


7-9-2013

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION: THE LIFE SPAN OF A UNITED STATES SUPREME COURT CITATION CONTAINING AN INTERNET LINK (1996-2010)

Raizel Liebler

June Liebert

Follow this and additional works at: <http://digitalcommons.law.yale.edu/yjolt>

 Part of the [Computer Law Commons](#), [Intellectual Property Commons](#), and the [Science and Technology Commons](#)

Recommended Citation

Liebler, Raizel and Liebert, June (2013) "SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION: THE LIFE SPAN OF A UNITED STATES SUPREME COURT CITATION CONTAINING AN INTERNET LINK (1996-2010)," *Yale Journal of Law and Technology*: Vol. 15: Iss. 2, Article 2.

Available at: <http://digitalcommons.law.yale.edu/yjolt/vol15/iss2/2>

This Article is brought to you for free and open access by Yale Law School Legal Scholarship Repository. It has been accepted for inclusion in Yale Journal of Law and Technology by an authorized administrator of Yale Law School Legal Scholarship Repository. For more information, please contact julian.aiken@yale.edu.

**SOMETHING ROTTEN IN THE STATE OF LEGAL
CITATION: THE LIFE SPAN OF A UNITED STATES
SUPREME COURT CITATION CONTAINING AN INTERNET
LINK (1996-2010)**

Raizel Liebler & June Liebert[†]

15 YALE J.L. & TECH. 273 (2013)

ABSTRACT

Citations are the cornerstone upon which judicial opinions and law review articles stand. Within this context, citations provide for both authorial verification of the original source material at the moment they are used and the needed information for later readers to find the cited source. The ability to check citations and verify that citations to the original sources are accurate is integral to ensuring accurate characterizations of sources and determining where a researcher received information. However, accurate citations do not always mean that a future researcher will be able to find the exact same information as the original researcher. Citations to disappearing websites cause serious problems for future researchers.

Our present mode of citing websites in judicial cases, including within U.S. Supreme Court cases, allows such citations to disappear, becoming inaccessible to future scholars. Without significant change, the information in citations within judicial opinions will be known solely from those citations. Citations to the U.S. Supreme Court are especially important of the Court's position at the top of federal court hierarchy, determining the law of the land, and even influencing the law in international jurisdictions.

Unfortunately and disturbingly, the Supreme Court appears to have a vast problem with link rot, the condition of internet links no longer working. We found that number of websites that are no longer working cited to by Supreme Court opinions is alarmingly high, almost one-third (29%). Our research in Supreme Court cases also found that the rate of disappearance is not affected by the type of online document (pdf, html, etc) or the sources of links (government or non-government) in terms of what links are now dead. We cannot predict what links will rot, even within Supreme Court cases.

[†] Raizel Liebler, Research Services Librarian, John Marshall Law School; June Liebert, Library Director, John Marshall Law School. The authors thank Jessica de Perio Wittman, Ramsey Donnell, William Ford, Tobe Liebert, David Schwartz, Clarissa Simon, Corey Yung, and participants at the Law & Society Annual Meeting and the Midwest Law & Society Retreat for helpful comments. We would like to thank Fang Han, Anna Rostovtsev, and Kimberly Wise for their excellent research assistance as Research Fellows.

TABLE OF CONTENTS

| | |
|--|-----|
| INTRODUCTION | 275 |
| I. OVERVIEW OF THE SUPREME COURT’S CITATION TO WEBSITES | 278 |
| II. LINK ROT & DEAD LINKS | 283 |
| A. <i>Link Rot Generally</i> | 283 |
| B. <i>How Link Rot Affects Law Generally</i> | 286 |
| C. <i>How Link Rot Is Related to Legal Citation</i> | 289 |
| D. <i>How Link Rot Affects Judicial Opinions</i> | 290 |
| E. <i>Prior Studies of Link Rot in Legal Citation</i> | 292 |
| III. CURRENT STUDY—PURPOSE AND RESEARCH QUESTION.... | 293 |
| IV. METHODOLOGY | 293 |
| A. <i>General Methodology</i> | 293 |
| B. <i>Citation Format Problems</i> | 295 |
| V. RESULTS..... | 296 |
| VI. POSSIBLE SOLUTIONS | 299 |
| A. <i>The U.S. Supreme Court Clerk of Courts</i> | 300 |
| B. <i>The Judicial Conference of the United States</i> | 301 |
| C. <i>Non-Court Based Alternative Solutions</i> | 302 |
| 1. <i>The Archival Approach</i> | 302 |
| 2. <i>The Permanent URL Approach</i> | 305 |
| 3. <i>Other Projects of Note</i> | 306 |
| CONCLUSION..... | 307 |
| APPENDIX | 309 |

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

*“If permanence of legal thought is important to legal scholarship then it must be preserved consciously.”*¹

INTRODUCTION

Citations are the cornerstone upon which judicial opinions and law review articles stand. Citations provide both authorial verification of the original source material at the moment they are used and the needed information for readers to later find the cited source.² The ability to check citations and verify that citations to the original sources are accurate is integral to ensuring accurate characterizations of sources and determining where a researcher found information.³ However, accurate citations do not always mean that a future researcher will be able to find the exact same information as the original researcher.⁴ Citations to disappearing websites cause serious problems for legal researchers.⁵ In this

¹ Howard A. Denemark, *The Death of Law Reviews has Been Predicted: What Might be Lost When the Last Law Review Shuts Down?*, 27 SETON HALL L. REV. 1, 12 (1996).

² Daniel J. Baker, *A Jester’s Promenade: Citations to Wikipedia in Law Reviews, 2002-2008*, 7 I/S J.L. & POL’Y 361, 364-65 (2012) (“Generally speaking, all citations serve two main functions. First, citation identifies the resources that the author examined in developing and writing his work; this is the documentary function of citation The second purpose of citation is to help the reader locate (and access) the same resources the author used by providing the necessary bibliographic information.”) (internal citations omitted).

³ Mary I. Coombs, *Lowering One’s Cites: A (Sort of) Review of the University of Chicago Manual of Legal Citation*, 76 VA. L. REV. 1099, 1105-06 (1990) (describing one function of a citation as either “‘to lead its reader to the work cited’ or ‘to give credit for borrowed material [and] to provide the reader with access to research materials.’”) (quoting M. PRICE, A PRACTICAL MANUAL OF STANDARD LEGAL CITATIONS III (2d ed. 1958); and William R. Slomanson, *Footnote Logic in Law Review Writing: Previously Unaddressed in the Criminal Justice System*, 9 CRIM. JUST. J. 65, 68 (1986)).

⁴ Baker, *supra* note 2, at 378 (“However, citations to Wikipedia, like all Internet citations, lack permanence, and [t]his lack of permanence undermines one of the goals of a citation, to provide “‘the information necessary to find and read the cited material.’”) (quoting Paul Axel-Lute, *Legal Citation Form: Theory and Practice*, 75 LAW LIBR. J. 148 (1982); see, e.g., *id.* at 379 (“The content of a Wikipedia entry can change multiple times in one day.”)).

⁵ Lee F. Peoples, *The Citation of Blogs in Judicial Opinions*, 13 TUL. J. TECH. & INTELL. PROP. 39, 71, 72-73 (2010) (“A recent study on blog preservation found that ninety six percent of blog posts are changed after they are initially posted. The study revealed that forty eight percent of bloggers have deleted posts because they no longer held an opinion expressed in the post, and twenty three percent have intentionally deleted an entire blog. . . . While seventy one percent of bloggers surveyed believed their blog should be archived, only twenty-four percent admitted archiving their blog on a regular basis.”).

article, we follow the popular practice of studying legal citations,⁶ including law review citations⁷ and citations in empirical legal analysis,⁸ focusing specifically on the use of Internet links in Supreme Court citations.

As this article will demonstrate, our field's present method of citing websites in judicial cases, including within U.S. Supreme Court cases, allows such citations to disappear, becoming inaccessible to future scholars. Without significant change, the information in citations within judicial opinions will be known solely from those citations. Citations by the U.S. Supreme Court are especially important, given the Court's position at the top of

⁶ See, e.g., Richard Posner & William Landes, *Legal Precedent: A Theoretical and Empirical Analysis*, 19 J.L. & ECON. 249 (1976). Judge Posner states that this research is important:

Citations analysis is growing mainly because it enables rigorous, quantitative analysis of elusive but important social phenomena such as reputation, influence, prestige, celebrity, the diffusion of knowledge, the rise and decline of schools of thought, *stare decisis* (that is, the basing of judicial decision on previous decisions—precedents), the quality of scholarly output . . . and the productivity of scholars, judges, courts, and law schools.

Richard Posner, *An Economic Analysis of the Use of Citations in the Law*, 2 AM. L. & ECON. REV. 381, 382-83 (2000).

⁷ See, e.g., FRED R. SHAPIRO, COLLECTED PAPERS ON LEGAL CITATION ANALYSIS (2001). Fred R. Shapiro has also published a series of studies examining law review citations over the years. See, e.g., Fred R. Shapiro & Michelle Pearse, *The Most-Cited Law Review Articles of All Time*, 110 MICH. L. REV. 1483 (2012); Fred R. Shapiro, *The Most-Cited Law Review Articles Revisited*, 71 CHI.-KENT L. REV. 751 (1996); Fred R. Shapiro, *The Most-Cited Law Review Articles*, 73 CALIF. L. REV. 1540 (1985). For a detailed recent analysis of the use of legal scholarship in intellectual property cases at the Supreme Court, see David L. Schwartz & Lee Petherbridge, *The End of an Epithet? An Exploration of the Use of Legal Scholarship in Intellectual Property Decisions*, 50 HOUS. L. REV. 523 (2012). For a detailed recent analysis of the use of legal scholarship in the Federal Courts of Appeals, see David L. Schwartz & Lee Petherbridge, *Use of Legal Scholarship by the Federal Courts of Appeals: An Empirical Study*, 96 CORNELL L. REV. 1345 (2011); David L. Schwartz & Lee Petherbridge, *Legal Scholarship and the United States Court of Appeals for the Federal Circuit: An Empirical Study of a National Circuit*, 26 BERKELEY TECH. L.J. 1561 (2011).

⁸ See, e.g., David G. Post & Michael B. Eisen, *How Long is the Coastline of Law? Thoughts on the Fractal Nature of Legal Systems*, 29 J. LEGAL STUD. 545 (2000) (study analyzing data from the New York Court of Appeals and the Seventh Circuit to determine the number of times judicial opinions cite and are cited as precedent); Montgomery N. Kosma, *Measuring the Influence of Supreme Court Justices*, 27 J. LEGAL STUD. 333 (1998) (empirical study measuring the influence of 99 retired Supreme Court Justices by analyzing over 1.2 million citations in over 24,000 Supreme Court opinions written between 1793 and 1991); David J. Walsh, *On the Meaning and Pattern of Legal Citations: Evidence from State Wrongful Discharge Precedent Cases*, 31 LAW & SOC'Y REV. 337 (1997) (examining citation data obtained from state wrongful discharge cases to determine whether judges use citations as indicators of substantive influence).

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

the federal court hierarchy, its role in determining the law of the land, and even its ability to influence the law in international jurisdictions. Studying the Supreme Court's citations has become increasingly prevalent, beginning with the first comparative analysis of Supreme Court citation practices of secondary sources in 1986⁹ to more recent empirical studies of citations to law review articles.¹⁰ However, ours is the first detailed study analyzing Supreme Court citations to websites, and it is the first paper considering the topic since a 2006 study published by the National Center for State Courts.¹¹ Considering that every Justice on the Court from Rehnquist through Kagan has cited a website in a majority opinion, studying these citations of Internet sources has become critical to understanding Supreme Court citation patterns and the future of legal information.¹²

Our findings parallel results of previous studies in other academic areas and specific aspects of law, studies that focused on Internet citations in law review articles generally, and one study detailing their use in Washington state cases. We cannot predict what links will rot, even within individual Supreme Court cases. The Internet's ephemeral nature means websites can be available today—and gone tomorrow.¹³ Unfortunately and disturbingly, the

⁹ Louis Sirico & Jeffrey Margulies, *The Citing of Law Reviews by the Supreme Court: An Empirical Study*, 34 UCLA L. REV. 131 (1986).

¹⁰ Lee Petherbridge & David L. Schwartz, *An Empirical Assessment of the Supreme Court's Use of Legal Scholarship*, 106 NW. U. L. REV. 995 (2012).

¹¹ William R. Wilkerson, *The Emergence of Internet Citations in U.S. Supreme Court Opinions*, 27 JUST. SYS. J. 323 (2006).

¹² See *infra* notes 24 & 25.

¹³ See generally Martha Anderson, *A Tale of a Disappearing Website*, LIBRARY OF CONGRESS: THE SIGNAL: DIGITAL PRESERVATION (Jan. 17, 2012), <http://blogs.loc.gov/digitalpreservation/2012/01/a-tale-of-a-disappearing-website> (discussion of the importance of and the difficulty of archiving the National Biological Information Infrastructure (NBII) website and issues with using the archive version).

For a recent example of a popular website archived by archive.org whose archived version is highly limited and not user friendly, see Everyblock.com. Everyblock was an “experiment in online journalism, offering a news feed for every city block in [at the time of closure—nineteen] cities. Enter any address, neighborhood or ZIP code in those cities, and the site shows you recent public records, news articles and other Web content that’s geographically relevant to you. To our knowledge, it’s the most granular approach to local news ever attempted.” *The EveryBlock FAQ*, EVERYBLOCK.COM, available in archived format at <http://web.archive.org/web/20100322230422/http://www.everyblock.com/about/faq> (last visited June 1, 2013). The last crawled version of this FAQ is available at <http://web.archive.org/web/20130113004735/http://www.everyblock.com/faq>, while the latest FAQ may still be available at <http://www.everyblock.com/about>. However, the website was shuttered suddenly on February 7, 2013, by its owner, NBC Universal, with all of the search functionality missing from the archived version. The EveryBlock Team, *Farewell, Neighbors*, EVERYBLOCK.COM (Feb.

Supreme Court appears to have a vast problem with “link rot,” the condition of Internet links no longer working. We found an alarmingly high number of websites cited by the Supreme Court that are no longer working, almost one-third (29%) of the time. Our research also found that the type of online document (pdf, html, etc.) and the sources of links (government or non-government) do not affect the rate of link rot. These two aspects in combination are highly problematic for efforts to craft solutions to this problem.

In this article, we give an overview of the Supreme Court and its citation of websites in Part I; define and explain link rot in Part II; and present our study, methodology, results, and potential solutions in Parts III, IV, V, and VI, respectively. Finally, we conclude in the final Part. Considering interest in individual Supreme Court Justices, we provide an Appendix including Internet citations by Justice.

I. OVERVIEW OF THE SUPREME COURT’S CITATION TO WEBSITES

Citations to links by the Supreme Court started slowly. However, they have increased greatly over time. As part of her study of the past fifteen years of Supreme Court opinions, Allison Orr Larsen found that Justices were looking further afield for sources to support fact-finding, outside of any of the party briefs, amici briefs, or the joint record,¹⁴ including citations to websites.¹⁵ Larsen states that “it is quite common for [Supreme Court opinions] to cite raw statistics of all types—collected from websites, solicited from agencies, or found in a journal—about a huge range of prevailing practices or social norms.”¹⁶ Her study found that of the over one hundred “most important Supreme Court cases” from 2000 to 2010, 56% include mentions of facts the Justices did not find in the record and instead found independently.¹⁷ She found specifically that:

7, 2013), <http://blog.everyblock.com/2013/feb/07/goodbye>; Jeff Sonderman, *NBC Closes Hyperlocal, Data-driven Publishing Pioneer EveryBlock*, POYNTER.ORG (Feb. 8, 2013), <http://www.poynter.org/latest-news/top-stories/203437/nbc-closes-hyperlocal-pioneer-everyblock>; Raizel Liebler, *EveryBlock No More: Why Dynamic Archiving of Websites is Needed*, THE LEARNED FANGIRL (Mar. 1, 2013), <http://thelearnedfangirl.com/2013/03/01/everyblock-no-more-why-dynamic-archiving-of-websites-is-needed>.

¹⁴ See Allison Orr Larsen, *Confronting Supreme Court Fact Finding*, 98 VA. L. REV. 1255, 1261 (2012) (including discussion of the Justices’ use of websites to conduct research during oral argument and for opinions).

¹⁵ *Id.* at 1274, 1276, 1302.

¹⁶ *Id.* at 1279.

¹⁷ *Id.* at 1262.

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

[I]t was quite common for Justices to demonstrate the prevalence of a practice through statistics they found themselves. And, at a fairly high rate these statistics were supported by citations to websites—I found seventy-two such citations in my non-exhaustive search. Importantly, statistics were independently gathered from websites with widely ranging indicia of reliability.¹⁸

The Court’s practice of citing websites has become common—“the practice of Justices searching online for authorities to support factual assertions is not rare in the least.”¹⁹

This additional research leading to citations to Internet websites does have the potential to be positive, helping to shape the decision through additional information that can be much more up-to-date than the briefs could be, considering that there can be significant delays between the time of brief submission and the time of the decision.²⁰ In some cases, the information in Internet sources helps to fill out the analysis in a similar way as any other secondary source cited, such as a law review article or a treatise. In some cases, the information from Internet sources can serve at the heart of disputes between the majority and the dissent.²¹

Citing websites has become common practice for the Supreme Court Justices, but it started with a mere trickle of citations. Justice Souter was the first Justice to cite the Internet in a 1996 opinion, when he provided two URLs for sources he relied on in a portion of his concurring opinion.²² Then, in 1998, Justice Ginsburg used the Internet for sources to demonstrate different meanings of the word “carry” in her dissent.²³ In the following years, other Justices joined the trend of citing websites and eventually began including them in majority opinions. Justices’ citations to websites are discussed further in the Appendix.

By the time John Roberts became Chief Justice, the Court commonly cited websites. Between 2000 and 2006, all of the Justices serving on the Rehnquist Court cited at least one website

¹⁸ *Id.* at 1288.

¹⁹ *Id.* at 1289.

²⁰ See Kedar Bhatia, *Update on October Term 2012 and an Interim Stat Pack*, SCOTUSBLOG (Feb. 8, 2013, 3:47 PM), <http://www.scotusblog.com/2013/02/update-on-october-term-2012-and-an-interim-stat-pack>.

²¹ See *infra* notes 25-35 and accompanying text.

²² *Denver Area Educ. Telecomm. Consortium, Inc. v. FCC*, 518 U.S. 727, 777 n.4 (1996); John J. Hasko, *Persuasion in the Court: Nonlegal Materials in U.S. Supreme Court Opinions*, 94 LAW LIBR. J. 427, 440 (2002).

²³ *Muscarello v. United States*, 524 U.S. 125, 143 n.6 (1998) (Ginsburg, J., dissenting).

in a majority opinion for the first time.²⁴ Following their appointment to the Court, Chief Justice Roberts and Justices Alito, Sotomayor, and Kagan also cited websites in majority opinions.²⁵ We found that within the time period we examined (1996-2010), 114 majority opinions of the Supreme Court included links. Since the time those links were cited, some of the cases with links cited have been superseded or modified, such as *Hamdi v. Rumsfeld*.²⁶

In some cases, the information in the now-dead links is essential to understanding the analysis of the Court. These include citations to an order by a lower court and a video containing a victim impact statement. The loss of secondary support, the type that would be found in an extensive string of “*but see*” citations, can matter; however, the loss of primary source documentation is more significant. For example, *Hollingsworth v. Perry*’s majority opinion relies, in its analysis of the legality of a federal court policy allowing cameras in the courtroom, on a now-dead link to that policy on the Ninth Circuit Court of Appeals website.²⁷ The

²⁴ *Burlington N. & Santa Fe Ry. v. White*, 548 U.S. 53, 65 (2006) (Justice Breyer’s first majority opinion with an Internet citation); *Wagnon v. Prairie Band Potawatomi Nation*, 546 U.S. 95, 108 n.3 (2005) (Justice Thomas’s first majority opinion with an Internet citation); *Smith v. Doe*, 538 U.S. 84, 105 (2003) (Justice Kennedy’s first majority opinion with an Internet citation); *EEOC v. Waffle House, Inc.*, 534 U.S. 279, 290 n.7 (2002) (Justice Stevens’s first majority opinion with an Internet citation); *Becker v. Montgomery*, 532 U.S. 757, 764 (2001) (Justice Ginsburg’s first majority opinion with an Internet citation); *Cal. Democratic Party v. Jones*, 530 U.S. 567, 578 (2000) (Justice Scalia’s first majority opinion with an Internet citation); *Crosby v. Nat’l Foreign Trade Council*, 530 U.S. 363, 383 n.20 (2000) (Justice Souter’s first majority opinion with an Internet citation); *Green Tree Fin. Corp.-Ala. v. Randolph*, 531 U.S. 79, 95 n.2 (2000) (Chief Justice Rehnquist’s first majority opinion (and first overall opinion) with an Internet citation); *Roe v. Flores-Ortega*, 528 U.S. 470, 474 (2000) (Justice O’Connor’s first majority opinion with an Internet citation).

²⁵ *Matrixx Initiatives, Inc. v. Siracusano*, 131 S. Ct. 1309, 1320 (2011) (Justice Sotomayor’s first majority opinion with an Internet citation); *Ransom v. FIA Card Servs., N.A.*, 131 S. Ct. 716, 722 n.2 (2011) (Justice Kagan’s first majority opinion with an Internet citation); *MeadWestvaco Corp. v. Ill. Dep’t of Revenue*, 553 U.S. 16, 20 n.1 (2008) (Justice Alito’s first majority opinion with an Internet citation); *DaimlerChrysler Corp. v. Cuno*, 547 U.S. 332, 350 (2006) (Justice Roberts’s first majority opinion with an Internet citation).

²⁶ 542 U.S. 507 (2004).

²⁷ *Hollingsworth v. Perry*, 558 U.S. 183, 188-89 (2010) (“On January 8, 2010, Chief Judge Kozinski issued an order approving the District Court’s decision to allow real-time streaming of the trial to certain federal courthouses listed in a simultaneously issued press release. Five locations had been selected: federal courthouses in San Francisco, Pasadena, Seattle, Portland, and Brooklyn. The press release also indicated that ‘[a]dditional sites may be announced.’ Federal Courthouses to Offer Remote Viewing of Proposition 8 Trial, online at http://www.ca9.uscourts.gov/datastore/general/2010/01/08/Prop8_Remote_Viewing_Locations.pdf (as visited Jan. 13, 2010, and available in the Clerk of Court’s case file).”). The Internet Archive does have a copy of this order. See INTERNET ARCHIVE WAYBACK MACHINE (Jan. 19, 2010),

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

opinion of the Court relies on a change to the local rules:

Our review is confined to a narrow legal issue: whether the District Court's amendment of its local rules to broadcast this trial complied with federal law. We conclude that it likely did not and that applicants have demonstrated that irreparable harm would likely result from the District Court's actions. We therefore stay the court's January 7, 2010, order to the extent that it permits the live streaming of court proceedings to other federal courthouses.²⁸

A researcher looking for the ways the Supreme Court has or has not upheld local rule changes would not be able to rely on the original link in the Supreme Court decision, despite the fact that, as of 2013, the oldest link would be only three years old. This document does not appear to be anywhere on the new Ninth Circuit website—and the closest source seems to be an incomplete link resolver stating only the following rather than the actual PDF:

This page redirects to http://www.ca9.uscourts.gov/datastore/general/2010/01/08/Prop8_Remote_Viewing_Locations.pdf on the public site.²⁹

The most egregious examples of dead links in Supreme Court opinions refer back to dead links on the Supreme Court website. In two cases, the dead link plays an important role in the Supreme Court's decision-making process. In *Kelly v. California*, the Supreme Court cites a video at a now-dead link on its own website, a video used in the majority's denial of a petition for a writ of certiorari regarding inclusion of video victim impact evidence.³⁰ Justice Stevens states: "Although the video shown to each jury was emotionally evocative, it was not probative of the culpability or character of the offender or the circumstances of the offense."³¹ The same website is also cited extensively in Justice Breyer's dissent, including the statements that:

http://wayback.archive.org/web/*/http://www.ca9.uscourts.gov/datastore/general/2010/01/08/Prop8_Remote_Viewing_Locations.pdf.

²⁸ *Hollingsworth*, 558 U.S. at 189.

²⁹ UNITED STATES COURTS FOR THE NINTH CIRCUIT, http://www.ca9.uscourts.gov/content/view.php?pk_id=0000000397 (last visited Apr. 10, 2013).

³⁰ *Kelly v. California*, 555 U.S. 1020, 1021 n.1 (2008) ("The full video is available online at http://www.supremecourtus.gov/opinions/video/kelly_v_california.html and in Clerk of Court's case file.").

³¹ *Id.* at 1025.

The film, in my view, is poignant, tasteful, artistic, and, above all, moving. . . . On the other hand, the film's personal, emotional, and artistic attributes themselves create the legal problem. They render the film's purely emotional impact strong, perhaps unusually so. That emotional impact is driven in part by the music, the mother's voiceover, and the use of scenes without victim or family (for example, the film concludes with a clip of wild horses running free). Those aspects of the film tell the jury little or nothing about the crime's "circumstances" . . . but nonetheless produce a powerful purely emotional impact. It is this minimal probity coupled with the video's *purely emotional* impact that may call due process protections into play.³²

In *Scott v. Harris*, a video with a dead link was cited extensively by both the majority and minority opinions, serving as the focal point of a serious disagreement in the case.³³ The majority opinion states, "We are happy to allow the videotape to speak for itself."³⁴ Additionally, the majority used the citation to the video to disagree with the dissent, stating that "Justice Stevens suggests that our reaction to the videotape is somehow idiosyncratic, and seems to believe we are misrepresenting its contents."³⁵ While the information cited in these cases is still available on the Supreme Court website, it is now located someplace new—and the old links cited by the Court do not point to the new location.³⁶

Considering that the information cited in Supreme Court decisions is so important—in these instances the linked information was at the heart of disputes between the majority and the dissent—having working links in opinions is essential. Nonetheless, our study found that the percentage of websites that are no longer working within Supreme Court cases is alarmingly high: almost one-third (29%).

³² *Id.* at 1026-27 (Breyer, J., dissenting) (Justice Breyer concluding that, "A review of the film itself, http://www.supremecourtus.gov/opinions/video/kelly_v_california.html, along with the sources to which Justice Stevens refers, makes clear that the due process problem of disproportionately powerful emotion is a serious one.").

³³ *Scott v. Harris*, 550 U.S. 372, 376, 381 (2006).

³⁴ *Id.* at 379 n.5 ("See Record 36, Exh. A, available at http://www.supremecourtus.gov/opinions/video/scott_v_harris.html and in Clerk of Court's case file.").

³⁵ *Id.*

³⁶ *Video Resources*, SUPREME COURT OF THE UNITED STATES, <http://www.supremecourt.gov/media/media.aspx> (last visited Apr. 12, 2013).

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

II. LINK ROT & DEAD LINKS

A. *Link Rot Generally*

Online research is distinct from traditional print research. One of the most powerful attributes of online research is the ability of a web-based document to link to other online sources, allowing the researcher to instantly access a cited authority. This process of linking, which permits an immediacy of access never available in the print world, has a very considerable downside, however, which Ian Gallacher describes as follows:

Books are inert information repositories, and are therefore also immune to information retention problems. By contrast, the Internet is a volatile environment and information can be added or removed without any notice to the end user. One aspect of this problem is familiar to anyone who has clicked on a link to an apparently interesting website only to discover that the site is no longer available. This phenomenon [is] appropriately termed “link rot.”³⁷

With the advent of the Internet, academic research has been able to blossom in unexpected ways, allowing for new means of research, collaboration, and study. However, like Pandora’s Box, the gift of a new way of conducting research via the Internet comes packaged with a burden: academics are forced to deal with the fallout from link rot. Link rot has been described in a variety of ways, but the most concise definition indicates that link rot is “the tendency of hyperlinks to become invalid over time due to sites changing or vanishing.”³⁸

The resulting no-longer-working links are the end result of the link-rot phenomenon.³⁹ Uniform Resource Locators (URLs), or links, fall victim to rot through addition, deletion, or alteration of site material by site owners without notice to users. Links rot for various reasons, including fear of litigation, copyright issues, desire to remain current (by moving information to another

³⁷ Ian Gallacher, *Forty-Two: The Hitchhiker’s Guide to Teaching Legal Research to the Google Generation*, 39 AKRON L. REV. 151, 187 (2006). We do however disagree with Gallacher’s statement that a good legal researcher can usually find other versions of information cited from a website.

³⁸ Kurt Schiller, *301works.org: Seeking an End to ‘Link Rot,’* 27 INFO. TODAY 13 (Jan. 2010).

³⁹ Benjamin J. Keele & Michelle Pearse, *How Librarians Can Help Improve Law Journal Publishing*, 104 LAW LIBR. J. 383, 391 (2012) (explaining that link rot results in “uniform resource locators (URLs) that no longer direct researchers to the correct online resource.”).

location), a lack of interest in maintaining a link into the future, a server shut down caused by business failure, or the erroneous citation of URLs within sources.⁴⁰

Often websites restructure to better serve their present users without thinking about the impact of lost information that occurs during the restructuring process. Some websites use URL redirection, a coding technique to make an older website available within a new website by forwarding users to the correct new location.

Cached websites can exist for short amounts of time. Through Google cache, websites can exist as long as six months, but most often they are preserved for a shorter time frame.⁴¹ However, when a website is completely replaced, there is no forwarding address for the older information. For example, every four or eight years (depending on whether the president serves for one or two terms), the entire White House website is replaced on the date of the Inauguration, scrubbing away all previous links.⁴²

⁴⁰ Michelle M. Wu, *Why Print and Electronic Resources are Essential to the Academic Law Library*, 97 LAW LIBR. J. 233, 238-39 (2005) (explaining the desire to remain current, a lack of interest in maintaining a link into the future, fear of litigation, and copyright issues, noting, “They may be removed from databases by the publisher for reasons ranging from desire to keep the database current, to disinterest in maintaining a low-use resource, to fear of litigation over database copyright issues, to objections of political nature.”); Erick Ducut, Fang Liu & Paul Fontelo, *An Update on Uniform Resource Locator (URL) Decay in MEDLINE Abstracts and Measures for its Mitigation*, 8 BMC MED. INFORMATICS & DECISION MAKING 23 (2008), <http://www.biomedcentral.com/1472-6947/8/23> (Under ‘Background’ the authors state, “In general, URL decay occurs for several reasons: servers may shutdown because of business failures; URL content may change or reconfigured by the Web site owners; and errors in URL citing may occur.”); see Wallace Koehler, *An Analysis of Web Page and Web Site Constancy and Permanence*, 50 J. AM. SOC. INFO. SCI. & TECH. 161, 172, 174 (1999).

⁴¹ *Requesting Removal of Content from Our Index*, GOOGLE WEB MASTER CENTRAL BLOG, <http://googlewebmastercentral.blogspot.com/2007/04/requesting-removal-of-content-from-our.html> (last visited Apr. 12, 2013) (“We’ll automatically make the latest cached version of the page available again after six months (and at that point, we likely will have recrawled the page and the cached version will reflect the latest content.”); *Recrawling Sites*, GOOGLE HELP HOME, <http://support.google.com/adsense/bin/answer.py?hl=en&answer=10534> (last visited Apr. 10, 2013) (“At this time, we’re unable to control how often our crawlers index the content on your site. Crawling is done automatically by our bots. If you make changes to a page, it may take up to 1 or 2 weeks before the changes are reflected in our index.”).

⁴² Compare the White House website on the last full day of the Bush Administration (<http://web.archive.org/web/20090120010611/http://www.whitehouse.gov>) with the White House website after Obama’s inauguration (<http://web.archive.org/web/20090122232441/http://www.whitehouse.gov>). The fact that the earliest versions of the Obama White House website do not look fully formed and functional demonstrates one of the issues with archiving

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

In 1995, some of the first researchers to “document the impact of fleeting web content” found that one-third of URLs in e-journal articles from 1993-1995 were no longer available.⁴³ While all types of web material fall prey to link rot, a 2003 study showed that different types of web content have different life spans.⁴⁴ That study found that on average, legal citations have a half-life—the time required for half of a defined selection of web literature to disappear—of 1.4 years, scholarly articles have a half-life of 1.5 years, and random web pages have a half-life of 2 years.⁴⁵ Notably, 98.3% of web pages show some changes after 6 months, which increases to 99.1% after 1 year.⁴⁶

Given the importance of the Internet to all academic disciplines today, link rot is being studied extensively, but the study of its effects on the legal discipline lags behind. A number of studies, articles, and notes are dedicated to determining the reach of link rot’s effects within given areas,⁴⁷ including medicine and science.⁴⁸ Although link rot may be a mere annoyance to the casual consumer, academia is taking a deep interest in the problem because it has been documented to limit progress in medicine and science.⁴⁹ Similar to legal researchers, researchers in other fields need to be able to easily access past studies for the purpose of replication.

One study investigating the incidence of link rot in the three “highest circulation US journals with scientific impact” (*Journal of the American Medical Association* (JAMA), *New England Journal*

websites—if the CSS (a type of formatting language) is not also archived, the website will not look as it did at the time of archiving.

⁴³ Wallace Koehler, *A Longitudinal Study of Web Pages Continued: A Consideration of Document Persistence*, 9 INFO. RES. 2 (2004), available at <http://InformationR.net/ir/9-2/paper174.html>.

⁴⁴ *Id.*

⁴⁵ *Id.*

⁴⁶ Mary Rumsey, *Runaway Train: Problems of Permanence, Accessibility, and Stability in the Use of Web Sources in Law Review Citations*, 94 LAW LIBR. J. 27, 37 (2002).

⁴⁷ See Daniela Dimitrova & Michael Bugeja, *Consider the Source: Predictors of Online Citation Permanence in Communication Journals*, 6 PORTAL: LIBR. & ACAD. 269, 270-71 (2006) (providing an overview of studies conducted on link rot in various genres).

⁴⁸ See, e.g., Robert P. Dellavalle et al., *Going, Going, Gone: Lost Internet References*, 302 SCI. 787 (2003) (examining citation practices and link rot in over 1,000 articles published between 2000 and 2003 in *The New England Journal of Medicine*, *The Journal of the American Medical Association*, and *Science*); John Markwell & David W. Brooks, “Link Rot” Limits the Usefulness of Web-based Educational Materials in Biochemistry and Molecular Biology, 31 BIOCHEMISTRY & MOLECULAR BIOLOGY EDUC. 69 (2003) (documenting link rot issues in the biochemistry field); Andrea W. Thorp & Lance Brown, *Accessibility of Internet References in Annals of Emergency Medicine: Is It Time to Require Archiving?*, 50 ANNALS EMERGENCY MED. 188 (2007).

⁴⁹ Ducut, Liu & Fontelo, *supra* note 40.

of *Medicine* (NEJM), and *Science*) found that for issues from 2003, the average rate of link rot increased from 3.8% at 3 months post-publication to 10% at 15 months and 13% at 27 months.⁵⁰ Another study, conducted on MEDLINE databases, found that for records from 1994-2006 including 10,208 URLs, 81% were generally available during a 30-day period, but only 78% of the URLs contained the information discussed in the corresponding citing article.⁵¹ These numbers illustrate that link rot has infiltrated the professional scientific disciplines.

B. *How Link Rot Affects Law Generally*

The phenomenon of link rot in law is troublesome because citations are the cornerstone upon which both judicial opinions and law review articles stand. Paul Axel-Lute stated that “[a] legal citation serves two purposes. First, it indicates the nature of the authority upon which a statement is based. Second, it contains the information necessary to find and read the cited material.”⁵² Citations are used extensively in law review article footnotes and judicial opinions. The ability to confirm citations and to ensure that they are accurate is essential to ensure that precedents are indeed cited correctly.

As in other professional fields, Internet citations have become commonplace in the legal realm. In fact, the use of Internet citations in law review articles has increased more than a thousand-fold since the early nineties. According to Susan Lyons, in 1994, there were four Internet citations in three law review articles.⁵³ By 2003, however, there was an explosion in the total number of Internet citations within the entire corpus of law reviews—up to 96,946.⁵⁴ In another study, Mary Rumsey looked at the number of law reviews with Internet citations, finding that the percentage of law review articles including Internet citations increased from 0.57% in 1995 to 23% of law review articles in 2000.⁵⁵

Ideally, every Internet citation would provide parallel print sources. However, this is impossible for several reasons, including partly that not every document has an equivalent print source, especially for “born digital material.”⁵⁶ When citing a book first published in print, it is relatively easy for a researcher to cite the print version as well as the electronic version, but many online

⁵⁰ Dellavalle et al, *supra* note 48, at 787.

⁵¹ Ducut, Liu & Fontelo, *supra* note 40.

⁵² Rumsey, *supra* note 46, at 28 (citing Paul Axel-Lute, *Legal Citation Form: Theory and Practice*, 75 LAW LIBR. J. 148, 148 (1982)).

⁵³ Susan Lyons, *Persistent Identifiers of Electronic Documents and the Future of Footnotes*, 97 LAW LIBR. J. 681, 681 (2005).

⁵⁴ *Id.*

⁵⁵ Rumsey, *supra* note 46, at 32-33.

⁵⁶ See Martin A. Kesselman & Sarah Barbara Watstein, *Creating Opportunities: Embedded Librarians*, 49 J. LIBR. ADMIN. 383, 392 (2009).

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

documents are cited precisely because they are online only. Additionally, due to budget constraints, there is an increasing amount of information from governmental agencies and other organizations that is only disseminated online.⁵⁷ Because Internet citations are now the norm, it is critical to have a firm understanding of the consequences when such citations rot.

Access to the sources in citations is critical in legal scholarship due to the doctrine of *stare decisis* and the development of the common law. Without reliable access to cited materials, subsequent researchers or lawyers cannot examine the bases for the original author's conclusion.⁵⁸ Generally, "legal researchers have traditionally looked for information that is more than just informative; they have looked for information that is unquestionably authoritative."⁵⁹ Law relies on fixed records; it is difficult to create meaningful precedents based upon something impermanent.⁶⁰ The constantly changing nature of the Internet threatens the authority of scholarly research.⁶¹ If all content were capable of manipulation and sudden disappearance, it stands to reason that unpopular theories and opinions vital to understanding the evolution of legal thought may vanish to save reputations.⁶²

⁵⁷ Michael Whiteman, *The Impact of the Internet and Other Electronic Sources on an Attorney's Duty of Competence Under the Rules of Professional Conduct*, 11 ALB. L.J. SCI. & TECH. 89 (2000).

⁵⁸ Rumsey, *supra* note 46; see Keele & Pearse, *supra* note 39, at 391 ("At worst, broken links undermine an article's soundness by removing support for its assertions.").

⁵⁹ Colleen M. Barger, *On the Internet, Nobody Knows You Are a Judge: Appellate Courts' Use of Internet Materials*, 4 J. APP. PRAC. & PROCESS 417, 419 (2002).

⁶⁰ Denmark, *supra* note 1, at 12-13 ("Making articles infinitely malleable can decrease judges' ability to rely on them. Law relies on fixed written records. Allowing authors the unrestrained power to update their work means that every article will forever be a mere draft and never a final product. This is not the stuff on which courts will gladly rest precedents.").

⁶¹ Lyons, *supra* note 53; Dimitrova & Bugeja, *supra* note 47, at 270 (noting that "the erosion of Internet footnotes—the phenomenon of inaccessible online footnotes—undermines the standards of scholarship and the methods of research, primarily because it destabilizes fixed language and original source.").

⁶² Denmark, *supra* note 1, at 10-11 ("Authors would also maintain the liberty to update their articles in response to readers' comments, later-decided cases, articles, or reports of new events. This ability holds serious dangers for the history of ideas and the usefulness of legal articles. Articles published on a home page are infinitely malleable. They can be updated with the speed of hands on a keyboard and provided someone had not downloaded the article or kept a private copy, the words can be made to vanish altogether. . . . Without the permanence of paper copies, legal scholarship will become fragile knowledge, dependent on the willingness of authors to be associated with their previously held opinions. . . . If the work underwent multiple revisions, authors might not get the version containing the crucial part that formed the basis of scholarly interest. This could be due to dishonesty or uncooperativeness on the part of the author, but might also be a common side effect of working on a word processor.").

Traditional sources—or as Lee Peoples states, the “stable universe of settled sources” utilized by lawyers, judges, and scholars have been stable in nature: once something is printed in an official (print) Reporter, it remains there always unchangeable and can be accessed the same way by a researcher pulling a book from the shelf exactly the same on the day of publication as a hundred years later.⁶³ Now, legal professionals are using potentially unstable Internet sources that may be impossible to retrieve in the future.⁶⁴

One of the issues raised with potential solutions for archiving or preserving materials cited in judicial opinions is copyright in those documents. Copyright in court filings has become an issue recently because these documents are available on PACER and other sources. The concept of copyright limiting the use of publicly available documents, such as litigation documents or court filings, may seem to be settled law, or at least settled practice. However, it is not.⁶⁵ Our solutions below regarding archiving judicially cited information includes sources that are definitively protected by copyright, such as video files, databases, or websites that are cited and are even occasionally the focal point of Supreme Court cases. Copyright is not waived by placement in the public record, such as open access websites and audio and video recordings. According to Pamela Samuelson:

⁶³ Lee F. Peoples, *The Citation of Wikipedia in Judicial Opinions*, 12 YALE J.L. & TECH. 1, 36 (2009) (“The traditional sources that common law judges, lawyers, and academics cite come from a ‘stable universe of settled sources.’ For example, once a case appears in a reporter it is essentially fixed for all time.”) (quoting Robert C. Berring, *Legal Information and the Search for Cognitive Authority*, 88 CAL. L. REV. 1673, 1675 (2000)).

⁶⁴ *Id.* at 36-37 (“Traditional citation methods work perfectly for these sources. In recent years lawyers and judges have begun citing less traditional sources like websites, blogs, and of course Wikipedia entries. Once cited these sources can be difficult to locate in the future.”).

⁶⁵ Davida H. Isaacs, *The Highest Form of Flattery? Application of the Fair Use Defense Against Copyright Claims for Unauthorized Appropriation of Litigation Documents*, 71 MO. L. REV. 391, 393 (2006) (“Breaking from centuries-old tradition, attorneys have recently begun threatening one another with claims of copyright infringement based on the unauthorized appropriation and adaptation of their legal documents, particularly litigation documents.”); Michael Whiteman, *Appellate Court Briefs on the Web: Electronic Dynamos or Legal Quagmire?*, 97 LAW LIBR. J. 467, 478-79 (2005) (discusses copyright issues regarding legal briefs). *Nimmer on Copyright* describes one of the only two cases on point that attempted to limit the placement of litigation documents into the public record as “appear[ing] erroneous” and based on “the desire to save copyright owners from adverse effects based on its having released a few stray copies of a work otherwise maintained securely.” 2 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 8.12 (B)(1)(c)(ii) (2012). *See also* *White v. West Pub. Corp.*, No. 12 Civ. 1340(JSR), 2013 WL 544057 (S.D.N.Y. Feb. 11, 2013) (dismissing on summary judgment motion plaintiff’s copyright complaint against Lexis and Westlaw for making available plaintiff’s copyrighted briefs).

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

Although there is no caselaw involving judicial, legislative, executive, or administrative uses of copyrighted materials, these uses should similarly be accorded broad fair use privileges insofar as copyrighted materials are relevant inputs to legitimate governmental decision-making and other activities. Consider, for example, fair use as a justification for court and West Publishing Co. reproductions of the texts of copyrighted works, such as the Supreme Court's recitation of the contested song lyrics in *Campbell* and the writings at issue in *Harper & Row*. Those decisions are more informative and precise in their holdings because they reproduce the contested uses.⁶⁶

C. How Link Rot Is Related to Legal Citation

Citation rules also affect how the Internet is used within the context of legal sources, including cases and law reviews. In Rule 18.2.1(a), *The Bluebook* allows for online sources to be cited as if they are print sources, without including the direct URL. This rule permits this despite the fact that the source may not necessarily be found in print so a link would ease the process of finding the source for future researchers.⁶⁷ Still, Rule 18.2.1(c) makes the inclusion of a citation to an online source discretionary, “[e]ven if a printed source is available, a parallel citation to an electronic source as related authority [] **may be appropriate where it would substantially improve access to the relevant information.**”⁶⁸

For some material, the online source may merely provide access to information or a publication also available in print. Rule 18.2.2 of *The Bluebook* states that when there is a choice between electronic and print media, the print source should be used: “An Internet source may be cited directly when it does not exist in a traditional printed format or when a traditional printed source, such as a letter or unpublished dissertation, exists but cannot be found or

⁶⁶ Pamela Samuelson, *Unbundling Fair Uses*, 77 *FORDHAM L. REV.* 2537, 2596 (2009) (citation omitted).

⁶⁷ See Nicholas Taylor & Susan Manus, *The Value of a Broken Link*, THE LIBRARY OF CONGRESS: THE SIGNAL: DIGITAL PRESERVATION (Mar. 28, 2012), <http://blogs.loc.gov/digitalpreservation/2012/03/the-value-of-a-broken-link> (Nicholas Taylor, Information Technology Specialist for the Repository Development Group notes that “there are at least two good reasons to conserve broken links: facilitating discovery of archived versions of a resource and providing metadata about the resource, whether or not it’s available.”).

⁶⁸ THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION, R. 18.2.2, at 166 (Columbia Law Review Ass’n et al., eds., 19th ed. 2010) (emphasis added) [hereinafter THE BLUEBOOK].

is so obscure that it is practically unavailable.”⁶⁹

Despite this general citation rule, the use of electronic sources has mushroomed in the last decade.⁷⁰ *The Bluebook* does acknowledge that not all citations are equally useful. In this vein, Rule 18.2.2 states, “All efforts should be made to cite to the most stable electronic location available. The Internet citation should include information designed to facilitate the clearest path of access to the cited reference.”⁷¹

D. How Link Rot Affects Judicial Opinions

Some courts have traditionally chosen to use Internet citations sparingly because of a lack of confidence in their reliability and accuracy and their thus not meeting traditional citation rules.⁷² Some Internet citations in opinions are used for substantial factual information, while others are used for context or clarification.⁷³ Websites have been used for facts that are judicially noted.⁷⁴ Still, for some controversial online sources such as Wikipedia, courts struggle to determine whether to accept citations as valid or reject them altogether.⁷⁵ A study of federal appellate opinions found that,

⁶⁹ *Id.*

⁷⁰ Peoples, *supra* note 63, at 3 (“Citations to Wikipedia in judicial opinions first appeared in 2004 and have increased steadily ever since.”); *id.*, at 6 (“The ALLCASES database includes all United States federal and state cases available on Westlaw from the year 1658 to present. This returned 407 cases with some reference to a wiki or Wikipedia article. Four hundred and one cases referenced a Wikipedia article and six cases referenced a wiki other than Wikipedia. Interestingly, Wikipedia contains two pages listing judicial opinions citing Wikipedia entries. One page lists thirteen opinions citing a Wikipedia entry and another lists ninety-eight United States judicial opinions citing a Wikipedia entry.”); *id.*, at 28 (“Citations to Wikipedia entries in judicial opinions have been steadily increasing since the first citation appeared in 2004.”); *id.*, at n.174 (“Wikis or Wikipedia were cited in 4 cases in 2004, 18 cases in 2005, 80 cases in 2006, 136 cases in 2007, and 169 cases in 2008.”); *id.*, at 57 (“In my research I discovered that since July of 2007, twenty-six judicial opinions have cited blogs for factual information.”). For a discussion on the motivations behind citations in general and web citations specifically, see Alistair G. Smith, *Web Links as Analogues of Citations*, 9 INFO. RES. 188 (2004), <http://InformationR.net/ir/9-4/paper188.html>.

⁷¹ THE BLUEBOOK, *supra* note 68.

⁷² Barger, *supra* note 59, at 425.

⁷³ Arguably, Internet sources do not meet the standard for judicial notice. *See id.*, at 433 (“merely citing to a web site and inviting others to visit the site does not satisfy [Federal Rule of Evidence 201’s] requirement that the fact be ‘capable of accurate and ready determination’—at least where the pro se prisoner is denied any access to the web site, much less ‘ready’ access.”).

⁷⁴ *Id.*

⁷⁵ Peoples, *supra* note 63, at 45. (“[A]ppellate courts have reversed or found error in lower court decisions relying on Wikipedia entries for psychological research in a child custody case, for attempting to refute expert medical testimony with a Wikipedia entry, and perhaps most egregiously for denying an asylum seeker’s request based on information obtained from Wikipedia.”); *see*

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

in 2002, 84.6% of Internet citations in cases from 1997 were inaccessible; moreover, 34% of citations in cases from 2001 were already inaccessible by 2002.⁷⁶ In another study specifically examining Supreme Court opinions from 1996-2001, fourteen cases cited the Internet (twenty-five citations), and by the 2001/2002 term when the study was conducted, 36% of these Internet citations had become unavailable.⁷⁷

In May of 2009, to help resolve some of these link rot issues, the Judicial Conference of the US released guidelines for *Citing To, Capturing, and Maintaining Internet Resources in Judicial Opinions/Using Hyperlinks in Judicial Opinions*.⁷⁸ A policy adopted by the conference dictates that all Internet materials cited in final court opinions should be considered for preservation, albeit within the discretion of the judge.⁷⁹ These guidelines provide judges with criteria for assessing Internet sources for accuracy, scope, objectivity, timeliness, authority, and verifiability.⁸⁰ These guidelines suggest capturing and preserving any source that is “fundamental to the reasoning of the opinion and refers to a legal authority or precedent that cannot be obtained in any other format.”⁸¹ Further, preservation should be considered when there is a chance the source will be changed or removed.⁸² The Secretary of the Judicial Conference, Jim Duff, wrote to chief judges that, “[u]nlike printed authority, Internet information is often not maintained at a permanent location, and a cited webpage can be changed or deleted at any time. Obviously, this has significant implications for the reliability of citations in court opinions.”⁸³ Unfortunately, this report was not made available to the general public by request—or on the Internet.

also Noam Cohen, *Courts Turn to Wikipedia, but Selectively*, N.Y. TIMES, Jan. 29, 2007, <http://www.nytimes.com/2007/01/29/technology/29wikipedia.html>.

⁷⁶ Barger, *supra* note 59, at 448.

⁷⁷ *Id.*

⁷⁸ *Internet Materials in Opinions: Citations and Hyperlinking*, UNITED STATES COURTS: THE THIRD BRANCH (July 2009), http://www.uscourts.gov/News/TheThirdBranch/09-07-01/Internet_Materials_in_Opinions_Citations_and_Hyperlinking.aspx.

Interestingly, during the research process for this paper, we needed to look at this source again, and due to a failed link resolver, we were unable to access this source for some time. The irony of being unable to access a website we wanted to cite in an article about the ephemeral nature of websites, including discussion of reasons to avoid citing websites, was not lost on us.

⁷⁹ *Id.*

⁸⁰ *Id.*

⁸¹ Peoples, *supra* note 63, at 43 (citing JUDICIAL CONFERENCE OF THE U.S., GUIDELINES ON CITING TO, CAPTURING, AND MAINTAINING INTERNET RESOURCES IN JUDICIAL OPINIONS/USING HYPERLINKS IN JUDICIAL OPINIONS 2 (2009)).

⁸² *Internet Materials in Opinions: Citations and Hyperlinking*, *supra* note 78.

⁸³ *Id.*

E. *Prior Studies of Link Rot in Legal Citation*

Prior studies demonstrate the overall increase in the raw number and general upward trajectory of website citations. In 1995, 0.57% of law review articles contained at least one web citation, while in 2000, 23% contained at least one.⁸⁴ Mary Rumsey's 2002 study examining link rot in law review articles from 1997-2001 found that for URLs appearing in articles published in 2001, at the point of testing in June 2001, only 61.8% of the URLs were operational, and for citations from 1997, only 30.27% of the links were operational.⁸⁵ A 2004 study looked at three Washington law reviews containing Internet citations over a three-year period (2001-2003) and found high levels of link rot. Specifically, 441 of the 1,104 URLs (40%) in the sample were broken when tested.⁸⁶ These earlier studies demonstrate that citations to websites that existed at the moment they were used are disappearing later on.

In some studies, the rate of link rot increased over time; in other words, the older a link is, the more likely it is to no longer work. In Rumsey's study, 38% of the websites cited in law review articles did not work even within the year they were cited.⁸⁷ But she also found that older websites were more likely to no longer work: 70% of the websites from 1997 (the earliest date studied) were invalid.⁸⁸ Additional studies have found similar rates of link rot. Helene Davis's 2006 study of law reviews found that for a body of law review articles from 2001 to 2003, 40% of the web citations within those articles no longer worked.⁸⁹ In Tina Ching's study of Washington state court opinions published from 1999-2005, of the websites that were found (132 websites in 84 cases), 35% were determined to be inaccessible.⁹⁰

Other studies echo this frightening trend of link decay. The Chesapeake Project, sponsored by the Legal Information Preservation Association, collects born-digital documents. This collection of government, policy, and legal information is archived by the Georgetown Law Library, Harvard Law School Library, Maryland State Law Library, and Virginia State Law Library. In

⁸⁴ Rumsey, *supra* note 46, at 32-33.

⁸⁵ *Id.* at 35.

⁸⁶ Helene E. Davis, *Keeping Validity in Cite: Web Resources Cited in Select Washington Law Reviews, 2001-03*, 98 LAW LIBR. J. 639, 646 (2006).

⁸⁷ Rumsey, *supra* note 46, at 35.

⁸⁸ *See id.*; *see also* David C. Tyler & Beth McNeil, *Librarians and Link Rot: A Comparative Analysis With Some Methodological Considerations*, 3 PORTAL: LIBR. & ACAD. 615 (2003) (also finding that rates of link rot increased with age of publication).

⁸⁹ Davis, *supra* note 86.

⁹⁰ Tina Ching, *The Next Generation of Legal Citations: A Survey of Internet Citations in the Opinions of the Washington Supreme Court and Washington Appellate Courts, 1999-2005*, 9 J. APP. PRAC. & PROCESS 387, 395 tbl.7 (2007).

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

2008, the Chesapeake Project conducted a link rot study of a random sample of 579 items that were solely available in digital format from 2007, and discovered rot in 48 items (8.3%).⁹¹ The same sample has been annually reexamined since 2008. In 2009, link rot was found in 83 of the 579 items (14.3%). In 2010, link rot was found in 160 items (27.6%).⁹² In 2011, link rot was found in 176 items (30.4%). In 2012, link rot was found in 218 items (37.7%).⁹³ Their findings documented increasing rates of link rot for what are considered top-level domains, like those designated for state and federal government, which were thought to be less likely to suffer link rot.

Even commercial databases are not the data safe havens users assume them to be because they do not always include the exact copies of URLs when items are entered into the database, as found in earlier research.⁹⁴ This is a finding that we also discovered in completing our study.

III. CURRENT STUDY—PURPOSE AND RESEARCH QUESTION

Our purpose was to study Supreme Court citations to Internet sources, asking whether the links in Supreme Court opinions rot, and if so, if the rate of link rot was analogous to the rate found in previous studies of legal citations.

Our initial hypothesis was that the Supreme Court would use very few Internet citations in their opinions due to the generally ephemeral and unreliable nature of those sources. Additionally, if the Supreme Court did cite Internet sources, we hypothesized that the links would be to government websites and therefore be less susceptible to link rot.

IV. METHODOLOGY

A. General Methodology

To address these questions empirically, we searched the Lexis Supreme Court case database for citations to websites and related Internet sources for all terms through the 2009-2010 term. Our

⁹¹ Sarah Rhodes, *Breaking Down Link Rot: The Chesapeake Project Legal Information Archive's Examination of URL Stability*, 102 LAW LIBR. J. 581, 589 (2010), <http://www.aallnet.org/main-menu/Publications/lj/LLJ-Archives/Vol-102/publljv102n04/2010-33.pdf>.

⁹² *Id.* at 589-90.

⁹³ The Chesapeake Digital Preservation Group, *"Link Rot" and Legal Resources on the Web: A 2012 Analysis by the Chesapeake Digital Preservation Group*, CHESAPEAKE DIGITAL PRESERVATION GROUP fig.1, tbls.1, 5, <http://legalinfoarchive.org/cdm/linkrot2012> (last visited Apr. 10, 2013).

⁹⁴ Peoples, *supra* note 5, at 69-70 (2010).

reasons for using Lexis rather than Westlaw are discussed in section *Citation Format Problems, infra*.

Considering the varied ways Internet sources have been cited in legal opinions over time, the search we used to find all possible citations was:

(http! or www) or ((available or found or visited or online or Internet) w/20 (com or org or mil or edu or gov))⁹⁵

Cases were then coded, as described below, excluding cases that did not actually refer to websites—such as cases from the 1800s that appeared in the results due to abbreviations, such as citations to reporters and Blackstone’s commentaries. The coding of data was piggybacked on the basic demographic data on Supreme Court cases, such as case name and citation, available through the Spaeth Supreme Court database, to avoid coding errors.⁹⁶ For all terms up to and including the 2009-2010 term, 430 website citations were identified, starting with a citation in the 1995-1996 term.

Cases were coded based on the following factors:

Judicial Factors:

- Justice (name of Justice),
- Where in opinion (majority, concurrence, dissent)

Link Factors:

- Link (entire link),
- Link suffix (com, org, mil, edu, gov, other),
- Link type (htm, html, pdf, other),
- Government website (yes/no),
- Government information from non-government site (yes/no),
- Database search (whether the link was leading to a query of a database, meaning that each time the link was clicked, the result would be different) (yes/no)

Usability and Reliability:

- Bluebook signal,
- Direct quote,

⁹⁵ The search query was constructed this way after trying other potential searches to capture all possible links without missing any, though we did have some false positives. We used Lexis rather than Westlaw due to transcription errors for websites within Westlaw.

⁹⁶ *The United States Supreme Court Judicial Database, 1953-2010 Terms*, THE SUPREME COURT DATABASE, <http://scdb.wustl.edu> (originally compiled by Harold J. Spaeth).

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

- Transcription error (within the Lexis results for the link).

B. Citation Format Problems

Westlaw and Lexis use different methods of handling Internet citation formatting. One of the biggest problems we encountered when we tried to use Westlaw to find cases was that Westlaw inserts a space after every period—even within a URL.⁹⁷ For example, the following URL can be found within *Illinois Tool Works Inc. v. Independent Ink, Inc.*⁹⁸:

<http://www.usdoj.gov/atr/public/guidelines/0558.pdf>

This is how it is actually displayed within Westlaw—with additional spaces added in:

<http:// www. usdoj. gov/ atr/ public/guidelines/0558.pdf>

It is certainly possible for a researcher to discern what the actual Internet citation should be, but he/she must copy, paste and correct the address first. This adds an extra step and assumes the researcher knows how to make such a correction. We therefore used Lexis to compile the cases for our study.

The citation formats varied greatly from one opinion to another, which added to the difficulty in finding the original source material. This demonstrates that the Supreme Court does not always follow strict Bluebook citation format. For example, one of the most unusual citations was this one⁹⁹:

CMS, State Medicaid Manual § 3907, available at [https://www.lexis.com>Legal>Secondary Legal>CCH>Health Law>CMS Program Manuals>CCH CMS Program Manuals P 3907](https://www.lexis.com/Legal/Secondary Legal/CCH/Health Law/CMS Program Manuals/CCH CMS Program Manuals P 3907)

Rather than citing directly to the material needed, this citation takes the reader through a breadcrumb trail to find the source. Should an Internet source be included in a citation, hopefully it will lead the researcher to the source in its present

⁹⁷ Peoples, *supra* note 63, at 37, 38 (“When Westlaw adds the text of judicial opinions to its database, URLs included in the opinions are not always exact copies of the URLs as they appear in the print reporter.” “In contrast to Westlaw, LexisNexis includes active hypertext links in their opinions which take the researcher directly to any web resource cited in a judicial opinion.”).

⁹⁸ *Illinois Tool Works Inc. v. Indep. Ink, Inc.*, 547 U.S. 28, 45 (2006).

⁹⁹ *Ark. Dep’t of Health and Human Servs. v. Ahlborn*, 547 U.S. 268, 282 n.11 (2006).

location. Unfortunately, should a researcher look for this source following the crumbs, it will lead nowhere. Subscription-based commercial vendors, such as Lexis and BNA, alter their internal functionality frequently and remove sources, such as in this citation. Additionally, there was likely a version of this cited document on a government specific website—specifically on the website for the agency at issue, the Centers for Medicare & Medicaid Services, or on the specific Medicaid government website.¹⁰⁰ However, the Court chose not to cite a government website version of this document.

The ready availability of draft versions of documents through the Internet creates another interesting problem. For example, this citation, from *Baze v. Rees*, 553 U.S. 35 (2008)¹⁰¹ contains a link to a version of an article that was subsequently published within a law journal:

Alper, Anesthetizing the Public Conscience: Lethal Injection and Animal Euthanasia, 35 Ford. Urb. L. J. 817, 840 (2008), online at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1109258

There is no simple way to determine what the differences are between the pre-publication version of the article and the final, published version within the journal. In this case, the passage relied upon in the Supreme Court citation, “the field of animal euthanasia has reached a unanimous consensus that neuromuscular blocking agents like pancuronium have no legitimate place in the execution process,” does appear in both versions. However, that will not always be the case with draft or interim versions of articles posted on websites. Even for this citation, the version posted on the link is not the exact same version cited by the Supreme Court because this opinion was issued in April 2008 and the SSRN article was last updated later, in November 2008.

V. RESULTS

Our initial hypothesis was that the Supreme Court would use very few Internet citations in their opinions due to the ephemeral and unreliable nature of those sources. However, there were

¹⁰⁰ CENTERS FOR MEDICARE & MEDICAID SERVICES, <http://www.cms.gov> (last visited Apr. 10, 2013); MEDICAID.GOV, <http://www.medicaid.gov> (last visited Apr. 10, 2013). We looked into whether this exact document was available to the Supreme Court on these websites at the time of citation in March 2006, but due to limitations of *archive.org*, discussed *infra* in Section VI, we were not able to verify that there was a government website version of this document.

¹⁰¹ *Baze v. Rees*, 553 U.S. 35, 73 n.1 (2008).

**SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)**

actually a significant number of Internet citations. Table 1 below reports descriptive information about the Supreme Court's use of Internet citations. We found that since 1996, the first year citations were used, through the 2009-2010 term, there were 430 Internet citations from 144 cases. The average number of Internet citations in a case using Internet citations was 2.56. The case with the most Internet citations was *Arkansas Department of Health & Human Services v. Ahlborn*, 547 U.S. 268 (2006) with 23 Internet citations.

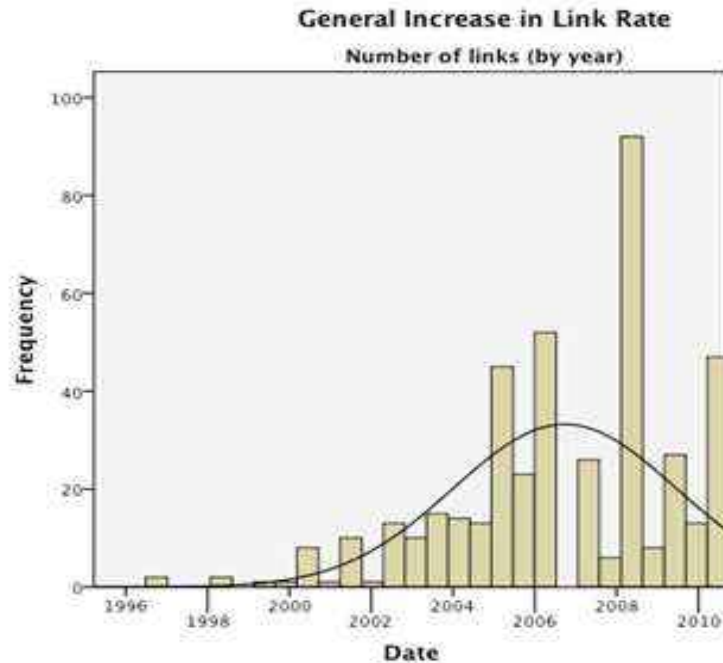
Table 1: Supreme Court Internet Citations

| | |
|---|------|
| Year first URL cited | 1996 |
| Total number of Internet citations | 430 |
| Total cases with Internet citations | 144 |
| Percentage of cases with Internet citations | 14% |
| Average # Internet citations per case with citation | 2.56 |
| Most Internet citations within a single case ¹⁰² | 23 |

In the histogram in Figure 1 below, plotting frequency of URLs in Supreme Court opinions by year of use, URLs are used beginning in 1996, with increases around 2000, and the largest number in 2008. In 2008, there were over 90 links in Supreme Court opinions.

¹⁰² Crawford v. Marion County Election Bd., 553 U.S. 181 (2008).

Figure 1: Number of URLs Cited Each US Supreme Court Term¹⁰³



Of the URLs used within the U.S. Supreme Court opinions during our study period, we found that 29% of them were invalid. Our results are similar to the findings of others, including the Chesapeake Project, which found a link rot rate of 27.6% in a study of solely digital materials over three years,¹⁰⁴ Tina Ching, who found a link rot rate of 35% in a study of Washington state cases,¹⁰⁵ and Helene Davis, whose study of Washington state law review articles revealed a link rot rate of 40%.¹⁰⁶ The only previous study on link rot with a dissimilar conclusion was Wilkerson's 2006 study, which found a link rot rate in Supreme Court opinions of 15.5%.¹⁰⁷ Considering the preeminence of the United States Supreme Court, a link rot rate of almost one-third during our study time frame is quite shocking.

Figure 2 shows our invalid links over time. Based on statistical tests, we found no clear relationship between the time elapsed

¹⁰³ Due to each Supreme Court term taking place within two calendar years, each term receives two possible bars in our figure.

¹⁰⁴ Rhodes, *supra* note 91.

¹⁰⁵ Ching, *supra* note 90.

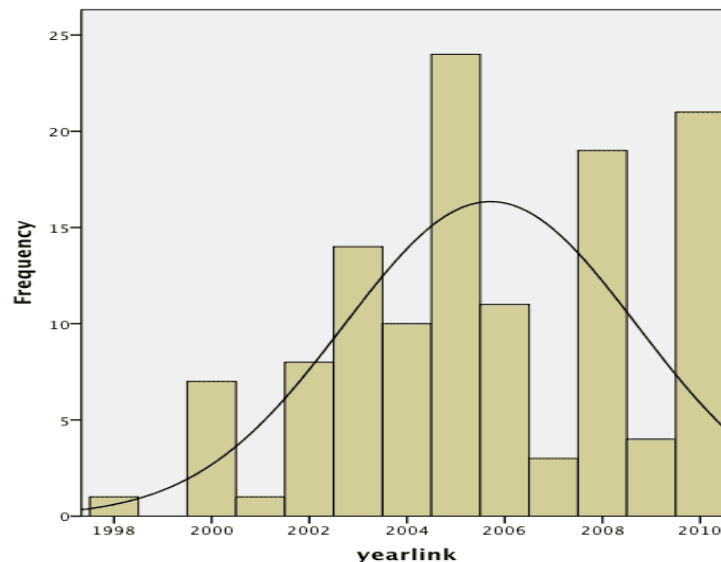
¹⁰⁶ Davis, *supra* note 86.

¹⁰⁷ Wilkerson, *supra* note 11, at 334.

**SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)**

since a link was cited and whether the link still works. If there was a relationship between time and invalidity, we would have many more invalid links occurring within earlier cases, and likely few or no invalid links in more recently decided cases. Additionally, we found no association between types of URLs (pdf vs. html) and their sources (government vs. non-government) and validity/invalidity; therefore we cannot predict whether certain types of webpages are more likely to become invalid than others.¹⁰⁸

Figure 2: Frequency of Invalid URLs



VI. POSSIBLE SOLUTIONS

Measures should be taken to safeguard and preserve the Internet content cited within Supreme Court opinions. It is encouraging that the Supreme Court itself has already taken steps to preserve the Internet content cited. The Judicial Conference is also aware that an issue with preservation exists, and they are taking steps to preserve the Internet resources cited.

Ideally, every court should digitally archive all materials cited within an opinion, regardless of the format. This would be of tremendous benefit to future researchers if access to such materials were readily and freely available online, particularly for materials

¹⁰⁸ *Contra* Wilkerson, *supra* note 11, at 334 (finding differences in link rot based on type of website. However, he lists percentages only and does not cite any statistical tests run).

that are updated frequently, such as loose-leaf publications. Loose-leaf publications are continuously updated through the replacement of existing pages with new ones in a binder. If a particular page in a loose-leaf were to be cited, that page may be discarded when the next update arrives. Rule Nineteen in the Bluebook,¹⁰⁹ which includes loose-leaf citations, does not provide citation guidelines for materials that are no longer available in print.

A. *The U.S. Supreme Court Clerk of Courts*

The U.S. Supreme Court retains a print copy of the cited Internet materials with the Clerk of Court's case file. The Court also takes the extra step of including the date in the copies of each cited Internet resource within the opinions. This is important because such resources can change over time, and researchers need to be able to pinpoint the correct version of the cited resource.

There are limitations to the approach taken by the Supreme Court, and limited access is one of the most significant issues. The case files are only available to those with sufficient means to go to Washington, DC, and visit the office of the Clerk of the Supreme Court. The case files are eventually sent to the National Archives and Records Administration (NARA),¹¹⁰ and access can be obtained by contacting or visiting NARA offices.¹¹¹ This approach is fine for simple text files, but what if images, sound, video, or software files are included as a significant aspect of the Internet page itself? A printed copy of the Internet resource cited may not be sufficient to capture all of the content on the page. The problem is further complicated when the cited source is an online audio, video, or software file—none of which can be printed and viewed on paper. For at least two cited videos, the Supreme Court has chosen to store and provide access to a copy of the videos on the Supreme Court website¹¹² but not for any other file formats. These files also have technology migration issues because they are not

¹⁰⁹ THE BLUEBOOK, *supra* note 68, R. 19, at 177-78.

¹¹⁰ *Records of the Supreme Court of the United States*, NATIONAL ARCHIVES, <http://www.archives.gov/research/guide-fed-records/groups/267.html> (last visited Apr. 10, 2013). As of Apr. 10, 2013, the National Archives and Records Administration website states that they have the case files for US Supreme Court cases up to 1997. However, NARA is not the most permanent of solutions, as they are now destroying documents from some federal cases. *See Making Room, Saving History*, UNITED STATES COURTS: THE THIRD BRANCH (May 2011), http://www.uscourts.gov/News/TheThirdBranch/11-05-01/Making_Room_Saving_History.aspx; William K. Ford, *Copy Game for High Score: The First Video Game Lawsuit*, 20 J. INTELL. PROP. L. 1, 7, 7 nn.25-27 (2012).

¹¹¹ *See Record Reproductions & Microfilm*, THE NATIONAL ARCHIVES, <https://eservices.archives.gov/orderonline/start.swe?SWECmd=Start&SWEHo=eservices.archives.gov> (last visited Apr. 10, 2013).

¹¹² *Video Resources*, *supra* note 36.

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

text-based.¹¹³

The Court could create its own archive of cited content by uploading the material to its website. This would not be technically difficult to accomplish, but the question is whether the Court has the staff and expertise available to handle such an endeavor. A partnership with an organization such as the Chesapeake Digital Preservation Group would allow the Court to share knowledge and technical information.

The Ninth Circuit Library may be able to provide some guidance because it has created its own archive of websites cited within Ninth Circuit opinions, dating back to 2008.¹¹⁴ The library “saves a copy of the cited material as a PDF file and adds a watermark to denote the document’s archived status.”¹¹⁵ This is an improvement over the Internet Archive’s complete lack of authentication.

B. The Judicial Conference of the United States

In 2009, the Judicial Conference of the United States created a report titled *Internet Materials in Opinions: Citations and Hyperlinking* that recommended two primary solutions to the broken Internet link problem:

- Clerks should download any cited Internet resources and include them with the opinions.
- The downloaded Internet resources should be included as attachments on a non-fee basis in each court’s Case Management/Electronic Case Files System, such as PACER.¹¹⁶

These recommendations are similar to those employed by the Supreme Court, so the same limitations mentioned earlier would apply here. It is unknown how many courts have adopted these recommendations. Additionally, there has been criticism for the idea of placing Internet sources on PACER, such as this argument by Lee Peoples:

Making Internet sources available through PACER
is a logical choice but not the most efficient way to

¹¹³ For examples of the difficulties regarding video files and archiving for the Supreme Court, see *supra* Section I.

¹¹⁴ *Websites Cited in Ninth Circuit Opinions, 2013*, UNITED STATES COURTS FOR THE NINTH CIRCUIT <http://www.lb9.uscourts.gov/webcites/2013.php> (last visited Apr. 10, 2013).

¹¹⁵ *Id.*

¹¹⁶ *Internet Materials in Opinions: Citations and Hyperlinking*, *supra* note 78.

make them accessible to the public at large. Most lawyers and legal researchers do not search for case law on the PACER system but use LexisNexis, Westlaw, or free alternatives. The PACER system has come under criticism recently for being difficult to use and expensive.¹¹⁷

C. Non-Court Based Alternative Solutions

There are two primary approaches to solving the dead links problem in court citations. The first approach would be to permanently archive a copy of the underlying materials. The second approach would be to assign a permanent URL to the materials, so they can be found regardless of where the content actually resides.

1. The Archival Approach

As discussed above, the Supreme Court is already archiving the Internet resources that are cited, but primarily in print format, which means that access to that archive is limited. Instead, we recommend that the Court consider archiving the materials on its website by partnering with existing online archival resources, such as the Internet Archive (also known as the Wayback Machine).

The Internet Archive is a nonprofit organization that runs an online library dedicated to permanently storing digital “snapshots of the World Wide Web” (copies of web pages taken at different times). The archive was founded in 1991 and is funded through grants, donations, foundations, and partnerships with outside organizations, including the Library of Congress.¹¹⁸ As of March 2013, the archive contained approximately 10 petabytes of data (or 10,000 terabytes).¹¹⁹

The Supreme Court website has been visited by the Internet Archive with increasing frequency starting in 2010.¹²⁰ On March 11, 2013, we ran our list of invalid URLs from Supreme Court opinions in the Internet Archive and found that 68% were accessible there. The success rate is significant enough that

¹¹⁷ Peoples, *supra* note 63, at 44.

¹¹⁸ FAQs, ARCHIVE.ORG, <http://archive.org/about/faqs.php#31> (last visited Apr. 10, 2013).

¹¹⁹ E-mail from Jeff, Internet Archive Team, to Fang Han, Research Fellow, The John Marshall Law School (Mar. 1, 2013, 17:01 EST) (on file with authors).

¹²⁰ Snapshots of the number of times <http://www.supremecourt.gov> was crawled by the Wayback Machine. INTERNET ARCHIVE WAYBACK MACHINE, http://wayback.archive.org/web/*/http://www.supremecourt.gov (last visited Apr. 10, 2013).

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

researchers can rely upon it to find at least a portion of the broken Internet links cited by the Supreme Court.

The success rate may be increased over time if the US Supreme Court or the Judicial Conference were to ask the Internet Archive to more thoroughly and frequently visit all court websites. Although the Internet Archive does not typically take suggestions for what websites to visit, it seems likely that the Archive would consider such a request. The key step would be to ask the Internet Archive to archive the web pages cited within the opinions themselves, which would ensure that at least one other copy of the cited documents would be freely available on a more permanent basis.

If the Internet Archive were unable to accommodate the Court, another option would be for the Court to subscribe to the Archive-it service¹²¹ from the Internet Archive. This service would allow the Court to create its own archive within the Internet Archive space. The National Archives and Records Administration (NARA) is already working with the Internet Archive to create specific snapshots of some federal websites, so NARA may be willing to expand this program and include the US Supreme Court website.¹²²

Whether or not materials found within the Internet Archive can be authenticated is another area of concern. The Internet Archive offers the following explanation for legal use of materials contained within the archive:

The Wayback Machine tool was not designed for legal use. We do have a legal request policy found at our legal page. Please read through the entire policy before contacting us with your questions. We do have a standard affidavit as well as a FAQ section for lawyers. We would prefer that before you contact us for such services, you see if the other side will stipulate instead. We do not have an in-house legal staff, so this service takes away from our normal duties. Once you have read through our policy, if you still have questions, please contact us for more information.¹²³

¹²¹ ARCHIVE-IT, <http://www.archive-it.org> (last visited Apr. 10, 2013).

¹²² *Web Harvest and Web Snapshot Information*, NAT'L ARCHIVES, <http://www.archives.gov/records-mgmt/policy/web-harvest-snapshot.html> (last visited Nov. 15, 2012).

¹²³ *FAQs*, ARCHIVE.ORG, <http://archive.org/about/faqs.php#274> (last visited Apr. 10, 2013).

If the Internet Archive is not a viable option, there are other Internet archiving projects that focus on specific subject areas. For example, the Chesapeake Digital Preservation Group,¹²⁴ discussed *supra*, is a consortium of state and academic law libraries, which may be better suited to accommodate such a request. The Chesapeake Group has been archiving web-based legal information since 2007, and as of 2012, the collection contains more than 8,600 digital items and 3,700 titles.¹²⁵ The Chesapeake Group could actually collect the Internet resources cited by the courts and work with the Internet Archive to then store them.

WebCite,¹²⁶ a free, on-demand archiving service that allows users to enter a specific URL to be archived, presents another approach. WebCite would allow court clerks to automatically archive the cited materials in each written opinion. The WebCite service can archive a specific page or it can “comb” through a web page and archive all the URLs listed within it. The process requires the clerk to enter the URL for the opinions into a web form, and the Internet page or site is automatically archived. Users can then reference the newly created, archival version of the web page. The archived pages are also passed along and archived in other online sites, such as the Internet Archive.

Here is an example of what a WebCite citation would look like:

Lessig, Lawrence: “this is a fantastically cool idea” (Blog). Sept. 8, 2006. <http://lessig.org/blog/2006/09/> Archived by WebCite at <http://www.webcitation.org/5UzgHmsS7> on 20-01-2008.

The above citation (in which Professor Lawrence Lessig discusses the WebCite service) contains both the original URL and the archived version of the URL, so researchers can use either version to find the original content.

Although the WebCite service shows a lot of promise, we have concerns about the viability of the organization. WebCite has apparently withdrawn its membership from the International Internet Preservation Consortium.¹²⁷ In fact, according to their website, “WebCite will stop accepting new submissions end of

¹²⁴ THE CHESAPEAKE DIGITAL PRESERVATION GROUP, <http://cdm266901.cdmhost.com/cdm/> (last visited Apr. 10, 2013).

¹²⁵ “Link Rot” and Legal Resources on the Web: A 2012 Analysis by the Chesapeake Digital Preservation Group, THE CHESAPEAKE DIGITAL PRESERVATION GROUP, <http://legalinfoarchive.org/cdm/linkrot2012> (last visited Apr. 10, 2013).

¹²⁶ WEBCITE, <http://www.webcitation.org> (last visited Apr. 10, 2013).

¹²⁷ Members, INTERNATIONAL INTERNET PRESERVATION CONSORTIUM, <http://www.netpreserve.org/about-us/members> (last visited Apr. 10, 2013).

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

2013, unless we reach our fundraising goals to modernize and expand this service.”¹²⁸

2. *The Permanent URL Approach*

Another solution would be to create “persistent identifiers” for cited Internet resources.¹²⁹ The transitional nature of Internet resources makes it extremely difficult to rely on them within court opinions. It is easy to move entire folders of information to new locations and new servers, but not every website will leave a forwarding address (or a redirect). The cited resources may still exist online, but they could have moved to a different URL.

In order to provide improved address reliability, the persistent identifier approach would point to the actual document itself, instead of the location of the document. This is accomplished by using a centrally managed registry of documents each with a digital object identifier (DOI). The registry would then automatically direct users to the location of the Internet resource. There are many challenges to maintaining such a registry, but the most significant drawback is the requirement that content owners notify the registry every time the location of the content changes.

However, “No legal citation guide requires using persistent identifiers []although both the *Bluebook* and the *ALWD Citation Manual* recommend using unique identifiers in commercial databases[], and law journals generally have not used DOIs in footnotes, even when DOIs exist for cited articles.”¹³⁰ Therefore, the chances of law review articles including DOIs without greater acceptance by the two major citation manuals are close to nil.¹³¹

¹²⁸ WEBCITE, <http://www.webcitation.org> (last visited Apr. 10, 2013).

¹²⁹ See Lyons, *supra* note 53.

¹³⁰ Keele & Pearse, *supra* note 39, at 392.

¹³¹ Legal citations already include less information than many other frequently used citation forms by deliberately excluding information such as last page number and publisher, unlike APA, MLA, and Chicago Style formatting.

Here is an example of a citation using a persistent identifier:

Blanchi, Christophe, Sandra Payette, Carl Lagoze, Edward A. Overly, "Interoperability for Digital Objects and Repositories: The Cornell/CNRI Experiments," *D-Lib Magazine*, May 1999.
[doi:10.1045/may99-payette]

LegisLink.org¹³² is an attempt to create a citation-based URL that allows users to link directly to sublevels within legislative documents. The service is a collaborative effort from the Open Source Software Institute, Cornell Legal Information Institute, and the eCitizen Foundation. The LegisLink service will save a copy of document and insert anchors at the appropriate places.

Here is an example of a LegisLink URL that links to Title 3, section 302 of the USC.

<http://legislink.org/us/USC-3-302>

3. *Other Projects of Note*

The Memento Project¹³³ calls itself time travel for the Web because users can seamlessly jump in time from one snapshot of a web page to another. Users who install the MementoFox add-on for the Firefox browser can look up a web page and then use the slider on the Memento add-on to view different versions of the same web page at different points in time. The Memento Project requires a digital archive, such as the Internet Archive, to use a standard format, which allows Memento to access and use content from many different archives at once. There is even a Memento browser app for mobile devices running the Android operating system.¹³⁴

The Legal Information Preservation Alliance (LIPA)¹³⁵ is a nonprofit organization of academic law libraries dedicated to the preservation of print and electronic legal information. LIPA sponsors a wide variety of projects including:

- The Law Review Preservation Program¹³⁶ will archive law reviews published in the bePress Digital Commons system.

¹³² LEGISLINK, <http://www.legislink.org> (last visited Apr. 10, 2013).

¹³³ MEMENTO, <http://mementoweb.org/> (last visited Apr. 10, 2013).

¹³⁴ MEMENTO-BROWSER, <http://code.google.com/p/memento-browser> (last visited Apr. 10, 2013).

¹³⁵ LEGAL INFORMATION PRESERVATION ALLIANCE, <http://www.aallnet.org/lipa> (last visited Apr. 10, 2013).

¹³⁶ *Law Review Preservation Program*, LEGAL INFORMATION PRESERVATION ALLIANCE, http://www.aallnet.org/lipa/law_review_pres.asp (last visited Apr. 10, 2013).

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

The articles will be stored within CLOCKSS—an international dark archive run as a not-for-profit organization that implements a Creative Commons licensing scheme for articles that are no longer available from the publisher.

- LIPA sponsors Archive-It access for consortia members. This is a subscription service from the Internet Archive to harvest and preserve born-digital documents.
- The Chesapeake Project Legal Information Archive (discussed *supra*) was sponsored by LIPA.

The LIPA organization is another group that may be able to provide the Supreme Court and other federal courts with assistance with their archiving efforts.

CONCLUSION

As demonstrated through our research, the number of Supreme Court citations to Internet links that are no longer working is alarming. Of the 430 links contained in Supreme Court opinions from 1995-2010, 29% are invalid, closely following the results of earlier, related studies. It is disturbing that even at the Supreme Court, where creating and citing precedent is of the utmost importance, citations often fail to point the researcher to the authority on which the court based its decision.

Because our research also found that, within the context of Supreme Court opinions, there is no clear connection between types or sources of links and the links' validity, we cannot reasonably predict which individual links will rot and thereby become inaccessible. This uncertainty gives greater weight to the need to preserve all sources cited within Supreme Court cases in a format that can be found and accessed by future lawyers and researchers. Technology will continue to change, but our study helps to demonstrate that the ephemeral, ever-changing nature of the Internet also leads to unpredictability of which links will rot. Therefore, because any link has the potential to rot, any citation to an Internet website might not include a valid link when needed—today, tomorrow, or at any point in the future. The fact that the Supreme Court itself has links to its own website that no longer function shows the depth of the link rot problem.¹³⁷

Lee Peoples details the problems with the interaction between the need for permanence in legal information with the ephemeral nature of websites. An example is given where:

¹³⁷ See *supra* Section I.

A court cites a blog post discussion of a substantive legal issue to support its reasoning on that issue. The court does not cite any other authority in support of its reasoning. Several weeks after the court's opinion is published, the blogger changes her mind on the issue and edits the blog post cited in the opinion. The issue for which the blog was cited may not have been significant at the time the opinion was published, but imagine several years pass and the issue becomes significant. Researchers, lawyers, and judges begin to examine the court's opinion for guidance on the issue. They pull up the blog post cited in the opinion and are puzzled to find that the post no longer includes the information originally attributed to the blog. Or imagine that instead of the post being updated after the opinion is published, the blog post or the entire blog has been deleted. What will researchers do when trying to follow the logical steps of the court's argument? When a blog post has changed significantly after being cited or disappears entirely, researchers are likely to lose confidence in the court's opinion.¹³⁸

While his example seems like only a thought experiment, actual examples of permanently lost websites cited by the Supreme Court already exist. But even more disturbing, at least some of the websites cited by the Court are searches themselves—meaning that if the website still exists, we will never see the website as the Justice saw the website.

Our overall recommendation is twofold: first, the Supreme Court should follow the lead of the Ninth Circuit to archive cited websites on the Court's own website; and second, as a longer-term solution, the Supreme Court and the Judicial Conference should form partnerships with existing Internet archiving organizations, such as LIPA and the Internet Archive, in order to provide easier access to the Internet resources cited within each court opinion.

¹³⁸ Lee F. Peoples, *supra* note 5, at 71, 72-73.

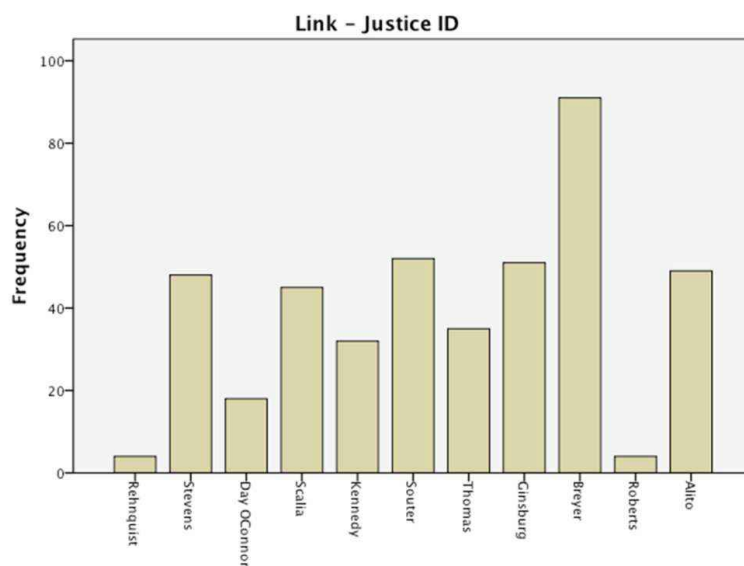
SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

APPENDIX

Though we did not focus on individual Justices in our article, early commenters were interested in differences between Justices. Therefore, below is an appendix regarding individual Justices for our study. Overall, Justices Alito and Breyer are the most likely to use links, with Chief Justices Roberts and Rehnquist the least likely. The present Justices' citation to Internet sites compared to their citation of scholarship is quite different.¹³⁹ Only Justice Alito heavily cites both websites and scholarship, while Justice Breyer heavily cites websites but not scholarship, compared to other Justices.¹⁴⁰

In the bar graph in Figure 3 below, the frequency of Internet resources cites (y axis) is plotted by Supreme Court Justice (x axis). Justice Breyer is the most frequent user of URLs in his opinions, with around ninety uses, while Chief Justices Roberts and Rehnquist were the least frequent users, with fewer than ten uses.

Figure 3: Total Number of URLs per Justice (1995-2010)

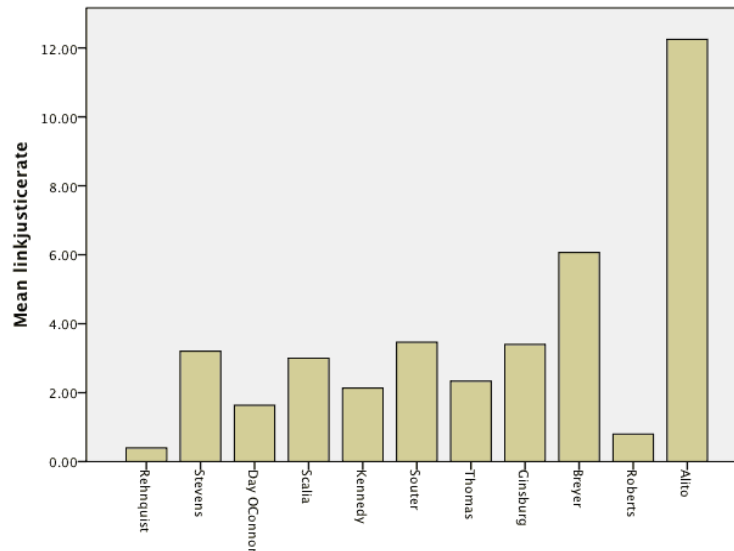


¹³⁹ Petherbridge & Schwartz, *supra* note 10, at 1025 tbl. 6 (study finding that the percentage of cases citing legal scholarship by the present Supreme Court Justices through 2010 ranged from the highest for Justice Ginsberg (29.08%) to the lowest for Justice Thomas (10.53%)).

¹⁴⁰ *Id.* (study finding that the percentage of cases citing legal scholarship by the present Supreme Court Justices through 2010 was 27.27% for Justice Alito (second highest percentage) and 19.55% for Justice Breyer (second lowest percentage)).

The bar graph in Figure 4 below shows that Justice Alito has used a disproportionately high number of URLs in his opinions when compared to the other Justices, while once again, Chief Justices Roberts and Rehnquist were the least likely to use links.

Figure 4: Average Number of URLs per Year Served (2005-2010)



Again, in the bar graph in Figure 5 below, we looked at the use of Internet links in citations in five-year increments per Justice. Still, Justice Alito has a high number of URLs in his opinions, while Chief Justices Roberts and Rehnquist had low incidences of citations to links. Within our study period, individually, two of the Justices, Breyer and Alito, used far more URLs from 2006-2010 than in previous years.

SOMETHING ROTTEN IN THE STATE OF LEGAL CITATION:
THE LIFE SPAN OF A UNITED STATES SUPREME COURT
CITATION CONTAINING AN INTERNET LINK (1996-2010)

Figure 5: Number of Links in Five-Year Increments

