> Benn then purchased the domain name at the public auction.<sup>167</sup> The Alabama trial court subsequently reversed Novak's default judgment and the writ of execution, leading Benn to transfer the domain name back to Novak.<sup>168</sup> Novak then sued the domain name registrars involved in the transfers (Tucows, Inc. and Nitin Networks, Inc.) for trademark dilution, unfair competition, and cyberpiracy, alleging that the registrars played an illegal role in the destruction of his business.<sup>169</sup> As proof, Novak sought to admit several printouts obtained by way of the Wayback Machine.<sup>170</sup> To fulfill the foundation requirement, Novak included his own declaration attesting to the authenticity of the pages printed from the Wayback Machine.<sup>171</sup>

The defendants moved to strike the exhibits as inadmissible hearsay.<sup>172</sup> The court concluded that the exhibits were, in fact, hearsay,<sup>173</sup> but went on to expound on the Rule 901 authentication inquiry. Judge Bianco concluded that Novak could not authenticate the printouts by his testimony because he lacked the personal knowledge to prove that the content of the printouts obtained from the Internet Archive website was what he asserted it to be, namely, the content that was posted on the original website.<sup>174</sup> The court concluded that it could not properly authenticate the Internet Archive printouts absent testimony or a sworn statement by an employee of the company that hosted the original version of the site because

the information posted on the Wayback Machine is only as valid as the third-party donating the page decides to make it—the authorized owners and managers of the archived websites play no role in ensuring that the

---

164. *Id.*

165. *Id.*

166. *Id.*

167. *Id.*

168. *Id.* at \*5.

169. *Id.* at \*5–6. Domain name registrars have the power to “delete” domain names at any given time. For a discussion of the impact of such private powers on democratic governance and law, see Michael Froomkin, *Wrong Turn in Cyberspace: Using ICANN to Route Around the APA and the Constitution*, 50 DUKE L.J. 17 (2000). See, e.g., *Hawes v. Network Solutions, Inc.*, 337 F.3d 377 (4th Cir. 2003); *GlobalSantaFe Corp. v. Globalsantafe.com*, 250 F. Supp. 2d 610 (E.D. Va. 2003).

170. *Novak*, 2007 U.S. Dist. LEXIS 21269, at \*5. The proffered exhibits included several printouts from Nitin's website and an online summary of Novak's past and present lawsuits. *Id.*

171. *Id.* at \*15.

172. *Id.* at \*14.

173. The court concluded that the proffered printouts were not statements made by declarants testifying at trial and therefore constituted inadmissible hearsay. *Id.* at \*15–16.

174. *Id.* at \*17.

material posted in the Wayback Machine accurately represents what was posted on their official websites at the relevant time.<sup>175</sup>

Soon after *Novak*, the U.S. District Court for the Southern District of New York likewise found foundation lacking for Internet Archive printouts in *Chamilia, LLC v. Pandora Jewelry, LLC*.<sup>176</sup> There, the plaintiff in a false advertising lawsuit sought to introduce a series of archived Web pages procured through the Wayback Machine to prove that the defendant engaged in false advertising.<sup>177</sup> The court determined that the printouts failed to prove anything and so denied the motion to strike as moot.<sup>178</sup> However, the court also cited *Novak* with approval in concluding that the printouts suffered fatal authentication problems.<sup>179</sup>

The Court of Appeals for the Second Circuit has since affirmed the outcome of *Novak*.<sup>180</sup> The panel concluded in an unpublished opinion that “the District Court did not err, much less abuse its discretion” in denying admission of the exhibits.<sup>181</sup>

While the holding in *Telewizja* suggests that the Internet Archive and services like it stand to play a significant role in trials, *Novak*, *Chamilia*, and the recent endorsement of their outcome at the appellate level bring this potential into question. Little has been written addressing the different viewpoints with which these courts have approached the authentication of printouts obtained using the Wayback Machine. However, “[t]he consistency of these cases begs the question of the circumstances under which Internet evidence, particularly printouts from the Internet archive, would be admissible.”<sup>182</sup> Beryl A. Howell, a former New York federal prosecutor and General Counsel of the U.S. Senate Committee on the Judiciary, suggested shortly after *Telewizja* that proponents could authenticate archived websites by “producing the testimony, either orally or in written form, of the person who copied or supervised the copying of the archived website and the process followed to accomplish this task. In addition, the proponent must establish the general reliability of the

---

175. *Id.* at \*17–18.

176. 85 U.S.P.Q.2d 1169 (S.D.N.Y. 2007).

177. *Id.* at 1175 n.4.

178. *Id.* The court found the plaintiff’s legal argument unconvincing. In dispute was whether the defendant falsely claimed to have a patent. The plaintiff asserted that even if the defendant had claimed to have a “patent pending,” the infirmity of defendant’s patent application rendered that statement untrue. *See id.* The Wayback Machine printouts purportedly proved the infirmity of the patent application by proving that the defendant advertised the product prematurely in contravention of patent application guidelines. *Id.*

179. *Id.*; *see also* Audi AG v. Shokan Coachworks, Inc., No. 1:07-CV-00173, 2008 WL 4911730, at \*23 (N.D.N.Y. Nov. 13, 2008) (citing *Chamilia*, *Novak*, *St. Luke’s*, and *Telewizja* for the proposition that Internet Archive evidence “may only be authenticated by a knowledgeable employee of the website,” but failing to clarify which website the employee must have knowledge of—the archived version or contemporaneous version).

180. *Novak v. Tucows, Inc.*, No. 07-2211-cv, 2009 U.S. App. LEXIS 9786, at \*6 (2d Cir. May 6, 2009).

181. *Id.*

182. Johnson & Francis, *supra* note 27, at 5.

copy.”<sup>183</sup> To accomplish this end and resolve the current conflict, Part III of this Note urges courts to approach admission of archived Web pages as implicating best evidence and, accordingly, to address the reliability of the crawling process apart from and in addition to the Rule 901 authentication inquiry.

### III. ARCHIVED WEB PAGES SHOULD BE EVALUATED AS AN ISSUE OF BEST EVIDENCE

Part II of this Note explored the current conflict in the case law as it pertains to authentication of archived Web pages. While *Novak*, *Telewizja*, and the cases that followed reached different conclusions about the foundation necessary to authenticate archived Web pages, they all rested on the same premise: courts should evaluate whether or not a printout of an archived Web page accurately represents the contents of its contemporaneous counterpart as part of the Rule 901 authentication inquiry. *Telewizja* concluded that an affidavit by an Internet Archive employee would establish a sufficient foundation to authenticate the printouts,<sup>184</sup> while *Novak* concluded that only testimony or an affidavit from the original content provider would prove sufficient.<sup>185</sup> Part III moves away from the rationales of these holdings and asserts that the question of whether or not an archived Web page may stand in properly for a contemporaneous version should not get bundled into the Rule 901 foundation inquiry; instead, it is best approached as an issue of best evidence. Under this approach, the authentication of archived Internet content would consist of two separate questions. First, the judge would evaluate whether the court may admit the archived Web page fairly in lieu of the original page as an issue of best evidence pursuant to Rules 104(a), 1001(4), and 1003.<sup>186</sup> Second, the proponent would need to authenticate the archived Web page as exactly what it purports to be, namely, a contemporaneous Web page stored on an Internet Archive server.

#### A. Channeling Xerox: Copies on the Web

This section concludes that the best evidence rule provides the most suitable framework in which to consider the admission of printouts from the Wayback Machine because archived Web pages constitute digital copies of the HTML code that content providers publish. Specifically, this section asserts that, because caching is a form of copying, courts should evaluate Web pages that are stored by the Internet Archive in the same fashion as other types of duplicates in addition to requiring their authentication. The subsequent sections lay out recommendations to guide this two-part inquiry.

---

183. Howell, *supra* note 17, at 8.

184. See *supra* note 141 and accompanying text.

185. See *supra* notes 174–75 and accompanying text.

186. See *supra* note 87 and accompanying text.

Printouts obtained via the Wayback Machine can be conceptualized in relation to two separate loci on the Internet: first, in relation to the original content provider's website; second, in relation to the Internet Archive server on which the archived copy is stored. This section urges courts to treat the proffer as implicating a best evidence issue with respect to the page stored by the Internet Archive vis-à-vis the page as originally published by the content provider.

The role of HTML code in digital archiving resembles the role of negatives in photography and photographic reproduction. The content that gets displayed when users view contemporaneous Web pages derives from the interaction between the HTML code and users' browsers.<sup>187</sup> The code and the browser act in concert; the browser caches the content components, and the code instructs the browser how to reassemble the page on the screen.<sup>188</sup> Digital crawlers such as Alexa and Googlebot use the same caching technology involved in this process. In this respect, every Web page, whether contemporaneous or archived, is a copy reproduced according to the template of the original HTML file.<sup>189</sup> Therefore, the viewable archived Web page that ensues from the reproduction of the HTML code likewise is "a counterpart produced by the same impression as the original, or from the same matrix."<sup>190</sup>

Admittedly, archived Web pages are imperfect reproductions.<sup>191</sup> While the date listed in the URL of the archived page corresponds to the date on which the original Web page was archived, it will not necessarily correspond to images and links, which in turn may never be archived.<sup>192</sup> However, to the extent that users may not be able to view links and images, archived Web pages are merely incomplete; this aspect does not render the pages outside the ambit of the duplicate definition. Instead, the archived

---

187. See *supra* notes 98–99 and accompanying text.

188. See *supra* notes 98–99 and accompanying text.

189. Conceptually, the repeated copying of the HTML file resembles the process of producing prints from a single negative. For purposes of best evidence, however, prints from negatives generally are regarded as originals. See *supra* note 79. In the case of archived Web pages, practicality may warrant distinguishing between Web pages viewed contemporaneously as originals and archived Web pages as duplicates. See *supra* note 83 and accompanying text. Regardless of the distinction that courts choose to draw, however, the principle would remain the same.

190. *Supra* note 75 and accompanying text. Taken to its logical conclusion, treating caching as a method of duplication points to an alternative approach for the authentication of contemporaneous Web pages generally. Under this approach, the HTML code promulgated by a content provider would constitute the original. To the extent that individual computers access this code, cache the components of the website, and reassemble them on a user's screen pursuant to the code, printouts of the contemporaneous Web pages would then amount to duplicates. Cf. *supra* note 69 and accompanying text. Therefore, a court could theoretically treat the issue of whether the printout reflects the contents of the Web page as a matter of best evidence for the judge under Rules 104(a), 1001(4), and 1003. Ultimately, whether or not the content could in fact be attributed to the content provider would still be a question for the jury.

191. See *supra* notes 120–26 and accompanying text.

192. See *supra* notes 122–25 and accompanying text.



Web page is akin to a recording that duplicates sound but not sight.<sup>193</sup> Because the HTML file that gets copied during the crawl includes the text that appears on-screen, the text in the ensuing digital copy falls within the definition of a duplicate under Rule 1001(4) because it “accurately reproduces the original.”<sup>194</sup> In many cases, the text will be the content at issue in trial, and the images and links on a page will be irrelevant. Where the images themselves are at issue, whether or not an archived page qualifies as a duplicate will likely vary depending on the circumstances and the dates the page purports to represent. Where the image files are not stored by the Internet Archive or are sourced to a since-expired link, a page offered to prove the content of the images will not qualify. If anything, however, digital copies are less prone to the sorts of errors that arise during the duplication process precisely because the copies are produced by way of an automated process.<sup>195</sup> The prevalence of caching technology in search engines and consumers’ everyday reliance on this process further substantiate the reliability of archived copies.<sup>196</sup>

Consequently, the best evidence rule and its allowance for the admission of duplicates provide a superior framework in which to address concerns that may arise with respect to the uniformity between the original Web page and the archived copy. In conjunction with already established standards for the authentication of contemporaneous Web pages, best evidence stands to resolve concerns about the fairness of admitting archived Web pages when their original counterparts are unavailable.<sup>197</sup> The following sections of this Note suggest a two-part inquiry whereby courts would first decide using evidence sufficient for the purpose—but not necessarily admissible at trial—whether the archived page qualifies as a duplicate, then would apply already developed authentication standards to decide whether a reasonable jury could find, based only on admissible evidence, that the proffered printout accurately represents the page stored on the Internet Archive server.

#### *B. Step One: Whether the Archived Web Page Qualifies as a Duplicate*

This section recommends that courts resolve the question of whether the archived Web page accurately reflects the content of the original page as an issue for the judge pursuant to Rule 104(a).<sup>198</sup> Under this approach, the judge would evaluate as a preliminary question whether the archived page

---

193. See *supra* note 82 and accompanying text.

194. *Supra* note 75 and accompanying text.

195. See *supra* note 80 and accompanying text.

196. Of course, this is not to say that that digital copies are infallible simply because their creation is automated. Programmers make code errors, and opponents should have the opportunity to question the virtue of the copying process. However, inaccuracies are the exception to the norm and best resolved in the same fashion as analogous errors that occur during the photocopying process. See *supra* notes 39, 50.

197. See *supra* notes 70–71 and accompanying text.

198. See *supra* Part I.A.3.

in fact qualifies as a duplicate under Rule 1001(4).<sup>199</sup> In this framework, the opponent would bear the burden of introducing specific facts or circumstances to raise a reasonable question regarding the authenticity of the original Web page or, alternatively, of the archived version itself.<sup>200</sup> Treatment under Rule 104(a) would permit courts to resolve this question by use of affidavits without hearsay problems.<sup>201</sup> As archived Internet evidence becomes more commonplace, proponents should no longer always have to enlist affidavits absent legitimate concerns regarding, for example, third-party interference, hacking, or infirmities with the archived version in and of itself.

Under this approach, if a website is altered or removed prior to the discovery phase of trial, fairness might weigh in favor of permitting admission of the archived version, given the absence of an opportunity for discovery.<sup>202</sup> Questions pertaining to completeness, passage of time, or the potential of intermeddling with the third party server would be left to the jury to consider in its role as fact finder when it resolves issues of the weight and credibility to afford the printout.<sup>203</sup> In addition, if an archived copy does not qualify as a duplicate under Rule 1003, it may still qualify for admission as secondary evidence of the website's contents.<sup>204</sup>

Absent a credible challenge to the quality of the digital copy, a proponent would next need to establish a foundation to admit the printout under Rule 901. The next section of this Note advocates authentication of printouts from the Internet Archive in the same fashion as contemporaneous websites.

*C. Step Two: Whether the Proffered Printout Shows What Actually Appeared On-screen at the Internet Archive Website*

To satisfy the Rule 901 threshold, a proponent needs to provide evidence sufficient to support a finding that the proffered printout in fact shows the archived version of the Web page as it appeared on a given date at the Internet Archive URL listed on the printout.<sup>205</sup> This section urges courts to allow proponents to establish this foundation in the same manner as contemporaneous Web pages because, with respect to the showing required

---

199. See *supra* Part I.A.3.

200. See 5 SALTZBURG ET AL., *supra* note 29, § 1003.02.

201. See *supra* notes 84–88 and accompanying text. Even so, the affidavit used in *Telewizja* would not resolve this threshold question because it did not attest to the congruity between the archived page and the original. See *supra* note 139 and accompanying text. Instead, if the court requires an affidavit, it should come from a representative of the third-party crawler that performed the copying, such as Alexa, Inc. This approach comports with that recommended by the Board of Appeal of the European Patent Office. See *supra* note 145.

202. See 5 SALTZBURG ET AL., *supra* note 29, § 1003.02. For a discussion of a closely related issue, the duty to preserve, and its application to ephemeral data see SCHEINDLIN ET AL., *supra* note 62, at 129–36.

203. See *supra* note 78 and accompanying text.

204. See *supra* note 77 and accompanying text.

205. See *supra* notes 53–55 and accompanying text.

by Rule 901, any distinction between archived and contemporaneous sites proves largely artificial. Like contemporaneous Web pages, archived Web pages are accessible by way of a unique URL address.<sup>206</sup> A printout of either a contemporaneous or archived Web page would include that unique URL and the date on which the Web page was printed.<sup>207</sup>

Consequently, any court that has authenticated contemporaneous Web pages based on the circumstantial indicia of a dated printout could apply the same standard to authenticate archived Web pages.<sup>208</sup> Likewise, a court that generally requires testimony by the content provider or webmaster could require an affidavit or testimony from an Internet Archive representative.<sup>209</sup> For example, like those courts that require testimony by a party with personal knowledge of a contemporaneous website, when called upon to authenticate Internet Archive evidence the *Telewizja* court called for testimony by a party with personal knowledge of the website from which the proffered printout was retrieved.<sup>210</sup> Therefore, this approach would comport with the increasingly well-established standards that have already emerged for authenticating Internet evidence generally.

#### CONCLUSION

The skepticism that courts have displayed with respect to archived Web content appears to result from a combination of reluctance to attribute archived content to a content provider and concerns about accuracy and the potential for intermeddling.<sup>211</sup> However, treated as a question about the quality of duplication, these concerns would be addressed as a threshold matter by the judge. It follows that printouts from the Wayback Machine could then be authenticated under Rule 901 in a manner commensurate with

---

206. See *supra* Part I.B.3.

207. See *supra* Part I.B.3.

208. See *supra* notes 58–59 and accompanying text.

209. See *supra* notes 61–62 and accompanying text.

210. This approach does raise hearsay concerns. Where an affidavit would be inadmissible, testimony from an Internet Archive employee would be necessary—but perhaps impracticable. While this Note does not purport to recommend a uniform standard for the authentication of Web pages generally, it would seem that circumstantial indicia of authenticity should constitute sufficient foundation to verify that a printout from the Wayback Machine shows what was actually on the Internet Archive server. *Cf. supra* note 46. Given the absence of control that Internet Archive employees exert over the content of archived Web pages and the automated nature of archiving, the issue of whether content may be attributed properly to the website owner is not as substantial for the Internet Archive as it is in the case of contemporaneous websites, and therefore testimony from Internet Archive employees lacks the same importance.

211. Under the framework of the current Federal Rules of Evidence, the best evidence rule presumably should confront concerns about whether the archived version may stand in for the original. However, the varying treatment afforded archived Internet evidence falls within the ambit of the broader ongoing debate about the adequacy of the best evidence framework as applied to digital evidence generally. See *supra* note 74. While this Note does not weigh in on the broader discussion, archived Internet evidence serves as a useful case study that tests the boundaries of the current standard. The fact that the best evidence rule may appear to some an unsatisfactory response to the authentication challenges posed by archived Internet evidence is perhaps indicative of the limits of the status quo.

contemporaneous websites. When conceptualized properly, archived Internet evidence fits squarely within this framework. By including this additional step, reliable evidence from the Wayback Machine stands to become as easily admitted as any other Internet-derived proof. Together with the standards heretofore employed to authenticate Internet evidence, the approach advocated by this Note would enable archived Internet evidence to play a useful and substantial role at trial.

## *Notes & Observations*