# DOES ROCKY & BULLWINKLE HOLD THE KEY TO UNLOCKING THE MYSTERY OF FAIR USE IN THE AGE OF INTERNET ARCHIVING?

KELLY JINES-STOREY\*

Take not from others to such an extent and in such a manner that you would be resentful if they so took from you.

—Joseph A. McDonald on non-infringing uses<sup>1</sup>

#### Introduction

"Sherman and I entered the WABAC and were immediately wisked back through time," said Mr. Peabody.<sup>2</sup> Although that scenario was presented by a genius, talking dog and his pet boy in the beloved Rocky & Bullwinkle cartoon, a type of time travel is exactly the goal of the Wayback Machine, the Internet Archive's database of web pages that date back to 1996.<sup>3</sup> With the Wayback Machine, its name homage to the WABAC of cartoon fame, digital history and culture can be explored along a space-time continuum.<sup>4</sup> This is possible because the Internet Archive views the web as a cultural artifact and has declared as its mission the preservation of the web's digital history.<sup>5</sup>

With this twenty-first-century luxury of viewing the digital past, as it once existed, comes the uncharted legal territory of the digital age. Even

Copyright © 2007, Kelly Jines-Storey

\* J.D. 2007, Capital University; B.A. 1997, Loyola University Chicago. I thank attorney and professor Christen Millard for sparking my interest in this area of law and for the clarity in guidance, attorney and friend Kim Finley for her brilliant writing expertise, and attorney Melanie Martin Jones for providing the basis for this comment.

<sup>&</sup>lt;sup>1</sup> Joseph A. McDonald, Non-infringing Uses, 9 Bull. Copyright Soc. 466, 467 (1962).

<sup>&</sup>lt;sup>2</sup> DVD: *The Best of Mr. Peabody & Sherman, Volume 1, Chapter 6: Ferdinand Magellan* (Sony Wonder 2005) (on file with author).

<sup>&</sup>lt;sup>3</sup> See Internet Archive, About The Internet Archive, http://www.archive.org/about/about.php (last visited Nov. 22, 2007) [hereinafter About IA].

<sup>&</sup>lt;sup>4</sup> See Internet Archive, Frequently Asked Questions, http://www.archive.org/about/faqs.php (last visited Nov. 22, 2007) [hereinafter FAQs] ("Imagine surfing circa 1999 and looking at all the Y2K hype, or revisiting an older version of your favorite Web site. The Internet Archive Wayback Machine can make all of this possible.").

<sup>&</sup>lt;sup>5</sup> See About IA, supra note 3.

though copyright law provides the precedent for handling some of the larger legal issues facing a digital archive, distinct factual situations still arise which are unique to the digital infrastructure of the particular archive. The real dilemma lies in deciding whether copyright law can handle the combination of these legal issues and unique factual situations or whether other areas of the law need to be extended.

This Comment will show why courts must provide a definitive ruling declaring that the fair use exception to copyright infringement encompasses a digital archive making periodic copies of publicly accessible websites. This exception provides that the use of a copyrighted work will not be considered an infringing use if it is done for certain purposes or, if after consideration of several factors, the use is found to further the purpose of making copyrighted works available for use by the public.<sup>6</sup>

Part I of this Comment explores the background of the Internet Archive. In Part IA, the Internet Archive's creation and mission are explained. The first section of Part IA describes the historical development and purpose of the Wayback Machine. Section two discusses the pertinent issues of a copyright owner who does not want a website archived by the Wayback Machine. Section three delves into how users have employed the Wayback Machine and the legal consequences of those uses. The fourth section explains the issues in a lawsuit currently pending before the federal District Court in the Eastern District of Pennsylvania and describes how the issues further highlight why courts must decide whether fair use protects digital archives.

In Part IB, the fair use exception is fully explained. Section one provides a historical perspective on copyright law and the fair use doctrine. The second section describes the statutory law with a detailed look at the factors of fair use. In conclusion to Part IB, section three reviews the major, relevant case law developments that address fair use.

Part II presents the analysis of possible results and the implications associated with courts declaring a digital archive's activities to be within the realm of fair use. Parts IIA through IID analyze the Internet Archive's activities in terms of the four statutory factors, respectively. The case law interpretations of these factors are woven into the analyses as well as the

<sup>&</sup>lt;sup>6</sup> 17 U.S.C. § 107 (2000).

individual facts that make the Internet Archive the perfect example for looking at digital archives.<sup>7</sup>

#### I. BACKGROUND

#### A. The Internet Archive and its Wayback Machine

#### 1. Understanding the Internet Archive and the Wayback Machine

The Internet Archive is a website that catalogues, or archives, almost the entire World Wide Web. In conjunction with the Library of Congress, the Smithsonian, and the Alexa Internet, the Internet Archive is working to preserve periodically obtained copies of publicly accessible websites as they appear over time. It is also making a record of the Internet and other items such as movies, books, live concerts, and freeware through the visual images, print, audio, and software collections available on the site. To make this archive of web pages available, the Internet Archive provides a search tool on its website called the Wayback Machine. Aptly, this search tool is named after a fictional invention that seeks the truth of history, not merely the version proliferated by spin doctors. It is named after the "Rocky and Bullwinkle" cartoon's time-travel machine in which the characters "Sherman and Peabody would use Peabody's 'WABAC machine' (pronounced 'way-back' . . . ) to go back in time to discover the real story behind historical events" (albeit the real story may go comically awry). Users of this free service can search previous versions

<sup>&</sup>lt;sup>7</sup> In trying to restrict the discussion in this Comment to digital archives, namely the Internet Archive, there will not be any discussion of related topics, such as music downloads and the debate on whether online linking is copying.

<sup>&</sup>lt;sup>8</sup> See FAQs, supra note 4.

<sup>&</sup>lt;sup>9</sup> *Id*.

<sup>&</sup>lt;sup>10</sup> *Id.* Internet Archive excludes websites that are not publicly accessible from the Wayback Machine. *Id.* If the websites are password protected, use the standard for robot exclusion, require a form for access, or have pages on secure servers, then the Internet Archive will not copy the sites. *Id.* Additionally, a website owner specifically requesting that the Internet Archive not make copies of the site will also insure that a site is excluded. *Id.* 

<sup>&</sup>lt;sup>11</sup> See About IA, supra note 3.

<sup>&</sup>lt;sup>12</sup> See FAQs, supra note 4.

Wikipedia, The Rocky and Bullwinkle Show, http://en.wikipedia.org/wiki/ The Rocky and Bullwinkle Show (last visited Nov. 22, 2007).

Wikipedia, Internet Archive, http://en.wikipedia.org/wiki/Wayback\_Machine (last visited Nov. 22, 2007) [hereinafter Wikipedia, Internet Archive]. Federal Rule of Evidence (continued)

of web pages, including those which are no longer supported or which no longer exist. By simply typing a website address, or Uniform Resource Locator (URL), into the Wayback Machine's text box, a user can access one of "85 billion web pages archived from 1996 to a few months ago." The Wayback Machine contains almost two petabytes of data, which is more information than that available in the Library of Congress. However, keyword searching is not available at this time.

Creating the archive requires the Internet Archive to amass digital copies of web pages over a period of time, which is primarily accomplished through donations of such data.<sup>20</sup> The Internet Archive specifically uses the Wayback Machine as the database, or "three dimensional index," for storing the donated digital copies of web pages.<sup>21</sup> This donated digital data comes from The Alexa Internet,<sup>22</sup> a company that was begun with the purpose of gathering information about users visiting its site and then providing the users with other links related to their interests.<sup>23</sup> The Alexa Internet designed and built the Wayback Machine database, in conjunction with the Internet Archive, and continues to donate "crawls" of the Web.<sup>24</sup>

602 has an analogous requirement, which states that "[a] witness may not testify to a matter unless evidence is introduced sufficient to support a finding that the witness has personal knowledge of the matter." *Id.* 

<sup>&</sup>lt;sup>15</sup> Reg P. Wydeven, *Old Web Sites Can Be Forgotten but Aren't Legally Gone*, APPLETON POST-CRESCENT, Oct. 17, 2005, at 12D, *available at* http://www.mccarty-law.com/resources/101505forgottenbutnotgonegatheringevidenceontheint.pdf.

<sup>&</sup>lt;sup>16</sup> OXFORD DICTIONARY OF ENGLISH 1942 (Catherine Soanes & Angus Stevenson eds., 2d ed. 2003).

<sup>&</sup>lt;sup>17</sup> Internet Archive, The Wayback Machine, http://www.archive.org/web/web.php (last visited Nov. 22, 2007) [hereinafter Wayback Machine].

<sup>&</sup>lt;sup>18</sup> FAOs, *supra* note 4.

<sup>&</sup>lt;sup>19</sup> See Wayback Machine, supra note 17.

<sup>&</sup>lt;sup>20</sup> See FAQs, supra note 4.

<sup>&</sup>lt;sup>21</sup> *Id*.

<sup>&</sup>lt;sup>22</sup> *Id*.

<sup>&</sup>lt;sup>23</sup> See Wikipedia, Alexa Internet, http://en.wikipedia.org/wiki/Alexa\_Internet (last visited Nov. 22, 2007).

<sup>&</sup>lt;sup>24</sup> FAOs, *supra* note 4.

#### 2. Copyright and the Internet Archive

Some copyright owners, usually website owners, prefer that the Wayback Machine not archive their sites. The Internet Archive addresses these situations by providing a variety of options to the copyright owner. As a general matter, "[t]he Internet Archive is not interested in preserving or offering access to Web sites or other Internet documents of persons who do not want their materials in the collection." As a first way of showing its commitment to exclusion, the Internet Archive posts a copyright policy. Second, an exclusion policy link is also provided on the website. Third, directions are offered on how to exclude a website from being archived. Fourth, there is always the option of emailing the Internet Archive and directly communicating the desire to opt out. Each of these systems, which will be examined in turn, is designed to eliminate a copyright owner's material from being accessible to the public from the Wayback Machine site.

The Internet Archive posts a copyright policy in at least two places on the website.<sup>32</sup> The policy includes introductory remarks acknowledging the rights afforded intellectual property owners, after which a disclaimer is displayed granting the Internet Archive discretion to remove apparently infringing content.<sup>33</sup> As further evidence of its recognition of and commitment to copyright law principles, the Internet Archive designates a Copyright Agent to handle infringement claims.<sup>34</sup> Any copyright owner who believes a copyright violation has occurred can contact the Copyright Agent and provide basic contact information, the location within the Internet Archive of the alleged infringing material, and a statement "under penalty of perjury" that accurate information has been provided and that

<sup>&</sup>lt;sup>25</sup> *Id*.

<sup>&</sup>lt;sup>26</sup> *Id*.

<sup>&</sup>lt;sup>27</sup> *Id*.

<sup>&</sup>lt;sup>28</sup> See id.

<sup>&</sup>lt;sup>29</sup> *Id*.

<sup>&</sup>lt;sup>30</sup> Internet Archive, Removing Documents From the Wayback Machine, http://www.archive.org/about/exclude.php (last visited Nov. 22, 2007) [hereinafter Removing Documents].

<sup>&</sup>lt;sup>31</sup> See id.; see also FAQs, supra note 4.

<sup>&</sup>lt;sup>32</sup> Internet Archive, Terms of Use, http://www.archive.org/about/terms.php (last visited Nov. 22, 2007) [hereinafter Terms of Use]; FAQs, *supra* note 4.

<sup>&</sup>lt;sup>33</sup> Terms of Use, *supra* note 32.

<sup>34</sup> See id.

the complainant is the owner of the copyright or is otherwise authorized to defend the alleged infringement.<sup>35</sup> When the Copyright Agent receives such a complaint, the allegedly infringing material is presumably investigated, as the policy indicates the Internet Archive "may, in appropriate circumstances and at its discretion, remove certain content or disable access to content that appears to infringe the copyright or other intellectual property rights of others."<sup>36</sup>

A link to the Internet Archive's exclusion policy is also posted on the website.<sup>37</sup> After clicking on this link, the user is redirected to another website where a list of anticipated removal requests and the sources from which they might originate are listed.<sup>38</sup> Basically, the Internet Archive is predicting who will send removal requests and how to handle such requests.<sup>39</sup> The first potential source of a removal request, the owner of a private (non-governmental) website, addresses the fair use inquiry.<sup>40</sup> The policy for handling this request states, "Archivists should provide a 'self-service' approach site owners can use to remove their materials based on the use of the robots.txt standard."<sup>41</sup> The other six sources delve into fringe issues, such as a third party's removal request and how to handle verifying the authority of the third party.<sup>42</sup> This list of sources and responses was compiled with input from "commercial and non-commercial organizations"<sup>43</sup> concerned about such issues and by considering policies suggested by various associations' codes of conduct.<sup>44</sup>

Beyond the exclusion policy, the Internet Archive offers instructions on how to "self-service" the removal of a website from the Wayback Machine. <sup>45</sup> Basically, a website owner adds a robots.txt protection

<sup>&</sup>lt;sup>35</sup> *Id*.

<sup>&</sup>lt;sup>36</sup> See id.

<sup>&</sup>lt;sup>37</sup> FAOs, *supra* note 4.

<sup>&</sup>lt;sup>38</sup> See id.; University of California, Berkley, School of Information Management and System, The Oakland Archive Policy, http://www.sims.berkeley.edu/research/conferences/aps/removal-policy.html (last visited Nov. 22, 2007) [hereinafter Oakland Archive Policy].

<sup>&</sup>lt;sup>39</sup> FAQs, supra note 4.

<sup>&</sup>lt;sup>40</sup> See id.

<sup>&</sup>lt;sup>41</sup> *Id*.

<sup>42</sup> See id.

<sup>&</sup>lt;sup>43</sup> *Id.* These organizations included the Electronic Frontier Foundation, Chilling Effects, The Council on Library and Information Resources, Berkeley Boalt School of Law, and various other organizations. *Id.* 

<sup>&</sup>lt;sup>44</sup> Id.

<sup>&</sup>lt;sup>45</sup> See Removing Documents, supra note 30.

measure to her website.<sup>46</sup> This entails literally putting "robots.txt" at the end of a domain name (e.g., www.domainname.com/robots.txt) and then submitting the updated website to a specific link at the Internet Archive.<sup>47</sup> By adding the robots.txt string to the end of the domain name, a website can alert a search engine crawling the web, such as the Alexa Internet, that all documents associated with the particular domain name should be removed from the Wayback Machine and that there should be no future crawls of the website.<sup>48</sup>

# 3. Using the Wayback Machine and the Legal Ramifications

The Wayback Machine is becoming an oft-used tool for lawyers. "'It's becoming almost an automatic," stated a trademark and domainname lawyer with Ballard Spahr Andrews & Ingersoll LLP. 49 Playboy's senior counsel for intellectual property issues said that Playboy uses the Wayback Machine every month or so to search for potentially infringing use of the company's images, including its trademark bunny. 50 According to an August 2005 magazine article, "Intellectual property lawyers have been using these techniques for a few years to locate old websites that demonstrate that their clients' trademark or domain name has been misused."51 Additionally, the use of the Wayback Machine is becoming common in other areas of law, such as in evaluating the credibility of witnesses.<sup>52</sup> Speaking on the Wayback Machine and other digital archives, Ken Strutin of the New York State Defenders Association said, "This is going to spread quickly." A criminal defense lawyer in Arkansas said he now asks each client, "'Have you ever said anything on the Internet you would regret?" A trademark lawyer with the Austin, Texas firm of Fulbright & Jaworski, Stephen Meleen, agreed with this approach.<sup>55</sup>

<sup>&</sup>lt;sup>46</sup> See id.

<sup>&</sup>lt;sup>47</sup> *Id*.

<sup>&</sup>lt;sup>48</sup> See FAQs, supra note 4.

<sup>&</sup>lt;sup>49</sup> David Kesmodel, *Not Fade Away, Lawyer's Delight: Old Web Material Doesn't Disappear*, WALL St. J., July 27, 2005, at A4 (quoting Rich Peirce).

<sup>50</sup> Id

<sup>&</sup>lt;sup>51</sup> Reni Gertner, *Lawyers Are Turning To Old Websites For Evidence*, LAW. USA, Aug. 15, 2005, *available at* http://www.lawyersusaonline.com/subscriber/archives\_FTS.cfm? page=USA/05/815052.htm&recID=347182&QueryText=wayback%20and%20machine.

<sup>&</sup>lt;sup>52</sup> See, e.g., id.; Kesmodel, supra note 49.

<sup>&</sup>lt;sup>53</sup> Gertner, *supra* note 51 (quoting Ken Strutin).

<sup>&</sup>lt;sup>54</sup> *Id.* (quoting John Wesley Hall).

<sup>&</sup>lt;sup>55</sup> *Id*.

With the increased use by lawyers of the database housed by the Wayback Machine comes increased litigation. The following cases demonstrate how the Internet Archive and its Wayback Machine are sometimes caught in the middle of a lawsuit and are occasionally sued.

The first example involves a pornography website that likely misled consumers to think that it was associated with the Sex Court television show produced by Playboy. Sex Court is a cable television series produced by Playboy Enterprises, Inc., through one of its wholly owned subsidiaries. The show is broadcast in the United States on "Playboy TV" and internationally on "Playboy TV Networks." Playboy also produces video collections of the show for sale through its playboy.com website. These videos display the Sex Court mark. The defendant, On Line Entertainment, Inc., is a pornography website that registered sexcourt.com as its domain name. Then, as alleged by Playboy, the pornography website tried to bait consumers into visiting the site and buying the contents it offered for sale by trading on the goodwill of the Sex Court trademark. Before the jury trial concluded, the parties entered into a settlement agreement.

One of the major legal issues for the jury would have been to determine whether there was trademark infringement based on *when* each party began using the trade name Sex Court.<sup>64</sup> Playboy's internet research manager stated in a 2003 news article that the pornography website first used the term in "January 1999, four months after Playboy aired the first installment of its cable show." However, the pornography website's attorney submitted evidence that his client first displayed its website in May 1998. <sup>66</sup>

<sup>&</sup>lt;sup>56</sup> See Playboy Enters. Int'l, Inc. v. On Line Entm't, Inc., No. CV 00-6618(DGT), 2004 U.S. Dist. LEXIS 5145, at \*2 (E.D.N.Y. Mar. 30, 2004).

<sup>&</sup>lt;sup>57</sup> *Id*.

<sup>&</sup>lt;sup>58</sup> *Id*.

<sup>&</sup>lt;sup>59</sup> *Id*.

<sup>&</sup>lt;sup>60</sup> Id.

<sup>&</sup>lt;sup>61</sup> See id. at \*3.

<sup>&</sup>lt;sup>62</sup> *Id*.

<sup>&</sup>lt;sup>63</sup> *Id.* at \*4.

<sup>&</sup>lt;sup>64</sup> See Anthony M. DeStefano, An Un.com-mon Revelation in Porn-Name Case, NEWSDAY (New York), Apr. 10, 2003, at A23.

<sup>&</sup>lt;sup>65</sup> *Id*.

<sup>&</sup>lt;sup>66</sup> Id.

The Wayback Machine was used by Playboy to show that the pornography website did not use the term Sex Court until after Playboy had begun airing its show by the same name. Unfortunately, the tool used by the Playboy attorney was quite the surprise to the pornography website's attorney: "[He] seemed stunned by the revelation about the huge archive as he cross-examined a Playboy official." However, this quote is from 2003, when use of the Internet Archive was not as prolific as it is now. While the judge in the Sex Court case "quipped that [the Wayback Machine] sounded like something from the film 'Back to the Future,'" the legal field is becoming more aware of the Wayback Machine's potential uses.

A forum for disgruntled consumers provides the second example of the Wayback Machine's involvement in litigation. Innervision Web Solutions, a relatively unknown home-based computer company, used DellComputersSuck.com as a domain name to redirect consumers to the Innervision website promoting sales of other computers. Dell has been in the business of marketing and selling computers for more than twenty years. It began using the registered Dell trademark in 1987. Dell sent a cease and desist letter to Innervision in March 2005, requesting the company transfer the DellComputersSuck.com domain name to Dell. Three days later Ed Ziejka, a representative for Innervision, left a voicemail for the Dell lawyer who sent the letter. Ziejka indicated that Dell could purchase the domain name from him. However, the next day

<sup>&</sup>lt;sup>67</sup> Kesmodel, *supra* note 49.

<sup>&</sup>lt;sup>68</sup> DeStefano, supra note 64.

<sup>&</sup>lt;sup>69</sup> *See supra* notes 49–52 and accompanying text; *cf.* Wikipedia, Internet Archive, *supra* note 14 (describing the rapid growth of the amount of data contained on the Wayback Machine).

<sup>&</sup>lt;sup>70</sup> DeStefano, *supra* note 64.

<sup>&</sup>lt;sup>71</sup> See supra note 69 and accompanying text.

<sup>&</sup>lt;sup>72</sup> See Dell Inc. v. Innervision Web Solutions, Claim No. FA0503000445601 (National Arbitration Forum, May 23, 2005), http://www.arb-forum.com/domains/decisions/445601.htm.

<sup>73</sup> See id.

<sup>&</sup>lt;sup>74</sup> *Id*.

<sup>&</sup>lt;sup>75</sup> Id.

<sup>&</sup>lt;sup>76</sup> See Wydeven, supra note 15.

<sup>&</sup>lt;sup>77</sup> Dell Inc., supra note 72.

<sup>&</sup>lt;sup>78</sup> *Id*.

<sup>&</sup>lt;sup>79</sup> *Id*.

Dell saw that the DellComputersSuck.com website was no longer redirecting consumers, but was providing a forum for disgruntled consumers to lodge their complaints against different computer manufacturers. Diejka contended that the previous redirection of consumers to the Innervision page had been a temporary situation for testing purposes. He said the intent for the domain name was for elegitimate, noncommercial, and fair use... to express criticism and commentary pursuant to his First Amendment rights, primarily the right of free speech.

Dell submitted additional evidence alleging this contention was untrue, based on its research on the Internet Archive's Wayback Machine. Bell used the Wayback Machine to prove that DellComputersSuck.com had been redirecting traffic to the Innervision website from the time of registering the domain until responding to Dell's lawyer. The arbitrators determined that Innervision had registered and used the domain name in bad faith and ordered it transferred to Dell.

The third case highlighting the Wayback Machine's use in lawsuits is the *EchoStar Satellite Corporation* case. <sup>86</sup> In this suit, Telewizja Polska USA, Inc. sued EchoStar Satellite Corporation for unfair competition and breach of contract, among other claims. <sup>87</sup> Polska provides Polish-language television and radio to United States customers. <sup>88</sup> EchoStar broadcasts television and radio programs via satellite throughout the United States. <sup>89</sup> The two companies had an agreement that EchoStar would distribute Polska's programming to the EchoStar network. <sup>90</sup> EchoStar was permitted to use Polska's trademark in conjunction with the advertising and selling of subscriptions to the EchoStar network. <sup>91</sup> The agreement provided for

```
80 See id.
```

<sup>81</sup> See id.

<sup>&</sup>lt;sup>82</sup> *Id*.

<sup>&</sup>lt;sup>83</sup> *Id*.

<sup>&</sup>lt;sup>84</sup> *Id*.

<sup>&</sup>lt;sup>85</sup> Id

 $<sup>^{86}</sup>$  Telewizja Polska USA, Inc. v. EchoStar Satellite Corp., 69 F. App'x 793 (7th Cir. 2003).

<sup>&</sup>lt;sup>87</sup> See id. at 794.

<sup>88</sup> Id.; Kesmodel, supra note 49.

<sup>&</sup>lt;sup>89</sup> Telewizja Polska USA, Inc., 69 F. App'x. at 794.

<sup>&</sup>lt;sup>90</sup> Id.

<sup>&</sup>lt;sup>91</sup> *Id*.

stages of expiration. <sup>92</sup> The marketing and selling rights were to end after the three-year contract expired, and then the programming would conclude up to a year later in order to service customers who purchased multi-month subscriptions at the end of the three-year contract. <sup>93</sup> However, EchoStar continued to market with the Polska trademark and sell subscriptions after the three-year contract expired. <sup>94</sup>

What is interesting in this case is that it is one of the underlying decisions on whether information from the Wayback Machine is admissible in court. EchoStar used the Wayback Machine to counter Polska's attack. EchoStar found after expiration of the three-year contract that Polska had continued to tout its offerings of Polish programming through EchoStar on the Polska website. Polska then countered with various motions to suppress the evidence on grounds that it was hearsay and that it was from an unauthenticated source. The magistrate found that the "statements," or the previous versions of the websites, did not fall under the hearsay rules for exclusion. He also "accepted an affidavit from an Internet Archive employee as sufficient to authenticate the snapshots for admissibility."

# 4. The Wayback Machine in Current Litigation

One of the most recent legal battles involving the Wayback Machine raises another issue that circles around the question of whether the Internet Archive's activities fall under the fair use doctrine. The following question is posed in the pending *Healthcare Advocates*, *Inc.* case<sup>100</sup>: does the

<sup>92</sup> See id.

<sup>&</sup>lt;sup>93</sup> *Id*.

<sup>&</sup>lt;sup>94</sup> *Id*.

<sup>&</sup>lt;sup>95</sup> See Lauren Gelman, Internet Archive's Web Page Snapshots Held Admissible as Evidence, 2 STAN. CENTER FOR INTERNET & SOC'Y (2004), http://cyberlaw.stanford.edu/taxonomy/term/185.

<sup>&</sup>lt;sup>96</sup> See id.

<sup>&</sup>lt;sup>97</sup> *Id*.

<sup>&</sup>lt;sup>98</sup> *Id*.

<sup>&</sup>lt;sup>99</sup> Id.

Healthcare Advocates, Inc. v. Harding, Earley, Follmer & Frailey, No. 2:05-CV-03524-RK, 2005 WL 2151678 (E.D. Pa. July 8, 2005). The Internet Archive was also a defendant in this case. Complaint at 1, *Healthcare Advocates, Inc.*, No. 2:05:CV-03524-RK, 2005 WL 215678. On August 23, 2006, Healthcare Advocates, Inc. and the Internet Archive announced they had reached a settlement. Internet Archive, Forums (Aug. 24, (continued))

Internet Archive's unilateral use of the robots.txt mechanism make it liable for breach of contract, if the mechanism fails or if someone circumvents the mechanism? Healthcare Advocates was involved as the plaintiff in an underlying lawsuit where the defendant law firm, Harding, Earley, Follmer & Frailey, used the Wayback Machine to access historical copies of the Healthcare Advocates' website. 101 The firm began trying to access the website through the Wayback Machine the day after Healthcare Advocates initiated use of the robots.txt mechanism to block the historical copies of its site. 102 Multiple attempts to access the old versions of the website were made by employees of the law firm. 103 For whatever reason, some of these attempts were successful, despite the robots.txt prevention mechanism. 104 Healthcare Advocates then initiated the lawsuit against the law firm and the Internet Archive. 105 One of the basic claims against the law firm is that it "intentionally circumvented" the robots.txt prevention mechanism with the multiple attempts to access the website. 106 Prior to settlement, the major claim by Healthcare Advocates against the Internet Archive was for breach of contract. 107 Healthcare Advocates claimed that the Internet Archive "failed to perform its duty" of blocking access to the Healthcare Advocates website via the Wayback Machine, in light of the fact that Healthcare Advocates installed the robots.txt blocking mechanism. 108 Because of the settlement, a court will not have a chance to issue a ruling on whether Internet Archive breached a contract with this business. However, the original case against the law firm is still in the pretrial stage. 109 The suit against Internet Archive was dismissed with prejudice, but the suit against the law firm is still calling for dispositive motions to be

2006), http://www.archive.org/iathreads/post-view.php?id=67826. The suit against the law firm of Harding, Earley, Follmer & Frailey is still pending. *Id*.

<sup>&</sup>lt;sup>101</sup> See Complaint at 2, Healthcare Advocates, Inc., No. 2:05:CV-03524-RK, 2005 WL 215678.

<sup>&</sup>lt;sup>102</sup> *Id.* at 8–10.

<sup>&</sup>lt;sup>103</sup> *Id.* at 9–11, 17.

<sup>&</sup>lt;sup>104</sup> *Id.* at 11, 17.

<sup>&</sup>lt;sup>105</sup> See id. at 2.

<sup>106</sup> Id. at 20.

<sup>&</sup>lt;sup>107</sup> Id. at 28.

<sup>&</sup>lt;sup>108</sup> *Id.* at 28–29.

<sup>&</sup>lt;sup>109</sup> See Second Amended Complaint, Healthcare Advocates, Inc., No. 2:05-CV-03524-RK, 2006 WL 1784161.

filed.<sup>110</sup> The suit against the law firm focuses on the issue of whether the firm "intentionally circumvented" a prevention mechanism.<sup>111</sup> If and when this case goes to trial and the court issues an opinion, it may or may not provide clues as to how it would have decided the breach of contract claim against Internet Archive. However, William Patry, former Congressional copyright counsel and current intellectual property lawyer, has said that "no real contract exists between the nonprofit Internet Archive and any of the historical Web sites it preserves." He emphasizes the fact that the robots.txt prevention mechanism is a completely "voluntary system." Regardless of whether a decision in the suit against the law firm allows reasonable inferences and deductions about the breach of contract claim, the overarching issue of whether the Internet Archive's copying of publicly accessible web pages is fair use will still not have been specifically addressed by a court.

#### B. Fair Use

#### 1. Historical Development

In 1709, the Statute of Anne created the copyright in the Act for the Encouragement of Learning. Soon after the creation of the copyright, courts recognized the concept of fair abridgment, the predecessor to fair use. What the courts described as fair abridgment was the idea that certain instances of unauthorized reproduction of copyrighted material . . . would not infringe the author's rights. In the United States, this right was first encompassed by Article I, Section 8, of the Constitution, which provides that Congress has the power [t]o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and

<sup>&</sup>lt;sup>110</sup> See Healthcare Advocates, No. 2:05-CV-03524-RK, Docket Proceedings, as of February 2007.

<sup>&</sup>lt;sup>111</sup> See Complaint at 20, Healthcare Advocates, Inc., No. 2:05:CV-03524-RK, 2005 WL 215678

<sup>&</sup>lt;sup>112</sup> 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.

<sup>&</sup>lt;sup>113</sup> Id

<sup>&</sup>lt;sup>114</sup> Pierre N. Leval, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105, 1105 (1990).

<sup>&</sup>lt;sup>115</sup> *Id*.

<sup>&</sup>lt;sup>116</sup> *Id*.

Discoveries."<sup>117</sup> Common law copyright developed for almost 200 years before it was codified by an act of Congress.<sup>118</sup> Deriving its power from the Constitution, Congress codified the right in the Copyright Act of 1976.<sup>119</sup>

"The Copyright Act represents the culmination of a major legislative reexamination of copyright doctrine." However, neither this statutory codification nor the common law decisions spanning centuries have clearly delineated the application of the fair use doctrine to a copyright. The problem of discerning fair use is seen in the lack of identifying characteristics in the law that would help a court determine where the facts of a case fall within the spectrum of each fair use factor. In addition, there is no "consensus on the meaning of fair use" among judges.

#### 2. Interpretation of the Statutory Law

Fair use is a statutory limitation on what constitutes copyright infringement. Title 17 of the United States Code provides the statutory law concerning copyrights. Section 102 describes the subject matter of copyrightable materials. This section states, in pertinent part, "Copyright protection subsists, in accordance with this title, in original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device." Section 106 explains the exclusive rights of a copyright owner, such as the right to reproduce copies of the work and the right to distribute

<sup>&</sup>lt;sup>117</sup> U.S. CONST. art. I, § 8, cl. 8.

The Constitution was signed on September 17, 1787, and it became effective in 1789. The Library of Congress, Primary Documents in American History, http://www.loc.gov/rr/program/bib/ourdocs/Constitution.html (last visited Nov. 22, 2007). Codification occurred in the Copyright Act of 1976, thus the two hundred years of development before codification. *See* Leval, *supra* note 114, at 1105.

<sup>&</sup>lt;sup>119</sup> See id.

<sup>&</sup>lt;sup>120</sup> Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 552 (1985).

<sup>&</sup>lt;sup>121</sup> Leval, *supra* note 114, at 1105.

<sup>122</sup> See id. at 1106.

<sup>&</sup>lt;sup>123</sup> *Id*.

<sup>124 17</sup> U.S.C. § 107 (2000).

<sup>125</sup> See id. §§ 101-1332.

<sup>&</sup>lt;sup>126</sup> Id. § 102.

<sup>&</sup>lt;sup>127</sup> *Id.* § 102(a).

copies through sale or other means.<sup>128</sup> Then, section 107 provides the limitations on these exclusive rights; specifically, section 107 defines fair use:

Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include--

- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
- (2) the nature of the copyrighted work;
- (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- (4) the effect of the use upon the potential market for or value of the copyrighted work.

The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors. 129

The four factors, combined with the facts of the case, are the key to the fair use analysis. However, there is no bright-line rule for courts to follow either in a standard copyright case or in one involving digital issues. Courts can go either way in deciding whether fair use applies, as it is very case-specific and fact-specific. Speaking specifically about fair use,

<sup>129</sup> *Id.* § 107.

<sup>&</sup>lt;sup>128</sup> Id. § 106.

<sup>&</sup>lt;sup>130</sup> See Leval, supra note 114, at 1106–07.

<sup>&</sup>lt;sup>131</sup> See Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 560 (1985); Melanie Martin-Jones, Intellectual Property and Technology Attorney for Porter Wright Morris & Arthur, LLP, Lecture at Capital University Law School: Fair Use: *Grokster*, Google and Beyond (Nov. 3, 2005).

Judge Pierre Leval, of the Southern District of New York, <sup>132</sup> noted that "historians, publishers, and their legal advisers can only guess and pray as to how courts will resolve [these] disputes." Leval suggests that "a cogent set of governing principles [already] exists" and using these would make the fair use inquiry less mysterious. <sup>134</sup>

The first factor, "purpose and character of the use," is applied by determining whether the use of the copyrighted material is for commercial gain or for nonprofit educational reasons. If the use is primarily for commercial gain, then it will most likely be found not to be a fair use. The closer the use is to a nonprofit educational use, the more likely a fair use will be found to exist. However, "[t]he crux of the profit/nonprofit distinction is not whether the sole motive of the use is monetary gain but whether the user stands to profit from exploitation of the copyrighted material without paying the customary price." 138

The court must also consider the justification for finding the use to be a fair use, which can be revealed by answering the following question: "Does the use fulfill the objective of copyright law to stimulate creativity for public illumination?" Courts must also consider that the more transformative the work, the more likely it is fair use. The underlying policy consideration is to protect the copyright owner's rights, yet still allow the public the benefit of creative stimulation. A truly transformative work uses the copyrighted work as raw material and transforms it in any number of ways, including producing "new information, . . . new insights, . . . critici[sm of] the quoted work, [or] parody." However, finding a use transformative does not guarantee a successful fair use defense. It is the use takes too much of the copyrighted

<sup>&</sup>lt;sup>132</sup> Leval, *supra* note 114, 1105 n.\*.

<sup>&</sup>lt;sup>133</sup> *Id.* at 1107.

<sup>134</sup> Id. 1105, 1107.

<sup>&</sup>lt;sup>135</sup> See Martin-Jones, supra note 131.

<sup>&</sup>lt;sup>136</sup> 4 DAVID NIMMER, NIMMER ON COPYRIGHT § 13.05 [A][1][c], at 13-170 (2006).

<sup>137</sup> See id. at 13-176.

<sup>&</sup>lt;sup>138</sup> Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 562 (1985) (citing Roy Exp. Co. Establishment v. Columbia Broad. Sys., Inc., 503 F. Supp. 1137, 1144 (S.D.N.Y. 1980)), *cited in* NIMMER, *supra* note 136, at 13-175.

<sup>&</sup>lt;sup>139</sup> Leval, *supra* note 114, at 1111.

<sup>&</sup>lt;sup>140</sup> *Id.*; see Martin-Jones, supra note 131.

<sup>&</sup>lt;sup>141</sup> See Leval, supra note 114, at 1110–11.

<sup>&</sup>lt;sup>142</sup> *Id.* at 1111.

<sup>&</sup>lt;sup>143</sup> *Id*.

work and only minimally transforms it, the court will likely find the justification "outweighed" and will be less likely to find it a fair use. 144

The second factor, "nature of the . . . work," is a determination of whether the work is published or unpublished, and whether it is creative or factual in nature. This determination is foreshadowed by the purpose of the work. For example, an individual can claim a copyright in a diary entry, a grocery list, or even an extortion note. However, simply claiming a copyright does not insulate the author from facing a fair use defense. However, whether the second s

When considering a fair use defense, there is a "meaningful difference" among works created for private communication and works created for publication. Authors of works intended for publication will seek copyright protection for the creative, artistic, or factual nature, whereas the authors of private communications may not contemplate such issues. The heart of the issue for determining whether fair use exists for either unpublished or published works lies in the right of first publication. This right belongs to the original author and spans the

<sup>&</sup>lt;sup>144</sup> See id. at 1112.

<sup>&</sup>lt;sup>145</sup> See Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 563 (1985); Leval, *supra* note 114, at 1117–18; Martin-Jones, *supra* note 131.

<sup>&</sup>lt;sup>146</sup> See Leval, supra note 114, at 1116.

<sup>&</sup>lt;sup>147</sup> Id. at 1116–17, 1120.

<sup>148</sup> See id. at 1117.

<sup>&</sup>lt;sup>149</sup> *Id*.

<sup>150</sup> See id.

<sup>&</sup>lt;sup>151</sup> See Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 564 (1985). In this case, former President Gerald Ford had an agreement with publisher Harper & Row to produce and market his memoirs of the Watergate scandal and the subsequent pardon of former President Richard Nixon. Id. at 542. The agreement also allowed the publisher to license excerpts from the book, prior to its release. Id. Shortly before the release, a license was granted to *Time* magazine for the right to be the exclusive publisher of these excerpts. Id. at 542-43. However, an unknown person took a copy of the manuscript to the editor of The Nation magazine. Id. at 543. The editor quickly assembled an article that consisted of material copied directly from the manuscript and void of independent editorial comments. Id. The article was published right before Time was scheduled to release the excerpts, so Time cancelled its agreement with the publishers. Id. The publisher sued The Nation. Id. Ultimately, the publisher's claims were addressed by the Supreme Court. Id. at 542. The claim to the right of first publication became the pivotal issue for the Court. Id. at 551, 553. In analyzing this issue, the Court focused heavily on the policy underlying a fair use defense for unpublished versus published works. Id. at 549-51. As a whole, "the Nation decision communicates a concern for protection of unpublished works that were created for (continued)

decision of whether to publish and the decision of how the publication will be formatted and released. 152

For unpublished works, "the scope of fair use is narrower." <sup>153</sup> Traditionally, common law recognized the author's intellectual property right in the work "until he voluntarily part[ed] with the same." <sup>154</sup> Over time this rule was recognized as not being an absolute rule. <sup>155</sup> Congress realized that secondary use of unpublished works that were put forth in some manner other than what typically represented a publication might be defended with the fair use doctrine. <sup>156</sup> The author's "implied consent through *de facto* publication on performance or dissemination of a work may tip the balance of equities in favor of prepublication use." <sup>157</sup> However, the general rule remains that unpublished works that are proliferated without the author's consent effectively deny the author the right of first publication. <sup>158</sup>

Copyright was created for works intended for publication, so as to encourage production of the creative, artistic, and factual works society seeks to preserve. Even though most works will benefit from copyright, the thrust of copyright is to protect works created with the intent of publication. Copyright law encourages protective rights for original creators so that they do not hesitate from illuminating the public with their creations, whether artistic or factual, through publication. 161

Given these underlying public policies, the use of a simple published or unpublished inquiry for determining fair use can thwart the purpose of copyright and fair use. <sup>162</sup> If courts always found published works could be

publication, or on their way to publication, and not for unpublished matter created for private ends and held in secrecy." Leval, *supra* note 114, at 1120. However, the fine line between works destined for publication and for communications for private means can be difficult to distinguish. *Id.* 

 $^{154}$  Id. at 551 (alteration in original) (quoting Am. Tobacco Co. v. Werckmeister, 207 U.S. 284, 299 (1907)).

n

<sup>&</sup>lt;sup>152</sup> Harper & Row, 471 U.S. at 553.

<sup>&</sup>lt;sup>153</sup> *Id.* at 564.

<sup>155</sup> See id.

<sup>156</sup> See id.

<sup>&</sup>lt;sup>157</sup> *Id*.

 $<sup>^{158}</sup>$  See id.

<sup>&</sup>lt;sup>159</sup> See Leval, supra note 114, at 1107, 1119.

<sup>&</sup>lt;sup>160</sup> *Id.* at 1119.

<sup>&</sup>lt;sup>161</sup> See id. at 1111.

<sup>&</sup>lt;sup>162</sup> See id. at 1118–19.

reused under fair use and unpublished works could never be reused under fair use, then the creative, artistic, and factual works are always fair game for copyright infringement, but the diary entries, grocery lists, and extortion notes are vigorously protected from any claim of fair use. 163 The result is opposite of the purposes of copyright law. 164 The second factor is meant to help courts determine the nature of the work as a whole and not merely as an item that may receive protection depending on its publication status.

The third factor, "amount" of the work used, seeks to determine if the heart of the work, or just something from the peripheral edges, has been used. Generally, the larger the amount of the work taken or the greater the importance of the portion taken, the more likely a fair use defense will fail. Courts must balance the amount and importance of the use with the goals of copyright. If a great amount of the work was taken and it was an important piece of the work, then the fair use doctrine would not be a defense. However, if a small, unimportant portion of the work was used, then a fair use defense would probably succeed. Each of these premises must be weighed against the risk of injury to the owner.

Judge Leval suggests that this third factor should be considered with the "purpose and character of the secondary use" factor and the "effect . . . upon the potential market" factor. When combined with the purpose factor, the amount and importance factor should be judged as either reasonable or unreasonable "in relation to the purported justification." In other words, if a secondary use copies the heart of an original work for criticism purposes (a legitimate fair use defense) then such copying *might* be allowed under fair use. The decision will depend on a balancing of this factor with the others. The

<sup>&</sup>lt;sup>163</sup> See id. at 1119. "To suggest that simply because a written document is unpublished, fair use of that document is forbidden . . . has no logical support in . . . copyright law." *Id*.

<sup>&</sup>lt;sup>164</sup> See id.

<sup>&</sup>lt;sup>165</sup> See id. at 1122–23; Martin-Jones, supra note 131.

<sup>&</sup>lt;sup>166</sup> Leval, *supra* note 114, at 1122.

<sup>&</sup>lt;sup>167</sup> See id. at 1124.

<sup>&</sup>lt;sup>168</sup> See NIMMER, supra note 136, § 13.05 [A][3], at 13-192.

<sup>&</sup>lt;sup>169</sup> See id

<sup>&</sup>lt;sup>170</sup> See Leval, supra note 114, at 1124.

<sup>&</sup>lt;sup>171</sup> Id. at 1123–24.

<sup>&</sup>lt;sup>172</sup> *Id.* at 1123.

<sup>173</sup> However, it is unlikely that this individual factor would claim a fair use victory for the secondary author. *See, e.g.*, Suntrust Bank v. Houghton Mifflin Co., 268 F.3d 1257, (continued)

The final factor, "effect . . . upon the potential market," is an inquiry into whether the copyright owner would lose revenue if the use is found to be fair; if the copyright owner loses money, the use is likely not fair. 176 The Supreme Court has declared the fourth factor "undoubtedly the single most important element of fair use." It should be noted that while the importance of the market's reaction to a secondary use is significant, this seems to only be the case when the secondary use does indeed interfere with the market for the original copyright owner and provides a strong incentive against fair use. 178 However, if the secondary use does not interfere with the market, then there is still no guarantee that the court will find fair use. 179 The reasoning for the decision will likely hinge on the justification of the purpose and character of the use. 180 The importance of viewing the factors together is underscored by the intricacies courts face in each of the factors. These factors are typically not examined in isolation though. 181 The outcome still turns on the balancing of all four factors, although other factors may enter the decision. 182

### 3. Case Law Interpreting Fair Use

In an 1841 case regarding copyright infringement of a book detailing the life and providing the documents of former President George Washington, the plaintiff-publishers sued the defendant-booksellers for compiling and selling their own similar book, of which over three hundred pages were copied out of the book the publishers had undertaken to publish. The narrow legal issue centered on the fact that the defendant's book contained large amounts of Washington's documents, from the

1268–69 (11th Cir. 2001) (finding a broad interpretation of what constitutes parody helps determine the criticism issue for fair use).

<sup>&</sup>lt;sup>174</sup> ROCHELLE COOPER DREYFUSS & ROBERTA ROSENTHAL KWALL, INTELLECTUAL PROPERTY: TRADEMARK, COPYRIGHT AND PATENT LAW 412 (2d ed. 2004).

<sup>&</sup>lt;sup>175</sup> Leval, *supra* note 114, at 1124.

<sup>&</sup>lt;sup>176</sup> See Martin-Jones, supra note 131.

<sup>&</sup>lt;sup>177</sup> Harper & Row, 471 U.S. 539, 566 (1985).

<sup>&</sup>lt;sup>178</sup> Cf. Leval, supra note 114, at 1124 (claiming that the Supreme Court has overstated the importance of the fourth factor).

<sup>&</sup>lt;sup>179</sup> *Id*.

<sup>&</sup>lt;sup>180</sup> See id.

<sup>&</sup>lt;sup>181</sup> DREYFUSS & KWALL, *supra* note 174, at 412.

<sup>&</sup>lt;sup>182</sup> Id.

<sup>&</sup>lt;sup>183</sup> See Folsom v. Marsh, 9 F. Cas. 342 (Mass. Dist. Ct. 1841) (No. 4901).

collection in the plaintiff's book. 184 Those documents had been published for the first time in the plaintiff's book which contained a narrative of Washington's life and volumes of his private documents, with the occasional explanatory phrase or editorial note. 185 The defendant's book contained a narrative of Washington's life, obviously written by a different person, which interspersed the narrative with pirated copies of Washington's documents, to make it look like Washington was telling the story himself through his letters and documents. 186

The problem arose because the plaintiff had a copyright in the Washington documents through the editor. The editor compiling the book in which the plaintiff had an interest, had acquired a copyright interest from Washington's nephew. The defendant claimed that a right existed under fair abridgement (the predecessor to fair use) to copy and publish the Washington documents. The claim rested on the fact that Congress had purchased the documents from the two people with a copyright interest in the documents. Justice Story dismissed this claim because the government's purchase was "subject to the copyright already acquired by the plaintiffs in the publication thereof." Later in his opinion, he provided a now-famous quote that became the basis of the 1976 Copyright Act: 192

In short, we must often, in deciding questions of this sort, look to the nature and objects of the selections made, the quantity and value of the materials used, and the degree in which the use may prejudice the sale, or diminish the profits, or supersede the objects, of the original work. <sup>193</sup>

This was essentially a summary of the weighing process Justice Story used to arrive at the decision that this case presented a copyright infringement and that fair abridgement did not constitute a defense. 194

```
<sup>184</sup> Id. at 345.
```

<sup>&</sup>lt;sup>185</sup> *Id*.

<sup>&</sup>lt;sup>186</sup> Id.

<sup>&</sup>lt;sup>187</sup> See id.

<sup>&</sup>lt;sup>188</sup> *Id*.

<sup>&</sup>lt;sup>189</sup> See id.; Leval, supra note 114, at 1105.

<sup>&</sup>lt;sup>190</sup> See Folsom, 9 F. Cas. at 345.

<sup>&</sup>lt;sup>191</sup> *Id.* at 347.

<sup>&</sup>lt;sup>192</sup> Leval, *supra* note 114, at 1105.

<sup>&</sup>lt;sup>193</sup> Folsom, 9 F. Cas. at 348.

<sup>&</sup>lt;sup>194</sup> See id. at 348–49.

The next truly significant case for modern jurisprudence that involved the development of fair use is the *Sony* case. The problem was videotape recorders, a technological advancement that made it possible for people to record television programs and view them at a later time. Sony was selling the recorders, and consumers were using them to record copyrighted television programs. Universal and Disney owned the copyright to some of these television programs. In the lawsuit, Universal claimed that its exclusive copyrights were being infringed. Essentially, Universal was trying to impose infringement liability upon the distributor of the recorders, not upon the people actually recording the programs. The problem was videotape and view that its exclusive copyrights.

"Time-shifting," the process of recording a program to view at a later time, <sup>201</sup> became an important phrase during the litigation of this case and would also be considered in future cases. <sup>202</sup> The Supreme Court found "time-shifting" to be the primary use of the recorders, based on a study by both Universal and Sony. <sup>203</sup> Turning to the next section of its analysis, the Court found that this case was really about contributory infringement, because Sony was merely the distributor of the allegedly infringing, time-shifting device. <sup>204</sup>

In this case, fair use was forced into the framework of the opinion.<sup>205</sup> One of the issues underlying fair use was whether Sony had control over what the infringers were doing.<sup>206</sup> Discussing a case from the first half of

```
<sup>195</sup> See Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417 (1984).
```

<sup>&</sup>lt;sup>196</sup> Id. at 422–23.

<sup>&</sup>lt;sup>197</sup> *Id.* at 423.

<sup>&</sup>lt;sup>198</sup> Id. at 421, 423 n.3.

<sup>&</sup>lt;sup>199</sup> *Id.* at 420.

<sup>&</sup>lt;sup>200</sup> Id.

<sup>&</sup>lt;sup>201</sup> See id. at 423.

<sup>&</sup>lt;sup>202</sup> See, e.g., MGM Studios Inc. v. Grokster, Ltd., 545 U.S. 913 (2005).

<sup>&</sup>lt;sup>203</sup> Sonv, 464 U.S. at 423.

<sup>&</sup>lt;sup>204</sup> *Id.* at 434–35. Contributory infringement is not specifically mentioned in the Copyright Act of 1976. *Id.* However, "[t]he absence of such express language in the copyright statute does not preclude the imposition of liability for copyright infringements on certain parties who have not themselves engaged in the infringing activity." *Id.* at 435. An infringer is usually thought of as one who uses a work without the permission of the creator or owner. *Id.* at 435 n.17. This is an overly simplistic idea of an infringer, as infringers can also be those who encourage third parties to commit infringement. *Id.* 

<sup>&</sup>lt;sup>205</sup> See Martin-Jones, supra note 131.

<sup>&</sup>lt;sup>206</sup> See Sony, 464 U.S. at 437–38.

the century, the Court noted that a producer, who was found to be a contributory infringer, had control over the use of his film because he supplied not only the means, but also the infringing work—the film.<sup>207</sup> Then, distinguishing this early case, the Court said Sony's "control" was limited to the contact with the consumers at the point of sale, and therefore Sony did not have control over the infringing activities.<sup>208</sup> The Court declared that because the recorders were "capable of substantial noninfringing uses," Sony could not be liable for contributory infringement.<sup>209</sup> It would seem this decision would be the final holding and bring closure to the case. However, the Court needed to address fair use in its analysis, so the Court tacked this analysis on to the end of the case.<sup>210</sup>

In beginning to explain the doctrine, the Court was quick to point out that the applicable statutory section of the Copyright Act, § 107, is not a rigid test but a factoring test that essentially allows a Court to decide a fair use question on the facts of the case.<sup>211</sup> The Court then addressed each of the four statutory factors of fair use. 212 For the first factor, purpose of the use, the Court focused more on the commercial-noncommercial dichotomy, rather than the commercial-nonprofit educational dichotomy. <sup>213</sup> The Court held that if the recorders were being used for a profitable purpose, such as making copies of movies and shows for commercial profit, then the use would definitely be unfair.<sup>214</sup> Because it had been factually established that the recorders were being used for time-shifting purposes in the home, to watch programs that could not be viewed during the broadcast time, the use was declared nonprofit. 215 The Court addressed the second factor—the nature of the copyrighted work—with the admonition that it was a "televised copyrighted audiovisual work";<sup>216</sup> the viewers had been invited to see it.<sup>217</sup> The Court let this description stand

<sup>&</sup>lt;sup>207</sup> See id. at 436 (discussing Kalem Co. v. Harper Bros., 222 U.S. 55 (1911)).

<sup>&</sup>lt;sup>208</sup> See id. at 437–38.

<sup>&</sup>lt;sup>209</sup> *Id.* at 442.

<sup>&</sup>lt;sup>210</sup> See id. at 447-56.

<sup>&</sup>lt;sup>211</sup> Id. at 447–48.

<sup>&</sup>lt;sup>212</sup> See id. at 448–56.

<sup>&</sup>lt;sup>213</sup> See id. at 448–49.

<sup>&</sup>lt;sup>214</sup> *Id.* at 449.

<sup>&</sup>lt;sup>215</sup> *Id*.

<sup>&</sup>lt;sup>216</sup> Id. at 449–50.

<sup>&</sup>lt;sup>217</sup> *Id*.

on its own, without further explanation.<sup>218</sup> Quickly moving to the third factor as an extension of the second, the Court found that because viewers were invited to see the program in its entirety, recording the program in its entirety did not offend the amount factor.<sup>219</sup>

Finally, considering the effect on the market, the Court stated that even copying done for noncommercial reasons could jeopardize the copyright owner's potential rewards for the creation. The Court reiterated that the purpose of copyright is to spur further creative work and that sometimes even noncommercial copying may obstruct this purpose. However, the Court held that because there was "no demonstrable effect upon the potential market[,]... the copyrighted work need not be prohibited."

After going through these four factors, the Court returned to the commercial versus nonprofit comparison and emphasized its importance in determining the effect on the market factor. The Court concluded that without showing a likelihood of substantial harm to the market for the original use, the secondary use of the video recorders was noninfringing and therefore fell under fair use. Looking at all of the factors together, the Court held that time-shifting in one's home is fair use. It further held that Sony's distribution of recorders did not constitute contributory infringement because the primary use was a fair use. This case has survived many attempts to weaken its position of fair use.

The next case to add to the fair use analysis is one that added to the jurisprudence on contributory infringement, an issue where fair use is typically an underlying premise. A Hollywood studio and others sued Grokster and StreamCast Networks, Inc., the distributors of free software that aided customers in their quest to download copyrighted music, thereby creating significant infringements on the copyrights. Like *Sony*, this case also focused on contributory infringement, but then it expanded upon

```
<sup>218</sup> See id.
```

<sup>&</sup>lt;sup>219</sup> *Id*.

<sup>&</sup>lt;sup>220</sup> Id. at 450.

<sup>&</sup>lt;sup>221</sup> *Id*.

<sup>&</sup>lt;sup>222</sup> Id.

<sup>&</sup>lt;sup>223</sup> *Id.* at 451.

<sup>&</sup>lt;sup>224</sup> See id. at 451–55.

<sup>&</sup>lt;sup>225</sup> *Id.* at 454–55.

<sup>&</sup>lt;sup>226</sup> See id. at 456.

<sup>&</sup>lt;sup>227</sup> See Martin-Jones, supra note 131.

<sup>&</sup>lt;sup>228</sup> See MGM Studios Inc. v. Grokster, Ltd., 545 U.S. 913, 919–21 (2005).

the legal understanding of fair use as applied to a product with infringing and noninfringing capabilities. <sup>229</sup>

Grokster and StreamCast distributed free peer-to-peer network software. This software allowed multiple users to share files directly among their computers without having to use a central server. Even though this technology had become indispensable to the government and researchers, the technology had also made it easier for copyright infringers to download music and video files. The plaintiffs sued Grokster and StreamCast for contributorily infringing on their copyrights by freely providing software that they knew was capable of such infringement. The software that they knew was capable of such infringement.

The Court added to the legal analysis of contributory infringement and the underlying fair use considerations by introducing the inducement doctrine. The *Sony* Court's view of contributory infringement was that "the sale of . . . articles of commerce[] does not constitute contributory infringement if the product is widely used for legitimate, unobjectionable purposes." In this case, the Court went a step further with the idea that even if the product is capable of legitimate noninfringing uses, and therefore not presumptively infringing, the intentional inducement to commit infringement will still make a party liable for contributory infringement. The Court then began looking at the intent of Grokster and StreamCast to induce infringing activities and found such intent to be prevalent in the business models of each company. Therefore, the Court held that the distribution of a device that can infringe upon copyrights, coupled with the promotion of such capability, makes the distributor liable for third-party infringement. The court held that the distribution of such capability, makes the distributor liable for third-party infringement.

The final case to highlight fair use, the autobiography case of former President Gerald Ford, showcases modern jurisprudence in analyzing the four factors. The former President was working with Harper & Row, Publishers, Inc. to publish his autobiography.<sup>239</sup> Ford gave Harper & Row

```
<sup>229</sup> See id. at 928-41.
```

<sup>&</sup>lt;sup>230</sup> *Id.* at 919.

<sup>&</sup>lt;sup>231</sup> Id. at 919–20.

<sup>&</sup>lt;sup>232</sup> See id. at 920–21.

<sup>&</sup>lt;sup>233</sup> Id.

<sup>&</sup>lt;sup>234</sup> See id. at 936–37.

<sup>&</sup>lt;sup>235</sup> Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 442 (1984).

<sup>&</sup>lt;sup>236</sup> MGM, 545 U.S. at 934–35.

<sup>&</sup>lt;sup>237</sup> See id. at 938.

<sup>&</sup>lt;sup>238</sup> *Id.* at 941.

<sup>&</sup>lt;sup>239</sup> Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 542 (1985).

ownership of the copyright in exchange for the publisher's work to produce and advertise the book. Time Magazine purchased an exclusive deal with the copyright holders, Harper & Row, to print excerpts from the book before the book was published. Somehow, The Nation magazine obtained Ford's manuscript and an editor created a similar excerpt piece, quoting verbatim approximately three hundred words, which ran prior to Time's scheduled release. Time cancelled the agreement with Harper & Row. Harper & Row then sued The Nation, claiming copyright infringement. The Nation claimed the copyrighted material fell under fair use.

The Supreme Court first explained that the statutory factors are a codification of the common law and that the statute was not meant to expand or narrow the common law. The opinion of the Court also mentioned several times that fair use cases must be decided on a case-specific basis. Once beginning the factor analysis, the Court noted that news reporting is one of the suggested reasons the fair use doctrine should be asserted. But it warned against declaring these reasons as indicative of a presumption of fair use. The Nation went beyond simply reporting uncopyrightable information and actively sought to exploit the headline value of its infringement, making a 'news event' out of its unauthorized first publication of a noted figure's copyrighted expression.

Expanding upon the previous analyses of the purpose factor, the Court explained that the commercial-noncommercial dichotomy cannot be decided based upon whether the intent was to make a profit.<sup>251</sup> Rather, the issue turns on whether a potential profit will be realized "from exploitation of the copyrighted material without paying the customary price."<sup>252</sup> The

<sup>&</sup>lt;sup>240</sup> *Id.* at 542–43.

<sup>&</sup>lt;sup>241</sup> *Id*.

<sup>&</sup>lt;sup>242</sup> See id. at 542–45.

<sup>&</sup>lt;sup>243</sup> *Id.* at 542.

<sup>&</sup>lt;sup>244</sup> *Id*.

<sup>&</sup>lt;sup>245</sup> *Id*.

<sup>&</sup>lt;sup>246</sup> *Id.* at 549.

<sup>&</sup>lt;sup>247</sup> See, e.g., id. at 549, 561.

<sup>&</sup>lt;sup>248</sup> See id. at 561.

<sup>&</sup>lt;sup>249</sup> See id.

<sup>&</sup>lt;sup>250</sup> Id.

<sup>&</sup>lt;sup>251</sup> *Id.* at 562.

<sup>&</sup>lt;sup>252</sup> *Id*.

propriety of the secondary user's conduct should also be considered under this factor. <sup>253</sup> Good faith is expected in a fair use defense. <sup>254</sup>

As to the second factor—the nature of the copyrighted work—the Court characterized the nature of the book to be published and then discussed how even though it is factual, thereby implying a greater chance of a fair use defense, there are levels of factual works. A basic map may be completely factual, but a stately autobiography may contain "expressive language." The Court analyzed the extent to which the expressive language detracted from the fair use defense and found that small quotes and phrases essential to the expression of the facts might be permitted, but copying the descriptive work went beyond simply communicating the facts. 257

The third factor, amount and substantiality of the use, provided a clear example of when use of a small amount can be outweighed by conveying "the heart" of the work. In actuality, the number of words copied from the original work "were an insubstantial portion." But, the words copied were the most pertinent and intriguing passages of the book. The Court declared that the passages used by The Nation were used "precisely because they qualitatively embodied Ford's distinctive expression." This resulted in a decision that the excerpts used in the infringing work were substantial.

The fourth factor, the effect on the market, was manifested in actual damages incurred by the publisher.<sup>263</sup> When Time cancelled the agreement to pay the remaining installment of funds to the publisher for the exclusive right to print the excerpts, the publisher suffered actual damages.<sup>264</sup> The Court stated that "once a copyright holder establishes with reasonable probability the existence of a causal connection between the infringement

<sup>&</sup>lt;sup>253</sup> *Id*.

<sup>&</sup>lt;sup>254</sup> See id.

<sup>&</sup>lt;sup>255</sup> See id. at 563.

<sup>&</sup>lt;sup>256</sup> *Id.* (quoting Robert A. Gorman, *Fact or Fancy? The Implications for Copyright*, 29 J. COPYRIGHT SOC'Y 560, 563 (1982)).

<sup>&</sup>lt;sup>257</sup> See id. at 563–64.

<sup>&</sup>lt;sup>258</sup> *Id.* at 564–65.

<sup>&</sup>lt;sup>259</sup> *Id.* at 564.

<sup>&</sup>lt;sup>260</sup> See id. at 564–65.

<sup>&</sup>lt;sup>261</sup> *Id.* at 565.

<sup>&</sup>lt;sup>262</sup> See id. at 566.

<sup>&</sup>lt;sup>263</sup> *Id.* at 567.

<sup>&</sup>lt;sup>264</sup> *Id*.

and a loss of revenue, the burden properly shifts to the infringer to show that this damage would have occurred had there been no taking of copyrighted expression."<sup>265</sup> Because The Nation failed to even rebut this information, the Court held that actual damages were the effect on the market.<sup>266</sup> The Court also found that there was an adverse effect on the market because The Nation "directly competed" for the market of people interested in these excerpts.<sup>267</sup>

#### II. ANALYSIS

# A. Application of the Combined Statutory and Case Law Surrounding Fair Use

The historical development of the fair use doctrine, both through case law and the statutory law, has grown without providing future courts specific guidelines on how to analyze a fair use claim. This is the result of a history plagued with vague applications of the doctrine. The prevailing cases in this area of law demonstrate the difficulty courts have had in interpreting and applying the statutory factors surrounding fair use. However, the courts can arrive at a definitive ruling about whether digital archives meet the fair use exception requirements by taking a careful look at all of the legal issues courts have analyzed in the past surrounding fair use. By comparing how each of the statutory factors has been applied in all of the major cases, courts can arrive at a consensus on how to analyze the fair use issue for new technologies such as digital archives.

However, before reaching the analysis of a fair use defense, there must be a violation of law to defend. In the case of the Internet Archive, the archive must have infringed a copyright before a discussion of whether its activities fall under the fair use doctrine can be initiated. This is because the fair use doctrine is a limitation on what can be found to be infringing. In other words, it is an exception to activity that constitutes copyright infringement. Because the Internet Archive makes or buys

<sup>&</sup>lt;sup>265</sup> *Id*.

<sup>&</sup>lt;sup>266</sup> See id.

<sup>&</sup>lt;sup>267</sup> *Id.* at 568.

<sup>&</sup>lt;sup>268</sup> See Leval, supra note 114, at 1105–06.

<sup>&</sup>lt;sup>269</sup> See, e.g., MGM Studios Inc. v. Grokster, Ltd., 545 U.S. 913, 933–34 (2005); *Harper & Row*, 471 U.S. at 569; Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 420–21 (1983).

<sup>&</sup>lt;sup>270</sup> See 17 U.S.C. § 107 (2000).

copies of websites, regardless of whether they are publicly accessible or not, the Internet Archive is infringing on the copyright of the website owners. The United States Code specifically describes copyrightable materials as "original works of authorship fixed in any tangible medium of expression." The websites that the Internet Archive copies are "original works of authorship" because they display material that, as a whole, constitutes the website or the original work. Further, a website qualifies as a "tangible medium of expression" because it is part of the web, a common and global medium. Because the Internet Archive has technically infringed on copyrights, it can properly raise the fair use defense.

To determine whether a secondary use constitutes a fair use, courts may consider any factors but they "shall include" the four main factors discussed previously.<sup>272</sup> This must be done regardless of whether the infringement meets one of the purposes specifically delineated in the statute, such as criticism or scholarship.<sup>273\*</sup> For example, the Internet Archive prominently places notice on its main web page that it "is opening its collections to researchers, historians, and scholars,"<sup>274</sup> Section 107 of the statute describing fair use can easily be misconstrued as absolving any copyright infringement when done for scholarly purposes, <sup>275</sup> such as work that might be done by researchers, historians, scholars, and the general public. However, even if a use is deemed to fall into one of the enumerated categories like scholarship, research, criticism, or news reporting, the use must still undergo analysis using the four factors.<sup>276</sup> Therefore, regardless of the intent that might be attributed to the Internet Archive concerning potential uses of the material it stores, an analysis of the four fair use factors must still be conducted.

## B. The First Factor as Applied to the Internet Archive

The first factor, the purpose and character of the use, has been considered in the contexts of the commercial versus nonprofit educational purposes, <sup>277</sup> the commercial versus noncommercial dichotomy, <sup>278</sup> the

<sup>&</sup>lt;sup>271</sup> 17 U.S.C. § 102(a).

<sup>&</sup>lt;sup>272</sup> 17 U.S.C. § 107.

<sup>&</sup>lt;sup>273</sup> See NIMMER, supra note 136, § 13.05[A].

<sup>&</sup>lt;sup>274</sup> About IA, *supra* note 3.

<sup>&</sup>lt;sup>275</sup> See 17 U.S.C. § 107(1).

<sup>&</sup>lt;sup>276</sup> See NIMMER, supra note 136, § 13.05[A].

<sup>&</sup>lt;sup>277</sup> See Martin-Jones, supra note 131.

<sup>&</sup>lt;sup>278</sup> See Sony Corp. of Am. v. Universal Studios, Inc., 464 U.S. 417, 449 (1984).

justification for the use,<sup>279</sup> the transformative nature of the work,<sup>280</sup> and whether the profit will be realized from inappropriate exploitation.<sup>281</sup> A continuum is perhaps the best way to visualize a guideline combining all of these contexts. One side of the continuum represents uses that will not be found to be a fair use. The most extreme example is a use for pure commercial activity that results solely in commercial profit.<sup>282</sup> Included on this far side of the continuum are instances where the copyright owner was never compensated and is therefore being exploited.<sup>283</sup> The opposite side of the continuum, where fair use will be found to exist, represents the uses that are for nonprofit educational purposes, completely void of commercial activity, profit, and exploitation. 284 The area in the middle of the continuum is where the analysis of the justification for the secondary use and of the transformative nature of the use should occur. The justification for finding a secondary use to be a fair use is supplied when the use meets the purpose of copyright law by stimulating creativity.<sup>285</sup> transformative nature of the secondary use will equate to a fair use when the original work has been used as a foundation for new insights. criticisms, or any other substantive alteration of the work.<sup>286</sup> To the contrary, justification and the transformative nature of a work will gravitate toward the other end of the spectrum when creativity is not stimulated and transformation of the work has not occurred.<sup>287</sup>

Applying the secondary use claimed by the Internet Archive to this continuum for the first factor helps provide a very clear picture as to how courts should rule on digital archives claiming fair use. The Internet Archive is a registered nonprofit organization. Its goal is to provide an "internet library" that "foster[s] education and scholarship in this era of digital technology." Therefore, the Internet Archive's first placement on the continuum is to the far side with the nonprofit educational purpose that

<sup>&</sup>lt;sup>279</sup> See Leval, supra note 114, at 1111.

<sup>&</sup>lt;sup>280</sup> See id.; Martin-Jones, supra note 131.

<sup>&</sup>lt;sup>281</sup> See Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 562 (1985).

<sup>&</sup>lt;sup>282</sup> See NIMMER, supra note 136, § 13.05[A][1][c].

<sup>&</sup>lt;sup>283</sup> See Harper & Row, 471 U.S. at 562.

<sup>&</sup>lt;sup>284</sup> See generally Martin-Jones, supra note 131; cf. NIMMER, supra note 136, § 13.05[A][1][c]; Harper & Row, 471 U.S. at 562.

<sup>&</sup>lt;sup>285</sup> See Leval, supra note 114, at 1111.

 $<sup>^{286}</sup>$  See id.

<sup>&</sup>lt;sup>287</sup> Cf. id.

<sup>&</sup>lt;sup>288</sup> See About IA, supra note 3.

<sup>&</sup>lt;sup>289</sup> Id.

definitively meets the fair use exception. Determining whether its use stays at that point on the continuum requires an inquiry into the potential commercial activity, profit, and exploitation by the Internet Archive. The funding for the organization comes from "in-kind and financial donations" by well-known foundations and from the general public.<sup>290</sup> There are no commercial advertisements on the website, for the purposes of financial profit.<sup>291</sup> The Internet Archive is not selling the information to third parties.<sup>292</sup> It is not receiving revenues from advertisers that typically display their ads on websites, because there are no advertisers. Rather, the Internet Archive's only source of income is that which flows through its nonprofit organization.<sup>293</sup> It does not exploit the information it offers for the purposes of profit. Therefore, the Internet Archive's place on the continuum does not move from its original position under fair use.

The justification of the use is the next issue to consider when determining the Internet Archive's place on the continuum. The Internet Archive justifies its use of copyrighted material through its mission statement. It states that it "was founded to build an Internet Library, with the purpose of offering permanent access for researchers, historians, and scholars." Courts would need to decide if this justification rises to the level of supporting public creativity. It would likely be difficult for courts to find that providing access to copyrighted material for such lofty purposes is not a justification. On the other hand, good faith requires admitting that easy internet access to copyrighted material is merely that: easy access. Therefore, the Internet Archive's position on the continuum moves slightly away from the perfect example of a fair use.

Finally, the transformative nature of the use is the last issue to consider under this first factor. Here, the requirement is that the original work is used as the basis for the new, different work.<sup>296</sup> The Internet Archive really acts as more of a conduit, providing the original work to other authors that might create a new work. The Internet Archive is not creating a new work, it is simply copying the original work and making it available

<sup>&</sup>lt;sup>290</sup> See FAQs, supra note 4.

<sup>&</sup>lt;sup>291</sup> See generally Internet Archive, http://www.archive.org/index.php (last visited Nov. 22, 2007).

<sup>&</sup>lt;sup>292</sup> Cf. About IA, *supra* note 3 ("The Archive has no vested interest in the discoveries of the users of its collections . . . .").

<sup>&</sup>lt;sup>293</sup> See id.

<sup>&</sup>lt;sup>294</sup> Id.

<sup>&</sup>lt;sup>295</sup> See Leval, supra note 114, at 1111.

<sup>&</sup>lt;sup>296</sup> See id.

to third parties. Through the Wayback Machine, where the URL for a website is simply typed into a text box, the public can retrieve all the dates on which the website and its pages were copied by the Internet Archive. <sup>297</sup> Then, by clicking on a date link, web pages can be viewed as they existed on that date. <sup>298</sup> This does not amount to a transformative use. Therefore, for this one issue under the first factor, the Internet Archive's position on the continuum moves further away from the ideal fair use claim. However, combining all of the issues analyzed under the first factor indicates that the Internet Archive's use probably does not cross the middle point of the continuum which would put it in the realm of not being able to claim fair use.

#### C. The Second Factor as Applied to the Internet Archive

The second factor, the nature of the work, is determined by issues that cover whether the original work is published or unpublished, <sup>299</sup> whether the right of first publication has been exercised, <sup>300</sup> whether it is creative or factual in nature, <sup>301</sup> whether others were invited to use it, <sup>302</sup> and whether the use contains expressive language. <sup>303</sup>

The published-unpublished inquiry is merely a perfunctory designation of the work as being either one or the other. A copyright can be claimed in either published or unpublished works. The publication issue really turns on the right of first publication. This is the author's decision as to whether or not publication should be pursued and, if so, in what format. If the author has the chance to exercise this right, then the fair use defense may well turn on other issues and factors. However, if the author never has the chance to exercise this right, because the original work was

See, e.g., www.archive.org/index.php. Once on this web page, type www.loc.gov into the Wayback Machine text box and click the "Take Me Back" button. The results page will display dates beginning with December 11, 1997. By clicking on any of these dates, a view of the web pages for the Library of Congress is available as they existed on that date.

<sup>&</sup>lt;sup>298</sup> Id

<sup>&</sup>lt;sup>299</sup> See Martin-Jones, supra note 131.

<sup>&</sup>lt;sup>300</sup> See Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 564 (1985).

<sup>&</sup>lt;sup>301</sup> See Martin-Jones, supra note 131.

<sup>&</sup>lt;sup>302</sup> See Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 449–50 (1984).

<sup>&</sup>lt;sup>303</sup> See Harper & Row, 471 U.S. at 563-64.

<sup>&</sup>lt;sup>304</sup> See Martin-Jones, supra note 131.

<sup>&</sup>lt;sup>305</sup> See Harper & Row, 471 U.S. at 564.

 $<sup>^{306}</sup>$  See id.

proliferated without the author's consent, then a fair use defense will likely be scrutinized more harshly. 307

The nature of the work is the copying of already published and publicly accessible web pages to make them available in a historical collection. The right of first publication becomes a moot issue because the material has already been published, in a sense. Therefore, a fair use defense can easily be raised. The material has not only been published on the web pages, it has been made easily accessible to the entire world. The Internet Archive is cataloguing these web pages, for preservation. Under these issues revolving around publication, the Internet Archive can raise the fair use defense and should prevail.

The next issue addresses whether the nature of the work is more creative or more factual. Because the Internet Archive simply makes copies, the nature of its work is neither creative nor factual. The Internet Archive does copy whole web pages without filtering content, so some of those works are creative and some are factual. However, these labels are the nature of the original works, not the nature of the work by the Internet Archive. There is really no claim as to whether the Internet Archive's "work" is creative or factual in nature. The Internet Archive is, again, acting as a conduit so that others may have easier access to works that will encourage such illumination of the public. The Internet Archive's fair use claim should still succeed because there is no way to judge the nature of the work as either creative or factual.

The fact that others were previously invited to view the web pages that were later copied by the Internet Archive is the next issue. Much of the recent litigation involving the Internet Archive and its Wayback Machine stems from this small issue. Playboy and Dell used the Wayback Machine to look at old versions of the opposing party's website and prove the opposing party wrong. EchoStar was able to get a judge to declare that copies of the old versions of web pages are admissible as evidence. The currently pending *Healthcare Advocates* case involves the fight to keep

<sup>&</sup>lt;sup>307</sup> See id. at 551.

<sup>&</sup>lt;sup>308</sup> See About IA, supra note 3.

<sup>&</sup>lt;sup>309</sup> See id.

<sup>&</sup>lt;sup>310</sup> See supra notes 50 and 83 and accompanying text.

<sup>&</sup>lt;sup>311</sup> See Internet Archive's Web Page Snapshots Held Admissible as Evidence, http://cyberlaw.stanford.edu/packets002728.shtml (Nov. 17 2004, 11:09 EST).

previously viewed web pages from being copied by the Internet Archive.<sup>312</sup> The Sony case barely gave mention to this issue, simply noting that viewers were invited by the television broadcast<sup>313</sup> thereby implying that later claims that viewers should not view the broadcast were not worth the Court's time. Even though the medium is different in the case of the Internet Archive, the reasoning is the same. The global community was invited to view these web pages previously, without restriction. The public was impliedly invited to view the web pages as they existed on the date the Internet Archive made a copy of them. Viewing these web pages at a later date does not change the fact that originally the material on the web pages was made available for public consumption. The owners of the web pages may claim that they reserve the right to withhold the use of such copyrighted material at a later date, say after they have changed the web pages. This is true if someone wanted to copy the material and distribute it for a reason other than what is acceptable under fair use. However, the Internet archive is copying this information and making it available for purposes such as research and commentary. Under this issue, the Internet Archive's use supports a finding of fair use.

The final issue for the second factor is whether the use contains expressive language. In the *Harper & Row* case, small amounts of expressive language that were essential to the expression of the idea were permitted under fair use.<sup>314</sup> Here, the Internet Archive cannot argue that its use qualifies as such. The Internet Archive copies the whole web page. It does not act as the editor did in the *Harper & Row* case. Under this issue, the Internet Archive's claim cannot support a fair use determination.

Overall, under the factor of nature of the use, the Internet Archive can prevail in terms of publication issues and previous invitations to view the web pages. The other two issues of creative or factual nature and the use of expressive language may or may not detract greatly from the other two issues. This would be an area for the trier of fact.

#### D. The Third Factor as Applied to the Internet Archive

The third factor, the amount and substantiality of the work used, has been construed as an inquiry into whether the heart of the work has been

<sup>&</sup>lt;sup>312</sup> See Complaint at 2, Healthcare Advocates, Inc. v. Harding, Earley, Follmer & Frailey, No. 05-CV-03524, 2005 WL 2151678, (E.D. Pa. July 8, 2005).

<sup>&</sup>lt;sup>313</sup> See Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 449–50 (1984).

<sup>&</sup>lt;sup>314</sup> See Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 563 (1985).

copied,<sup>315</sup> regardless of size,<sup>316</sup> and whether previous viewers had viewed the work in its entirety.<sup>317</sup> The heart of the work can be described as a piece of great importance, if not the greatest, to the original work.<sup>318</sup> Courts will typically find that the greater the importance of the piece taken from the work, the less likely the fair use defense will succeed.<sup>319</sup> Such importance can be found in large and small amounts.<sup>320</sup>

The Internet Archive has taken the entire work, so the heart of the work is not at issue. However, courts must really focus on striking a balance between the importance of the use and the goals of copyright.<sup>321</sup> Even though the Internet Archive has taken the whole of the work, the importance in doing this is to create an internet library where digital works previously made available to the world are preserved for future generations. This will really be a public policy decision for the courts.

The use of this factor in the *Sony* case buttresses a conclusion that a policy decision in favor of the Internet Archive would not stray from the law. In *Sony*, the Court found that inviting the viewers to see the program in its entirety did not pose a problem for recording the program in its entirety. Likewise, the website owners' implied invitations to the global community should not pose a problem for the Internet Archive's copying of entire websites.

#### E. The Fourth Factor as Applied to the Internet Archive

The fourth factor, the effect on the marketplace, has been declared the most important factor by the Supreme Court.<sup>323</sup> At its most basic function, this factor represents an inquiry into whether the copyright owner would lose revenue should a fair use defense succeed.<sup>324</sup> One of the concerns is the market's reaction to the secondary use and whether the reaction interferes with the original copyright owner's market.<sup>325</sup> The Internet

<sup>&</sup>lt;sup>315</sup> See Martin-Jones, supra note 131.

<sup>&</sup>lt;sup>316</sup> See Harper & Row, 471 U.S. at 564–65.

<sup>&</sup>lt;sup>317</sup> See Sony, 464 U.S. at 449–50.

<sup>&</sup>lt;sup>318</sup> See Leval, supra note 114, at 1123.

<sup>&</sup>lt;sup>319</sup> See id. at 1122.

<sup>&</sup>lt;sup>320</sup> See id. at 1124; Harper & Row, 471 U.S. at 565.

<sup>&</sup>lt;sup>321</sup> See Leval, supra note 114, at 1124.

<sup>&</sup>lt;sup>322</sup> See Sony, 464 U.S. at 449-50.

<sup>323</sup> See Harper & Row, 471 U.S. at 566.

<sup>&</sup>lt;sup>324</sup> See Martin-Jones, supra note 131.

<sup>&</sup>lt;sup>325</sup> See Leval, supra note 114, at 1124.

Archive is only copying publicly accessible web pages.<sup>326</sup> These pages have been displayed freely to the public. The Internet Archive's copy merely provides a look at how the web page looked and functioned in the past. This service is not offered for a price; rather, it is free. The copy does not interfere with the market for the original copyright owner, as the original owner made the web page publicly accessible in the first place. In *Sony*, the Court held that without a "demonstrable effect upon the potential market" there was no effect in prohibiting the secondary use.<sup>327</sup> In the *Harper & Row* case, the Court looked only at actual damages.<sup>328</sup>

The Internet Archive would have to create a demonstrable effect on the primary market or cause actual damages with its actions before this factor could be used against it. Because the Court has declared this factor the most important of the four, great weight should be given to the fact that the Internet Archive has not caused an effect on the market with its copying of publicly accessible web pages.

#### **CONCLUSION**

The Internet Archive and its Wayback Machine have preserved many of the web pages that have been or are accessible by the public. The goal is to preserve historical information. Because disagreement exists as to whether this goal should be allowed under copyright law, courts must conduct a detailed fair use analysis that incorporates the four factors and all of their applications in the seminal cases. By doing so, there is an opportunity for courts to definitively declare whether digital archives fall within the fair use exception.

The first factor indicates that the Internet Archive's activities fall closer to a designation of fair use than not when thought of on a continuum. The second factor requires the trier of fact to assign a weight to the issues of publication and previous viewing as opposed to the more abstract issues like creativeness. The third factor requires a policy decision, but it likely will be decided in favor of the Internet Archive because of major holdings in two of the prominent cases. The fourth and most important factor shows that the Internet Archive's activities fall within the fair use exception as there has been no demonstrable effect on an original market.

<sup>326</sup> See FAQs, supra note 4.

<sup>&</sup>lt;sup>327</sup> Sony, 464 U.S. at 450.

<sup>328</sup> See Harper & Row, 471 U.S. at 567.

These factors must be taken as a whole and not viewed in isolation because of their intricacies. When taken as a whole, the Internet Archive should prevail on its fair use claim. The courts should apply the factors with the combined interpretations and find that the use is a fair use. Making such a definitive ruling will prevent future cases from leading us comically awry, as did Mr. Peabody and Sherman.

<sup>&</sup>lt;sup>329</sup> See Dreyfuss & Kwall, supra note 174, at 412.