

# BAR BULLETIN

## The “Wayback Machine” – Yes, Virginia, You Can Time Travel

**By Larry G. Johnson**

Thomas Wolfe famously wrote that “you can’t go home again,” but with the “Wayback Machine,” maintained by the nonprofit Internet Archive at [www.archive.org](http://www.archive.org), you can do just that.

Billions of website pages have routinely been “visited” by the Wayback Machine through what are known as “bots” that suck up all of a website’s content into a giant database that is so huge it has to be measured in “petabytes” (a million billion bytes, or a quadrillion bytes). As of December 2014, it “contained almost nine petabytes of data and was growing at a rate of about 20 terabytes each week.”<sup>1</sup>

The Wayback Machine has been randomly archiving cached pages of websites since 1996, revisiting sites every few weeks or months to archive a new version to add to previous ones.

The [archive.org](http://archive.org) tool is very easy to use. At the top of the home page under “Wayback Machine” is an input box where you can type in the web address (known as a URL<sup>2</sup>) for any website of interest to you, and up pops a timeline with vertical bars indicating the dates on which the website was visited.<sup>3</sup>

For example, my website, [www.e-dataevidence.com](http://www.e-dataevidence.com), has been visited 16 times between August 29, 2007, and October 2, 2016.

Beneath this timeline is a standard calendar, with the dates highlighted by a blue circle when my website was “crawled.” Clicking on the blue circle reveals not only the date, but also the time of day the bot was there. After a short pause you are then taken to a rec-

reation of my website as it looked on that date. Magic!

### How Can Lawyers Use the Wayback Machine?

With data preserved from days gone by, the Wayback Machine has the potential of offering a number of useful opportunities for lawyers in different practice areas. Here are just a few examples that come readily to mind:

- Copyright violations: finding instances of unauthorized use of copyrighted works. I personally benefitted from the Wayback Machine in two instances where I prevailed in proving copyright infringements of a book I wrote entitled *The Deposition Guide, a Practical Handbook for Witnesses*.<sup>4</sup> In both cases, a law firm had purloined most of the book’s content and passed it off as its own.<sup>5</sup> The Wayback Machine showed when the firm had first posted the material on its website, and then when it was removed from the website after I sent a cease and desist letter.

- Prior false statements in misrepresentation and fraud cases.

- Patent law: “Provided some additional requirements are met (e.g., providing an authoritative statement of the archivist), the United States patent office and the European Patent Office will accept date stamps from the Internet Archive as evidence of when a given Web page was accessible to the public. These dates are used to determine if a Web page is available as prior art for instance in examining a patent application.”<sup>6</sup>

- Divorce: Where the issue is a

spouse’s ability to earn a sufficient living on his/her own, the spouse’s glowing CVs found on prior websites where he/she used to work could be helpful.

- For litigators, archived websites are valuable in finding former employees of an enterprise whom one would be ethically free to interview directly and who may be disgruntled and willing to help you in your case. Also, in planning for e-discovery of electronically stored data from a particular timeframe, the Wayback Machine can retrieve a website’s past open job postings to determine what kinds of skill sets are sought for its IT personnel. This can reveal the kinds of technology and software that have been used by an enterprise, such as the email software product used, thus making it possible to narrow your e-discovery requests.

Additionally, for your own firm’s HR purposes you can use a job applicant’s stated prior employment history and the Wayback Machine to verify times and places where that person formerly worked and to substantiate what that person’s duties were.

### Some Evidentiary Caveats

While the Wayback Machine is an amazing tool, I have found it occasionally to have some drawbacks. Some old websites are not completely archived, i.e., you may find a number of places on a web page reserved for graphics where the images are missing.<sup>7</sup> I am not sure why that is, but if you encounter that phenomenon, you may have the same evidentiary problems you’d have with a

paper document fragment.

There is also at least one case that found evidence from the Wayback Machine to be inadmissible. In *Telewizja Polska USA, Inc. v. Echostar Satellite*,<sup>8</sup> the plaintiff sought to exclude snapshots of web pages from the Wayback Machine as hearsay and unauthenticated. The trial judge held that neither the affidavit of the Internet Archive employee nor the underlying pages (i.e., the Telewizja Polska website) were admissible as evidence. The judge “reasoned that the employee’s affidavit contained both hearsay and inconclusive supporting statements, and the purported web page printouts were not self-authenticating.”<sup>9</sup>

I think a different result would have been obtained if the defendant had offered sufficient expert witness testimony and forensic proof to authenticate the

screenshots and their archived website origins. ■

---

*Larry G. Johnson is a lawyer in Newcastle, and has been a member of the Washington bar since 1974. He recently served on the E-Discovery Subcommittee of the WSBA Escalating Cost of Civil Litigation (ECCL) Task Force. Besides being a litigator, for the past 20 years he has served as a consultant and expert witness in e-discovery matters. He does business as Electronic Data Evidence (www.e-dataevidence.com).*

---

<sup>1</sup> [https://en.m.wikipedia.org/wiki/Wayback\\_Machine](https://en.m.wikipedia.org/wiki/Wayback_Machine). **Editor’s Note:** For those of a younger generation, the original “WABAC Machine” was a feature of the “Peabody’s Improbable History” cartoons on the “Rocky and Bullwinkle Show.”

<sup>2</sup> URL = Uniform Resource Locator. I bet you’ve always wanted to know that.

<sup>3</sup> Though a website is not necessarily archived at each “visit,” as explained by the Internet Archive: “This calendar view maps the number of times http://

www.e-dataevidence.com was crawled by the Wayback Machine, not how many times the site was actually updated.” [www.archive.org](http://www.archive.org). More detailed information about how the Wayback Machine technology works can be found by clicking on the FAQ link on the archive.org home page.

<sup>4</sup> Shameless plug: The e-book version is available on Amazon at [https://www.amazon.com/Deposition-Guide-Practical-Handbook-Witnesses-ebook/dp/B001TOCFZY/ref=sr\\_1\\_2?ie=UTF8&qid=1481069780&sr=8-2&keywords=the+deposition+guide](https://www.amazon.com/Deposition-Guide-Practical-Handbook-Witnesses-ebook/dp/B001TOCFZY/ref=sr_1_2?ie=UTF8&qid=1481069780&sr=8-2&keywords=the+deposition+guide).

<sup>5</sup> Another valuable site to help you find copyright infringements of content from a website is [www.copyscape.com](http://www.copyscape.com). A video on the home page describes its many useful features.

<sup>6</sup> [https://en.m.wikipedia.org/wiki/Wayback\\_Machine](https://en.m.wikipedia.org/wiki/Wayback_Machine). See also *Telewizja Polska USA, Inc. v. Echostar Satellite*, discussed *infra*.

<sup>7</sup> Even so, there will often be text still at that location in lieu of the missing graphic — text included in the HTML language that generated that page that will identify the name of the missing graphic, e.g., “newbuilding.jpg.” A discovery request may then find that graphic among the ESI stored elsewhere by a party on its servers or backup media.

<sup>8</sup> No. 02 C 3293, 65 Fed. R. Evid. Serv. 673 (N.D. Ill., Oct. 15, 2004).

<sup>9</sup> [https://en.m.wikipedia.org/wiki/Wayback\\_Machine](https://en.m.wikipedia.org/wiki/Wayback_Machine)