

## Keeping Validity in Cite: Web Resources Cited in Select Washington Law Reviews, 2001–03\*

Helane E. Davis\*\*

*Ms. Davis analyzes three years of URLs cited in Washington law reviews. The sample indicated high invalidity rates, that certain domains had higher invalidity rates than others, and that shorter URLs fared better than longer ones. These rates validate ongoing concern about the impermanence of Web sources.*

Technology has invaded the world of legal information. . . . It is commonplace to complain about the proliferation of judicial decisions, but the far larger universe of information that is available to the online legal researcher today is not primary source legal material. It is all of the world's information that lies at the other end of a hypertext link. Legal researchers are awash in judicial reports, but that is only a drop in the ocean of information that lies a keypad away.<sup>1</sup>

¶1 While the author of this passage was commenting on the entire world of legal information, nowhere is the truth of this statement more evident than in the citation of Web resources in legal scholarship. Law reviews have not been immune to the increasing overlap of the print and electronic worlds. A venue long respected because publication is equated with accomplishment finds itself increasingly populated by citations with the capacity to be as transitory as print is enduring.

¶2 The most in-depth look at Web citations in law reviews to date was done by Mary Rumsey.<sup>2</sup> Focusing on the permanence of these citations, Rumsey began her study with a presumption that “link rot”<sup>3</sup> would be significant and only increase as documents continued to age. Rumsey sampled a wide variety of publications

---

\* © Helane E. Davis, 2006.

\*\* Director and Assistant Professor of Law, University of Kentucky College of Law Library, Lexington, Kentucky. The author wishes to acknowledge her gratitude to Stephen Burnett, Associate Dean for Business Development, Concord Law School; Penny Hazelton, Associate Dean for Library and Computing Services, University of Washington School of Law; the librarians at the Seattle University Law Library; and the librarians at the Marian G. Gallagher Law Library at the University of Washington for their support, encouragement, and assistance.

1. Robert C. Berring, *Legal Research and the World of Thinkable Thoughts*, 2 J. APP. PRAC. & PROCESS 305, 311–12 (2000).
2. See 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, 2002 LAW LIBR. J. 2.
3. Rumsey adopts Howard Denemark's definition for link rot. *Id.* at 30 n.21. Denemark defined it as “the problem of links that lead the reader to defunct sites.” Howard 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, 32 n.77. Ironically, the Web source Denemark cited in support of his definition was an invalid link by the time the Rumsey study was published in 2002, a fact acknowledged by Rumsey, *supra* note 2, at 30, n.21. As of this writing, the link is still broken.

and a large number of Web citations, examining a hundred documents a year over a five-year period. She also compared link stability in legal scholarship to a similar study examining undergraduate work, and reviewed content stability, that is, whether a particular Web address contained the same information when tested as it did when cited.

¶3 In contrast, this study is comprised of a much narrower sample, originally necessitated by the fact that it was conceived as a class-based project focused purely on comparing how, when, and by whom domains were cited in specific law reviews. A smaller sample size also allowed for more in-depth analysis of a single set of comparative content issues, a manageable way of revisiting the issue of URL impermanence to see if the issues Rumsey identified still persisted. And although trends in three publications can be extrapolated only so far, I believed that if the trend identified by Rumsey had been reversed, even a small sample would indicate such. Therefore, I limited my study to examining three years of URLs cited in Washington law reviews and focusing on overall validity rates. Although distinguishable in scope and method from Rumsey's study, both resulted in high URL failure rates, validating the ongoing concerns about the impermanence of Web sources.

¶4 Having studied how users approach the Web and their expectations about the information they will locate there, library and information professionals often question whether it is wise for users to apply the same expectations they have for print sources to Web sources. Many of these professionals presume that users do just that, while believing that they should not.<sup>4</sup> One reason for that presumption is a laissez-faire attitude about credibility that many users have about Web content. This begs one particular question: If the Web does not engender the same sets of expectations for permanence, will anyone care when Web resources disappear? The question helps frame this study, by evaluating the data users see and one example of the context in which they encounter it.

### **Initial Assumptions**

¶5 I undertook this study with certain assumptions firmly in place about the number of Web resources cited in law reviews, the types of resources cited, and certain trends.

#### ***Number of URLs Cited***

¶6 I fully expected to find a significant number of URLs included in law reviews, although I did not have a specific percentage in mind at the outset of the study. Given the nature of this particular type of legal scholarship, I did not think that most of the legal authority cited would be electronic or digital. Legal scholarship

---

4. In fact, concerns about how students treat information that they locate on the Web is evidenced by the proliferation of guides on evaluating Web sources. Such guides are found on nearly every college library Web site.

has long been and will continue to be dominated by citations to print materials. This is true for two primary reasons. The first relates to the nature of legal authority and how lawyers approach it. Legal scholarship may have long been a place for lawyers and academics to explore new (even revolutionary) avenues and ideas about the practice and purposes of the law, but the way that scholarship is created and produced is an outgrowth of a system that has been very slow to change.<sup>5</sup> The second reason is that law reviews overwhelming use *The Bluebook*, the most widely accepted guide to legal citation, which clearly states that print formats should be used whenever they are available.<sup>6</sup>

¶7 Nevertheless, I expected to find that the increase in Internet awareness and use had permeated law reviews.<sup>7</sup> Consequently, I expected an increase each year, with more articles per issue containing URLs and more URLs per article the more recently an article was published. I also expected that the status of a writer would not create any discernible pattern to the use of URLs or their validity.

### *URL Domains and Validity*

¶8 My expectations regarding which types of Web resources were being cited and where they are located began with a presumption that most of the URLs would correspond to .gov and .org domains.<sup>8</sup> I thought it unlikely that most commercial Web sites<sup>9</sup> (i.e., .com domains) would contain the primary and secondary legal sources traditionally cited to in law review articles. I also expected authors not to cite to Web versions of easily obtainable primary sources. I did not expect international or foreign sources to be heavily cited. This hypothesis was a function of my understanding of how the “main” law reviews work at most law schools, favoring core substantive issues and leaving specialized issues to subject-oriented journals.

¶9 However, if there was a pattern about reliability of URLs, I expected the .com domain Web addresses to be more consistently stable and valid than other

- 
5. This is also why it seemed unnecessary to compare the age of URLs to the age of the print sources cited in any given article. Since age alone never disqualifies the value of legal authority, and since subject matter can have a significant impact on the types of sources cited in an article, this type of comparison did not hold much value.
  6. THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION rule 18, at 151 (Columbia Law Review Ass'n et al. eds., 18th ed. 2005) (“*The Bluebook* requires the use and citation of traditional printed sources unless (1) the information cited is unavailable in a traditional printed source; or (2) a copy of the source cannot be located because it is so obscure that it is practically unavailable.”).
  7. This is hardly a great leap given the ways in which computer-assisted legal research, specifically Westlaw and LexisNexis, now permeate legal research and legal practice, especially in academia where the real-world cost barriers do not exist. The Internet and the Web (not including the “deep web”), however, are generally free. Their resources are open to anyone who can locate them.
  8. For an explanation of domains, see the Internet Corp. for Assigned Names & Numbers, What is the Domain Name System? <http://www.icann.org/faq/#dns> (last visited Apr. 19, 2006). For a complete list of top-level domains, see Internet Corp. for Assigned Names & Numbers, Registry Listing, <http://www.icann.org/registries/listing.html> (last visited Apr. 19, 2006).
  9. With the exception of the commercial computer-assisted legal research services, e.g., Westlaw, LexisNexis, Versuslaw, or Loislaw. All of these have .com Internet interfaces but they are not the type of .com sites included in the study.

domains. This was based on the premise that owners of commercial sites have greater incentive (and financial means) to keep their Web sites in good working order since those sites represent a part of their overall moneymaking apparatus. At the outset, I had no preconceptions about the percentage of URLs that would still be valid. I did, however, hope that the number would be significantly high (e.g., more than two-thirds).

### *Publication Comparisons*

¶10 Regarding the three publications included in the study, I wondered if usage and validity trends would differ in any way along the lines of the schools' relative rankings.<sup>10</sup> Arguably, ranking *should not* have any impact on these factors since all law schools generally enforce rigorous editorial and technical policies when it comes to student-edited publications.<sup>11</sup> Their incentive in so doing is significant, given two factors. First, law schools are always happy to see healthy subscription rates for their student-edited journals. Subscription rates can be indicative not only of the success of the publication but also its reputation within the legal academy. Second, the reputations that student members of these publications take out into the world with them after graduation reflect on the school's reputation and standing. My expectations notwithstanding, I did wonder what the data would show, and if the publication policies regarding the use of URLs played any role.

## **Methodology**

### *About the Publications*

¶11 Student-edited (and usually student-run) publications serve a unique role in legal scholarship. They provide a unique academic experience for law students. They vet the work of the school's brightest students (thereby providing an essential credential for later career use). They provide the primary publication medium for legal scholarship for the entire legal academy—both faculty and students.<sup>12</sup> Each of the publications in this study serves these functions for its school.

- 
10. Temporarily discounting complaints about law school rankings, my source was, for better or worse, the oft-criticized *U.S. News & World Report* law school rankings. The 2005 rankings consulted in the study remain on file with the author.
  11. I still believe this to be the case despite occasional dissatisfied faculty rumblings about the editing standards of student-run publications. For an example of the context of this debate, see Phil Nichols, *A Student Defense of Student Edited Journals: In Response to Professor Roger Cramton*, 1987 DUKE L.J. 1122.
  12. Legal scholars have, from time to time, looked upon the practice of students controlling publication of faculty scholarship with disdain. See, e.g., Arthur D. Austin, *The "Custom of Vetting" as a Substitute for Peer Review*, 32 ARIZ. L. REV. 1, 4 (1990) ("The use of student edited journals as the main outlet for legal writing is an embarrassing situation deserving the smirks of disdain it gets from colleagues in the sciences and humanities."); John Kester, *Faculty Participation in the Student-Edited Law Review*, 36 J. LEGAL EDUC. 14, 14 (1986) ("[S]tudents without law degrees set the standards for publication in the scholarly journals of American law—one of the few reported cases of the inmates truly running the asylum.").

¶12 The *Washington Law Review (WLR)* is the oldest of the three journals in the study, having begun publication in 1925.<sup>13</sup> Like most law student-edited publications, it publishes a combination of professional<sup>14</sup> and student legal scholarship. Publication occurs four times a year. Students at the University of Washington Law School become members of the law review via one of four routes: outstanding first-year students are invited to join based on their first-year grades; first-year students can participate in a “write-on” competition; a combination of these two methods; or selective invitation based on other “analytical” work during the first year.<sup>15</sup> Articles are selected for publication based on interest, timeliness, and quality of work.<sup>16</sup>

¶13 The *Seattle University Law Review (SULR)* originally began as the *University of Puget Sound Law Review (UPSLR)* and, in that form, it dates back to an “intramural” issue published in 1975. The first issue of the official *University of Puget Sound Law Review* was published in 1977. When the law school was purchased by Seattle University in 1994, *UPSLR* became the *Seattle University Law Review*. *SULR* publishes a combination of professional and student materials four times a year. The student membership is comprised of both students who were extended membership based on their first-year grades, and those who qualified based on a written competition.

¶14 The *Gonzaga Law Review (GLR)* was started in the mid-1960s (its first issue was published in March 1966) by the then dean of the Gonzaga Law School.<sup>17</sup> Following the typical model of publishing a mix of professional and student scholarship, *GLR* publishes three times a year and focuses on “current legal issues.”<sup>18</sup> Gonzaga students become members of *GLR* based on first-year grade point average and performance on a mandatory written competition.

¶15 These three journals were selected because they represent the main student-run publications of each of the parent institutions. In each case, I presumed that their longevity would provide the best foundation for drawing conclusions about the patterns of use and characteristics of Web resources in law review articles, given their long histories and the established practice of using print resources in legal scholarship.

¶16 It should be noted that each of these institutions has specialized journals that were not included in this study. In addition to *WLR*, the University of

---

13. Wash. Law Review, A Brief History of the *Washington Law Review*, <http://www.law.washington.edu/WLR/history.html> (last visited Apr. 19, 2006).

14. “Professional” in this context refers to works by nonstudent authors: faculty, practitioners, and judges.

15. Wash. Law Review, Membership, <http://www.law.washington.edu/WLR/membership.html> (last visited Apr. 19, 2006).

16. Wash. Law Review, About the *Washington Law Review*, <http://www.law.washington.edu/WLR/aboutWLR.html> (last visited Apr. 19, 2006).

17. Gonzaga Law Review, History, <http://www.law.gonzaga.edu/Academics/Law+Reviews/Gonzaga+Law+Review/History.htm> (last visited Apr. 19, 2006).

18. Gonzaga Law Review, <http://www.law.gonzaga.edu/Academics/Law+Reviews/Gonzaga+Law+Review/Gonzaga+Law+Review.htm> (last visited Apr. 19, 2006).

Washington publishes two specialized law journals: *Pacific Rim Law & Policy Journal* and *Shidler Journal of Law Commerce & Technology*.<sup>19</sup> Gonzaga publishes *Gonzaga Journal of International Law*.<sup>20</sup> Seattle University's specialized journal is the *Seattle Journal for Social Justice (SJSJ)*.<sup>21</sup> These journals were excluded from the study for various reasons. *Shidler* and *SJSJ* are relatively new publications that were not published during the entire run of the study. The Gonzaga journal was an online-only journal during the period studied, and I wanted to limit the study to citations to online resources in traditionally print publications.

### *Defining the Sample*

¶17 The sample for this study includes all articles—nonstudent-authored and student-authored—published in the *Washington Law Review*, *Gonzaga Law Review*, and *Seattle University Law Review* during 2001, 2002, and 2003 that cite to at least one URL. Since the study's primary focus is comparing domains cited and their validity, articles without any URL citations were omitted. The total sample contained 119 articles citing 1104 URLs.<sup>22</sup> The sample was derived with an online search of a computer-assisted legal research system performed in early 2004.<sup>23</sup> Due to the volume of documents originally retrieved, the initial search was refashioned into individual searches, by year, for better manipulation of the data set. Documents were stripped down to the fields containing the search terms and then printed.<sup>24</sup> For each article, the following was recorded: article citation, title, author(s), and publication date. Not only did this information provide a reference for individual URLs, it was used for verification purposes and to revisit articles, if necessary, to problem-solve. For each URL, I recorded the article number, URL, domain, validity, inclusion of a non-HTML document format (e.g., .doc, .pdf, .txt, or .rtf), and any unusual characteristics.

19. For more information on these journals, see *Pacific Rim Law & Policy Journal*, <http://www.law.washington.edu/PacRim/> (last visited Apr. 19, 2006); *Shidler Journal of Law, Commerce & Tech.*, <http://www.lctjournal.washington.edu/> (last visited Apr. 19, 2006).
20. For more information, see *Gonzaga Journal of Int'l Law*, <http://www.gonzagajil.org> (last visited Apr. 19, 2006).
21. For more information, see *Seattle Journal for Social Justice*, <http://www.law.seattleu.edu/sjsj> (last visited Apr. 19, 2006).
22. These 119 articles represent 40.34% of the total number of articles (295) that were published from 2001 to 2003 by the three journals in the study, supporting the notion that print sources still predominate. However, there was a marked increase in the total number of URLs across the sample each year, with 217 URLs cited in 30 articles in 2001, 345 URLs cited in 39 articles in 2002, and 542 URLs cited in 50 articles in 2003.
23. While the search statement was the result of several fine-tunings, the final search run in Westlaw's Washington Journals & Reviews database (WA-JLR) was **ftp http www\* gopher nntp wais telnet & so(seattle gonzaga "washington law review") & da(200x)**.
24. This process introduced a handful of errors whenever using the **Locate** command produced a page with the term but not necessarily the complete URL. In the few instances where this happened, the actual document title was used to later revisit the database and retrieve the complete URL from the document.

¶18 During compilation, I made minor corrections to URLs. Several of the URLs appeared in online versions of articles with additional spaces that usually caused a hyperlink to fail when followed. Spaces inserted after the slashes and before *www* were stripped. Likewise, additional spaces added around file extensions at the end of URLs were also stripped. However, I did not correct obvious typographical or spelling errors. Percent signs in URLs were left undisturbed. I counted these links as working or not depending on what happened when I tested them.

### *Testing Methodology*

¶19 My initial validity testing of URLs was automated, but unfortunately unusable. Using a list of combined URLs by school, I used the built-in hyperlink verification function of MS Front Page, an HTML editing application, to test the validity of the URLs. Unfortunately, not only was the list difficult to use when MS Front Page reordered the URLs, but spot-checking revealed that the editor was returning links as broken that were not.<sup>25</sup> As a result, I reverted to checking all of the URLs in the sample manually. All were checked for validity within a forty-eight-hour period in October 2004.

¶20 Redirection to pages containing password prompts happened several times during the testing process. URLs that link to news stories on commercial newspaper Web sites often expire if the site does not permit unlimited access to its archives.<sup>26</sup> In some instances, access is restricted to registered users. Short of establishing site registration for every newspaper cited in the study, it was impossible to determine if certain links were valid. Consequently, URLs that linked to items since archived or otherwise made unavailable were considered invalid for the purposes of this study based on the premise that they no longer provided access—on their face—to the information cited.

### *General Sample Characteristics*

¶21 Several general trends were evident before I had performed any in-depth analysis of the sample. For instance, while the search I wrote to create the sample group of articles was designed to cull URLs using any of several protocols, of the 1104 URLs only *one* was for a protocol other than the hypertext protocol, or *http*. An overwhelming number of the URLs cited by the documents in the sample were to *HTM* or *HTML* documents. However, a significant number did cite to another

---

25. It also truncated URLs that included certain characters. For instance, if a URL included a question mark, an octathorp, or an equal sign, the HTML software often truncated the URL to either the next closest subdirectory or the original domain. Depending upon how a site is constructed, this may cause a URL to fail.

26. This period varies by newspaper and is often anywhere from seven to sixty days. For example, Web sites for the *Washington Post* ([www.washingtonpost.com](http://www.washingtonpost.com)), the *New York Times* ([www.nytimes.com](http://www.nytimes.com)), and the *Seattle Times* ([www.seattletimes.com](http://www.seattletimes.com)) all require site registration to perform certain functions, including retrieving older articles. Searching and accessing the *Seattle Times* archives is free, once registered. Searching is free for the other two newspapers but there is a fee to retrieve *New York Times* articles older than seven days and *Washington Post* articles older than fourteen days.

document type. Additionally, a wide range of top-level domains (TLDs) was represented in the final sample. Not only did the articles cite to URLs with .com, .gov, and .org TLDs, but there were also .state, .edu, and several country TLDs found in the sample.

¶22 To ensure that my online retrieval methodology located all the applicable articles in the publications comprising the sample for the specified years, I verified that each article appearing in the law review's printed version also appeared in the online database against which I searched.<sup>27</sup> By comparing the table of contents for each journal to the applicable list of electronically retrieved documents, I was able to verify that every article printed appeared on my Westlaw citation list and every document on my Westlaw citation list appeared in the print table of contents.

## Findings and Analysis

### *URL Validity*

¶23 The focus of this study was to determine the extent to which the URLs cited in articles published in 2001, 2002, and 2003 were still valid when tested in October 2004. Since a citation is designed to lead the user to the source it identifies, this is quite important. For those articles published in 2001, this statistic holds particular interest given findings from Rumsey's study regarding the time it takes for links to expire or break.<sup>28</sup> A URL was considered invalid if: it generated a "404 not found" message, defaulted to a login screen requiring either a password or site registration, redirected the user to a generic Web hosting service, or produced an error message. Error messages usually reported that the page was no longer available,<sup>29</sup> that it had been moved, or that the site had been redesigned. If testing a URL led to automatic redirection that didn't require user intervention, that URL was considered valid for purposes of the study as long as it redirected to appropriate content (in contrast to a generic placeholder).

¶24 Of the 1104 URLs in the study sample, 441, or 40%, were broken links when tested. As is to be expected, the number of invalid URLs rises as the total number cited in a publication rises. However, although the data confirm my expectation that more recent articles have a better chance of containing valid URLs than older articles, a 37% failure rate for the most recent year in the study is still a remarkably high invalidity rate. Table 1 illustrates the invalid links by publication and by year.

---

27. Database information for Westlaw's Washington Journals & Reviews database (WA-JLR) notes that full-text coverage is available within a certain date range to the extent that permissions were obtained to include articles. Although the recentness of the years in the sample led me to believe that the likelihood was quite high that every article printed had been included in Westlaw, this characteristic of the database (and most databases containing similar material) made verification necessary.

28. Rumsey, *supra* note 2, at 35, ¶ 23.

29. The difference between such a message and a "404 not found" message is the difference between the page no longer connecting to its domain versus the page making the connection but linking to different content.

**Table 1***Number of Invalid URLs by Publication; Percentage of That Year's Total*

Publication	2001		2002		2003	
	No. of Bad Links	% of Total Links	No. of Bad Links	% of Total Links	No. of Bad Links	% of Total Links
Washington Law Review	23	55	33	38	53	32
Seattle Univ. Law Review	70	47	49	36	110	41
Gonzaga Law Review	13	48	54	44	36	32
TOTAL	106	49	136	39	199	37

¶25 How specific domains perform when evaluating URL validity is also of interest. The highest percentage of invalid URLs came with one of the least cited top-level domain groupings: state. Since this is actually a combined category representing more than jurisdiction, it is impossible to draw any specific conclusions about the validity of URLs from a particular state.<sup>30</sup>

¶26 Among the remaining TLDs, the percentages are more evenly distributed. There is no pattern across publications or years for any given domain except in the most general terms. Each publication had at least one domain with an invalidity rate of over 60%; however, it was not always the same domain. Table 2 details invalid URLs by domain and year.

### *Domains*

¶27 Nine top-level domains or TLDs were identified among the URLs in the sample. They are .com, .org, .edu, .net, .gov, .int, and .info. In addition, the few individual country-related domain endings found in the data were combined under a general “ctry” category, and the few state-related domain endings were combined under a general “state” category. URLs were only categorized in these two groups in the absence of another domain within the URL.<sup>31</sup> Table 3 shows how the 1104 URLs in the sample break down by domain and by year.

- 
30. A “by state” measure would probably be more valuable in a larger sample. Such a sample would include a broader subject matter base and would include law reviews published in multiple states. Theoretically, such a measure would provide a snapshot of which states do the best job at maintaining their Web sites.
31. In some instances, a URL included a .gov domain and other identifying subdomains that clearly indicated that it pointed to an official state Web site or subsite. In other instances, URLs pointing to official state sites were configured with .state.xx.us. subdomain endings. For data analysis purposes, I did not conform these URLs to one pattern or another: .gov domains were counted in that category and .state.xx domains were counted in that category only if they did not also include a .gov domain.

**Table 2***Invalid URLs by Domain and Year (Including Percentage of Domain Total)*

Domain	2001		2002		2003	
	Total	% of Domain	Total	% of Domain	Total	% of Domain
com	34	40	34	31	68	39
org	17	35	40	36	46	32
edu	14	67	12	48	20	43
net	1	33	4	57	1	11
gov	30	61	39	48	44	37
state	9	100	2	50	6	46
ctry	1	100	5	63	10	38
int	0	0	0	0	3	38
info	0	0	0	0	1	100

**Table 3***Number of Top-level Domains by Year*

Domain	2001	2002	2003	Total
com	85 (39.2%)	109 (31.6%)	176 (32.5%)	370 (33.5%)
org	49 (22.6%)	110 (31.9%)	142 (26.2%)	301 (27.3%)
edu	49 (22.6%)	82 (23.8%)	120 (22.1%)	251 (22.7%)
net	21 (9.7%)	25 (7.2%)	47 (8.7%)	93 (8.4%)
gov	1 (.46%)	8 (2.3%)	26 (6.6%)	35 (3.2%)
state	9 (4.1%)	4 (1.2%)	13 (2.4%)	26 (2.4%)
ctry	3 (1.4%)	7 (2.0%)	9 (1.7%)	19 (1.7%)
int	0	0	8 (1.5%)	8 (.72%)
info	0	0	1 (.18%)	1 (.09%)
TOTAL	217	345	542	1104

¶28 Consistent with the general pattern of an increase in the number of citations over the three-year span of the sample, these numbers also increase over that span. Most noteworthy are the domains most frequently occurring in the URLs cited, and the comparisons between the publications over the three-year span. With the exception of the 2003 *WLR* and 2003 *SULR* URLs, .com domains are the second most frequently cited domain—after .org—for each publication and each year. This was an unexpected finding because I had assumed that academic scholarship, with its heavy emphasis on primary law or the implications of same, wouldn't include as many cites to commercial sites as to governmental, educational, or noncommercial entities. Notably, .gov domains are cited almost as frequently as .org domains.

¶29 With aggregated domain percentages for all years, table 4 provides an overview of all the domain-related trends, including the fact that after accounting for the “big three” domains, the next most frequently cited domain is not the same for all publications.

### *URLs Cited*

¶30 As noted previously, the sample contained 1104 URLs cited in 119 articles from 2001–03, for an overall average of 9.3 URLs cited per article. Table 5 breaks this down by publication and by year. Consistent with my initial presumption, both the total number of URLs cited and the number of articles citing to them rose steadily from year to year. Table 6 represents the percentage of change from year to year, per publication.

**Table 4**

#### *Individual Domain Percentages by Publication*

<b>Domain</b>	<b>Washington Law Review</b>	<b>Seattle Univ. Law Review</b>	<b>Gonzaga Law Review</b>
com	27.4	42.4	21.8
org	28.8	23.1	34.4
edu	22.3	21.6	25.6
net	5.1	7.3	14.5
gov	5.8	1.8	3.1
state	6.2	1.45	0
ctry	3.8	1.1	.76
int	.68	1.1	0
info	0	.18	0

**Table 5**

#### *Number of URLs Cited by Year and by Publication*

<b>Publication</b>	<b>2001</b>	<b>2002</b>	<b>2003</b>	<b>TOTALS</b>
Washington Law Review				
No. of Articles	12	10	14	36
No. of URLs	42	86	164	292
Average URLs/Article	3.5	8.6	11.1	8.11
Seattle Univ. Law Review				
No. of Articles	15	15	24	54
No. of URLs	148	136	266	550
Average URLs/Article	9.87	9.07	11.1	10.2
Gonzaga Law Review				
No. of Articles	3	14	12	29
No. of URLs	27	123	112	262
Average URLs/Article	9	8.79	9.33	9.03

Table 6

*Percentage Change in URLs by Year and Publication*

Publication	Percentage Increase or Decrease from 2001 to 2002	Percentage Increase or Decrease from 2002 to 2003
Washington Law Review		
Articles	decrease 16.67	increase 40
URLs	increase 104.76	increase 90.7
Seattle Univ. Law Review		
Articles	0	increase 60
URLs	decrease 8.11	increase 95.59
Gonzaga Law Review		
Articles	increase 366.67	decrease 14.29
URLs	increase 355.56	decrease 8.94

*URL Characteristics: Last Visited Notations*

¶31 Accepted *Bluebook* practice now mandates how authors address URL verification. Those rules have become progressively more specific with each edition.<sup>32</sup> Rule 18.2 expresses a preference for citation to paper formats whenever possible; including a parallel citation to a Web version is optional.<sup>33</sup> If a URL is being provided as a parallel citation, rule 18.2.2 requires the publication date of the print source but not the date that the URL was visited.<sup>34</sup> Rule 18.2.3(e) requires that a direct citation to an Internet source without a print equivalent should include a date.<sup>35</sup> If the material includes a date, that date should be used. If the material is undated, 18.2.3(e) requires a “last visited” date in parentheses.<sup>36</sup> With the eighteenth edition, the use of “last visited” is expressly preferred over “last modified” or “last updated.” However, since most of the articles in the study sample were published while the seventeenth edition of the *Bluebook* was in effect, for purposes of this study, the label “last visited” was taken to apply to any of these notations.

¶32 Not surprisingly, use of these notations has increased over the span of the study—irrespective of domain. This is consistent with the evolution of *Bluebook* requirements. Table 7 presents the breakdown for last visited notations by domain and year.

¶33 These very high rates indicate a conscious effort on the part of authors to conform to current citation practice. They can also be interpreted as a willingness to vouch for the authority of the Web resources by indicating the moment in time

32. Only with the seventeenth edition did the *Bluebook* explicitly address citation of Internet sources. The sixteenth edition, published in 1996, includes a much less explicit rule. THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION rule 17.3.3, at 124 (Columbia Law Review Ass’n et al. eds., 16th ed. 1996). That rule did, however, require a “last visited” type notation. The fifteenth edition of the *Bluebook*, published in 1991, was silent on this subject.

33. THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION, *supra* note 6, rule 18.2, at 153.

34. *Id.*, rule 18.2.2, at 155.

35. *Id.*, rule 18.2.3(e), at 157–58.

36. *Id.* at 158.

**Table 7***Total Number of URLs with “Last Visited” Notations, by Year and Domain*

<b>Domain</b>	<b>2001</b>	<b>2002</b>	<b>2003</b>	<b>Total</b>	<b>% of Domain</b>
com	39	63	124	226	61
org	21	88	112	221	73
edu	15	62	78	155	62
net	13	13	40	66	71
gov	0	7	19	26	74
state	9	4	8	21	82
ctry	2	7	7	16	84
int	0	0	7	7	88
info	0	0	1	1	100

when the author can say the hyperlinks were valid (if not permanent). It is distressing, then, to know so many URLs are so notated when compared to the invalidity percentages. A substantial majority of the 40% of links that were invalid when tested are .com, .org, and .gov domains, yet these are the three domains with the high rates of “last visited” parenthetical notations. This fact may be the strongest argument yet for the inherent impermanence of Web pages.

***URL Characteristics:  
Other Document Formats***

¶34 Another notable characteristic of URLs is the extent to which they actually link to material that has a life outside of the Web, that is, to documents created for uses other than the Web that have been made available via that medium. Several authors cited to full-text, non-Web documents ending in .pdf, .doc, .txt, or .rtf. Of the 1104 URLs cited in the study, 116, or 11%, cite to these documents. Portable document format (PDF) was cited to most frequently. This is consistent with *Bluebook* Rule 18.2.1(c), which states a preference for citing to PDFs over HTML documents.<sup>37</sup> Table 8 shows the data for URLs referring to non-HTML documents by year and domain.

## Discussion

### *Why the Web Matters*

¶35 Any new source of potential authority is a source that researchers will consider relevant for their work. This is as true for legal researchers as it is for other researchers, and true despite the unique nature of legal material—primary,

---

37. *Id.*, rule 18.2.1(c), at 154.

Table 8

## Number of Non-HTML URLs by Year and Domain

Domain	2001	2002	2003	Total
com	1	7	5	13
org	2	12	21	35
edu	4	12	32	46
net	2	1	8	11
gov	0	0	4	4
state	0	0	1	1
ctry	0	0	4	4
int	0	0	0	0
info	0	0	1	0

secondary, and otherwise.<sup>38</sup> It is natural to see the Web become an increasingly cited resource in legal scholarship.<sup>39</sup> This expansion is also a logical by-product of legal practitioners and members of the academy widening their own circle of interests. An additional factor is the rise of interdisciplinary research and scholarship performed by legal academics. Although legal sources are some of the best organized and most easily accessible, increasingly a lot of material of interest to legal scholars falls outside of the conventional panoply of sources.

¶36 The greatest challenge, however, is that while most legal information is carefully vetted and is produced by historically valued sources in formats well documented and documentable, the Web is a much more egalitarian medium. It does not contain only legislative, judicial, and academic legal information; it includes all this and much more. Moreover, anyone can create information avail-

38. It is widely accepted that West Publishing (now Thomson West) and Lawyers' Cooperative (whose former publications are now published by both Thomson and Reed Elsevier) have greatly influenced how we approach, access, and think about legal information through their creation of the Key Number System, the Total Client Service Library, legal encyclopedias, and annotated codes, among other publications or systems. In a broader discussion about the shifting nature of legal information in light of the rising influence of cyberspace, one author noted:

The whirlwind consolidation of the legal publishing industry into two large conglomerates, each centered on a full-text database, has shifted the tectonic plates of legal information. Part of the culture shock that will affect the legal profession will come from the changing role of the West Publishing Company. . . . It not only dominated American legal publishing, it established a special culture in doing so.

Robert Berring, *Chaos, Cyberspace and Tradition: Legal Information Transmogrified*, 12 BERKELEY TECH. L.J. 189, 191 (1997).

39. Some scholars have advocated that legal scholarship should be published online exclusively. See, e.g., Bernard Hibbits, *Last Writes? Re-accessing the Law Review in the Age of Cyberspace*, 71 N.Y.U. L. REV. 615 (1996), available at <http://www.law.pitt.edu/hibbits/lastrev.htm>. According to the printed version of this article, two iterations of it appeared on the Web before it was published in a print law review. *Id.* at 615 n.\*.

able via the Web. That information then cohabitates in cyberspace next to official or sanctioned legislative, judicial, and academic legal information. All of this can be accessed by anyone.

¶37 When it comes to legal scholarship, most of the burden for ensuring the legitimacy of cited sources falls on the author and those engaged in the editorial process. To date, the evolution of Web (or digital or electronic) resources has not changed this basic dynamic. However, legal scholarship is also subject to the expectations of the audience it engages, and this has traditionally influenced authors. The Web introduces an unanticipated impermanence to the author-reader relationship. After all, if a source is there when you retrieve it, and then again the next two, three, or four times you return to it for verification purposes, it is easy to assume that it will always be there. At a minimum, it can be difficult to imagine that if it is there the fifteenth time, it will not be there the thirtieth time.

¶38 Yet, the very nature of the Web makes this scenario much more plausible than the disappearance of a print resource used in the same law review article that cites to a Web site—any Web site. A case in *Federal Supplement* is not likely to disappear, nor is an article in a printed law journal. If you used a volume of the *Washington Law Review* in a particular library, even if that library no longer held that title (or held that title but not that volume), a user knows that the volume still exists and that *some* library *somewhere* still holds it. Once printed, most legal information never “disappears,” but a word-processed document on a federal government Web site can be gone in the few clicks it takes a webmaster to sign-on to a server and empty a directory.

¶39 For the authors who have determined that a Web-based source is valuable, credible, and timely, this can be a case of what they don’t know hurting them at some unknown point in the future. Or perhaps not. Researcher expectation is one of the most interesting aspects of how the Web is used to retrieve substantive information. Most researchers understand how easy it is to add and delete information from the Web, but they do not seem as concerned about the impermanence—or the potential for impermanence—of the information they find there. While it is hard to believe that those who publish in law reviews would not want their Web resources to be as permanent as their non-Web resources, it is less difficult to contemplate that expectations about print information have simply been transferred to this new medium and authors don’t yet really understand the differences between print and Web sources. Knowing in the abstract that anything can disappear is quite different from understanding that the one thing you are relying on might disappear.

### *Content Issues*

¶40 As noted earlier, the fullest exploration to date of the effects of citation impermanence on legal scholarship as a whole is a study by Mary Rumsey.<sup>40</sup> Differences

---

40. Rumsey, *supra* note 2.

between that study and the instant one notwithstanding, several of the overall trends Rumsey notes were also identified here. In both instances, the number of Web citations steadily increased from year to year. At the same time, the number of invalid links also increased. Rumsey notes that “even newly published articles display significant link rot.”<sup>41</sup> That is also true in the instant case. Ironically, not only are legal scholars “more compulsive”<sup>42</sup> about citations, but law review editorial practices are rigorous. This only adds to the significance of a 40% invalidity rate. In this study, the year with the least number of invalid URLs—2003—still produced a 37% invalidity rate.

¶41 Web content is only growing,<sup>43</sup> as are citations to that content in law reviews. Given the fact that most of the Web citations in the instant study are to nonlegal material, the type that is bound to become more readily available as the Web grows, there is no reason to think this issue will not remain a constant in legal scholarship. Yet, if most of what is cited in law reviews is still material that appears in non-Web formats (as based on the relatively small number of URL cites per article, i.e., slightly more than nine), there remains a question of how much authors and users will care about this growing impermanence.

¶42 One way in which this issue could resolve itself involves computer-assisted legal research systems and citations that not only identify a document but indicate a format. As nonlegal items proliferate across CALR systems, authors may opt to cite to more stable versions of nonlegal material on those systems versus citing to Web (e.g., HTML, PDF) versions. For example, if a *Washington Post* article disappears into the newspaper’s archive after fourteen days (and subsequently behind either a registration or purchase requirement) but is available in full text on Westlaw, an author might indicate this location as part of the citation. Since the primary audience for law reviews is other legal professionals, most readers would have access to this article in this format. Not only that, but an author could reasonably assume that *more* potential readers have access to it in that format than have access to a URL that would direct them to the newspaper’s Web site. It is now considered gospel that the legal research landscape has been forever changed by how legal researchers use electronic resources in their research. It is tempting to ponder how that same group might use their familiarity with proprietary databases to create more permanent citations to Web sources.

¶43 Traditionally, an author’s only responsibility is to cite to a document, leaving the reader to determine which source for the document is applicable and appropriate, depending on the reader’s resources. This is especially true for documents originally created in a nondigital format. Citations to cases published in specific reporters also serve a dual purpose of identifying that the resource is in

---

41. *Id.* at 35, ¶ 23.

42. *Id.* at 28, ¶ 4.

43. See Peter Lyman & Hal R. Varian, Sch. of Info. Mgmt. & Research, Univ. of Cal., Berkeley, How Much Information? 2003, <http://www.sims.berkeley.edu/research/projects/how-much-info-2003/index.htm> (last visited Apr. 19, 2006).

print, however. (The fact that these print citations are also used in online venues does not change this fact.) In the future, as concerns about impermanence grow, authors and readers might welcome exact format information, including citations that not only identify the document but indicate a permanent digital location. One example of this trend already exists, with the invention of persistent URLs (PURLs). Arguably, this invention is more necessary for documents that are “born digital” but there is nothing to keep reader expectations from transferring this expectation to all kinds of resources whether they are born digital or not.

### *Deep Linking*

¶44 Deep linking is the practice of linking to a page within a Web site and bypassing the site’s home page. Any URL with directional information after the top-level domain (excluding geographical suffixes) is deep linking into a Web site. Such linking can be controversial for several reasons. Commercial sites frown on deep linking because it allows the user to bypass whatever information the site owner wants them to see first (in many cases advertising or marketing content). It also thwarts any effort to “push” users through a Web site in a particular manner. Others have objected to deep linking on the ground that it violates their interest in the material being linked to, since they cannot control how users reach it. As technology has become more sophisticated, password protection for certain sections of Web sites has become more prevalent, even for sites that are not actually charging for their content. Not only does this allow them to form a profile of who is visiting the Web site, it exerts some control over random deep linking.

¶45 Given these objections, it would be surprising to note that nearly all of the 1104 URLs cited in this study are deep linking into Web sites were it not for two things. First, deep linking is just easier. Of the total in the sample, only 72,<sup>44</sup> or 6.5%, do *not* deep link. Nothing in the publications’ editorial policies addresses this practice, but certain conclusions can be drawn. One advantage to deep linking is that it loosely approximates the same function as a pinpoint citation. Not only do you know which Web site to go to, but also the specific page within that site. Given the complexity of some Web sites, this is not a trivial matter. Second, the *Bluebook* endorses the notion of providing as specific and direct a citation to a source as possible.<sup>45</sup> Few law review authors would cite to an entire treatise in lieu of citing to a particular page or section. Deep linking on a Web site is comparable to this practice.

---

44. This calculation does not count URLs mentioned more than once in a particular year; however, URLs were counted more than once if mentioned in more than one year for a given publication. It also includes URLs that have subdirectories or subdomains that appear *before* the top-level domain (e.g., <http://nasdaq.faa.gov>).

45. *Bluebook* rules recognize that HTML documents cannot yield page-level citations. The rule requires a pinpoint cite if the virtual document is a PDF. Otherwise, page citations are not expected and “screen page” citations are discouraged. *The Bluebook: A Uniform System of Citation* rules 18.2.1(a), 18.2.3(a), *supra* note 6, at 154, 156.

*Usage Issues: PDFs*

¶46 This example of how certain conventions do or do not become accepted practice outside of their original environment is also evident in the usage of cites to PDF documents available via the Web. Because PDFs have become ubiquitous, and because they exist outside of the Web and authors are used to encountering them in that way, it hardly seems a leap of logic to accept their use when they appear on the Web. PDF documents typically work the same way on the Web as they do outside of that environment. Yet, the use of PDFs raises another issue.

¶47 Does the fact that anyone with the requisite software<sup>46</sup> can create a PDF have any affect on the credibility of the documents that appear on the Web in this format? Arguably, PDF content is no less credible or reliable than anything else that exists on the Web. Also, it is possible that authors making determinations about the worthiness of Web sites automatically extend their analysis to PDFs when determining what to cite. An author might decide that a PDF on a .com site containing legitimate information, for example, is consistent with other information on that site. If site evaluation reveals the site is trustworthy, then that decision might extend to all information found there (and perhaps linked to from there).

¶48 Notwithstanding that the *Bluebook* prefers citations to PDFs over HTML, what is interesting about non-HTML documents on the Web and citations to them in law reviews is that there is no way to *know* what evaluation authors undertake regarding these documents or the criteria they use. With traditional legal resources, readers are already familiar with what the author is citing, including who creates those resources. So while a reader is depending upon the author's characterization of what he or she is citing, readers are also capable of making independent evaluations. This is true both because they have the skills to make that determination and the materials are easy to locate.

¶49 With Web sources, however, the same can be true but does not necessarily have to be true. If a Web citation is invalid by the time a reader follows up on it, the opportunity for independent evaluation is lost. Also, to the extent that the Web citation refers to nonlegal material with which readers may not be familiar, they don't approach it with the same level of expertise that they might have for traditional sources. This presumes several things, of course, including that most readers of law reviews are in fact legal professionals. But the main premise is that Web sources invite a certain amount of uncertainty. Law review authors citing to PDFs expose readers to a much broader base of content creators. The creator of a PDF can literally be anyone.

---

46. Adobe Acrobat is required to create PDFs. This is a separate application from Adobe Reader, which is used to read those documents. Adobe has long made Reader available for free, and any site featuring PDF documents can link to a free Reader download. Adobe Acrobat, on the other hand, is a commercially sold product.

### *Usage Issues: Web Literacy*

¶50 Web literacy, or more specifically, the notion of evaluating Web sources before relying on them, is not only an issue when choosing to cite PDFs, but for other resources as well. Law review authors are presumably evaluating Web sources before citing them. But as previously noted, there is no indication of the criteria they use in this process. There is also no indication that appears with the citations themselves, parenthetically or otherwise. Moreover, the standards that these authors apply to legal resources would be very hard to apply to Web sources.

¶51 Presumably, a reader must trust the author (and the editorial process) that a cited Web source is appropriate, credible, and valuable. While I don't mean to suggest that authors can't be trusted to vet these sources any less than they are trusted to vet other sources, it remains troublesome that a reader can't know for sure one way or the other. Moreover, independent verification might be difficult or impossible. The fact that a significant percentage of the URLs cited are likely to be invalid within as few as twelve months does not lend much assurance to this process.

## **Issues for the Future**

### *Paper Trails*

¶52 As part of this study, I contacted the editorial staffs of the three publications and inquired about their policies regarding URLs. Specifically, I asked if authors were required to submit paper copies of Web sources referenced in submissions. One publication requires authors to submit paper copies on an "as needed" basis, but typically not when the Web source cited "supports the author's assertions."<sup>47</sup> Another does not require authors to submit paper copies, but does suggest authors maintain them. The rationale for this is an acknowledgement that material at URLs can cease to be available. The law review does not keep source files, but if the author opts to, the law review will include a notation to this effect. In that case, the law review expects the author to produce copies when requested.<sup>48</sup> The third publication in the sample also does not require authors to submit paper copies. However, this publication will produce Web copies in response to reader requests.<sup>49</sup>

¶53 These policies highlight one of the main distinctions between print and Web resources as well as one significant similarity. Readers coming to a law review article after it's been published (either in a print or an online source) who need to find a print source referenced within the article will presumably know

---

47. E-mail from Pete Meyers, Editor-in-Chief, *Seattle University Law Review*, to author (Nov. 5, 2004 10:40:13 PST) (on file with author).

48. E-mail from Jungmin J. Yoo, Managing Editor, *Washington Law Review*, to author (Nov. 5, 2004 16:16:12 PST) (on file with author).

49. E-mail from Jason Wilcock, Executive Editor, *Gonzaga Law Review*, to author (Nov. 10, 2004 12:30:17 PST) (on file with author).

that finding it is as easy as finding a law library. (They may also use a specialized database, if they have access and it is available via that route.) As noted earlier, there is never any supposition by readers that the book *cannot* be found, even if there is a belief that any given library might not hold it. More likely than not the source is *findable*.

¶54 If readers need to find an online source cited in that same law review article, however, the process is both the same and different. If readers see a URL, know what it is, and have Web access, they can go directly to that source. But here is where Web sources are unique: what if that source is no longer located at the URL provided in the printed article? Unlike the print reference, there is every possibility that this resource *cannot* be found. It may have been pulled off the Web entirely or it may have been re-situated behind a password prompt. One last possibility, and perhaps the most worrisome, is that the document has significantly changed from the version cited and there is no way for the reader to know this. Some pages will indicate when they are modified, but very few indicate *how* they are modified. Once that URL deadends, no longer serving its directional function, there's no obvious path for finding the article. A reader's success at this point is probably dependent upon ingenuity, searching skills, or access to a librarian.

¶55 Of course, even if the author has kept paper files, accessing those files is not very practical. Would readers contact authors for this information, would authors have it, and would they gladly provide it? Many documents on the Web are quite long (which is why presenting them in a digital format is so attractive in the first place). How many authors would want to make copies of source material available to readers? Instead, the more likely scenario is that once readers determine that a URL does not function, they will simply ignore that reference and move on. For them, that reference essentially is out of reach. This phenomenon, if repeated throughout an article (how many bad URLs will make a reader question the overall value of a law review article? Two? Three?), could erode the appeal of that particular piece of scholarship.

### *The Value of Contents*

¶56 The legal scholarship that appears in student-edited publications has a long and well-earned reputation for being well referenced. It is common to have any given page of a law review article have a two to one or three to two proportion of article text to footnote text. It is also customary for law review articles to have hundreds of footnotes and average one source reference per footnote. Traditionally, authors decide what to cite depending on the vagaries of their subject matter, but for the most part these references have usually been comprised of some combination of primary sources, secondary sources, and specialized materials. The nature of interdisciplinary research has shifted these parameters, but for the most part law review articles still cite heavily to well-recognized legal materials or supplemental factual materials that either are related to the cited materials or support their arguments.

¶57 In the present study, the high number of citations to .com and .org domain URLs indicates that most of the Web resources cited in these publications during the designated period were to neither primary nor secondary legal sources. These sources, when not being retrieved from proprietary research systems, are usually found at Web sites owned and operated by governments, courts, or administrative agencies—all entities usually found at .gov or .state domains. Substantive secondary sources (i.e., encyclopedias and treatises as opposed to newspapers or nonjournal periodicals) are almost never available in full text on the nonproprietary parts of the Web.

¶58 The prevalence of certain domains that are less likely to be home to primary materials (and therefore materials traditionally relied upon heavily in law review articles) raises the specter that law review content generally will be devalued. As noted earlier, the possibility that anyone can create a document and upload it to the Web raises several issues. In this context, the critical question concerns the level of credibility that should be afforded such a document in comparison to more traditional sources. The editorial staffs for the publications in this study all affirmed that URLs are checked before publication for both operability and credibility.

¶59 There is every possibility that readers will not care, of course, especially for any given article. But over time, especially if citations to URLs continue to increase (and there is no reason to think they will not), what impact will this have on the concept of the law review article as a thoroughly (perhaps tediously) researched source of analysis, complete with copious references to other well-researched sources? The issue is not a matter of whether Web content is inherently inferior; it is a function of how the nature of legal scholarship might be changing. At a time when there is still great interest in the most cited law reviews,<sup>50</sup> ranking legal periodicals,<sup>51</sup> and the most cited legal scholars,<sup>52</sup> clearly there is interest in the value that the legal academy puts in the system it knows and respects. Yet, at the same time, citations to URLs only grow.

### *The Sway of the Heavyweights*

¶60 The present study looks at a very small slice of the larger law review pie in terms of publications and timeframes. Although including citations to URLs in law reviews is apparently a recent trend,<sup>53</sup> tracking which law reviews are cited, how often, and who is published in them has long been a spectator sport in legal academia. Aside from the published studies, there are well-understood presumptions

---

50. See, e.g., Fred R. Shapiro, *The Most-Cited Law Reviews*, 29 J. LEGAL STUD. 389 (2000).

51. See, e.g., John Doyle, *Ranking Legal Periodicals and Some Other Numeric Uses of the Westlaw and Lexis Periodical Databases*, LEGAL REFERENCE SERVICES Q., 2004, no. 2/3, at 1.

52. See, e.g., Fred R. Shapiro, *The Most-Cited Law Scholars*, 29 J. LEGAL STUD. 409 (2000).

53. Within the three publications in this study, as recently as 1997 there were no citations to URLs in at least one of the publications.

about the merits of publication in one of the main law reviews of the nation's top twenty or twenty-five law schools.

¶61 In that light, it would be instructive to look at how URLs (and which ones, how often, and their corresponding validity rates) are cited in these publications. Likewise, it would be instructive to compare URL citations across publications from the same institution and to examine various editorial policies. Several top law schools have three or more law reviews or journals.<sup>54</sup> In addition to the main publication, these other journals usually explore subject matter specialties, from civil rights to gender issues to the environment. In combination with these issues, or independently, a study that analyzed exactly *what* is being cited to and the value of that content compared to known primary and secondary sources would be valuable as well.

¶62 Given the pervasiveness of the Web, it is not likely that these publications currently ban URL citations or ever would. A question ripe for study, however, is what policies they have put in place to address URL validity and permanence issues, including the extent to which they request, require, or store paper or electronic copies of Web materials. A related question would focus on the in-house editorial URL verification practices of these journals.

### Conclusion

¶63 At the outset of this study, I hypothesized that the law review articles in the sample would cite to a significant number of URLs. While this presumption was borne out, with an average of slightly more than nine URLs per article, that number still comprised a very small percentage of the total sources cited in most instances.<sup>55</sup> Paired with this presumption was the idea that the number of articles citing to URLs, as well as the number of URLs they cited, would increase every year. This was, in fact, true. In the sample, there were significant increases each year, although the greatest jump was between 2002 and 2003.

¶64 The data did not completely support my initial hypothesis about domains that would be the most frequently cited. While citations to .org and .gov domains were plentiful, in each year there were more .com domains cited than any other group. This may argue for the idea that authors citing to Web sources are not citing to "high demand" materials, at least as defined by a legal audience. Primary and secondary legal materials are almost never found at these types of sites. I did not expect citations to primary legal sources, and the data confirmed this. It also

---

54. For example, Harvard Law School has thirteen student-edited publications, the University of California has eleven such publications, Northwestern University School of Law has five journals, Columbia Law School has ten, and Duke Law School has seven.

55. This is an observational finding rather than a statistical one, based on my review of articles while extracting URLs.

confirmed that most law review authors were not citing to foreign or international sites. These cites made up only 3.9% of the total sample.

¶65 In the area of URL invalidity, my hope was that most of the URLs would be stable and working, especially since the oldest of the documents in the sample was only three years old. However, the data showed a 40% invalidity rate, a finding consistent with the significant “link rot” identified in the leading study in this area. Given the predominance of .com, .org., and .gov URLs in the sample, it was not surprising that these groups also represented the highest invalidity rates. I expected commercial sites to have fewer broken links, but the study indicated parity between .com, .gov, and .net sites. The few international domains also returned an unusually high percentage of broken links.

¶66 I also wondered at the outset if any trends in the data would track to the relative rank of the three schools included in the study. The publication published at the highest ranked school, the University of Washington, actually includes the fewest number of URLs cited across all three years. That publication also has the lowest invalidity rate, but just barely, and the publication it squeaks past is affiliated with the lowest ranked school of those represented in the study, Gonzaga. That anomaly argues against stretching the analysis too far, and I have concluded that this is an empty measurement.

¶67 What remains to be seen is the extent to which these cites actually point readers to materials that are critical to their understanding of the article. There is no doubt that this can vary by article, but as long as most URLs don’t cite to primary and secondary legal materials, and as long as common practice prefers printed materials, there will always be a caveat about invalid URLs, no matter the percentage. At one level, law reviews cannot avoid being harmed by such a high rate of invalid sources, but at another level it is not clear how much their readership has noticed.

¶68 In time, it is likely that both authors and readers will start making judgment calls on whether a cited Web source is, in fact, the “best” source, regardless of format or medium. It is easy to suppose that the answer to this question is “sometimes yes, sometimes no.” Perhaps the most significant thing URLs have in common with traditional sources is that during the writing process, this is the author’s call. Once a law review article has been published, it is up to the reader to decide how deeply to delve into what the author has cited.