

**GOOGLE, LEGAL CITATIONS, AND ELECTRONIC FICKLENESS. LEGAL
SCHOLARSHIP IN THE DIGITAL ENVIRONMENT**

By Dana Neacșu

INTRODUCTION:

Law review articles which are published in law journals -- student-run publications -- represent the bulk of legal scholarship. A major part, and the scholarly backbone, of any law review article are the footnotes whose form follows the legal rules of citation established by the so-called Bluebook. While in the past it only allowed citation to print sources, for the last 10 years the Bluebook has encouraged citation to electronically available sources. Relaxing the rules of citation in that way can potentially threaten the development of legal scholarship in the classical sense--that is, scholarship which builds upon prior scholarship.

While law review articles are preserved in fee-based databases such as Westlaw and Lexis and thus are reliably accessible for the future, the footnotes usually refer to documents which far too often become inaccessible within a few months after their publication. Both government documents and documents privately published on the Internet have an unreliable life-span. This contradictory approach to digitization raises a large array of questions. Among them, is the following: How does this double digitization (that is, digitizing articles which refer to already-digitized, but unreliably retrieved, prior sources) affect the retrieval of legal information? Whose job is it to preserve legal information? Of course, I would like to say that it is the librarian's job. However, this is a more complex answer which I will not attempt to argue today, because if the publication of law review articles in commercial databases may only require reference librarians to

possess a heightened level of technological know-how, the inclusion of unreliable sources in their footnotes would require more than technological skills: it would require a substantive change in the institutional role of libraries in supporting, encouraging and archiving legal scholarship. Librarians would have to use both their legal and informational taxonomy knowledge to be able to retrieve the correct information their patrons require, and to understand (and explain) that a certain portion may become irretrievable. Here I will only attempt to show that digitization has created a different environment of legal information (which includes legal scholarship) and this new environment proves to be more elusive than we would like to think about it.

BRIEFLY ABOUT LEGAL SCHOLARSHIP AND DIGITIZATION:

As Professor Ira Katznelson reminds us scholarship, both in its creation and in its use, is a private and a collective act as well¹. At least in the tradition of Western civilization,² new knowledge cannot be produced without reference to past knowledge. Western scholarship relies on “a base of expertise, a ‘scholarly knowing’ that needs to and can be identified, made public and evaluated.”³ It requires accurate access – which

¹ *Marxism and the City* at ix (1992) (“scholarly work is balanced between private and collective acts”). For example, Descartes’ *Method* explains how knowledge is built on past knowledge that is clearly and distinctly perceived as true. See e.g., Descartes, René. *Discourse on Method and Related Writings*. London: Penguin Classics: 1999. Walsh, W. H.. “Knowledge in Its Social Setting.” *Mind, New Series*, 321 (1971): 321-336

² Of course, unlike Gandhi, this author believes that there is Western civilization. The saying goes that when asked what about Western civilization, Mahatma Gandhi replied that it was a good idea.

³ Paulsen, Michael and Feldman, Kenneth A.. “Toward a Reconceptualization of Scholarship: a Human Action System.” *Journal of Higher Education*. 16 (1995): 15-40.

emphasizes both access and accuracy – to sources. It requires reliability,⁴ which ensures that all users share either the same scholarly article,⁵ or its faithful reproductions,⁶ and that the sources used in the article are properly acknowledged.⁷

For example, Socrates himself was able to create, what the German philosopher Gadamer calls “knew knowledge,” because he had access to previously stored knowledge.⁸ And of course, any subsequent reference to one’s work assumes that what is accessed faithfully reproduces the original scholarly product.⁹ Like philosophy, any type of scholarship, including legal scholarship, requires reliable access to the product of earlier study.

Legal scholarship is mostly comprised of student-edited law review articles. There are more than 900 journals nationwide.¹⁰ While most of them are published both in

⁴ See. Neacșu, E.Dana. "Legal Scholarship and Digitization: Has Anything Changed in the Way We Do Legal Research?," *Legal Reference Services Quarterly*. 106 (2002): 105-122 (using the term “reliability” to include accurate access to sources that reproduce faithfully the original source).

⁵ The tragedy of the library of Alexandria was a trauma to the very nature of scholarship because it impeded scholars from having access to the stored product of earlier ages. See, Neacșu 2002, 105-122.

⁶ *Id.*, at 106-109.

⁷ *Id.*

⁸ Gadamer talks about knowledge stored within the moral tradition. And he further explains, “[s]tatements like ‘Virtue is knowledge’ do not come out of the blue, rather, their sense is determined in part by what they answer or respond to.” (emphasis added). Gadamer, Hans Georg. *The Idea of the Good in Platonic-Aristotelian Philosophy*. 61, translated by Christopher Smith. New Haven: Yale University Press (1986).

⁹ Socrates’ is a more complex issue as he did not commit his teaching to writing and what is generally known about his work comes through Plato’s *Dialogues*. So, in this case the access to his work is, in this case, access to Plato’s DIALOGUES. *The dialogues of Plato*, (trans. by B. Jowett) (1953). For a discussion of the problems oral cultures raised, see e.g., McLuhan, 1962, at 97 et seq.

¹⁰ The "Most-Cited Legal Periodicals: U.S. and selected non-U.S." site at <<http://law.wlu.edu/library/research/lawrevs/mostcited.asp>>.

print and electronically,¹¹ many are only available electronically.¹² As a result each law school and its law library have quite a few such journals edited by their student body. What is unique about this scholarship is that it heavily relies on pre-existing sources that are cited in numerous footnotes. Those footnotes cite their sources according to very well established rules which can be found in *A Uniform System of Citation (The Bluebook)*¹³. If before the advent of mass digitization relying and citing electronic sources was discouraged, starting with its 17th edition (2000), in Rule 18.2.1, *The Bluebook* allows citations to such sources.¹⁴ The new 2005 edition continues the trend. Rule 18.2.3(a) states:

An Internet source may be cited directly when it does not exist in a traditional printed format or on a widely available commercial database, or when a traditional printed sources, such as a letter or unpublished dissertation, exists but cannot be found and is so obscure that it is practically unavailable. In either case, the title, pagination, and publication date should be reported as they appear on the Internet. The Internet URL should then be appended directly to the end of the citation (i.e. not preceded by “*available at*” or “*at*”).¹⁵ (emphasis in the original)

¹¹ See e.g. *Duke Law Journal* may be also accessed at <<http://www.law.duke.edu/journals/dlj/>>. *Columbia Law Review* may also be accessed at <<http://www.columbialawreview.org/issues/>>.

¹² See e.g., *Columbia Science and Technology Law Review*, “a World Wide Web-based, solely online journal,” may be accessed at <www.stlr.org>. For more information about this publication see, <http://www.law.columbia.edu/current_student/student_service/Law_Journals/science_tech?#rtregion:mai>

¹³ *The Bluebook. A Uniform System of Citation*. Cambridge: The Harvard Law Review Association, 2000.

¹⁴ *Id* at 132 et seq.

¹⁵ *The Bluebook. A Uniform System of Citation*. at 156 (18th et).

Legal scholarship turning digital is only one aspect of the trend undertaken by the entire legal information world.¹⁶ And this trend cannot be satisfied only through fee-based commercial databases. A lot is left to the freely accessed and freely un For example, in *Scott v. Harris*, No. 05-1631 (April 30), the U.S. Supreme Court justices have joined the Internet age, and included digital access to videotaped evidence. The clip can be reached via a hyperlink on the Court's opinions Web page. The clip shows the view from the dashboard of a police car involved in a high-speed chase in suburban Atlanta.¹⁷ What the Justices do not seem to understand is that the clip, as well as the URL that identifies it, can disappear and thus, make it impossible for the future readers to understand the facts the Court used (by usurping the jury's role) in reaching that decision.¹⁸ Of course, unless the Court's library decides to create a mirror site.

BRIEFLY ABOUT GOOGLE:¹⁹

Search engines like Google made electronic searching infinitely easier. But Google and its competitors are not in the business of preserving scholarship. If their advertising money-making business requires them to amalgamate an amazing "library"

¹⁶ Martha Neil, "Supreme Court Meets YouTube. Camera-shy Justices Hyperlink a Video to Their Car-crash Opinion," <http://www.abanet.org/journal/ereport/my4video.html>.

¹⁷ Id.

¹⁸ A lone dissent by Justice John Paul Stevens said the case should have gone to a jury for decision. "The court has usurped the jury's fact-finding function," he wrote, since "whether a person's actions have risen to a level warranting deadly force is a question of fact best reserved for a jury."

Stevens challenged the majority's interpretation of what the video shows, and he referred to his fellow justices as "my colleagues on the jury."

¹⁹ More about Google, see e.g., Eben Moglen's speech on March 27, 2007, available at <http://video.google.com/videoplay?docid=-7477852615698435519>.

then they will do that as long as necessary. Google has assumed the role of running the information soup kitchen of optimal delivery deciding what people incorporate in their work and making heaps of money in the process,²⁰ with the air of running a charity. To let a very successful corporation which has little qualm about participating in censorship in China or limiting competition to the copying business, would seem poor judgment.

PROBLEMS RAISED BY DIGITIZATION AND CITATION TO ENCOURAGE ELECTRONIC DOCUMENTS:

Ordinarily internet sources change or disappear.²¹ As authors have recently noticed, “Uniform Resource Locators (URLs) experience ‘link rot,’ that is, over time the URL is more and more likely to become a dead link, making the footnote citation worthless or nearly so.”²²

In law review articles resources used to build legal arguments are identified in footnotes, according to rules of citation that are articulated in *The Bluebook*. Starting with its 17th edition, *The Bluebook* has relaxed its citation rules and allows the use of electronically identified sources, based on their URL.²³ As a result, due to their

²⁰ See endnote i above.

²¹ Rumsey, Mary. “Runaway Train: Problems of Permanence, Accessibility, and Stability in the Use of Web Sources in Law Review Citations.” *Law Library Journal* 29 (2002): 27-39. (After four years, only 30% of all URLs cited in law reviews still work.).

²² Lyons, Susan. “Persistent Identification of Electronic Documents and the Future of Footnotes” in *Law Library Journal*: 97 (Fall 2005): 681-694.

²³ While it distinguishes between the rules of citations covering commercial electronic databases (such as LEXIS and Westlaw) – Rule 18.1 – and those covering the rest of the online sources – Rule 18.2.1 – *The Bluebook* does not discourage online citations. Rule 18.2.1 at 132-33.

This rule requires the use and citation of traditional printed sources, except when the information is not available in a printed source, or if the traditional source is obscure or hard to find and when

convenience, more and more law review articles now cite information or texts from websites to support their arguments.²⁴ The citation used in the footnote is to the URL of the site, often with an additional note as to when the site was visited by the author or editor to confirm the citation is accurate (or, to be candid, which it was accurate at the time it was cited).

The percentage of law review articles that cite to at least one such source has increased from “half a percent in 1995 to 23% in 2000,”²⁵ and it is only logical that the percentage would increase even further likely because of the success of search engines such as Google and their daily insinuation in a researcher’s work. As Google makes it easier to produce scholarship – accessibility to sources has increased to unprecedented levels, the issue of the reliability of such sources becomes every day more meaningful.

As publishing on the Web, which search engines encourage, places information in a form where many facets of print works such as fixity, uniformity, and authenticity²⁶ are no longer automatically ensured, the concept of reliability has inevitably changed. However, because the fundamentals of Western scholarship are still the same, the need of accessibility of content-identical copies still remains.

There are many reasons why websites disappear or change. They may no longer be supported by the author or organization which created them. The text may be a fluid

the citation to an Internet source will substantially improve access to the same information contained in the traditional source. [...]

²⁴ Canick, Simon. "Availability of Works Cited in Recent Law Review Articles on Lexis, Westlaw, the Internet, and Other Databases." *Legal Reference Services Quarterly* 56 (2002): 55-65.

²⁵ Rumsey, 2002, 27-39.

²⁶ See generally, Katsh, M. Ethan. "Law Reviews and the Migration to Cyberspace." *Akron Law Review*. (1996): 115-121, and Neacșu, 2002, 105-122.

one as far as the parent organization is concerned, so the information as used by the article's author in writing may not be what a reader may find a year later. Some non-governmental organizations change their stance on different political issues quite often and do not archive the documents that would memorialize those changes. Electronic text may also be subject to malicious change, either from the outside or from internal mischief.

Irrespective of these reasons, when a law journal article references a text that does not exist or has been altered, and thus supplies "unreliable" footnotes, its scholarly value becomes questionable, and a solution needs to be advocated.²⁷ Such references are not minimal.

A June 14, 2005 search on Lexis/Nexis Academic Universe for law review articles published between March 13th and June 13th 2005, and citing at least three Web-based documents yielded 938 results. Twenty one of the first twenty five hits, a whopping 80%, cited at least one online source which had become publicly inaccessible.

Those online sources varied from federal and local governmental sites to self-published ones, from html to PDF documents, but invariably some of them proved evanescent.

A more recent empirical experiment produced equally telling results. A May 17, 2007 search on Westlaw-JLR for articles published from December 31st, 2006 citing at least three documents published on the Web produced 1789 results. Seventeen of the first 20 documents, or 85%, cited at least one online source which had become publicly

²⁷ Documents will change over time and, it is possible that the cited, original, URL address will no longer "point" towards the original document but its later modified version. Statutes become amended and administrative rules become repealed from one administration to another. In some cases, the URL may

inaccessible. Again those sources varied from federal and local governmental sites to self-published ones, from html to PDF documents.

Thus, it only makes sense that law reviews whose articles reference many electronic sources should take steps – perhaps under their librarians’ guidance – to reduce the vulnerability of such documents. Currently, the editorial boards of law reviews and journals are aware that documents which are identified only by URLs change or even disappear. Their current solution to the issues of stability and accessibility is to transform those electronic documents into the old “on file” class of print documents that are filed in some filing-cabinet.

For example, all the items published by *Columbia Law Review* in its May 2005 issue contain footnotes that reference electronic documents identified by their URLs.²⁸ All of them had at least one URL that referenced a document that no longer existed.²⁹ Being aware of that potential problem, the editorial board added next to the useless URL the following message “(on file *with the Columbia Law Review*)”. This message usually means that a hard copy of the electronic document cited in the footnote is stored in a file cabinet in the law review’s office.

Even the most recent issue of CLR, continues to offer readers such elusive information as "On file with Columbia Law Review"³⁰ rather than creating an electronic archive which can be reliably used by future readers.

“point” to nothing at all, the government agency having re-addressed some or all of its publications. For more on this see Neacșu. 2002, 120, 105-122.

²⁸ <http://tinyurl.com/88s7h>.

²⁹ Id.

³⁰ See 107 CLR 649-650, notes 82, 83 84, or 107 CLR 810, footnote 126.

Like *Columbia Law Review*, all other law reviews undertake this effort in order to preserve the need of reliability of legal scholarship. Unfortunately, their efforts are to change the nature of the document from electronic to print, which creates an often insurmountable obstacle to their accessibility. Instead, it is this author's opinion, the evanescent documents should be preserved in their original format – electronically – which would make them both stable and easily accessible by users around the globe. While often those “on file” documents are lost and never located, or even worse, the convoluted process such a request creates often reduces the number of requests to a minimum, such a scenario would never have to happen once those URL-identified documents are stored electronically in a legal archive accessible to all for research purposes.

While digitization is usually held as “heralding an era of equality of information and access,”³¹ except image-type documents, digital information, “by its very nature, it is open to manipulation, alteration, reformatting, and erasure.”³² Thus the question remains: how can we offer both electronic accessibility and reliability?

SOME SOLUTIONS:

This potential threat to reliability may be prevented. There are a few options librarians advocate.

³¹ Amiran, Eyal. “Electronic Time and the Serials Revolution.” *The Yale Journal of Criticism*, 448 (1997): 445-454. See also McLuhan, Marshall. *The Gutenberg Galaxy; The Making of Typographic Man*. 63. Toronto: University of Toronto Press: 1962 (emphasizing “the unified field of electric all-at-onceness”)

³² Mazikana, Peter. “The Challenges of Archiving Digital Information.” *International Information and Library Review*. 313 (1997): 307-317.

(1) One is by changing habits and use persistent identification of electronic documents, as Susan Lyons explains in her recent article with the same title. For example, PURLs (where PURL stands for Persistent Uniform Resource Locator—and it is developed by OCLC) provide a persistent way of identifying and locating electronic documents using the redirect feature built into the HTTP protocol. It requires that the curating institution maintain its own PURL server and update it regularly. As Lyons reminds us, the United States Government Printing Office is the largest user of PURLs.

(b) More solutions envision archiving. For example, Google, has undertaken a mammoth project in book archiving with the help of major libraries. Its scope though, limited to legal scholarship per-se and not documents referenced in legal scholarship, is totally different than the one of the URL Archive: it wants to block fair-use competition as Google is currently settling a lawsuit with publishers,³³ in which Google recognizes their copyright, and they, for billions of dollars allow Google to continue its copying. Why would law libraries think twice before becoming partners with Google? Because they are allowing a monopoly over information which may only currently be freely-available.

There is also the Internet Archive -- Wayback Machine³⁴ archive which offers glimpses to different versions of electronic documents which are not database driven, nor generate dynamic web pages. Again, the scope and the running of this archive does not answer the needs of legal scholarship: there is no reliable coverage of sources cited in law review articles and no reliable institutional overview.

³³ See New Yorker article.

³⁴ <http://www.archive.org/web/web.php>.

Similarly, the Andrew W. Mellon Foundation-sponsored e-journal archiving program under which seven major libraries have received grants from the Foundation “including the New York Public Library and the university libraries of Cornell, Harvard, MIT, Pennsylvania, Stanford, and Yale,”³⁵ does not focus on legal scholarship published in law review articles and the sources cited there.

More recently, a Toronto-based group, WebCite Consortium, has produced the very interesting WebCite site which is currently focused on science articles and journals. It can be found at <http://www.webcitation.org/index>. However, it does not seem to have any institutional overview and it rests on authors self-archiving their works: authors join in the consortium and download their sources used in their works, thus allow future researchers to access and verify their argument.

THE ROLES OF LIBRARIES:

One of the many functions of a law library is to support legal teaching and scholarship. As both teaching and scholarship rely increasingly on digital works, including those that are "born-digital" information³⁶ and have no print or analog equivalent,³⁷ libraries are increasingly required to accommodate those changes in their

³⁵ <http://www.diglib.org/preserve/ejp.htm>.

³⁶ Danner, Richard. “Issues in the Preservation of Born-Digital Scholarly Communications” in *Law Library Journal*: 96 (Fall 2004): 591-604.

³⁷ Kelly Russell. “Libraries and Digital Preservation: Who Is Providing Electronic Access for Tomorrow?” in *Libraries, the Internet and Scholarship. Tools and Trends Converging*, at 1-2. Edited by Charles F. Thomas. New York: Marcel Dekker, 2002.

services, and ensure that legal scholarship is enhanced by the digital age and not eroded by it.

Eventually, I argue, law libraries need to be involved at a national level and build a consortium charged with creating and maintaining a digital archives that will encompass copies of the electronic documents cited in all student-edited law journals. Of course, such a consortium has to start somewhere and major law libraries can engage in helping their schools' law review articles by storing copies of sources cited in those articles and identified only by their URL. Such electronic archives would replace the current version of in-print ad-hoc archives, known as "on file" with the editor or law journal. As shown below, such digital archives are not only feasible, but also perfectly defensible under copyright rules.

THE ROLE OF LIBRARIES IN THE NEW DIGITAL AGE:

To the extent that digitization threatens reliability, all scholarship – including legal scholarship-- is under the same threat posed by digital publishing. One of the demands posed by those changes in the information medium is to ensure that legal scholarship is not eroded by the potential threat of digitization, and that legal scholarship remains reliable in electronic format as well. As one of the many functions of a law library is to support legal teaching and scholarship, it becomes a vital problem for law libraries to help solve this problem.

Reference law librarians are faced with these issues on a regular basis as we all work both with scholarly patrons and with members of the law journals' editorial boards. With little effort, an electronic archive of snapshots of the cited websites would keep the

sources to a specific law review article accurate. A digital archive will answer this demand by encompassing those electronic documents cited in law review articles and identified only by their URL and that are most vulnerable to changes.

THE LEGAL URL ARCHIVE:

Such an archive – The “Legal URL Archive” – remains a new and unique project with legal content, despite this author’s many efforts to popularize it and make it attractive to librarians across the country. The main advantage of this archive is that it can be developed in two fairly easy steps: first, as a local pilot project, and second, as a centralized archive of all online sources cited in all domestic law review articles. Additionally, unlike other solutions that require student-editorial boards to change their habits and use persistent identification of electronic resources, the electronic archive proposed here relies on librarians to find and promote the solution to unreliable legal scholarship.

At the local level, for example, the pilot project would use the articles currently being produced by each law school’s journals. The project envisions a kind of mirror site in which a duplicate of the cited electronic documents would be created with two commitments:

- (a) to maintain the document exactly as it existed when it was cited – which would solve the issue of stability and
- (b) to keep the information available to the general public for research and educational purposes; this would solve the accessibility issue.

This on-line archival structure would archive and provide quick access to legal scholars to sources used in building a legal argument discussed in a law review article. Such a service would be available in what is known as a “one-click” manner. In many ways, this is an electronic version of the traditional “on file at” law journals use to archive all materials unavailable. This method has been useless as the sources so archived are rarely available for future consultation.

The “Legal URL Archive” proposes the establishment of a regulated but accessible database into which the original materials are entered and recorded in their original form, as a kind of “cached” document. At this incipient level a library employee under the supervision of a librarian will copy the majority of electronic documents into the database periodically, when the articles are ready for publication for the respective journals or immediately upon their publication.

The “Legal URL Archive” would store only those electronic documents that are most vulnerable to changes. Unlike electronic documents that are accessible through commercial databases, the free-of-charge electronically available documents, which often are in the public domain, are also most vulnerable.

For example, in footnote 55³⁸ of “Preclusion in Class Action Litigation,” an article that was just published by the *Columbia Law Review*, Professor Tobias Barrington Wolff cites to a PDF document,³⁹ which, as its URL shows, was originally posted on a governmental database. These documents are generally believed to be more stable than ordinary html documents. However, while the law review article was only published in

³⁸ Wolff, Tobias Barrington. “Preclusion in Class Action Litigation” *Columbia Law Review*, 808 (2005): 717-808.

May of 2005, the cited document is not accessible at its original URL anymore. Its previous universal accessibility, which made it both attractive and available to the author as well, is currently limited to the *Columbia Law Review* offices. Thus, except documents that can be accessed for a fee or are comprised in commercial databases,⁴⁰ the “Legal URL Archive” would store all cited documents which are freely accessible through the Internet, and which are rarely part of an organized database. As shown above many of them, if not most of them, by their very nature as governmental documents, are already in the public domain.

The “Legal URL Archive” envisages thus first the creation of a database that consists of an electronic archive at the local level. There, it would incorporate those electronic footnotes contained in the local law journals that the law school students manage as explained above. This database would be accessible on the Internet (through the World Wide Web, for example) to researchers that will have free access to it upon signing an agreement that limits their access to the database only for research purposes.

(Appendix 2)

The database would contain electronic documents, and its basic fields would be:

- a) the volume number of the law review or journal where the article that cites the electronic document was published;
- b) the date of the law review or journal
- c) the title of the law review or journal
- d) the title of the article
- e) the page number the article begins on
- f) the footnote number
- g) the page number the footnote appears on

³⁹ [http://www.fjc.gov/public/pdf.nsf/lookup/EmplDis4.pdf/\\$file/EmplDis4.pdf](http://www.fjc.gov/public/pdf.nsf/lookup/EmplDis4.pdf/$file/EmplDis4.pdf), at fn 55 in Wolff 2005, 717-808.

⁴⁰ Although it should be noted that there is no need for this restriction from a legal point of view, since that which is cited is only a small portion of any database, and is further almost inevitably itself in the public domain.

- h) the author's name
- i) the URL address of the cited material
- j) the downloaded html file, and
- k) the date when the document was harvested.

Tables could be created to translate the "Bluebook" format into the fields available in the database. For example, a citation search for "14 Colum. J. Asian L. 235 (2000)" should retrieve all of the footnotes that contain URLs or "on file" references in that article. (Appendix 1) Similarly a URL search for <<http://www.suite101.com/article.cfm/389/8554>> should retrieve the archived document behind it.

Of course, once a few law libraries are interested in co-sponsoring this project they could work with the journals' editorial boards and have either the libraries or the journals create a consortium that would host the database. At that point it is conceivable that even the rules of citation may easily be amended so the BLUEBOOK rules would demand that journals help libraries in gathering those electronic sources. Once an article is ready for publication, for example, its editors would give libraries a list of the URLs found in those articles, and through a joined effort those electronic sources could be archived before they have time to change in content or even disappear.

There is an existing site at <http://library.law.columbia.edu/urlmirror/> that offers a glimpse at what The "Legal URL Archive" proposes. The site contains a sample list of journal titles published at Columbia Law School; article titles in each of the journals; selected footnotes (indicated by number) from the articles that contain references to URLs; and links to the materials on the web pages cited as authority.

This is the result of the work of two people, this author – a reference librarian -- and an engineering student.⁴¹ This author searched LEXIS for all the footnotes containing online sources that were cited in recent issues of the Columbia Law School journals, and then created Word documents for each such issue. Then, the engineering student created electronic copies of the online sources and linked them to a web site on the library server. Of course, this is only a beginning, and has disadvantages at both the input and the output levels, but they are of such a nature that using a more sophisticated database and interface software can easily solve them.

In neither format, locally and nationally, is this project grandiose. As detailed above, the “Legal URL Archive” is quite modest. And more importantly, it is completely feasible.

Of course, even the archive this author advocates is not entirely safe. It runs the risk of becoming inaccessible if technological improvements are not ensured periodically. But what other entity will be more interested in ensuring the archive’s accessibility if not a national consortium of law libraries in charge of the reliability of the bulk of legal scholarship!

More and more law review articles cite to electronic-born documents because search engines such as Google are easy to use and their retrieval is highly responsive. Thus, the problems raised by the evanescent nature of digitized documents are only going to increase which makes the need for a solution imminent.

The “Legal URL Archive” will solve the vexing problem of storing the documents that come under the rubric of “on file at” by solving problems of access,

⁴¹ Nelson Brand, Columbia University graduate, 2002.

retrieval, and storage of that digital information. The materials thus accessed would be a copy of the original document and will be subject to influence or change neither by its own author nor by the author of the primary publication which refers to it, since the collection is done by the database. The material accessed would be available in the original form for research and educational purposes.⁴²

ANSWERING FALSE CONCERNS: FAIR USE AND THE “LEGAL URL ARCHIVE:”

Librarians are very copyright conscious. For reasons that escape this author, librarians seem more often worried about potential copyright infringement issues than about unconscious limitations to the public’s access to various works.

To assuage such worries this author will take a few moments to explain that the archive proposed here does not create any potential copyright infringement. Using search engines such as Google to access information freely available which allows researchers to then store it does not seem to create any copyright infringement issues regarding the authors’ distribution of rights. More so using the URL archive is copyright-infringement safe as it comes under the fair-use protection.

Under the copyright rules, 17 U.S.C. § 106, the author of a literary work⁴³ has well-established exclusive rights over her work – of course, to the extent that the work

⁴²Once the demonstration project successfully incorporates the electronic sources whose URLs are identified in the footnotes of one school’s journals, other law libraries will likely join in the effort and create similar databases for their law schools’ journals. Eventually, law libraries may join in a consortium that will create a centralized database that will cover the electronic sources cited in all the law journals.

⁴³ As legal scholarship consists of literary works, the majority of its sources consist of literary works archived in different formats by libraries.

has not fallen in the public domain,⁴⁴ or it does not consist of only facts and ideas.⁴⁵

Those rights are as follows:

- (1) the reproduction right (i.e., the right to copy, imitate, reproduce, duplicate or transcribe the work in fixed form);
- (2) the derivative work right (i.e., the right to modify the work to create a new work);
- (3) the distribution right (the right to distribute copies of the work to the public by sale or rental, lease, etc., often construed to be limited to the first sale only);
- (4) the performance right (the right to perform the work in public);
- (5) the display right (i.e., the right to show a copy of the work directly to the public by means of a web site, film, slide, or television image at a public place or to transmit it to the public.);⁴⁶ and
- (6) the digital transmission performance right.⁴⁷

In the case of the digital archive discussed here the potential copyright infringement⁴⁸ concerns the reproduction, distribution, and the display rights. It is obvious that the creation of a publicly accessible digital archive implies copying the

⁴⁴ To the original term limit of a copyrighted work THE SONNY BONO COPYRIGHT TERM EXTENSION ACT attached an additional twenty years of protection. Pub. L. No. 105-298, § 101, 112 Stat. 2827 (1998). The Act provides for a library exception under which libraries for purposes of preservation, scholarship or research may freely use works in the last 20 years of protection if the work is not subject to normal commercial exploitation or is not available at a reasonable price. It is worth noting that the Act does not define what is meant by "normal commercial exploitation," which allowed some publisher representatives to claim that "if a work is merely identified as "available for a license," that should qualify as normal commercial exploitation, even if the work has not been sold or licensed for decades and even if the publisher does not have a copy of the work." Id. See generally, Davis, Michael "Extending Copyright and the Constitution: 'Have I Stayed Too Long?'" *Florida Law Review* 990-993. (2000): 989-1037.

⁴⁵ For a definition of copyrightability, see Miller, Arthur, Michael Davis. *Intellectual Property. Patents, Trademarks and Copyright in a Nutshell*, at 292-94. St. Paul, Minn.: West Group (2000).

⁴⁶ For a brief introduction see e.g. Warwick, Shelly. "Copyright for Libraries, Museums, and Archives. The Basics and beyond." at 241-42 in *Libraries, Museums, and Archives: Legal Issues and Ethical Challenges in the New Information Era.* Edited by Lipinski, Tomas. Lanham, Md.: Scarecrow (2002).

⁴⁷ See also Miller and Davis.2000, 323.

electronic information, then storing and displaying it on the library server for reference use by legal scholars (both students and faculty members).

Indeed, in ordinary circumstances this might constitute a violation of copyright law. The library will be more than an ordinary online service provider (OSP). It will have control over the actual content of the transmissions. It will provide software to link users to the copy of the original sites. It will store the archived digital information on its server, and it will facilitate displays by subscribers (even if not further recordings through downloading).⁴⁹ But, as shown below, the library's archive does not constitute copyright infringement, as its use is covered by the fair-use doctrine.

The fair use doctrine is often raised in educational activities⁵⁰ like the one related to this digital archive, though it is certainly not limited to such uses. For example, under 17 U.S.C. § 107, one may reproduce the copyrighted work as fair use “for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research.”

⁴⁸ Although infringement is not statutorily defined, it occurs “whenever somebody exercises and of the rights reserved exclusively for the copyright owner without authorization.” *Id.*, at 340.

⁴⁹ Although some of the OSP's liability limitations under copyright law exist under the ONLINE COPYRIGHT INFRINGEMENT LIABILITY LIMITATION that was approved by Congress as part of the omnibus Digital Millennium Copyright Act of 1998 (DMCA) – Pub. L. No. 105-304, 112 Stat. 2860 (1998) – they do not interest us in this case. The type of activities covered by the DMCA are most transitory digital network communications that do not constitute the type of activity covered by the communications necessitated by the creation and the use of the digital archive which wants to become a perennial mode of storing transient web information.

⁵⁰ See Miller and Davis.2000, 354.

Fair use is almost invariably a conclusion based upon the weighing of competing policies⁵¹ (i.e., copyright monopoly versus public access). Thus, a finding of fair use depends on the existence of several factors:

- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purpose;
- (2) the nature of the copyrighted work (i.e., whether the work is fictional as opposed to factual);
- (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
- (4) the effect of the use upon the potential market for or value of the copyrighted work.⁵²

More importantly, it should be noted that the Supreme Court has emphasized on several occasions that the last factor -- the effect of the use upon the potential market of the work -- is the most important and even decisive one when finding fair use. It held it in *Harper & Row, Publishers, Inc. v. Nation Enterprises*.⁵³ “This last factor is undoubtedly the single most important element of fair use.”⁵⁴ Similarly, the Court held it again in *Stewart v. Abend*.⁵⁵ “The fourth factor is the ‘most important, and indeed, central fair use factor.’”⁵⁶

⁵¹ Dratler, Jay, Jr. 2. Intellectual Property Law: Commercial, Creative, and Industrial Property § 5A.02[2] (2001) (“it allows the courts to adjust the contours of copyright protection as necessary to balance the conflicting policies of copyright”).

⁵² Warwick, 2002, 244.

⁵³ 471 U.S. 539 (1985).

⁵⁴ 471 U.S. at 566.

⁵⁵ 495 U.S. 207 (1990).

⁵⁶ 495 U.S. at 238.

Fair use requires a specific factual inquiry and decisions are often contradictory and often apparently little guidance. For example, one case found no fair use when “a library user [copied] an entire book that is protected by copyright,⁵⁷” while another found the contrary, and it stated that a library can make copies of an entire work at the request of a faculty member, for in-class use.⁵⁸

However, fair use instances remain unquestioned⁵⁹ when the work is used for educational nonprofit archival purposes and when the archival has no economic impact upon the work’s marketability.⁶⁰ The goal of the URL archives fits squarely in this latter situation – it both educational and nonprofit.

Legal scholars can access the archived sources in order to verify a legal argument they found interesting in a law review article. The archive’s purpose is not to surreptitiously copy works against the will of their authors, its objective is to prevent changes in the content of such works once they became part of a legal argument published in a law review article.

Most importantly, unlike the cases in which remotely similar use was insufficient to establish the fair-use defense,⁶¹ because the copying (for “productive use”⁶²) adversely

⁵⁷ Warwick. 2002, 244.

⁵⁸ Of course, according to 37 C.F.R. § 201.14 a nonprofit library can make photocopies for their users consistent with the fair-use guidelines: one copy per use, and the copy becomes the user’s property. See also Warwick 2002, 247.

⁵⁹ See Miller and Davis. 2000, 359.

⁶⁰ The digital nature of the copyrighted work does not change the terms in which the fair use exemption applies here. See Litman, Jessica .*Digital Copyright* at 130-132, 180 (2001).

⁶¹ *Encyclopaedia Britannica Educ. Corp. v. Crooks*, 542 F.Supp. 1156 (W.D.N.Y. 1982) (large-scale reproduction by an educational service of copyrighted works originally broadcast on television did not constitute fair use).

affected the only market for that work, the digital archive stores only works that can be freely accessible on the web because their authors had posted them on the web with the intent that they be freely accessed.⁶³ Because the archived works are freely posted on the Internet by their authors, with no expectation of compensation, but with the understanding that they will be accessed freely by users by using such free search engines as Google, the Supreme Court's admonition that this factor is the most important and decisive could not be more fully satisfied. To that extent, the defense of fair use "seems preordained."⁶⁴

In addition, even a copying of the entire work is fair here, not only because "fair use is neither established nor precluded by a simple mathematical measurement,"⁶⁵ but most importantly because it is the document in its entirety which has been posted by the author herself for others to freely use. In fact in *Williams & Wilkins*⁶⁶ the Supreme Court affirmed the decision of the United States Court of Claims⁶⁷ which had rejected a straightforward quantitative measurement of proportional copying.⁶⁸

There is, in short, no inflexible rule excluding an entire copyrighted work from the area of "fair use." Instead, the

⁶² Id. See also Lape, Laura G. "Transforming Fair Use: The Productive Use Factor in Fair Use Doctrine." *Albany Law Review* 698 (1995): 677-723.

⁶³ It may also contain works that are not protected by copyright rules, such as U.S. Government works. See, e.g. Crews, Kenneth D. *Copyright Essentials for Librarians and Educators*. at 15-16.(2000).`

⁶⁴ Miller and Davis. 2002, 360.

⁶⁵ Id at 365.

⁶⁶ 420 U.S. 376 (1975).

⁶⁷ 203 Ct.Cl. 74 (1973).

⁶⁸ Miller and Davis. 2002, 364.

extent of the copying is one important factor, but only one, to be taken into account, along with several others.⁶⁹

In other words, it held that “an entire taking of all of the plaintiff’s work could be justified under fair use by other factors.”

CONCLUSION:

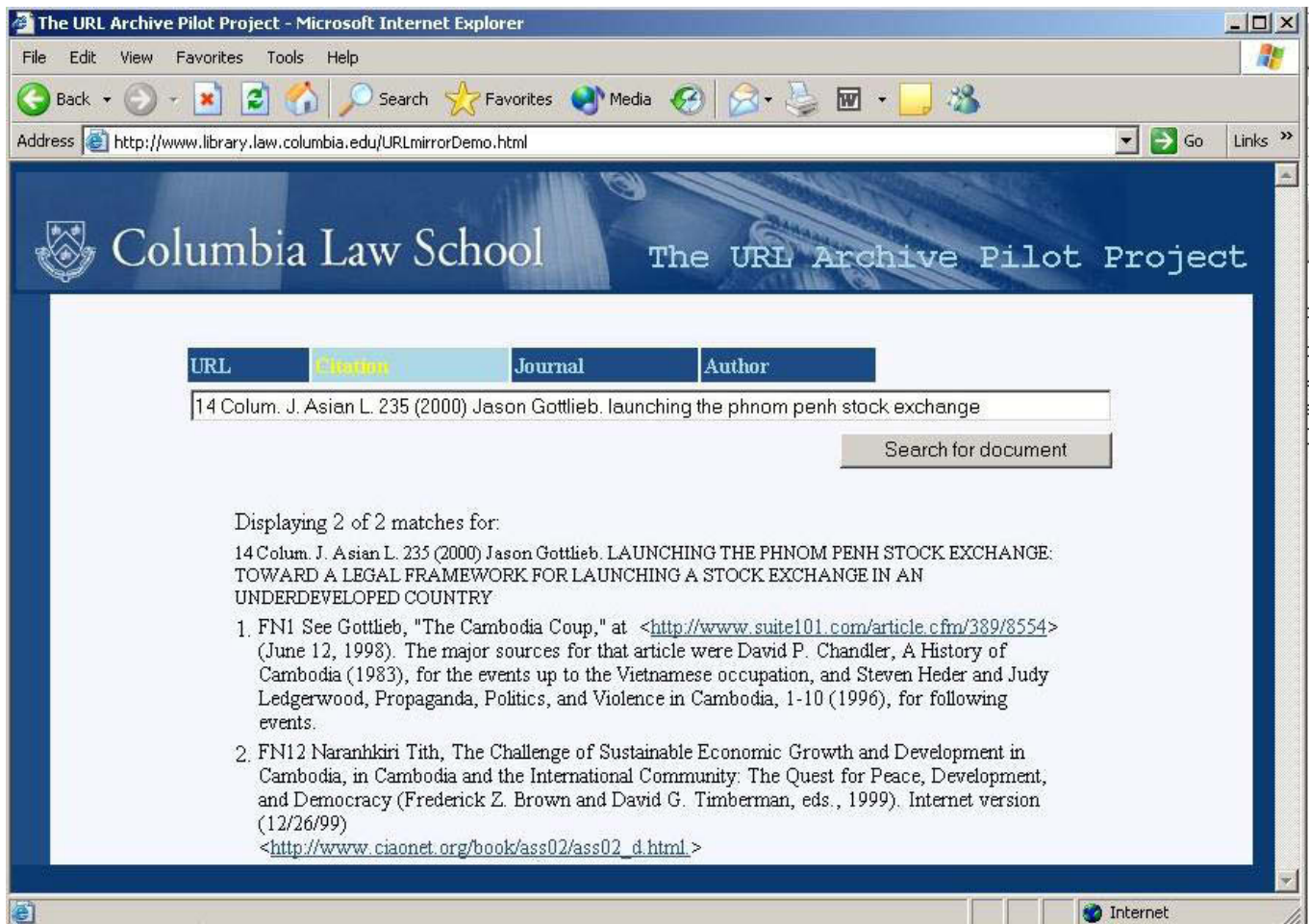
The advent of powerful search engines such as Google made electronic research an inescapable reality. However, while search engines have eased the access to electronic sources, and increased legal scholars’ reliability on electronic sources, Google has not solved the problems of accurate reliability for the future.

As one of the many functions of a law library is to support legal teaching and scholarship, it becomes a vital problem for law libraries to help solve this problem.

One way to answer these demands is by creating digital archives of those legal electronic documents cited in law review articles and identified only by their URL. The content of the archives will be limited to internet sources that are the most vulnerable to changes: documents in the public domain and those posted on the Web by their authors for free access through free search engines. As the archive will solve a most important aspect of the threat digital publishing, reliable access, it becomes an ideal candidate for a library project.

⁶⁹ 203 Ct.Cl. at 90.

Appendix 1:



The screenshot shows a Microsoft Internet Explorer browser window titled "The URL Archive Pilot Project - Microsoft Internet Explorer". The address bar contains the URL "http://www.library.law.columbia.edu/URLmirrorDemo.html". The page header features the Columbia Law School logo and the text "The URL Archive Pilot Project". Below the header is a search interface with a table of search criteria:

URL	Citation	Journal	Author
14 Colum. J. Asian L. 235 (2000) Jason Gottlieb. launching the phnom penh stock exchange			

A "Search for document" button is located to the right of the search criteria table. Below the search results, the text reads "Displaying 2 of 2 matches for:" followed by the search query. The results are listed as follows:

1. FN1 See Gottlieb, "The Cambodia Coup," at <http://www.suite101.com/article.cfm/389/8554> (June 12, 1998). The major sources for that article were David P. Chandler, A History of Cambodia (1983), for the events up to the Vietnamese occupation, and Steven Heder and Judy Ledgerwood, Propaganda, Politics, and Violence in Cambodia, 1-10 (1996), for following events.
2. FN12 Naranhkiri Tith, The Challenge of Sustainable Economic Growth and Development in Cambodia, in Cambodia and the International Community: The Quest for Peace, Development, and Democracy (Frederick Z. Brown and David G. Timberman, eds., 1999). Internet version (12/26/99) http://www.ciaonet.org/book/ass02/ass02_d.html

Appendix 2:

The screenshot shows a Microsoft Internet Explorer browser window titled "The URL Archive Pilot Project - Microsoft Internet Explorer". The address bar contains the URL "http://www.library.law.columbia.edu/URLmirrorFront.html". The page content includes the Columbia Law School logo and the title "The URL Archive Pilot Project".

The URL Archive

You can access and use The URL Archive if you agree with the terms of the agreement below ("Agreement"). If you agree with the terms of the Agreement, click on "I Agree" below and click on "GO" to submit the Agreement. If you do not agree with the terms of the Agreement click on "I Don't Agree." and you will not be provided access to The Archive.

The URL Archive comprises a mix of public domain materials (governmental web pages), and copyrightable material (privately authored web pages, mostly of factual compilations, freely posted on the Internet by their authors). Because this material is used here solely for purposes of comment, criticism, teaching, scholarship, or research, and otherwise substantially satisfies the requirements of 17 U.S.C. Section 107, its reproduction clearly falls within that section's exception of fair use.

Scholars may use **The URL Archive solely for purposes of comment, criticism, teaching, scholarship, or research.** Within those limits, readers are encouraged to fully explore the site and use it for reference purposes in their scholarly work.

Any further reproduction or modification of materials from this site is the sole responsibility of the user.

I Agree
 I Don't Agree

GO