

Yellow Flag Fever: Describing Negative Legal Precedent in Citators*

Aaron S. Kirschenfeld**

This study analyzes the accuracy with which descriptions of subsequent negative treatment are applied by online citators. A system making use of a hierarchical controlled vocabulary applying these descriptions appears marginally more accurate, but the citator's traditional role in legal research must be reconceptualized.

Flying the Flag: An Introduction	77
Looking Backward: Reviewing the Relevant Literature.	79
Methodology	83
Results	87
Looking Forward: A Discussion of the Citator's Place	93
Conclusion	98

Flying the Flag: An Introduction

¶1 Iconography without meaning is no iconography at all. Imagine if tomorrow we were to look at the Star Spangled Banner and think not of America but instead emotionlessly identify vaguely with thoughts of a shapeless, formless, but nonetheless large country in the Western Hemisphere. This scenario would make our flag's symbolic power next to nil because a symbol's sine qua non is to convey a coherent, readily grasped message. This much, without delving into the past half century of semiotics and linguistics, we should be able to say is true, especially when the object of the symbol is to communicate concrete meaning to subject-matter experts seeking answers. The flag's place at the center of a town's Fourth of July celebrations is discursively rich, but its use as a small digital image to be clicked on when setting up a new computer, indicating that U.S. English is the desired keyboard configuration, should not need to be.

¶2 This study began with a question: do the symbols and phrases used in commercial online citators adequately convey meaning to legal researchers? Based on the study's findings, I posit that serious questions surround the citator's place in legal research. I then provide suggestions for teaching students and practitioners how to best use citators. Among other things, I suggest that legal information

* © Aaron S. Kirschenfeld, 2016. This is a revised version of the winning entry in the student division of the 2015 AALL/LexisNexis Call for Papers Competition.

** Reference & Digital Initiatives Law Librarian, University of North Carolina at Chapel Hill, Kathrine R. Everett Law Library, Chapel Hill, North Carolina.

professionals bear the responsibility of questioning legal information vendors about the accuracy of their citator products. I also offer areas of further research for developing a more robust, rigorous literature on citators as legal research products. Finally, I present alternatives to traditional citators that might someday subsume their functions, including citation analysis tools based on visualization or on algorithmic extraction and presentation of subsequent negative treatment.

¶3 But let us start our inquiry with some first principles. The system of common law¹ as practiced in the United States depends on the doctrine of *stare decisis*, or in adhering to precedent established in earlier cases. Judges in state appellate courts rely on precedent from their own jurisdictions and are sometimes bound by it, but also consider precedent from other states, which can be seen as persuasive authority. On constitutional issues, state judges must defer to pronouncements of the U.S. Supreme Court. Judges in federal district courts look to precedent from the appellate court from the geographic circuit they practice within when interpreting federal law, and to the Supreme Court as well. Precedent from other circuits may also be persuasive. And state law precedent binds federal courts in diversity of jurisdiction actions.² In this system of overlapping, often complex, authority, lawyers must research and argue based on judicial precedent and its application to their particular clients' facts. Not only is finding precedent that supports clients' desired outcomes necessary, however; lawyers must also find precedent that calls into question their arguments. As the title suggests, this article focuses on the second type of precedent.

¶4 One tool in lawyers' arsenals for accomplishing this—specifically in determining cases' precedential value—is the citation index, or citator. Major commercial online citators signal negative treatment of a case using one general form: an introductory symbol—such as a colored flag or a colored shape—followed by an explanatory phrase. The appearance of both of these elements varies greatly among the two major citators, KeyCite (Westlaw) and Shepard's (LexisNexis). Some minor citators, such as Bad Law Bot (Fastcase), do not include both elements. These products rarely agree on the meaning of these explanations of subsequent negative treatment. But do these differences in characterization of a case's negative treatment affect how citators themselves, as well as the lawyers who use them, evaluate the case's significance in common-law precedent?

¶5 I propose that when a standardized and hierarchical controlled vocabulary is applied to a legal citation indexing system, evaluation of a case's subsequent negative treatment is more accurate. Knowing whether this proposition is valid will help inform lawyers, law librarians, and teachers of legal research in addition to legal information vendors. The LexisNexis Shepard's product is an example of a hierarchical controlled vocabulary of subsequent negative treatment. In contrast, KeyCite uses language from cases to produce explanatory phrases describing subsequent negative treatment. Both systems use symbols to represent subsequent negative treatment as well.

1. The term "common law" is used in an informal sense, given that so much law in the United States is statutory or regulatory in nature. However, judges do still make law by interpreting legislative actions and other cases, and their decisions can still be said to be meaningfully different from those in countries practicing civil law.

2. *Erie R.R. Co. v. Tompkins*, 304 U.S. 64 (1938).

¶6 After review of a set of cases in Shepard's and KeyCite, described in the methodology below, the results suggest, absent the action of other variables, that citators using a hierarchical controlled vocabulary do describe subsequent negative history more accurately than those that do not. Shepard's performs better than KeyCite by applying explanatory phrases more accurately in the limited sense tested. Shepard's also performs better in applying introductory symbols more accurately. To that end, this article introduces an operationalized definition of "accuracy" and a method for determining whether descriptions of negative precedent have been applied accurately.

Looking Backward: Reviewing the Relevant Literature

¶7 Legal research instructional materials and courses emphasize the importance of updating one's research before submitting a legal work to a publisher, court, or other lawmaking body.³ Updating involves revisiting selected documents to determine whether they still represent the most current law. (While citators are used to update statutory, regulatory, and case law materials, this article focuses on legal precedent found in case law.) Usually, a lesson on using citators involves a variation on a familiar theme in the law—fear of incompetence. Young lawyers who fail to update a case will find themselves in hot water, and fast. This is no doubt true—judges do not generally look kindly on counsel whose arguments rest on shaky precedent.⁴ Others have raised the question of whether the failure to update a case is a violation of a lawyer's professional ethical obligations.⁵

¶8 To date, academic research on citators has fallen into one of two general camps: (1) the history and development of citators, and (2) wholesale comparisons of Shepard's Citations and West's KeyCite.⁶ While interesting in its own right, a thorough review of the literature on the history and development of citators is unwarranted here, though it is treated in some depth in this article's discussion section.⁷ More significant are the attempts—made both by law librarians and by

3. See, e.g., J. ARMSTRONG & CHRISTOPHER A. KNOTT, *WHERE THE LAW IS: AN INTRODUCTION TO ADVANCED LEGAL RESEARCH* 143 (4th ed. 2013); SCOTT CHILDS & SARA SAMPSON, *NORTH CAROLINA LEGAL RESEARCH* 143 (2d ed. 2014); MORRIS L. COHEN & KENT C. OLSON, *LEGAL RESEARCH IN A NUTSHELL* 108 (10th ed. 2010); KENT C. OLSON, *PRINCIPLES OF LEGAL RESEARCH* 288 (2009); MARK K. OSBECK, *IMPECCABLE RESEARCH: A CONCISE GUIDE TO MASTERING LEGAL RESEARCH SKILLS* 51 (2010).

4. Lisa McElroy, *Motivating Students to Use Citators: Lessons from Litigation*, 18 *PERSP.: TEACHING LEGAL RES. & WRITING* 140 (2010).

5. Carol M. Bast & Susan W. Harrell, *Ethical Obligations: Performing Adequate Legal Research and Legal Writing*, 29 *NOVA L. REV.* 49 (2004).

6. See, e.g., Laura C. Dabney, *Citators: Past, Present, and Future*, 27 *LEGAL REFERENCE SERVS. Q.* 165 (2008); Michael J. Lynch, *Citators in the Early Twentieth Century—Not Just Shepard's*, 16 *LEGAL REFERENCE SERVS. Q.*, 1998, no. 3, at 5; Patti Ogden, *Mastering the Lawless Science of Our Law: A Story of Legal Citation Indexes*, 85 *LAW LIBR. J.* 1 (1993). For the institutional history, see FRANK SHEPARD CO., *A RECORD OF FIFTY YEARS OF SPECIALIZING IN A FIELD THAT IS OF FIRST IMPORTANCE TO THE BENCH AND BAR OF THE UNITED STATES: AN INSIGHT INTO AN ESTABLISHMENT THAT HAS GROWN FROM SMALL BEGINNINGS TO THE FIRST RANK IN THE LAW PUBLISHING FIELD* (1923). For the intriguing argument that citation indexing began with law, see Fred R. Shapiro, *Origins of Bibliometrics, Citation Indexing, and Citation Analysis: The Neglected Legal Literature*, 43 *J. AM. SOC'Y INFO. SCI.* 337 (1992).

7. Specifically, the claim that citation indexing was part of a late nineteenth-century classification trend as noted in GEOFFREY C. BOWKER & SUSAN LEIGH STAR, *SORTING THINGS OUT: CLASSIFICATION AND ITS CONSEQUENCES* 16 (1999).

employees of the legal information vendors—to differentiate the market-leading citators at different points in time.⁸

¶9 Perhaps the leading, albeit dated, comparison of citators (following the introduction of KeyCite in the mid-1990s) is Taylor’s 2000 article in *Law Library Journal*. He compared Shepard’s and KeyCite for three factors: completeness, currency, and accuracy. He defined accuracy as “whether the system correctly identifies all citing opinions that have a negative effect on the validity or persuasiveness of the cited opinion.”⁹ Because he was assessing the citator holistically, this definition suited his purposes. But it more describes the concept of “recall” than “accuracy.” Taylor’s study, however, supplies a framework for talking about the types of information found in citators, and his vocabulary is discussed in this article’s section on methodology.

¶10 That said, in other studies assessing the accuracy of a controlled vocabulary, the term is meant as a study in the accuracy of application—for example, were all articles about cholesterol-lowering drugs marked with one of the correct subject terms relating to cholesterol-lowering drugs?¹⁰ Again, this is sensible when using “accuracy” to judge an information system as whole, and roughly aligns with Taylor’s recall-heavy definition. This article’s operational definition of “accuracy” differs slightly due to the special nature and needs of a legal citation index and the quality this study purports to test.

¶11 In general, Taylor and similar older comparisons of citators conclude that each commercial system has its strengths and weaknesses, and that truly effective legal researchers ought to consult both Shepard’s and KeyCite to ensure correctness. The legal information vendors also had their say in response to Taylor.¹¹ More recent comparisons of citators have looked at the citator’s place within the online legal research product context—that is, how the citator behaves in relation to other parts of a unified tool containing hundreds of legal databases. Mart has studied both Shepard’s and KeyCite for their relative strengths (or, as it turns out, weaknesses) when accessed through case law subject headings, or “headnotes.”¹² The difference

8. And timing did have a lot to do with it. KeyCite was developed in the 1990s, and many of these comparative studies attempted to benchmark it to Shepard’s. See Elizabeth M. McKenzie, *Comparing KeyCite with Shepard’s Online*, 17 LEGAL REFERENCE SERVS. Q., 1999, no. 3, at 85 [hereinafter McKenzie, *Comparing KeyCite*]; Elizabeth M. McKenzie, *New Kid on the Block: KeyCite Compared to Shepard’s*, AALL SPECTRUM, Oct. 1998, at 8. For a more recent comparison, see Carol A. Levitt & Mark Rosch, *Are All Citator Services Created Equal? A Comparison of Google Scholar, Fastcase, Casemaker, LexisNexis, WestlawNext, and Bloomberg* (Internet for Laws white paper, 2012), <http://www.netforlawyers.com/content/caselaw-citators-comparison-Google%20Scholar-Fastcase-Casemaker-LexisNexis-WestlawNext-Bloomberg-0059>.

9. William L. Taylor, *Comparing KeyCite and Shepard’s for Completeness, Currency, and Accuracy*, 92 LAW LIBR. J. 127, 132, 2000 LAW LIBR. J. 13, ¶ 21. For an earlier and less comprehensive effort, see Fred R. Shapiro, *KeyCite and Shepard’s—Coverage and Currency of Citations to Recent Cases: A Comparative Study*, LEGAL INFO. ALERT, Apr. 1998, at 1.

10. Natalie Kupferberg & Bridget McCrate Protus, *Accuracy and Completeness of Drug Information in Wikipedia: An Assessment*, 99 J. MED. LIBR. ASS’N 310 (2011).

11. Jane W. Morris, *A Response to Taylor’s Comparison of Shepard’s and KeyCite*, 92 LAW LIBR. J. 143, 2000 LAW LIBR. J. 14; Daniel P. Dabney, *Another Response to Taylor’s Comparison of KeyCite and Shepard’s*, 92 LAW LIBR. J. 381, 2000 LAW LIBR. J. 33.

12. Susan Nevelow Mart, *The Case for Curation: The Relevance of Digest and Citator Results in Westlaw and Lexis*, 32 LEGAL REFERENCE SERVS. Q. 13 (2013).

that Mart fixes on is between human-generated headnotes and algorithm-generated headnotes, and she studies how inclusive each type of headnote is in listing cases cited. Mart's article does assume that the Shepard's explanatory phrases are both generated or selected and applied by human editors, while the KeyCite explanatory phrases are likely generated or selected and applied algorithmically.¹³ But this is merely an assumption based on a general impression of the data gathered, not on any otherwise verifiable information. And this point will not be a focus of the article, as it presents what is likely a distinction without a difference—the object of this study is the explanatory phrases themselves and whether or not they are accurately applied to the judicial disposition in a case, not how they were generated. Mart's conclusions ultimately echo those in earlier citator comparisons in the narrow sense: both products are flawed, but flawed in different ways.¹⁴

¶12 Some in the profession have questioned whether the companies that market citators ought to claim that their descriptions of subsequent negative treatment “validate” case law.¹⁵ Updating case law or checking citations is no simple task. Regardless of what the citator claims about the treatment of a case, there will be misapplications or gaps in coverage. Yet substantial anecdotal evidence suggests that lawyers accept the companies' claims about validity more or less unquestioningly.¹⁶ This perspective from actual lawyers adds urgency to our problem. Validity follows from perceived accuracy, and challenges to the role and uses of citators in attorney's legal research will be discussed following the presentation of this study's results.

¶13 Literature about standardized controlled vocabularies, judicial precedent, and citation indexing in general are most useful to this article's task. Studies of law-specific controlled vocabularies have focused on the difficult nature of selecting terms using literary warrant—or classification based on the content of the information resources being classified—in case law collections.¹⁷ (And, interestingly, none treat the distinction between literary warrant and end-user warrant—or classification based on how the information's users themselves conceptualize it—in the law.) On the whole, both the legal information literature and the academic literature concern the generation of controlled vocabularies or the appropriateness of terms chosen, but neither touches on something as specific as what is being studied here. In all, researchers seem to be concerned with studying the systems themselves or their effect on the legal profession or jurisprudence conceptually, rather than the effect that the system might have on particular users in particular instances, or on particular information objects. And of course, these are all worthy and interesting pursuits; some will even be taken up later in this article. But there are very few user studies of legal researchers and none, to my knowledge, focused solely on the use of

13. *But see* McKenzie, *Comparing KeyCite*, *supra* note 8, at 91.

14. Mart, *supra* note 12, at 43.

15. Alan Wolf & Lynn Wishart, *Shepard's and KeyCite Are Flawed (or Maybe It's You)*, N.Y. ST. B. ASS'N J., Sept. 2003, at 24.

16. *Id.* at 25.

17. *See, e.g.*, Robert C. Berring, *Legal Research and the World of Thinkable Thoughts*, 2 J. APP. PRAC. & PROCESS 305 (2000); Daniel P. Dabney, *The Universe of Thinkable Thoughts: Literary Warrant and West's Key Number System*, 99 LAW LIBR. J. 229, 2006 LAW LIBR. J. 14.

citators.¹⁸ And while this article is not a user study, it does aim to address this gap in the literature.¹⁹

¶14 The literature on the application of standards, specifically ANSI/NISO Z39.19 2005, to collections has been helpful, although ultimately limited. As with some of the studies mentioned earlier, the measure of “accuracy” is more about consistency in application rather than correctness or perceived usefulness. The studies of standards tend to consider issues of whether application of terms is “correct” rather than “useful.” Here, of course, we are concerned with the latter. That said, the standard is also relatively new, meaning that there have been few substantive studies on it since its publication and adoption by information professionals. The standard itself does, however, offer methods for assessing controlled vocabularies.

¶15 While the literature on Medical Subject Headings (MeSH) is well developed, it lacks relevance when applied to legal subject matter in one key respect: medical article and citation indexing does not, at least to this point, employ judgments of subsequent negative treatment.²⁰ This article might be able to shed light on an issue not yet taken up in other professional communities—which way is best to express when articles or hypotheses, like cases or points of law, are challenged, distinguished, or even discredited. And while law is not a science—and American writers in the field have to be given credit for acknowledging such, at least since the late nineteenth century—its peculiar crucible of precedent developed through trial and appeal can be seen as an analogue to work in more practice-oriented sciences like medicine.

¶16 Finally, a look into whether “subsequent negative treatment” has itself been satisfactorily defined is needed. In short, the literature on this topic is in one sense quite large, but for the most part focuses on the concepts of legal change or of “compliance,” that is, whether cases actually follow other cases, and to what degree.²¹ This literature tends to come from the political science community, and is probably not relevant here. However, compliance is a larger inclusive category for study of the common law doctrine of *stare decisis*, or judicial precedent. The legal

18. For an interesting recent survey on lawyer use of information, see ALL-SIS TASK FORCE ON IDENTIFYING SKILLS & KNOWLEDGE FOR LEGAL PRACTICE, A STUDY OF ATTORNEYS' LEGAL RESEARCH PRACTICES AND OPINIONS OF NEW ASSOCIATES' RESEARCH SKILLS (June 2013), <http://www.aallnet.org/sections/all/storage/committees/practicetf/final-report-07102013.pdf>.

19. That said, David L. Armond & Shawn G. Nevers, *The Practitioners' Council: Connecting Legal Research Instruction and Current Legal Research Practice*, 103 LAW LIBR. J. 575, 2011 LAW LIBR. J. 36, provides an interesting, if sobering, take on the difficulty of engaging practicing lawyers in a discussion of legal resources.

20. Mark E. Funk & Carolyn Anne Reid, *Indexing Consistency in MEDLINE*, 71 BULL. MED. LIBR. ASS'N 176, 183 (1983).

21. See, e.g., James F. Spriggs, II & Thomas G. Hansford, *Measuring Legal Change: The Reliability and Validity of Shepard's Citations*, 53 POL. RES. Q. 327 (2000). For particular studies, see Scott D. McClurg & Scott A. Comparato, *Rebellious or Just Misunderstood? Assessing Measures of Lower Court Compliance with U.S. Supreme Court Precedent*, in ANNUAL MEETING OF THE AMERICAN POLITICAL SCIENCE ASSOCIATION, PHILADELPHIA (2003); James F. Spriggs, II & Thomas G. Hansford, *The U.S. Supreme Court's Incorporation and Interpretation of Precedent*, 36 LAW & SOC'Y REV. 139 (2002); see also Sara C. Benesh & Malia Reddick, *Overruled: An Event History Analysis of Lower Court Reaction to Supreme Court Alteration of Precedent*, 64 J. POL. 534 (2002).

literature is low on systematic studies of *stare decisis*, and usually (and necessarily) focuses on one jurisdiction in depth. Alternatively, one could say that the entire focus of fields of jurisprudential study is the nature and meaning of *stare decisis*. I mention this because in each cited case studied below, an analysis of the case's meaning and what has happened to it, dispositively, after its decision and publication, is relevant to determining the accuracy of evaluation. Considerations of the notion of the rule of law and its relation to having publicized, well-described precedent deserve special mention, as well.²²

¶17 The scholarly and professional literature, then, has not touched directly on the issue at study here but does suggest appropriate methodological approaches.

Methodology

¶18 The story of the dataset used in this study is somewhat tortuous and deserves a clear explanation. In 2013, I worked as a summer research associate with Fastcase, Inc., the online legal research company. As part of my work, I performed a quality assurance analysis on the Bad Law Bot software, testing several features of the automated citator product.²³ My work on Bad Law Bot increased my interest in citators, and Fastcase CEO Ed Walters allowed me to keep a copy of the spreadsheet used in the quality assurance analysis and to use the data in this study.

¶19 Before I began my project at Fastcase, the following data had been collected: the Bad Law Bot algorithm then in existence identified 44,624 cases in the Fastcase database as negative. While the algorithm is proprietary, its construction at the time was quite simple. In essence, the algorithm searched for instances of negative subsequent case treatment using “explanatory phrases” from table 8 of the *Bluebook*.²⁴ Of those cases, 3800 were selected for hand-tagging by Fastcase staff.

¶20 Staff members searched for these cases in the WestlawNext database and recorded the flag color associated with the case from the integrated KeyCite product. KeyCite assigned Red Flags to 3465 of the cases. Of the cases remaining, approximately thirty-two percent (108) were not flagged, and sixty-eight percent (227) were assigned a Yellow Flag.

¶21 WestlawNext user guides explain that the Yellow Flag “warns that the case has some negative history, but hasn't been reversed.”²⁵ This independently conducted study begins with the set of 227 Yellow Flag cases, which, of course, had also been identified as “negative” by the Bad Law Bot algorithm. To study the phrases

22. See, e.g., Joseph Raz, *The Rule of Law and Its Virtue*, in *THE AUTHORITY OF LAW: ESSAYS ON LAW AND MORALITY* 210, 217 (1979), for one view on the importance of publicity and predictability in the law.

23. For a bit more information about how the product works, see *Introducing Bad Law Bot from Fastcase*, <https://www.youtube.com/watch?v=ZsKu7FoO2Ns> (last visited Dec. 7, 2015) [<https://perma.cc/5XK9-QCRN>].

24. *THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION* tbl. 8, at 500 (20th ed. 2015). The phrases, according to the *Bluebook*, are “commonly used to indicate prior or subsequent history and weight of authority of judicial decisions.” *Id.* at 500.

25. *KEYCITE ON WESTLAW NEXT* (2012), <https://info.legalsolutions.thomsonreuters.com/pdf/wln2/L-356347.pdf> (last visited Nov. 18, 2015) [<http://perma.cc/WS2F-C73Q>].

explaining negative subsequent case law treatment, 116 of the 227 Yellow Flag cases were selected for further analysis. These cases were selected because their KeyCite entries consisted only of what Taylor called “unrelated opinions,” those in which “the citing opinion is from a different legal matter but has some effect on the persuasiveness of the cited opinion.”²⁶ KeyCite calls these unrelated opinions “negative citing references” as opposed to negative direct history, which Taylor calls “related opinions . . . , where the citing opinion is a later phase of the same legal matter, such as appeals, rehearings en banc, substituted opinions, or supplemental opinions.”²⁷

¶22 To summarize, then, cases identified in WestlawNext’s KeyCite having direct negative history were excluded from the data examined here. This was done to keep the focus on the descriptions of precedent in later-citing cases. Another way to think of this is to imagine three hypothetical cases: Case A, Case B, and Case C. Case A states a proposition of law. Case B in some way acts on the proposition of law stated in Case A. (Case B, then, is an example of a related opinion.) Finally, in Case C, the action of Case B on Case A is somehow memorialized. Case C is an unrelated opinion, from a different legal matter, but describing the proposition of law stated in Case A.

¶23 These terms are useful ways of talking about the same phenomena. However, since there is no other vocabulary for discussing subsequent negative treatment, a few more terms of art may help in the study that follows. I call the case being updated or “Shepardized” the “Target Case.” The term is analogous to Case A, but feels somewhat less clinical and is more geared to a researcher’s actual needs. Locating the Citing Case, namely a relevant one, is the result of using a citator. We are primarily concerned with a Negative Citing Case, which is analogous to the idea of Case C. Finally, the Acting Case refers to the case commenting on the Target Case. This is the same as Case B. The relationship between Case B and Case A, the Acting Case and the Target Case, is the subsequent negative treatment and is expressed in a Negative Citing Case, or a Case C. Descriptions of subsequent negative treatment, and their accuracy, are the narrow focus of this study.

¶24 For the 116 cases analyzed, seventeen individual pieces of data were collected. Some data were collected to ensure that the proper case was being analyzed—identification data such as the case’s name, date of decision, and one or more citations associated with the case. Other data fields record the subsequent negative treatment in the databases studied. The seventeen data fields are listed in table 1, with brief descriptions of their content.

¶25 All data have been both collected or checked by me personally and, to the best of my knowledge, are current though at least January 2015 (though some data collection continued into February and March). To gather and check the data, I used my law student access to WestlawNext and LexisNexis (or “Lexis.com” linked from the Lexis Advance platform). And while the dataset is hosted in its entirety in .xlsx format in full on my personal website,²⁸ I offer a sample record in table 2 for a particular case to aid in understanding of why these data might be useful.

26. Taylor, *supra* note 9, at 131, ¶ 14.

27. *Id.*

28. Subsequent Negative Treatment Dataset, <http://perma.cc/XJB4-KZ7Z> (last visited Nov. 18, 2015).

Table 1

Descriptions of Data Collected

Data Field	Description of Data Field
Target Case Name	Name of the Target Case in short form
Decision Date	Year, month, and day of decision
Target Case Citations	Citations identifying the case from the West National Reporter System or Federal Reports (F. Supp. F., and S. Ct.), from official reports, or, for more recent unreported opinions, the case name, court of decision, and year (eschewing Westlaw or LexisNexis citations)
Jurisdiction	Federal or State
Flag Color on WLN	For all Target Cases analyzed in this study, the Flag Color on WestlawNext was Yellow
# of Neg. Treatment on WLN	The number of cases identified as “Negative Citing References” on WestlawNext’s KeyCite
WLN Neg. Treatment	The Citing Case identified as “Most Negative” among the Negative Citing References
Symbol on LEX	The graphical Shepard’s “signal” for the Target Case
# of Neg. Treatment on LEX	After selecting “All Neg” in the Shepard’s report, the number of “Citing Decisions” listed on the report
LEX Neg. Treatment	Identified by selecting the most negative Citing Case within the Shepard’s report
# of Neg. Treatment on FC	In the Fastcase Authority Check report, the number of cases flagged by the Bad Law Bot algorithm
FC Neg. Treatment	The top result in the list of negative Citing Cases identified by Bad Law Bot
FC Acting Case	The case identified by Bad Law Bot as the case acting on the Target Case
Classification on WLN	The phrase used in WestlawNext to describe the subsequent negative treatment in the “Most Negative” Citing Case
Classification on LEX	The phrase used in the Shepard’s report to describe the subsequent negative treatment in the most negative Citing Case
Classification on FC	The phrase describing the relationship between the Target Case and the Acting Case, as identified by Bad Law Bot and generally drawn from table 8 of the <i>Bluebook</i>
Quote from FC BLB	A full quote of the sentence that contains the description of the relationship between the Target Case and Acting Case on Fastcase

¶26 For the Target Case in table 2 we can reconstruct a picture in textual form from these data. The Target Case was decided in federal court in 1994. KeyCite applies the Yellow Flag, asserting that the most negative Citing Case “declined to follow” the Target Case. Shepard’s applies the Yellow Triangle symbol and asserts that most negative Citing Case “criticized” Target Case. The most negative Citing Case in both products was the same (787 F. Supp. 2d 961). The Bad Law Bot algorithm, however, identified that the Target Case was “overruled on other grounds” by the Acting Case, and this fact was recorded in a Citing Case different from the one selected by KeyCite and Shepard’s. Numerous questions possibly flow from these

Table 2

Sample of Data Collected for One Case

Data Field	Data Collected
Target Case Name	Mulhall v. Advance Sec., Inc.
Decision Date	4/22/1994
Target Case Citations	19 F3d 586 (note the absence of punctuation and spacing; entering this citation format into the databases tested works well)
Jurisdiction	Federal
Flag Color on WLN	Yellow
# of Neg. Treatment on WLN	14
WLN Neg. Treatment	787 F.Supp.2d 961
Symbol on Lex	Yellow Triangle
# of Neg. Treatment on Lex	8
Lex Neg. Treatment	787 F. Supp. 2d 961
# of Neg. Treatment on FC	2
FC Neg. Treatment	767 F.Supp.2d 1233
FC Acting Case	338 F.3d 1304
Classification on WLN	Declined to Follow by
Classification on Lex	Criticized by
Classification on FC	overruled on other grounds
Quote from FC BLB	"(citing Mulhall v. Advance Sec., Inc., 19 F.3d 586, 589 n.8 (11th Cir. 1994), overruled on other grounds, Manders v. Lee, 338 F.3d 1304, 1328 n.52 (11th Cir. 2003))"

data. The one we are concerned with, however, is: which phrase, if any, is an accurate description of subsequent negative treatment for the Target Case?

¶27 To analyze descriptions of subsequent negative treatment, Target Cases having matching Citing Cases were collected, read, and assessed.²⁹ Absent other methods of conducting this type of analysis, this seemed most likely to yield useful, practical results for lawyers and legal information professionals. If the most negative Citing Cases match on at least two of the citators tested, it can be assumed that users should presume these cases present the same subsequent negative treatment. The question here is whether a hierarchical controlled vocabulary term, like those employed by Shepard's, is a more accurate description of subsequent negative treatment than the uncontrolled systems used by KeyCite.³⁰

29. Note that the example from table 2 meets this criterion and was subject to the analysis described below.

30. *List of Analysis Definitions Grouped by Shepard's Signals*, LEXISNEXIS (available to LexisNexis subscribers at <https://www.lexisadvance.com/>; archived at <http://perma.cc/XK4R-U6WW>). The comparable WestlawNext document for KeyCite does not mention the explanatory phrases, let alone describe or classify them. See CHECKING CITATIONS IN KEYCITE (2010), <http://blog.legalsolutions.com/>

Results

¶28 In the 116-case dataset, 31 Target Cases had matching “most negative” Citing Cases in Shepard’s and KeyCite. That is, twenty-six percent of the time the two market-leading citator products agreed about which case best provided subsequent negative treatment. While this may seem surprising, I do not think the result is particularly meaningful as a measure of the reliability of either product or both products. First, “related cases” assigned direct or procedural negative history were excluded—nearly half of the original 227 Yellow Flag cases. I would hypothesize that there is greater alignment among related cases, though this question will not be considered further in this article. However, the seventy-four percent rate of disagreement between citators is taken up in its proper context in the discussion section.

¶29 What is more troubling to the analysis here is the “most negative” designation as a distinctive, meaningful metric. Only KeyCite explicitly identifies a most negative case for a given Target Case. Shepard’s does not make an explicit determination of which case in its report is most negative. That said, the decision of which case was chosen as most negative was far from arbitrary. The Shepard’s system is a hierarchy from most negative to least negative (Red Hexagon/Q in Orange Box/ Yellow Triangle), with each analysis definition falling under a particular visual signal.³¹ On the Shepard’s Report page, a box summarizes this ordering near the top of the page. The most negative case was selected by locating the first result in the most negative set of Citing Cases. Still, I do not expect a practitioner looks only at the case in a citator listed as most negative for the Target Case she is researching. As noted earlier, a user study of attorneys and other expert legal researchers would be needed to learn *how* these people use citators. We only know, of course, that researchers *do* use them.

¶30 These challenges aside, the thirty-one cases with identical most negative Citing Cases can tell us quite a bit about the language used to describe subsequent negative treatment in Shepard’s and KeyCite. First of all, twelve of the thirty-one cases, or thirty-nine percent, used identical language to describe subsequent negative treatment. Interestingly, these twelve instances all involved the same explanatory phrase: “distinguished by.” Two additional cases were assigned near-identical language: “disapproved by” (Shepard’s) and “disapproved of by” (KeyCite). That leaves seventeen cases in which different terms were used to describe subsequent negative treatment for a Target Case. This group of seventeen cases was analyzed for differences between each citator’s accuracy in applying subsequent negative treatment phrases.

¶31 Four of the Shepard’s phrases are present in this final collection of cases: “criticized by,” “distinguished by,” “overruled in part by,” and “questioned by.” Shepard’s defines these phrases as:

thomsonreuters.com/wp-content/uploads/2010/08/wlnkeyci.pdf [http://perma.cc/E634-NNSX]. Bad Law Bot, of course, does not generate its own explanatory phrases, but rather captures the ones used by courts in Citing Cases.

31. *List of Analysis Definitions Grouped by Shepard’s Signals*, *supra* note 30.

Criticized by—The citing opinion disagrees with the reasoning/result of the case you are *Shepardizing*, although the citing court may not have the authority to materially affect its precedential value.

Distinguished by—The Citing Case differs from the case you are *Shepardizing*, either involving dissimilar facts or requiring a different application of the law.

Overruled in part by—One or more parts of the decision you are *Shepardizing* have been expressly nullified by the subsequent decision from the same court, thus casting some doubt on the precedential value of the case you are *Shepardizing*.

Questioned by—The citing opinion questions the continuing validity or precedential value of the case you are *Shepardizing* because of intervening circumstances, including judicial or legislative overruling.³²

¶32 According to Shepard's, the "criticized by" and "distinguished by" phrases should be assigned the Yellow Triangle signal, meaning "caution."³³ "Overruled in part by" should be assigned the Red Hexagon signal, meaning "warning."³⁴ "Questioned by" should be assigned the Q in Orange Box signal, meaning "questioned."³⁵ "Criticized by," "distinguished by," and "questioned by" are considered to be "common analysis phrases."³⁶ Thus it can be inferred that, as common phrases, these would likely be encountered by a researcher while he conducts legal research.

¶33 Seven KeyCite analytical phrases are used in describing the subsequent negative treatment in the final seventeen cases studied: "called into doubt by," "criticized by," "declined to extend by," "declined to follow by," "disagreed with by," "not followed as dicta," and "rejected by." As mentioned earlier, WestlawNext does not define these phrases. The phrases, in the large dataset, appear in combination with either a Red Flag or Yellow Flag, the only two symbols used by KeyCite. These facts support two propositions. First, that KeyCite's system of applying analytical phrases to cases is not an example of a hierarchical controlled vocabulary. Second, it will be more difficult to assess whether the phrases have been applied to cases accurately because the definitions of the phrases are necessarily somewhat subjective.³⁷ Words such as "disagreed"³⁸ and "doubt"³⁹ suggest common definitions while "dicta" is a legal term of art.⁴⁰ "Followed,"⁴¹ "extend,"⁴² and "rejected"⁴³ seem to be

32. *Id.*

33. LEXISNEXIS, SHEPARD'S SIGNAL INDICATORS AND ANALYSIS PHRASES AT LEXIS.COM (2012), http://www.lexisnexis.com/documents/pdf/20120605102055_large.pdf [<http://perma.cc/YC7-NCLM>]. "Caution: Possible negative treatment indicated" is the full definition of the graphical signal.

34. *Id.* "Warning: Negative treatment is indicated" is the full definition of the graphical signal.

35. *Id.* "Questioned: Validity questioned by citing refs." is the full definition of the graphical signal.

36. *Id.*

37. I consulted two dictionaries in crafting the definitions that follow in this section: THE AMERICAN HERITAGE DICTIONARY OF THE ENGLISH LANGUAGE (2011); and BLACK'S LAW DICTIONARY (Bryan A. Garner ed., 10th ed. 2014).

38. As defined by BLACK'S LAW DICTIONARY, *supra* note 37, at 560, "disagreement" means "[a] difference of opinion; a lack of agreement," which is indeed close to the common use of the term.

39. "Reasonable doubt" is the only technical use of "doubt" in law I am familiar with; it has nothing to do a judge's determination, but rather a jury's.

40. "Dicta" is the colloquial plural form of "obiter dictum," which means "[a] judicial comment made while delivering a judicial opinion, but one that is unnecessary to the decision in the case and therefore not precedential (although it may be considered persuasive)." BLACK'S LAW DICTIONARY, *supra* note 37, at 1240.

41. "Follow," as defined by *Black's Law Dictionary*, means "[t]o conform to or comply with; to accept as authority." *Id.* at 759.

42. The closest definition of "extend" provided by *The American Heritage Dictionary of the English Language* is "to expand the influence of." AMERICAN HERITAGE DICTIONARY, *supra* note 37, at 626.

43. "Rejection" has technical legal meaning in the law of contracts, but I am not aware of a clear

hybrids of words used commonly and as legal terms of art; they are therefore the most difficult to assess. I define the seven KeyCite analytical phrases as follows:

1. **Called into doubt by**—The Citing Case expresses uncertainty about the precedential value of the Target Case.
2. **Criticized by**—The Citing Case disagrees with the reasoning or result of the Target Case, although the citing court may not have the authority to materially affect its precedential value.⁴⁴
3. **Declined to extend by**—The judge in the Citing Case has chosen to not increase the influence of the Target Case.
4. **Declined to follow by**—The judge in the Citing Case has chosen not to comply, conform with, or accept the Target Case as authoritative.
5. **Disagreed with by**—The Citing Case expresses a difference of opinion or lack of agreement with the Target Case.
6. **Not followed as dicta**—The court in the Citing Case will not accept the Target Case as authoritative because statements in the Target Case are considered to be unnecessary to the decision in the case and therefore not precedential.
7. **Rejected by**—The Citing Case declines to make use of reasoning from the Target Case.

¶34 With these definitions in place, we may finally assess how accurately the citators apply their descriptions of subsequent negative treatment. As described in the methodology, determining which system is more accurate is also necessarily subjective—a matter of reading the relevant portions of each case and determining whether one explanatory phrase (or both, or none) more clearly describes the nature of the judicial determination in the Citing Case as it relates to the Target Case. While this seems difficult in theory, there were relatively few challenging decisions to make when reading and classifying a statement of subsequent negative treatment.

¶35 Of the seventeen Target Cases (shown in table 3) with matching most negative Citing Cases, five had overlapping descriptions of subsequent negative treatment. This occurred when the explanatory phrase was equally accurate in describing the subsequent negative treatment. For example, for the three Target Cases labeled as “criticized by” in Shepard’s, the KeyCite phrases “rejected by,” “declined to follow by,” and “disagreed with by,” as defined above, were for all purposes identical. *United States v. Dorfman*, 542 F. Supp. 345, a Target Case, is described as “criticized by” in Shepard’s and as “rejected by” in KeyCite. The Citing Case comments:

The Court rejects the suggestion in *United States v. Dorfman*, 542 F.Supp. 345, 373 (N.D.Ill.), *aff’d, sub nom. United States v. Williams*, 737 F.2d 594 (7th Cir.1984), relied on by the government, that “[t]he test . . . is not the state of the mind of ‘the government,’ but the state of mind of the affiant.” This Court treats the affiant as a representative of the United States government as to the activities of any of its agents, whether or not the affiant has been made aware of those activities.⁴⁵

technical decision in terms of precedent. The common meaning from *The American Heritage Dictionary of the English Language* is “refuse[d] . . . to make use of.” *Id.* at 1482.

44. This definition is in effect identical to the Shepard’s definition.

45. Application for Interception of Wire Commc’ns, 2 F. Supp. 2d 177, 179 n.3 (D. Mass. 1998).

¶36 The Citing Case, from a federal district court in the First Circuit, chooses not to apply a particular doctrinal rule used in the Target Case, from a federal district court in the Seventh Circuit. It fits the definition of “criticized by” and the definition of “rejected by,” presented above. Therefore, the description of the subsequent negative treatment has been accurately applied.

Table 3

Seventeen Matching “Most Negative” Cases

Target Case Name	Target Case Citation	Flag Color on WLN	Symbol on LEX	Classification on WLN	Classification on LEX
Gregory v. United States	369 F.2d 185	Yellow	Q in Orange Box	Criticized by	Questioned by
McCurdy v. Steele	353 F. Supp. 629	Yellow	Red Hexagon	Declined to Follow by	Questioned by
United States v. DeCoster	487 F.2d 1197	Yellow	Red Hexagon	Called into Doubt by	Questioned by
Brandenburger v. Thompson	494 F.2d 885	Yellow	Q in Orange Box	Disagreed With by	Questioned by
Finley v. United States	404 F. Supp. 200	Yellow	Q in Orange Box	Disagreed With by	Questioned by
United States v. Dorfman	542 F. Supp. 345	Yellow	Red Hexagon	Rejected by	Criticized by
Melson v. Kroger Co.	578 F. Supp. 691	Yellow	Q in Orange Box	Disagreed With by	Questioned by
People v. Johnson	204 Cal. Rptr. 563	Yellow	Yellow Triangle	Not Followed as Dicta	Criticized by
In re A & C Properties	784 F.2d 1377	Yellow	Red Hexagon	Declined to Follow by	Questioned by
Barber v. National Bank of Alaska	815 P.2d 857	Yellow	Q in Orange Box	Called into Doubt by	Questioned by
Stafford v. Purofied Down Products Corp.	801 F. Supp. 130	Yellow	Red Hexagon	Declined to Follow by	Questioned by
Mulhall v. Advance Sec., Inc.	19 F.3d 586	Yellow	Yellow Triangle	Declined to Follow by	Criticized by
Doctor’s Associates, Inc. v. Distajo	66 F.3d 438	Yellow	Red Hexagon	Disagreed With by	Criticized by
Atchison, Topeka and Santa Fe Ry. Co. v. Brown & Bryant, Inc.	159 F.3d 358	Yellow	Red Hexagon	Not Followed as Dicta	Criticized by
Reed v. Reed	48 S.W.3d 634	No Flag	Red Hexagon	Declined to Follow by	Overruled in part by
Security Ins. Co. v. Old Dominion Freight Line	314 F. Supp. 2d 201	Yellow	Yellow Triangle	Declined to Extend by	Distinguished by
Pequignot v. Solo Cup Co.	640 F. Supp. 2d 714	Yellow	Red Hexagon	Declined to Follow by	Criticized by

¶37 But what about an instance from this group in which a descriptive phrase has been inaccurately applied? *Brandenburger v. Thompson*, 494 F.2d 885, a Target Case, is described in Shepard's as "questioned by" but as "disagreed with by" in KeyCite. Only the phrase from Shepard's accurately describes the negative precedent here:

A third exception, known as the private attorney general theory, has been recognized by several lower courts when the expense of litigation may act as a deterrent to the bringing of private litigation deemed necessary to enforce important public policies. See, e.g., *Taylor v. Perini*, 503 F.2d 899 (6th Cir. 1974); *Fowler v. Schwarzwald*, 498 F.2d 143 (8th Cir. 1974); *Brandenburger v. Thompson*, 494 F.2d 885 (9th Cir. 1974); *Knight v. Auciello*, 453 F.2d 852 (1st Cir. 1972); *Lee v. Southern Home Sites Corp.*, 444 F.2d 143 (5th Cir. 1971). In *Alyeska Pipeline* the Supreme Court expressly disapproved the use of the private attorney general exception in federal courts.⁴⁶

The Target Case, we can see, is included by the Citing Case as part of a string citation of cases that have been implicitly overruled by a Supreme Court case. This kind of treatment is accurately described by Shepard's "questioned by" phrase, but does not match at all with what "disagreed with by," the KeyCite term, seems plainly to mean.

¶38 In the set of cases studied, eleven of seventeen, or 64.7% of descriptions of subsequent negative treatment in Shepard's were applied accurately as shown in table 4. The KeyCite system applied descriptions of subsequent negative treatment accurately in nine of seventeen, or 52.9% of cases. Neither citator applied descriptions of subsequent negative treatment accurately in two of the seventeen cases, or 11.8% of the time for this set as shown in table 5. While the sample size studied is small and no conclusions can be reached about the ultimate comparative quality of this aspect of either citator, these data do suggest that each system describes the actions taken in identical Citing Cases with differing degrees of consistency.

¶39 The clear definitions in the hierarchical Shepard's system allowed for further comparison of the data from the comparative set described above to a Shepard's-only set. A systematic sample from the remaining 85 cases in the set of 116 described above was taken. Alternating fourth or fifth cases were selected from the table sorted chronologically to help mitigate any bias from the date of a case. In total, 19 cases were selected from the 85 and subjected to the same analysis as the set of 17 above. In 14 of the 19 cases in the Shepard's-only set, subsequent negative treatment from a Citing Case was accurately described by the analytical phrase. That is, 73.6% of the time, Shepard's, according to its own standards, described the subsequent negative treatment accurately. A similar test of the KeyCite system for the same set of cases was not performed since KeyCite gives no clear definitions for its explanatory phrases.

¶40 These results, it can be argued, offer no definitive conclusions of any kind about how accurately a given citator describes subsequent negative treatment in case law. This is an argument I ultimately accept—a truly systematic study of the topic is likely impossible given the number of variables that could affect its results. For instance, factors such as chronology, the presence of direct procedural history (Taylor's "related opinions"), consistent application of the phrases by LexisNexis or

46. *Clemons v. Runck*, 402 F. Supp. 863, 870 (S.D. Ohio 1975).

Thomson Reuters, and the case identified as “most negative” could each present different findings. Each variable might be controlled for on its own, but controlling for all at once seems exceedingly difficult. However, I do believe that by articulating a methodology for studying citators and offering a detailed picture of interpreting and operationalizing accuracy of application of subsequent negative treatment is valuable in its own right.

Table 4

Accurate Classification in Either or Both Citation Indexes

Target Case Name	Target Case Citation	Classification on WLN	Classification on LEX	Accurately Classified Within
Gregory v. United States	369 F.2d 185	Criticized by	Questioned by	KeyCite Only
McCurdy v. Steele	353 F. Supp. 629	Declined to Follow by	Questioned by	Shepard's Only
Brandenburger v. Thompson	494 F.2d 885	Disagreed With by	Questioned by	Shepard's Only
Finley v. United States	404 F. Supp. 200	Disagreed With by	Questioned by	Shepard's Only
United States v. Dorfman	542 F. Supp. 345	Rejected by	Criticized by	Both
People v. Johnson	204 Cal. Rptr. 563	Not Followed as Dicta	Criticized by	KeyCite Only
In re A & C Properties	784 F.2d 1377	Declined to Follow by	Questioned by	Both
Barber v. National Bank of Alaska	815 P.2d 857	Called into Doubt by	Questioned by	KeyCite Only
Stafford v. Purofied Down Products Corp.	801 F. Supp. 130	Declined to Follow by	Questioned by	KeyCite Only
Mulhall v. Advance Sec., Inc.	19 F.3d 586	Declined to Follow by	Criticized by	Both
Doctor's Associates, Inc. v. Distajo	66 F.3d 438	Disagreed With by	Criticized by	Both
Atchison, Topeka and Santa Fe Ry. Co. v. Brown & Bryant, Inc.	159 F.3d 358	Not Followed as Dicta	Criticized by	Shepard's Only
Reed v. Reed	48 S.W.3d 634	Declined to Follow by	Overruled in part by	Shepard's Only
Security Ins. Co. v. Old Dominion Freight Line	314 F. Supp. 2d 201	Declined to Extend by	Distinguished by	Both
Pequignot v. Solo Cup Co.	640 F. Supp. 2d 714	Declined to Follow by	Criticized by	Shepard's Only

Table 5

Inaccurate Classification in Both Citation Indexes

Target Case Name	Target Case Citation	Classification on WLN	Classification on LEX
United States v. DeCoster	487 F.2d 1197	Called into Doubt by	Questioned by
Melson v. Kroger Co.	578 F. Supp. 691	Disagreed With by	Questioned by

¶41 There are other indicia of accuracy in application descriptions of subsequent negative treatment: namely, the consistency with which cases are tagged with a citator's iconography. In the set of 116 cases, KeyCite assigns the Yellow Flag to 18 different explanatory phrases. The 18 explanatory phrases are not evenly distributed—two phrases, “declined to follow by” and “distinguished by,” account for 51 of the 116 cases. In the larger set of 227 Yellow Flag cases (containing related opinions and unrelated opinions), KeyCite assigns 57 different explanatory phrases.

¶42 Shepard's iconography was studied only in the set of 116 unrelated opinions. Seven of the 116 assigned no subsequent negative treatment and are excluded from this part of the analysis. Of the remaining 109 cases, 37 were assigned the Yellow Triangle, 20 were assigned the Q in Orange Box, and 52 were assigned the Red Hexagon. Of the 37 cases assigned the Yellow Triangle, four explanatory phrases were applied, with two phrases, “criticized by” and “distinguished by,” accounting for 35 of the 37 cases. Of the 20 cases assigned the Q in Orange Box, all 20 were given the “questioned by” analytical phrase. Finally, of the 52 Red Hexagon cases, Shepard's assigned fifteen analytical phrases. Some of the Red Hexagon cases included related opinions in the Shepard's report that were not captured by KeyCite, so for a meaningful comparison, it makes sense to look at the Yellow Triangle and Q in Orange Box cases. For the 57 such cases, only five phrases are applied, and three of the five phrases account for 55 of the 57, or 96.5%. Looking at this metric, the value of a controlled visual vocabulary is apparent when considering the accuracy of application of descriptions, albeit visual descriptions, of subsequent negative treatment. It seems fair to conclude that a user encountering the Shepard's Yellow Triangle and Q in Orange Box iconography for a Target Case would have a clear understanding of the type of subsequent negative treatment contained in the Citing Case. The user encountering a Yellow Flag on KeyCite—at least on a superficial level—would not likely know what meaning to make of the symbol.

Looking Forward: A Discussion of the Citator's Place

¶43 Questions about the use of online citators—particularly their presentation of graphical symbols as units of meaning—are not new among teachers of legal research.⁴⁷ A survey of the literature suggests that qualifications about the adequacy of these symbols form an integral part of lessons and texts on updating cases and statutes.⁴⁸ Even the most basic legal research text cautions that citators' “signals and editorial signposts are not authoritative statements of the law” and that there is “no substitute for reading a citing document and determining for yourself its scope and effect.”⁴⁹ These same texts, however, note that citators are “helpful as . . . starting place[s]”⁵⁰ in determining a case's continuing validity and should be considered “invaluable resources . . . [ensuring] that cases are still good law.”⁵¹

¶44 Marketing materials from LexisNexis and WestlawNext paint their citators as practically infallible. “Why do so many judges and lawyers rely on Shepard's

47. Kent C. Olson, *Waiving a Red Flag: Teaching Counterintuitiveness in Citator Use*, 9 PERSP.: TEACHING LEGAL RES. & WRITING 58, 58 (2001).

48. See, e.g., ARMSTRONG & KNOTT, *supra* note 3, at 135.

49. COHEN & OLSON, *supra* note 3, at 115.

50. ARMSTRONG & KNOTT, *supra* note 3, at 135.

51. COHEN & OLSON, *supra* note 3, at 109.

Citations Service, exclusively from LexisNexis®, when it comes to validating research?” The answer? “It’s a matter of trust.”⁵² Likewise, KeyCite materials tout “West’s 125-year tradition of editorial excellence” and the company’s “leading-edge technological expertise” in suggesting that users can “instantly verify whether a case . . . is good law.”⁵³ Of course, as Wolf and Wishart point out, these statements are largely meaningless, if not misleading, marketing attempts to differentiate products.⁵⁴ It can be tempting to believe these kinds of statements based on the credibility of the source,⁵⁵ not to mention the cost of the product.⁵⁶ But savvy legal researchers know—and legal research instructors profess—that there is no side-stepping the need to read a case and use professional judgment to determine the precedential effect of later-citing cases. Still, citators are seen as essential components of legal research databases.⁵⁷

¶45 In light of this study’s results, researchers have even greater reason to doubt the adequacy of citators’ determination. More troubling, perhaps, is the implicit boost that some instructional texts give to citators by identifying them as a “good starting place” or “invaluable resources.” Indeed, if the graphical signals and the explanatory phrases cannot be trusted to be applied accurately—at best, approximately seventy-five percent of the time, or at worst, less than fifty percent of the time—it is difficult to see how descriptions of subsequent negative treatment would reliably set the researcher down the correct path. It could be argued that the presence of *any* negative symbol or phrase is enough to alert the user that “here there be dragons”⁵⁸ and therefore, a concurrent need to read and understand the “flagged” case before citing it.⁵⁹ This argument is easily countered, however, by the fact that not reading a case before citing it is almost certainly a violation of Rule 1.1 of the ABA *Model Rules of Professional Conduct*, requiring a lawyer to provide “competent representation to a client[,]” a duty defined as consisting of, among other things, “thoroughness and preparation.”⁶⁰ So, if a tap on the shoulder from a

52. CITATION SERVICES: DEEP ANALYSIS AND UNIQUE PRODUCT DETAILS LEAD RESEARCHERS TO SHEPARD’S CITATIONS SERVICE AT LEXIS.COM (2012), <https://www.lexisnexis.com/pdf/Shep%20and%20KCite%20Comp%20Paper.pdf> [http://perma.cc/G4JV-5G2T].

53. KEYCITE ON WESTLAW NEXT, *supra* note 25.

54. Wolf & Wishart, *supra* note 15, at 27.

55. Bob Berring calls this the “Tinkerbell” effect. He writes: “In the Walt Disney animated feature ‘Peter Pan,’ Tinkerbell was a fairy. She only existed if children believed in her existence. This character, viewed by the author at an impressionable age, stands for the classic bootstrapping of authority-iveness.” Robert C. Berring, *Chaos, Cyberspace and Tradition: Legal Information Transmogrified*, 12 BERKELEY TECH. L.J. 189, 193 n.17 (1997). The term was later used by Mary Whisner. Mary Whisner, *Bouvier’s, Black’s, and Tinkerbell*, 92 LAW LIBR. J. 99, 2000 LAW LIBR. J. 8. The extension of the idea to citators specifically is my choice alone.

56. Hilke Plassmann et al., *Marketing Actions Can Modulate Neural Representations of Experienced Pleasantness*, 105 PROC. NAT’L ACADEM. SCI. U.S.A. 1050, 1051 (2008).

57. Greg Lambert, *Casemaker Unique Among Legal Research Providers*, 89 MICH. B.J. 54, 56 (2010).

58. This phrase comes to mind because of an article on a slightly different topic in legal research. See Peggy Roebuck Jarrett & Mary Whisner, “Here There Be Dragons”: *How to Do Research in an Area You Know Nothing About*, 6 PERSP.: TEACHING LEGAL RES. & WRITING 74 (1998).

59. More appropriately, the relevant portion of the case. See Robert C. Berring, *Unprecedented Precedent: Ruminations on the Meaning of It All*, 5 GREEN BAG 2d 245 (2002).

60. AM. BAR ASS’N, MODEL RULES OF PROFESSIONAL CONDUCT 11 (2011). This is assumed from the work done in Bast & Harrell, *supra* note 5. Likewise, there are no cases specifically on point in Kristina L. Niedringhaus, *Ethics Considerations Related to Legal Research Practices: A Selective Annotated Bibliography*, 31 LEGAL REFERENCE SERVS. Q. 104 (2012).

citator, even without any further intelligible explanation, means “read the case,” the question remains: where’s the value in that? A competent researcher will expect to do this anyways.

¶46 To justify their continued use, then, citators must add some value to the research process, not because they are thought of as “invaluable” but because they are precisely the opposite—important components of rather costly legal research databases. To explore this problem, the discussion first considers the citator’s origin and then how citators as currently constituted can be used beneficially (or, one could say, profitably) by researchers. Next I examine how citators might fit in a legal research environment with a multiplicity of tools for determining the meaning of subsequent negative treatment. Finally, I look to fruitful areas for further research on evaluating the quality of citators.

¶47 The roots of legal citation indexes in the United States stretch to the early nineteenth century, before they were known as citators.⁶¹ Ogden credits preeminent jurist and scholar Joseph Story with the idea of applying an “explanatory letter” such as D for doubted or O for overruled to a table of cases included in a reporter volume.⁶² The earliest popular citation index existed for much the same purpose as the contemporary citators—efficient identification of a case’s subsequent negative treatment. The title of the volumes produced over a lifetime by Simon Greenleaf, *Greenleaf’s Overruled Cases*, explains itself.⁶³ Other early citation indexes were also compiled by individuals, almost always at first for their creator’s personal use.⁶⁴ But the coming of the twentieth century brought a growth in the country’s urban population and commercial activity, which led to an increase in the number of courts (and reorganization of those courts), and eventually to a consequent explosion of the number of published precedential legal documents, such as appellate opinions.⁶⁵ In addition to figuring out how to organize these cases, interested parties in the profession, namely lawyers, law librarians, and the growing national publishers, needed to find a way to make sense of this ever-expanding stuff of law.⁶⁶

¶48 Until the digital era in legal publishing, and through its first generation, Shepard’s was the only national citator capable of accomplishing this comprehensive task.⁶⁷ While not originally conceived as a facsimile of the print version, Shepard’s online soon took on many of the functions, and idiosyncrasies, of the print volumes.⁶⁸ But its purpose remained the same—identifying precedential procedural and citing history in a comprehensive and current fashion. What digital publication of legal materials added, however, was the ability to generate a so-called table of authority for a given document, such as a case. That is, it was suddenly possible to extract information about which cases cited other cases and how often

61. Ogden, *supra* note 6, is the canonical history.

62. *Id.* at 4.

63. *Id.* at 6. Greenleaf was a close friend and colleague of Justice Story.

64. *Id.* at 23.

65. For an interesting, summary discussion of this problem, see Richard A. Danner, *The ABA, the AALL, the AALS, and the “Duplication of Legal Publications,”* 104 LAW LIBR. J. 485, 488, 2012 LAW LIBR. J. 35, ¶ 6.

66. This is an analogy to one of the main theses from BOWKER & STAR, *supra* note 7, at 16. In that work, the context is epidemiological classification.

67. Dabney, *supra* note 6, contains an excellent, detailed history of the development of KeyCite in addition to a history of Shepard’s in the late twentieth century.

68. *Id.* at 170.

a given case was cited, among other things. KeyCite includes graphical representations of a case's direct procedural history, while both services include extensive lists of Citing Cases, regardless of whether the treatments within are approving, negative, or neutral.

¶49 Given the change in the nature of citation analysis possible in the digital realm and the shaky performances of both systems in accurately describing substantive negative history, it is fair to ask whether the era of the citator *qua* citator has already passed. Instead of conceptualizing these indexes of citations as tools reliably identifying both the existence and nature of subsequent negative history, and therefore essential steps in "updating" legal research, it seems more appropriate to relieve these tools of the moniker they have acquired and the related heavy, if not impossible, burden associated with it. A legal citation index cannot, on its own, say whether a Target Case is bad law. It certainly cannot say if a case is good law. And, as discussed earlier, any legal researcher relying on a citator to do so is almost certainly professionally incompetent. This raises the question, hinted at earlier, of what legal citation indexes are still good for. To put it more circumspectly, what is the tool-known-as-citator's place in legal research now and in the future?

¶50 Currently, legal citation indexes can gather a wide range of subsequent citing authority. KeyCite is more explicit about displaying these sources—including statutes citing the Target Case as well as secondary sources like law reviews, treatises, or looseleaf series. The same variety of sources is also available in Shepard's, though the interface makes it a bit more difficult to find. Both products include a table of authorities cited within the Target Case. KeyCite also includes a unique feature, measuring the depth of treatment in citing court documents. This measure is likely calculated based on how often the Target Case is cited or mentioned in the citing document, and how many words separate the citation from the next different citation. Thomson Reuters (the parent company of WestlawNext) does not disclose how its depth-of-treatment algorithm functions, naturally, so the prior statement represents my best guess. Neither system provides other types of symbols or explanatory phrases to describe other later-citing authorities.

¶51 Legal research instructors need to emphasize these features of citation analysis currently available in both major citators while explaining potential problems with the descriptions of subsequent negative treatment in the later-citing cases when provided. Also, the notion of a citator as a tool for updating research must be expanded—the citation index, when integrated into the larger research platform, can serve a vital function throughout the research process. Examining the dates of later-citing cases (are there any that are current?) or the jurisdiction or court level of later-citing cases (will the decisions be binding precedent?) are two basic examples. Rethinking the citator as a tool for analyzing the influence of a case based on later citations in a variety of sources is needed, rather than calling it a final box to tick to ensure the validity of a case's proposition as good law.

¶52 Lawyers and legal information professionals also must continue to critique and qualify the marketing or training materials created by WestlawNext and LexisNexis about their respective citator products. As the literature cited in this article shows, much good work has already been done in this and related areas, and should continue. Law students and practicing attorneys need to be made aware of the

implicit overruling problem,⁶⁹ the inclusion or exclusion of unpublished opinions,⁷⁰ the misleading nature of certain graphical representations of precedent,⁷¹ and of course, the inaccuracy in how symbols and explanatory phrases are used to signal subsequent negative treatment. Additionally, it cannot be left to the information vendors to teach how these products work,⁷² especially if citators are being presented as tools for instantly confirming or disconfirming the validity of a point of law in a given Target Case.

¶53 Beyond teaching what citators can and should do well presently, looking to the future is necessary too. There, new tools will assume some of the traditional citator's functions, and market needs and competition will push WestlawNext and LexisNexis to reimagine their current products. Startup companies already are producing new-wave citation analysis tools based on algorithmic extraction of citation data, presenting information about case citation networks visually.⁷³ Researchers no longer need to rely solely on the textual representation of citation data when analyzing precedent.⁷⁴ Also intriguing is the possible application of machine-learning algorithms or artificial intelligence to legal systems. And while two recent articles on the topic do not mention the application to citators specifically, the authors do stress that teasing out the relationship between cases is a potential application of this technology.⁷⁵ While a fully automated citator with accurate descriptions of subsequent negative treatment may be far in the offing, more options for conducting citation analysis outside of the bounds of traditional citators have already arrived.

¶54 One hope for this article is that it will encourage more commentary on citators generally and on their descriptions of subsequent negative treatment specifically. The analysis here does not satisfactorily answer several questions it raises. For instance, does a hierarchical controlled vocabulary improve the accuracy in how citators apply their descriptions of subsequent negative treatment? Could the observed results instead be due to differences in how consistently individual services apply these descriptions? Or does the absence of any definitions of explanatory phrases in KeyCite make the idea of consistent application impossible? On the contrary, might the KeyCite terms actually represent narrower terms of broader

69. Wolf & Wishart, *supra* note 15, at 31.

70. Mart, *supra* note 12, at 44, n.3.

71. Olson, *supra* note 47, at 60 (arguing that the red hexagon should be thought to mean “go” rather than “stop”).

72. This, as it applies to law students, is not a new argument. See Shawn Nevers, *Candy, Points, and Highlighters: Why Librarians, Not Vendors, Should Teach CALR to First-Year Students*, 99 LAW LIBR. J. 757, 2007 LAW LIBR. J. 46.

73. The best example of this is Ravel Law. See *About Us*, RAVEL LAW, <https://www.ravellaw.com/about> (last visited Nov. 17, 2015) [<http://perma.cc/BTF5-884H>]. Fastcase also makes use of visual representations of precedent in its Authority Check tool.

74. Susan Azyndar, Katrina Lee & Ingrid Mattson, *A New Era: Integrating Today's “Next Gen” Research Tools Ravel and Casetext in the Law School Classroom*, 41 RUTGERS COMPUTER & TECH. L.J. 31, 69 (2015).

75. See Daniel Martin Katz, *Quantitative Legal Prediction—Or—How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*, 62 EMORY L.J. 909, 955 (2013); Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 109 (2014).

Shepard's analogues, which then form a kind of implicit hierarchical controlled vocabulary? Or do the phrases exist simply as snippets, used to distinguish among cases without actually categorizing the subsequent negative treatment contained within?

¶55 Other areas for research that I see as particularly helpful are exploring further what the idea of comprehensiveness in a citator means. To that end, researchers might estimate how many cases present in the case law databases have ever been cited by another case, and, if a case has been cited, what is the mean number of times it has been cited? Of this number, how many of those citing references would be negative? Knowing how to best ask and answer these types of quantitative questions would likely be of aid to researchers looking to benchmark observed qualitative phenomena. Would the necessarily subjective descriptions of negative subsequent precedent make sense to duplicate if one were to start building a citator today or, as I suspect, would the making of meaning take some other, more inferential path? That is, what would the modern-day Story emphasize as important to the modern-day Greenleaf?

¶56 More prosaically, it should be noted that Shepard's has a different iconography and user interface on the new Lexis Advance platform, which was not tested here. (KeyCite on WestlawNext was chosen because Thomson Reuters was beginning to phase out Westlaw Classic at the time this research began; it has now completed the transition.) Tests on this system, which seems on cursory examination to incorporate a depth-of-treatment element, could yield different results. Likewise, would an examination of, say, the three most negative cases lead to greater correspondence between the Citing Cases presented by each service?

¶57 Finally, a direct invitation to both LexisNexis and Thomson Reuters: please, critique my work. Present verifiable studies with publicly available data to justify advertising claims concerning subsequent negative treatment and the value a citator adds to the legal research platforms on the whole. Similarly, does the cost associated with producing descriptions of subsequent negative treatment make sense given the inaccuracy in application found here? If so, why should lawyers and legal professionals of all stripes—in large law firms, government, or legal education—continue to pay the ever-rising fees attached to your services? The call is not terribly complicated: show us why the systems, as currently construction, actually meet lawyers' research needs.

Conclusion

¶58 Citators are expected to accurately describe subsequent negative treatment, an expectation that dates to their beginning in the United States. This study sought to answer the question of whether a system, such as Shepard's Citations, employing a hierarchical controlled vocabulary, would apply descriptions of subsequent negative treatment more accurately than a system, such as KeyCite, lacking a hierarchical controlled vocabulary. To reach a meaningful conclusion, a large set of cases with later-citing history was narrowed until a fair comparison could be made. The results of this comparison suggest that the presence of a hierarchical controlled vocabulary produces more accurate descriptions of subsequent negative treatment.

¶59 To this point, the relevant scholarly literature has been largely silent on descriptions of negative precedent in law. Some comparisons of citators for other factors have been done, and considerable writing by legal information professionals has critiqued the overall trustworthiness of citators. This article addresses the gap and provides a framework for assessing particular elements of citators. But an important goal of scholarship in library and information science is to make concrete suggestions for improvement, which this article does in discussing the best use of citators given the results of the study. Citators remain valuable tools as citation indexes, gathering all later-citing history from a variety of primary and secondary sources. But in providing descriptions of subsequent negative treatment, citators seem tied to an antiquated historical model originating in the early nineteenth century. Their lackluster performance in this area suggests they are failing at the task deemed so central to their existence. Law librarians and legal educators and scholars must continue to examine the tools we use to evaluate the law so as best to communicate its meaning to those we serve.