

2017

The Algorithm as a Human Artifact: Implications for Legal [Re]Search

Susan Nevelow Mart

University of Colorado Law School

Follow this and additional works at: <http://scholar.law.colorado.edu/articles>

 Part of the [Computer Law Commons](#), [Legal Writing and Research Commons](#), and the [Science and Technology Law Commons](#)

Citation Information

Susan Nevelow Mart, *The Algorithm as a Human Artifact: Implications for Legal [Re]Search*, 109 LAW LIBR. J. 387 (2017), available at <http://scholar.law.colorado.edu/articles/755>.

Copyright Statement

Copyright protected. Use of materials from this collection beyond the exceptions provided for in the Fair Use and Educational Use clauses of the U.S. Copyright Law may violate federal law. Permission to publish or reproduce is required.

This Article is brought to you for free and open access by the Colorado Law Faculty Scholarship at Colorado Law Scholarly Commons. It has been accepted for inclusion in Articles by an authorized administrator of Colorado Law Scholarly Commons. For more information, please contact erik.beck@colorado.edu.

The Algorithm as a Human Artifact: Implications for Legal [Re]Search*

Susan Nevelow Mart**

The results of using the search algorithms in Westlaw, Lexis Advance, Fastcase, Google Scholar, Ravel, and Casetext are compared. Six groups of humans created six different algorithms, and the results are a testament to the variability of human problem solving. That variability has implications both for researching and teaching research.

Introduction	388
A Brief Discussion of Algorithms and Classification	391
What Legal Database Providers Say About the Search Experience	396
Westlaw	399
Lexis Advance	401
Fastcase	402
Ravel	403
Casetext	404
Google Scholar	405
Cost	405
The Empirical Study	406
Methodology	406
Hypotheses	411
Results	412
Uniqueness of Cases	412
Relevance	413
Relevant and Unique	415
Number of Results Returned by a Query	416
Age of Cases	416
Discussion	416
Conclusion	420
Appendix A	421

* © Susan Nevelow Mart, 2017.

** Associate Professor and Director of the Law Library, University of Colorado Law School, Boulder, Colorado.

An algorithm must be seen to be believed.¹

Research . . . is not a method, it is not an object, it is a behavior. . . . Research is the systematic indulgence of one's curiosity.²

"Legal research" is not merely a search for information; it is primarily a struggle for understanding.³

Introduction

¶1 In the twenty-first-century world of legal resources, finding the right information and turning that information into knowledge that can be used to solve a legal problem or advise a client requires confronting technology as a partner in the research enterprise. Having a relationship with a partner always requires an investment of time and energy, and partnering with technology is no different. Researchers need to acquire some expertise about the technology at the meta-level. If you are searching online, as all legal researchers do, you need to remember that an algorithm is being used to return your results, and that, as a Westlaw engineer once wrote to me, "all of our algorithms are created by humans."⁴ Those human creators made choices about how the algorithm would work that have implications for the search results returned to the researcher.

¶2 Those choices become the biases⁵ and assumptions that are built into systems. If the search entered into a legal database has five terms,⁶ and only four terms appear, how will the algorithm treat the search? If the algorithm is strict, it will return only results with exactly those five terms. But the algorithm can be adjusted so that results with four of the terms will appear in the results set. The algorithm is set to determine how close those words have to be to each other to be returned in the top results. The programming team decides which of the search terms entered are automatically stemmed⁷ and which are not. Only the team knows which legal phrases are recognized by the algorithm without quotation marks around the phrase and how many preexisting legal phrases are added to the search without

1. 1 DONALD E. KNUTH, *THE ART OF COMPUTER PROGRAMMING* 4 (3d ed. 1997).

2. Felix Frankfurter, Karl N. Llewellyn & Edson R. Sunderland, *The Conditions for and the Aims and Methods of Legal Research*, 6 AM. L. SCH. REV. 663, 664 (1930).

3. Michael J. Lynch, *An Impossible Task but Everybody Has to Do It—Teaching Legal Research in Law Schools*, 89 LAW LIBR. J. 415, 415 (1997).

4. E-mail from Mike Dahn, Senior Vice-President Prod. Dev., Thomson Reuters, Legal, to author (Oct. 1, 2012, 18:17 MST) (on file with author). The advent of coding algorithms is not an impossibility, and some sort of oversight beyond "algorithmic accountability" may be necessary. See, e.g., Anupam Datta, Shayak Sen & Yair Zick, *Algorithmic Transparency via Quantitative Input Influence: Theory and Experiments with Learning Systems*, 2016 IEEE SYMP. SECURITY & PRIVACY 598.

5. "Bias" is not used in this article in the usual pejorative sense. It is used to indicate a preference in a computer system. For a detailed analysis of bias in computer systems, see Batya Friedman & Helen Nissenbaum, *Bias in Computer Systems*, 14 ACM TRANSACTIONS ON INFO. SYS. 330 (1996), discussed *infra* ¶ 12.

6. "Term" is used throughout this article in its function in the research process, as a word used to query a database or search engine to retrieve relevant information. See JOAN M. REITZ, *DICTIONARY FOR LIBRARY AND INFORMATION SCIENCE* 641, 712 (2004).

7. To "stem" is to take the root of a word. *Id.* at 683. In information retrieval, there are many methods of using algorithms to stem words in an index and then look for the variants. See Anjali Ganesh Jivani, *A Comparative Study of Stemming Algorithms*, 2 INT'L J. COMPUTER TECH. & APPLICATIONS 1930 (2011).

user input. The researcher does not have access to the list of synonyms that are (or are not) added automatically to the search. The scope of any machine learning is not known to the researcher. Once these decisions have been made, searches are automatically executed; any bias is encoded into the system.⁸ For an example of how coding choices affect results, the effect when a technician working for Amazon changed the value of the metatag “adult” from “false” to “true” is illustrative;⁹ the change to “true” excluded 57,000 books with tags for “gay,” “lesbian,” “health,” “mind,” “body,” “sexual medicine,” and “reproductive” from appearing in the results.¹⁰ In the Amazon example, users, particularly authors, noticed. Legal researchers are not likely to be able to tell how the encoded biases and assumptions are affecting search results. Legal database providers have viewed their algorithms as trade secrets and so have been reluctant to discuss the algorithms.¹¹

¶3 This article argues that legal database providers can be much more transparent about the biases in their algorithms without compromising trade secrets. The article is, in part, a call for more algorithmic accountability. Algorithmic accountability in legal databases will help assure researchers of the reliability of their search results and will allow researchers greater flexibility in mining the rich information in legal databases. If researchers know generally what a search algorithm is privileging in its results, they will be better researchers. Law librarians will be better teachers of the kind of analysis researchers need to search in any new database. And in the likely event that researchers do not have access to all of the different databases studied in this article, knowledge about the variability of each database might mean that researchers will work search term and resource variability into their search strategies. More information about databases may also affect collection development decisions.

¶4 In the absence of transparency from the database providers themselves, there may still be things that can be learned about system biases. This article sets out the results of a study designed to reveal how hidden biases and assumptions affect the results provided by some of the major legal database providers.¹² While it is usually

8. Lisa Shay et al., *Do Robots Dream of Electric Laws? An Experiment in the Law as Algorithm 7* (2013), http://www.rumint.org/gregconti/publications/201303_AlgoLaw.pdf [<https://perma.cc/WMG8-QBYZ>]. Bias is used here in the usual sense; the coders for this project were creating algorithms that determined traffic violations for specific sets of circumstances.

9. Andrea James, *AmazonFail: An Inside Look at What Happened*, SEATTLE POST-INTELLIGENCER: AMAZON & ONLINE RETAIL BLOG (Apr. 13, 2009, 6:32 PM), <http://blog.seattlepi.com/amazon/2009/04/13/amazonfail-an-inside-look-at-what-happened/> [<https://perma.cc/M9R2-8JW9>].

10. Andrea James, *Amazon Calls Mistake “Embarrassing and Ham-fisted,”* SEATTLE POST-INTELLIGENCER: AMAZON & ONLINE RETAIL BLOG (Apr. 13, 2009, 2:43 PM), <http://blog.seattlepi.com/amazon/2009/04/13/amazon-calls-mistake-embarrassing-and-ham-fisted/> [<https://perma.cc/EDQ3-KX2M>]. The error in the system—changing the code—affected not only the sales rank of some books, which pushed them to the bottom of the list where they were unlikely to be found, and also had “the effect of removing the books from Amazon’s main product search.” *Id.*

11. The algorithms used by LexisNexis and Westlaw are trade secrets. See, e.g., JULIE E. COHEN, *CONFIGURING THE NETWORKED SELF: LAW, CODE, AND THE PLAY OF EVERYDAY PRACTICE* 209 (2012) (“Efforts to gain access to information about the algorithms that determine the order of online search results typically have been stymied by assertions of trade secrecy . . .”). The exact operation of a relevancy-ranked natural language algorithm is proprietary and usually not disclosed. See also Danny C.C. Poo & Christopher S.G. Khoo, *Online Catalog Subject Searching*, in 3 *ENCYCLOPEDIA OF LIBRARY AND INFORMATION SCIENCE* 2218, 2224 (Miriam Drake ed., 2d ed. 2003).

12. This article makes the following assumptions about the six database providers studied: they have access to and publish a similar corpus of published federal cases, and they are trying to accomplish a similar task: to return cases relevant to the researcher’s query with the algorithms they create.

difficult to know what documents are being searched in very large databases, using jurisdictional limits creates a unique opportunity to compare how different algorithms process the same search in the same set of documents. This study used Casetext, Fastcase, Google Scholar, Lexis Advance, Ravel, and Westlaw to look at the differences in results when six different sets of engineers set out to solve the same problem.

¶5 The results are a remarkable testament to the variability of human problem solving. There is hardly any overlap in the cases that appear in the top ten results returned by each database.¹³ An average of forty percent of the cases were unique to one database, and only about seven percent of the cases were returned in search results in all six databases. It is fair to say that each different set of engineers brought very different biases and assumptions to the creation of each search algorithm. The uniqueness of results may show something about the worldview of each database that suggests that searching in multiple databases may be the twenty-first-century version of making sure that multiple authorial viewpoints are highlighted in a library collection's holdings. One of the most surprising results was the clustering among the databases in terms of relevant results. The oldest database providers, Westlaw and LexisNexis, were at the top in terms of relevance, with sixty-seven percent and fifty-seven percent relevant results, respectively. The newer legal database providers, Fastcase, Google Scholar, Casetext, and Ravel, were clustered together at a lower relevance rate, each returning about forty percent relevant results.

¶6 Legal research has always required redundancy in searching; one resource does not usually provide a full answer, just as one search does not provide every necessary result. This study clearly demonstrates that this need for redundancy has not faded with the rise of the algorithm. From the law professor seeking to set up a corpus of cases to study, to the trial lawyer seeking that one elusive case, to the legal research professor showing students the limitations of algorithms, researchers who want full results need to mine multiple resources with multiple searches. An exemplar legal research problem illustrating what the human construction of algorithms means for the uniqueness and relevance of results in any given database is provided. Once researchers have determined on their own that every algorithm has a unique voice, they truly understand the need for and the usefulness of redundancy in searching.

¶7 Paragraphs 8–15 of this article discuss algorithms in the context of legal research and set the stage for today's research environment. Paragraphs 16–29 describe in general terms the types of search algorithms employed by legal database providers and discuss what each provider has revealed about its algorithms in promotional material. Paragraphs 30–56 discuss the empirical study, its protocols, the results of the empirical study, and some conclusions that can be drawn. Paragraphs 57–58 conclude by returning to the question of algorithmic accountability and the cognitive impact of algorithms on legal research strategies.

13. The full results and analysis discussed in this paragraph are presented *infra* ¶¶ 45–56.

A Brief Discussion of Algorithms and Classification

¶8 At the simplest level, an algorithm is “a set of step by step instructions, to be carried out quite mechanically, so as to achieve some desired result.”¹⁴ The Pythagorean theorem is an algorithm, and so is the set of instructions that Netflix uses to recommend movies. Although algorithms have always had a role in modern life,¹⁵ it is the role that algorithms play in selecting what legal information we see that is critical for legal researchers.¹⁶ As we increasingly rely on algorithms for the assessment of information, algorithms dominate in mediating our information environment.¹⁷ If researchers are not aware that the information they seek may be missing from a database, or that the results that might be helpful may not be privileged in the result set, or that the list of documents suggested may have been generated by a legal worldview that opposes the path the researcher is trying to forge, a research session may terminate with no helpful results when helpful results actually exist.¹⁸ So we have to have sets of questions like these to ask each algorithm: How is information included or excluded from a system? How does the resource use predictive algorithms to anticipate use? How is relevance evaluated? Does the “black box”¹⁹ of the algorithm’s work lend a seeming objectivity to the results? How does use of the system change result patterns?²⁰ For attorneys, learning to navigate black boxes is part of the ethical duty to do competent research: knowing something about why you received the results that you did is a critical skill. For legal research professors, teaching this skill may involve passing on some understanding of how the systems we use today evolved.

¶9 Online legal information systems did not arise as completely new structures. The initial transition to any new technology is frequently fairly literal. Think of the “horseless” carriage or the first bicycles. The first legal information that made the

14. Jean-Luc Chabert, *Introduction*, in *A HISTORY OF ALGORITHMS: FROM THE PEBBLE TO THE MICROCHIP* 1, 1 (Jean-Luc Chabert ed., 1999).

15. As just one example from industry, Gantt charts were simple instructions for scheduling that have been in use since the mid-1890s. Starting in 1958, those instructions were computerized with algorithmic instructions, using the Naval Ordnance Research Calculator, the most powerful computer in existence at the time. Jeffrey W. Herrmann, *A History of Production Scheduling*, in *HANDBOOK OF PRODUCTION SCHEDULING* 1, 11–12 (Jeffrey W. Herrmann ed., 2006).

16. See, e.g., Tarleton Gillespie, *The Relevance of Algorithms*, in *MEDIA TECHNOLOGIES: ESSAYS ON COMMUNICATION, MATERIALITY, AND SOCIETY* 167 (Tarleton Gillespie et al. eds., 2014) [hereinafter *MEDIA TECHNOLOGIES*].

17. *Id.* at 167–68. The study of algorithms as mediators of all public information is a rich field on its own, but one that is beyond the scope of this article, which will limit its focus to algorithms that mediate legal information systems.

18. See Gregory J. Downey, *Making Media Work: Time, Space, Identity, and Labor in the Analysis of Information and Communication Infrastructures*, in *MEDIA TECHNOLOGIES*, *supra* note 16, at 141; see also Nicholas F. Stump, *Following New Lights: Critical Legal Research Strategies as a Spark for Law Reform in Appalachia*, 23 *AM. U. J. GENDER SOC. POL’Y & L.* 573, 639 (2015) (stating that for the law review articles promoted to the researcher as “Context & Analysis,” the researcher has no way of knowing the criteria used by the publisher in picking those articles, and the work that went into making those decisions has a definite influence on the course of the research).

19. On one level, a black box is any “technical object that operates as it should. When this occurs, the complex sociotechnical relationships that constitute it are rendered invisible, or black-boxed.” Darryl Cressman, *A Brief Overview of Actor-Network Theory: Punctualization, Heterogeneous Engineering & Translation* 1, 6 (Ctr. for Policy Research on Sci. & Tech. No. 09-01, 2009), <http://summit.sfu.ca/item/13593>.

20. Gillespie, *supra* note 16, at 168.

transition online was the full text of cases, made searchable with Boolean logic.²¹ Headnotes, case summaries, statutes, news, business information, and finally law reviews were subsequently added to the systems.²² The freedom to search full text without the constraints of classification systems was supposed to unmoor the law from its structure.²³ But it turns out that trying to make sense of information without underlying ontologies or classification systems can impede automation practices.²⁴ Legal database providers may even make the human additives to their search explicit. LexisNexis boasts of the human indexing in Shepard's citations;²⁵ Westlaw is proud of its human-generated Key Numbers;²⁶ and Bloomberg BNA advertises that the human indexing in its BNA treatises significantly boosts search results.²⁷

¶10 The complexity of the source material may require classification to aid relevant search results. Even the current "Google-like" legal databases provide extensive prefiltering, postfiltering, and word wheel²⁸ options for granular classification by source, authority, jurisdiction, and content type, and by value-added indexing by humans.²⁹ Some of the changes in the levels of prefiltering and word wheel options have been in response to user demand; lawyers seem to need clas-

21. Robert C. Berring, *Legal Information and the Search for Cognitive Authority*, 88 CALIF. L. REV. 1673, 1696 (2000) ("LEXIS and WESTLAW built on the old foundations. They loaded the text of cases online, each word of each case . . . WESTLAW and LEXIS were brave pioneers, but one cannot build new information systems out of thin air. Both followed a predictable course. Like the first iteration of many systems, WESTLAW and LEXIS tried to use new technology to accomplish the old tasks. Since everyone was deeply immersed in the existing system, they aped the functions of the old system.").

22. Bernard J. Hibbitts, *Last Writes? Reassessing the Law Review in the Age of Cyberspace*, 71 N.Y.U. L. REV. 615, 657–58 (1996) (noting that this transition took until 1982).

23. Robert C. Berring, *Legal Research and Legal Concepts: Where Form Molds Substance*, 75 CALIF. L. REV. 15, 26–27 (1987) ("There is no underlying rational structure to the law other than what the positivists give it. Allowing people to go online in free text liberates them from any requirement to fit their thoughts into a pre-existing structure. Individual researchers are able to order legal doctrine as it suits their needs . . .").

24. Francisco Iacobelli et al., *Information Finding with Robust Entity Detection: The Case of an Online News Reader*, in HUMAN-COMPUTER INTERACTION: THE AGENCY PERSPECTIVE 375–87 (Marielba Zacarias & JoséValente de Oliveira eds., 2012). To create an automated news reader that worked to find sources for news reporting, among other inputs, the authors "manually built a higher-level classification system on top of Wikipedia's categories." *Id.* at 381. Netflix boasts that its "secret sauce" is human indexing. Netflix's chief content officer added that in its 70/30 mix of data and judgment, "the thirty needs to be on top." Tim Wu, *Netflix's Secret Special Algorithm Is a Human*, NEW YORKER (Jan. 27, 2015), <http://www.newyorker.com/business/currency/hollywoods-big-data-big-deal> [<https://perma.cc/A8CY-LH46>].

25. *Shepard's Citations Service*, LEXISNEXIS, <http://www.lexisnexis.com/Shepards> [<https://perma.cc/4TDA-FR85>] ("Experienced *Shepard's* attorney-editors read cases around the clock and make assessments according to strict standards and guidelines. They recognize implications, case subtleties and nuances that algorithms simply can't").

26. Westlaw is proud of its human-generated Key Numbers. *Topic and Key Number Overview*, WESTLAW, <https://lawschool.westlaw.com/marketing/display/RE/24> [<https://perma.cc/QFK8-HZ3L>].

27. *BNA's Law School Professional Information Center: Using Online Indexes*, BLOOMBERG BNA, <http://subscription.bna.com/pic2/lsl.nsf/id/DTRS-5L3RPC?OpenDocument> [<https://perma.cc/NH3H-7N5V>].

28. A word wheel describes the autocomplete list of choices generated by a legal database provider when the researcher starts typing in the search box. For example, a researcher typing in "American law" may get a list asking: "Looking for this? American Law Reports Digest; American Law Reports; American Law of Product Liability." Clicking on one of these options will take the researcher directly to the legal resource listed.

29. See the discussion of the value-added interfaces for six legal databases *infra* ¶¶ 20–28.

sification in the law.³⁰ Lawyers, after all, are human, and we are all hardwired to impose structure on the world.³¹

¶11 The transition to online searching has increased the complexity of the search task; as more information becomes available, more research is required.³² As is frequently the case, automation has made a task more complex for the humans involved.³³ In exchange for instant access, the user has to master increasingly complex tasks to recover information effectively. The human reasoning, classification schemes, design decisions, and other work that went into the creation of the systems the researcher is using are mostly hidden.³⁴ Going beneath the surface of research systems, even in the predigital search environment, has never been the norm.³⁵ There is a long history in legal research of researching with only a surface understanding of the underlying structure. Speaking of lawyers at the time of transition to online searching, Bob Berring has noted that most were unaware of the details of the classification systems imposed by the Key Number system,³⁶ and in the early days of online searching, most users were unaware of the structure underlying the system.³⁷ This is almost certainly still true. But that is not to say that some basic understanding of the forces at work would not be helpful to researchers in the

30. MAGGIE NELSON, *THE ARGONAUTS* 53 (2015) (positing an “Aristotelian, perhaps evolutionary need to put everything into categories”). Recent studies on the human mind illustrate the deep-seated desire to classify and categorize, and, in response, lawyers push online systems to recreate the systems and categories. DANIEL J. LEVITIN, *THE ORGANIZED MIND: THINKING STRAIGHT IN THE AGE OF INFORMATION OVERLOAD* 25 (2014) (“The formation of categories in humans is guided by a cognitive principle of wanting to encode as much information as possible with the least possible effort. Categorization systems optimize the ease of conception and the importance of being able to communicate about those systems.”). Even in the evolution of online databases, where the first databases were just the stripped-out text of cases, the momentum has always been toward more structure and classification in the online systems. F. Allan Hanson, *From Key Numbers to Keywords: How Automation Has Transformed the Law*, 94 *LAW LIBR. J.* 563, 569–72, 2002 *LAW LIBR. J.* 36, ¶¶ 19–25; see also William G. Harrington, *A Brief History of Computer-Assisted Legal Research*, 77 *LAW LIBR. J.* 543 (1984–1985).

31. LEVITIN, *supra* note 30, at 32. Lawyers may be singular in their need for control and order. See, e.g., Margaret Hagan, *Do Lawyers Want Bad Visual Design?*, OPEN LAW LAB (June 28, 2016), <http://www.openlawlab.com/2016/06/28/do-lawyers-want-bad-visual-design/> [<https://perma.cc/7FWQ-AFQ5>] (“Lawyers want maximum overload of information in response to queries they do; They want it listed out in detail, with lots of information packed onto the page; They don’t want white space, they want text covering as much of the screen as can fit. They want lots and lots of controls, all kinds of filters and sorting mechanisms.”).

32. Berring, *supra* note 21, at 1683–90 (tracing the differences between the forms and content of an 1891 Supreme Court case and a 1996 Supreme Court case).

33. ENGINEERING PSYCHOLOGY AND COGNITIVE ERGONOMICS: UNDERSTANDING HUMAN COGNITION 206–07 (Dan Harris ed., 2013).

34. See GEOFFREY C. BOWKER & SUSAN LEIGH STAR, *SORTING THINGS OUT: CLASSIFICATION AND ITS CONSEQUENCES* 33 (1999).

35. Berring, *supra* note 21, at 1694.

36. See *id.*; Robert Berring, *Chaos, Cyberspace and Tradition: Legal Information Transmogrified*, 12 *BERKELEY TECH. L.J.* 189, 210–11 (1997). In the pre-online research world, many lawyers were aware that the West classification system missed a lot. See Daniel P. Dabney, *The Curse of Thamus: An Analysis of Full-Text Legal Document Retrieval*, 78 *LAW LIBR. J.* 5, 14 (1986) (“This short review of ideas in indexing shows that the indexing process is prone to many sorts of errors and uncertainties. Manual indexing is only as good as the ability of the indexer to anticipate questions to which the indexed document might be found relevant. It is limited by the quality of its thesaurus. It is necessarily precoordinated and is thus also limited in its depth. Finally, like any human enterprise, it is not always done as well as it might be.”).

37. Berring, *supra* note 21, at 1697.

brave new world of information overload and satisficing³⁸ that we now live in. And that is what we need to teach legal researchers.

¶12 Lawyers are not alone, of course. Most people do not think about the underlying structures of the technologies they use.³⁹ But some inquiry into the forces at work in the legal research environment, at this moment when so much of the work is truly invisible, is certainly called for.⁴⁰ It is time to examine technical bias in legal computer systems.⁴¹ Technical bias is built into systems. We just don't see it because the systems we use are black boxes.⁴² The following attributes contribute to the biases that programmers embed in the black box:

- prioritization (“emphasiz[ing] . . . certain things at the expense of others,” like relevance ranking);
- classification (putting an “entity [in a] constituent . . . class.” Data training may import human biases);

38. Satisficing is a time-honored information-seeking activity; it means to settle for what is “most readily available with little or no regard for costs and benefits,” when you want quick answers and may not have time to optimize the search. BRIAN C. O’CONNOR ET AL., HUNTING AND GATHERING ON THE INFORMATION SAVANNA: CONVERSATIONS ON MODELING HUMAN SEARCH ABILITIES 131 (2003).

39. The classifications, design decisions, and choices made every day by information scientists in our technological environment frequently embody “moral and aesthetic choices” that impact our own decisions and thoughts. BOWKER & STAR, *supra* note 34, at 3–4. These hidden choices can have very deep effects, as the recent “great recession” has shown. The great recession was in part a failure of algorithmic oversight. See MORTON GLANTZ & ROBERT KISSELL, MULTI-ASSET RISK MODELING: TECHNIQUES FOR A GLOBAL ECONOMY IN AN ELECTRONIC AND ALGORITHMIC TRADING ERA 437–39 (2014).

40. There is a large literature on bias in databases. See, e.g., HUMAN VALUES AND THE DESIGN OF COMPUTER TECHNOLOGY (Batya Friedman ed., 1997); Tarleton Gillespie & Nick Seaver, *Critical Algorithm Studies: A Reading List*, SOC. MEDIA COLLECTIVE RES. BLOG (July 20, 2016), <https://socialmediacollective.org/reading-lists/critical-algorithm-studies/#4.3> [<https://perma.cc/4UQY-76C7>]. For a discussion of how algorithms may unintentionally encode bias on the basis of protected classes (ethnicity, gender, race, religion), see Michael Feldman et al., *Certifying and Removing Disparate Impact*, in PROCEEDINGS OF THE 21ST ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING 259 (2015), <http://dl.acm.org/citation.cfm?doi=2783258.2783311>. For the effect computational “nonreading” of texts—the pattern recognition that is the language of machine learning—may have on legal interpretation, see Mireille Hildebrandt, *The Meaning and the Mining of Legal Texts*, in UNDERSTANDING DIGITAL HUMANITIES 145, 148–49 (David M. Berry ed., 2012). There has been a broad discussion of bias in applications on the web. See, e.g., Lucas D. Introna & Helen Nissenbaum, *Shaping the Web: Why the Politics of Search Engines Matters*, 16 INFO. SOC’Y 169 (2000).

41. Friedman & Nissenbaum, *supra* note 5, at 330. Technical bias is one of the three biases that computer systems can display; the other two are emergent and preexisting. Emergent biases are those that arise in the actual use of the database; these biases can relate to new societal knowledge, differing expertise, differing values, or a mismatch between the user and the system design. *Id.* at 335. The preexisting bias in legal databases is, at a minimum, the structure of the law itself, and the content and classification systems that have been imposed on the law by legal vendors. *Id.* at 333; see also Hanson, *supra* note 30, at 569–72, ¶¶ 19–25. This was a gradual change, as the first computer retrieval systems included only the simple text of cases, with no structure or classification; the evolution to structure and classification was gradual. Harrington, *supra* note 30, at 543; see also Berring, *supra* note 21, at 1693, 1696. For an excellent history of the evolution of ideas about the effect of classification schemes on legal thinking, see Richard A. Danner, *Legal Information and the Development of American Law: Writings on the Form and Structure of the Published Law*, 99 LAW LIBR. J. 193, 2007 LAW LIBR. J. 13. The classification systems may also hide assumptions about the nature of the law that mask paths to justice. See Stump, *supra* note 18, at 573; see also Hildebrandt, *supra* note 40, at 148–49 (discussing the effect computational “nonreading” of texts—the pattern recognition that is the language of machine learning—may have on legal interpretation).

42. Cressman, *supra* note 19, at 6.

- association (“marks relationships between entities”); and
- filtering, which “includes or excludes information according to various rules or criteria.”⁴³

¶13 An interesting example of how assumptions or biases inform results is from a study on coding algorithms to enforce that exemplary seemingly simple rule of law, the speed limit.⁴⁴ What seems a relatively straightforward problem becomes dense with assumptions when one thinks about how to implement the law: do you enforce the letter of the law or the intent of the law; is every second you exceed the speed limit a separate violation; do weather or road conditions matter; how often should a driver be given a ticket; does context matter?⁴⁵ There were three groups of coders using actual driving data taken from a vehicle’s computer: the first group was asked to implement the letter of the law; the second group was asked to implement the intent of the law; the third group was asked to follow a detailed written specification.⁴⁶ The differences in the results of each coding scheme are stark: the number of tickets issued by the algorithms varied from zero to 661, for the same driving pattern.⁴⁷ All of the groups made assumptions independent of their instructions; for example, there were significant differences in how the first two groups coded tolerances for exceeding the speed limit, and all of the groups assumed, without instruction, perfect driving conditions.⁴⁸ As the authors of the study note, transparency about coding assumptions may be the only solution to success in implementing automated legal compliance or enforcement in a fair and open manner.⁴⁹

¶14 “Algorithmic accountability” is the term for disclosing prioritization, classification, association, and filtering.⁵⁰ What we need is a frank discussion with database providers about what it means to search in their databases. Trade secrets should not prevent algorithmic accountability. Some database providers do provide search tips that can help their users understand what happens between input and output;⁵¹ legal database providers also publish basic search information for their

43. Nicholas Diakopoulos, *Algorithmic Accountability: Journalistic Investigation of Computational Power Structures*, 3 DIG. JOURNALISM 398, 399 (2015). Black boxes result in technological opacity.

44. Shay et al., *supra* note 8, at 1.

45. See *id.* at 20 for a chart summarizing the differences in results for three groups of coders who were given three different assignments for coding violations.

46. *Id.* at 4–5.

47. *Id.* at 20.

48. *Id.* at 7–8, 14–15. The variations and subtle assumptions are quite varied, and the Shay article discusses only a few.

49. *Id.* at 30–31.

50. Diakopoulos, *supra* note 43, at 402.

51. For example, this information is from a help protocol from Summon, a discovery-layer search product, transmitted in an e-mail to the author because the information is behind the password-protected administrative module. E-mail from Joan Policastri, Collection Servs. & Research Librarian, Univ. of Colo. Law Sch., to author (Jan. 8, 2016, 12:05 PM MST) (on file with author). It would be more helpful if the information were readily available to users:

Boolean search and Summon relevancy algorithm: Boolean queries get processed by the same relevancy algorithm as any other query. This means relevancy enhancements that come from the application of stemming, character normalization, etcetera will apply in Boolean queries as well.

Applying the relevancy algorithm to Boolean queries is particularly helpful in Boolean searches using multiple search terms. For example: **paint drying time (glass OR wood)**.

users.⁵² The more we understand about the input into the black box, even without knowing the code for the algorithm, the more we can see how the system operates in practice. At the moment, we can really only see the output of the systems, and that is what this study investigates.

¶15 The need to know about the input, the paths that mark the way to the results, only increases as the amount of work being done by the algorithms increases. A case in point is the use of IBM's artificial intelligence program, Watson, by the medical community.⁵³ Watson is IBM's supercomputer, which uses artificial intelligence and machine learning to leverage large amounts of data.⁵⁴ Watson is better than humans at reading through documents and is starting to be used as a "quick-witted digital assistant" in oncology clinics, but with a caveat: doctors use it in conjunction with Watson Paths, a visual tool that allows a doctor to see the underlying evidence and inference paths Watson took in making a recommendation. "It's not sufficient to give a black-box answer," said Eric Brown, IBM's director of Watson technologies.⁵⁵ As decision makers, doctors want knowledge, not technological determinism.⁵⁶ Legal researchers need to demand the same kind of transparency.

What Legal Database Providers Say About the Search Experience

¶16 Some information on the search experience is available on each legal database. In terms of the basic types of search, researchers tend to refer to Boolean searching—meaning that the researcher uses terms and connectors, such as "and", "or", and "not",⁵⁷ to construct a search—and natural language searching—meaning

This Boolean query will provide a small relevancy boost for documents containing the words *paint*, *drying*, and *time* in close proximity to each other, which is a relevancy enhancement that previously did not apply to Boolean queries.

Id. Legal professionals who were expert searchers had compiled their own special tips and tricks for getting the best results, but those tips and tricks were for pre-2010 Boolean searching. New tips and tricks are necessary for keyword searching in the current set of search algorithms, but for the most part, these tips and tricks have not been revealed.

52. See *infra* ¶ 16.

53. Steve Lohr, *If Algorithms Know All, How Much Should Humans Help?*, N.Y. TIMES (Apr. 6, 2015), <http://www.nytimes.com/2015/04/07/upshot/if-algorithms-know-all-how-much-should-humans-help.html> [<https://perma.cc/48HA-N92U>].

54. Lauren F. Friedman, *IBM's Watson Supercomputer May Soon Be the Best Doctor in the World*, BUS. INSIDER (Apr. 22, 2014), <http://www.businessinsider.com/ibms-watson-may-soon-be-the-best-doctor-in-the-world-2014-4> [<https://perma.cc/UTD8-KQNR>].

55. *Id.*; see also *IBM Research: WatsonPaths*, IBM, <http://www.research.ibm.com/cognitive-computing/watson/watsonpaths.shtml> [<https://perma.cc/7FZZ-78JR>] (highlighting the visual image of the decision paths Watson shares with doctors to aid transparency and help doctors authenticate the value of the decision). IBM recently released Ross, the legal version of Watson, but nothing has been released about similar transparency in decision paths on the legal side. See *Watson Takes the Stand*, ATLANTIC, <http://www.theatlantic.com/sponsored/ibm-transformation-of-business/watson-takes-the-stand/283/> [<https://perma.cc/PQ5U-U4SQ>].

56. See Julie E. Cohen, *Network Stories*, LAW & CONTEMP. PROBS., Spring 2007, at 91, 92, ("What makes the network good can only be defined by generating richly detailed ethnographies of the experiences the network enables and the activities it supports, and articulating a normative theory to explain what is good, and worth preserving, about those experiences and activities.")

57. George Boole was a nineteenth-century mathematician, and his work on the analogies between algebraic symbols and symbols that represent logical forms and syllogisms resulted in the application of what is known as Boolean logic to searching. *George Boole*, BRITANNICA.COM, <http://>

that the researcher uses keywords without connectors.⁵⁸ However, the reference to natural language searching is frequently a misuse of a technical term that refers to a complex attempt to pattern-match speech or text “through references to a database with the aid of grammatical structures models”;⁵⁹ it does not technically refer to keyword searching where the terms entered into a text box are then ranked by algorithms for relevance, word count, citation count, or other nongrammatical structures.⁶⁰ This article will refer to searches entered into a legal database’s search box without terms and connectors as keyword searches, although efforts will be made to determine and then note whether there is natural language grammatical parsing utilized for a specific legal database. Even though this study focuses on one possible first step in the research process, where the researcher formulates a query, puts keywords in the search box, and looks at the first few results, the full research process is a more iterative, intuitive, and complex process⁶¹ that involves multiple

www.britannica.com/biography/George-Boole [<https://perma.cc/K6WK-9WQG>]; Gerard Salton et al., *Extended Boolean Information Retrieval*, 26 COMM. ACM 1022 (1983), <http://dl.acm.org/citation.cfm?id=358466>. This kind of searching is now more commonly referred to as “terms and connectors” searching. See, e.g., *Search Terms and Connectors*, UNIV. OF OXFORD, <https://www3.law.ox.ac.uk/lrsp/e/searchterms.php> [<https://perma.cc/8TBA-KRMN>].

58. *Searching Bloomberg Law, Lexis Advance and Westlaw: Natural Language v. Terms & Connectors Searching*, WASH U LAW, <http://libguides.law.wustl.edu/LRMSearchingIntro/SearchTypes> (“Natural Language Searching refers to the type of search you would do in Google: enter a few relevant terms in any order. The online service’s search algorithm takes control and delivers results it determines to be most relevant. Sometimes called Descriptive Term Searching, although technically Natural Language Searching is a different process. Terms & Connectors Searching refers to a targeted search strategy that instructs the computer to look for specific terms, often in a specific order and/or a specific proximity to one another. Also called Boolean Searching.”).

59. Takeshi Matsumoto et al., *Application of Search Algorithms to Natural Language Processing*, 1 PROC. AUSTRALASIAN LANGUAGE TECH. WORKSHOP U03-1003 (2003), <http://aclweb.org/anthology/U/U03/U03-1003.pdf>.

60. Natural language searching may include “vector space models, Bayesian inference net models, and language models.” Jack G. Conrad & Qiang Lu, *Next Generation Legal Search—It’s Already Here*, LEGAL INFO. INST.: VOXPOPULI (Mar. 28, 2013, 9:13 AM), <https://blog.law.cornell.edu/voxpath/2013/03/28/next-generation-legal-search-its-already-here/> [<https://perma.cc/4YPV-G8KM>]. Tamsin Maxwell, speaking of free text searching decontextualizing information, commented: “One thing to notice about current methods in open domain IR, including vector space models, probabilistic models and language models, is that the only context they are taking into account is proximate terms (phrases). At heart, they treat all terms as independent.” K. Tamsin Maxwell, *Pushing the Envelope: Innovation in Legal Search*, LEGAL INFO. INST.: VOXPOPULI (Sept. 17, 2009, 1:56 PM), <https://blog.law.cornell.edu/voxpath/2009/09/17/pushing-the-envelope-innovation-in-legal-search/> [<https://perma.cc/L8NQ-TR4G>]. But, Maxwell continues, “inference networks used in commercial legal [information retrieval] are not applied in the open domain,” and they can incorporate “index numbers, terms, phrases, citations, topics, and any other desired information” in a “directed acyclic graph (the network),” which can “then be used to estimate the probability of a user’s information need being met” by a specific document. *Id.* See also Staffan Malmgren, *Towards a Theory of Jurisprudential Relevance Ranking: Using Link Analysis on EU Case Law* (Sept. 4, 2011) (unpublished Master of Laws thesis, Stockholm University), <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.397.3802&rep=rep1&type=pdf> (discussing the hidden complication of natural language indexing of documents in databases: attempts to augment the authentic text of legal sources with semantic metadata introduce an interpretation of the legal sources, which get hidden in the system unbeknownst to the user, even if users’ interpretations may differ from the system’s).

61. Boulder Statement on Legal Research Education (June 2009), https://lawlibrary.colorado.edu/sites/default/files/images/docs/2009_boulder_statement_on_legal_research_education.pdf [https://web.archive.org/web/20150922115957/http://lawlibrary.colorado.edu/sites/default/files/images/docs/2009_boulder_statement_on_legal_research_education.pdf]; Marcia J. Bates, *The Design of Browsing and Berrypicking Techniques for the Online Search Interface*, 13 ONLINE REV. 407, 408 (1989) (“Each

methods of starting and continuing a search.⁶² What the study in this article illustrates is that, since every algorithm and database interface is a completely human construct,⁶³ and every search is a completely human construct, the researcher must view the search process as a human interaction, moderated by technology, and not as a technological interaction.

¶17 Keyword searching is just one of many modes of searching.⁶⁴ Each database provider seeks to provide access to some of these other modes of searching and to enhance the search options with different tools that are presented to the researcher on the results screen.⁶⁵ This article is not investigating these secondary prompts. It is investigating one possible first step in the research process and comparing the results, to see how database algorithms differ. No researcher should stop their inquiry after just looking at the top ten results from one keyword search.

¶18 Algorithms in legal databases process search terms, assess the information in their databases, and then represent a set of the “best results.”⁶⁶ Even so simple-seeming a task as assembling the information to be searched has important implications for understanding the context of the information we see as a result of a

new piece of information they encounter gives them new ideas and directions to follow and, consequently, a new conception of the query. At each stage they are not just modifying the search terms used in order to get a better match for a single query. Rather the query itself (as well as the search terms used) is continually shifting in part or in whole.”) See also Aaron Kirschenfeld, *Everything Is Editorial: Why Algorithms Are Hand-Made, Human, and Not Just for Search Anymore*, LEGAL INFO. INST.: VOXPOPULI (Nov. 20, 2013, 9:56 AM), <https://blog.law.cornell.edu/voxpath/2013/11/20/everything-is-editorial-why-algorithms-are-hand-made-human-and-not-just-for-search-anymore/> [<https://perma.cc/26VD-9X5Y>] (“Computer assisted legal research cannot be about merely returning ranked lists of relevant results, even as today’s algorithms get better and better at producing those lists. Search must be only one component of a holistic research experience in which the searcher consults many tools which, used together, are greater than the sum of their parts.”).

62. O’CONNOR ET AL., *supra* note 38, at 127. The techniques most utilized by the novice researcher would seem to be grazing and satisficing. Grazing is “foraging in a space where evaluation and supply are not issues,” and satisficing is a form of settling for “what is most readily available with little or no regard for costs and benefits.” *Id.* at 129, 131.

63. *Id.* at 106; Kirschenfeld, *supra* note 61 (discussing the “artisanal” quality of algorithms, as something “massaged and kneaded by caring craftsmen”).

64. Bates, *supra* note 61, at 408.

65. See *infra* ¶¶ 20–28 (discussing the information a searcher might find to enhance the search process in each of the six legal databases in this study).

66. Gillespie, *supra* note 16, at 167–68 (speaking generally of algorithms in search); see also Conrad & Lu, *supra* note 60 (discussing the actual separation of retrieval from ranking in Westlaw). The difficulty of getting from the input—the processing of search terms—to the output—documents that satisfy the researcher’s information need—is the age-old problem of information retrieval. Bates, *supra* note 61, at 407–08. This is the same philosophical problem the structuralists are dealing with. Even enhanced with machine learning and natural language processing, the match in the middle is difficult to achieve because of the “value” between the signifier and the signified articulated by the structuralists. Barton Beebe, *The Semiotic Analysis of Trademark Law*, 51 UCLA L. REV. 621, 626–42 (2004). In a structuralist model of search, the document (or text) would be represented in the search process (by the algorithm creators) as a signifier. The researcher with the information need creates a query that is intended to retrieve information matching the mental idea or the signified. But there is always a problem in understanding caused by the context-specificity of words, or their “value”; for structuralists, words only have meaning in relation to “that which exists outside” them, such as synonyms or context. *Id.* at 641–42. The structuralists have identified a problem that exists in all human communication; our own “natural language processing” when communicating with each other is not fool-proof. When communicating with a computer, the problem is certainly no easier.

search.⁶⁷ Tarleton Gillespie calls this dimension “patterns of inclusion.”⁶⁸ Before an algorithm is deployed, the dataset of information to search must be assembled; this involves choices as to what is collected, how it is readied for the algorithm, and what is excluded or demoted.⁶⁹ Legal databases use similar primary law, but how it is readied for the algorithm differs: by the elements of metadata, relational, or object-oriented database architecture, for example, or the categories of classification that are chosen. Historical notes can be included or excluded. For all secondary sources, the effect of inclusion and exclusion in a database has an obvious effect on search results. Exclusion may be based on copyright, licensing, or editorial concerns.

¶19 The following section is a brief review of the published information about each database provider’s search algorithm and the various mechanisms used to enhance the search experience beyond the return of “relevant” cases. Just as no good legal researcher stops with reviewing the results of just one search string in a case law database, so no legal database provider stops with returning cases as a dead end to research. Indeed, each provider adds enhancements to the search results page to facilitate connections, in ways that may encourage those serendipitous connections that expert researchers prize and that some expert researchers worry the online interface will prevent.⁷⁰

Westlaw

¶20 Westlaw describes its search functionality as a combination of methodologies, and the promotional material does not clarify whether the methodologies include true natural language searching.⁷¹ The Westlaw interface lists its search as a “plain language” search,⁷² while early WestlawNext promotional material calls it

67. Gillespie, *supra* note 16, at 167–68.

68. *Id.*; see also Richard Delgado & Jean Stefancic, *Why Do We Ask the Same Questions: The Triple Helix Dilemma Revisited*, 99 LAW LIBR. J. 307, 317–19, 2007 LAW LIBR. J. 18, ¶¶ 27–28 (finding the freedom from categories imposed by legal publishers in the online search environment somewhat illusory).

69. Gillespie, *supra* note 16, at 170–72; Jeffery Alan Johnson, *Representing “Inforgs” in Data-Driven Decisions*, in DIGITAL SOCIOLOGIES 163, 163–64 (Jessie Daniels et al. eds., 2017) (“Creating data requires some process that narrows the many possible representations of a given state of the world to a single data state. This process is carried out within translation regimes: systems of technical rules and social practices that establish a one-to-one correspondence between a given state of the world and a data state.”).

70. O’CONNOR ET AL., *supra* note 38, at 10–11. One dominant metaphor for searching has been the classic Aristotelian nature of access systems—which fails to address the whole of human engagement. The authors suggest the use of a bricoleur model. *Id.* Bricoleur is being used here in the Levi-Straussian sense of thinking and doing with the materials at hand. *Id.* For a full list of the kinds of search behavior that researchers engage in, many of which are not fully integrated into an online environment, see *id.* at 127–34. It is not really clear whether serendipity plays the same role in online searching as it does in print research or browsing the stacks; print resources have generally been felt to provide the most opportunities for analogic surprise. See, e.g., Robert J. Sheran & Douglas K. Amdahl, *Minnesota Judicial System: Twenty-Five Years of Radical Change*, 26 HAMLINE L. REV. 219, 365 (2003); Ryan Metheny, *Re-Searching: While Search Engines Have Made It Easy to Find Facts, Legal Research Still Benefits from a Methodological Approach*, L.A. LAW, Dec. 2014, at 27.

71. Thomson Reuters, *WestSearch: The World’s Most Advanced Legal Search Engine*, http://info.legalsolutions.thomsonreuters.com/pdf/wln2/l-355700_v2.pdf [<https://perma.cc/U8BP-NT37>].

72. Thomson Reuters, *Tips and Tricks to Get You Started on Westlaw*, http://static.legalsolutions.thomsonreuters.com/product_files/westlaw/wlawdoc/wlres/wln_tips_tricks_ug.pdf [<https://perma.cc/3FEQ-LGDV>].

“natural language” searching.⁷³ The search algorithm removes stop words; generates variations of words; identifies legal phrases, citations, topics, and key numbers; and then uses knowledge-based engineering, machine-learning techniques, and statistical classification to improve results.⁷⁴ What is clear from the promotional material is that Westlaw’s algorithm uses value-added content such as the human-generated Key Number System, notes of decisions and headnotes, and KeyCite’s citation networks, as well as commercial user document interaction history to return relevance-ranked results beyond the exact search terms entered.⁷⁵ Qiang Lu and Jack Conrad reveal that WestlawNext separates the function of document retrieval from document ranking, so that retrieval results in high recall (retrieving the highest possible number of relevant results of all the relevant results in the database), and then results are ranked, which “allows potentially dozens of weighted features to be taken into account and tracked as part of the optimal ranking process.”⁷⁶ This theoretically means that searching a small group of pre-selected documents, which used to be the preferred way to produce fewer results with higher recall, may not necessarily produce the same benefits in Westlaw anymore.⁷⁷

¶21 Once those results are returned,⁷⁸ Westlaw offers many options to enhance the search experience. The researcher can filter results by jurisdiction, topic,

73. Thomson Reuters, *WestSearch: WestlawNext Search Technology*, <https://docmh.com/embed/wlnsearch> [<https://perma.cc/S8GV-R5SL>] (“The WestSearch federated search employs natural language processing to retrieve relevant documents . . . [and] builds on the Westlaw Natural Language search method.”).

74. *Id.*; JIN ZHANG, *VISUALIZATION FOR INFORMATION RETRIEVAL* 25 (2008) (stop words are words that are not important within full text (e.g., “a,” “the,” and “and”); most search engines list such words).

75. Thomson Reuters, *supra* note 73 (“WestSearch combines sophisticated information retrieval, natural language processing, and machine learning to take maximum advantage of the editorial enhancements West editors have added to legal documents since 1876. . .”).

76. Conrad & Lu, *supra* note 60. Until recently, it has been an inviolable law of search that as recall goes up, precision goes down. Paul D. Callister, *Working the Problem*, 91 ILL. B.J. 43, 44 (2003). As far back as 1994, Westlaw’s own study of the relationship between precision and recall in the Federal Supplement database showed that as precision went up, recall went down at almost the identical rate. *Id.*

77. Research assistants were given sample instructions on limiting their search to a specific case database for each of the six databases in the study; in the example used below, the reverse was true for Westlaw; prefiltering produced *more* results. Here were the instructions:

- On the Start page, click on Cases.
- Click on Federal District Court Cases under “Federal Cases by court.”
- Click on the state where your district court is located (e.g., Michigan).
- Run your search (e.g., agency follows clear congressional intent).
- On the left, click open the District Court box and select your actual district (e.g., E.D. Mich.) ALSO use the filters on the left to limit your searches to reported cases and click APPLY FILTERS.

Note that in Westlaw, the results are different if you follow this path (2920 cases) than if you enter the search in the main search box first and then use the filters (clicking on cases/jurisdiction /reported) to limit the jurisdiction (4 cases).

78. For purposes of the discussion about the enhancements each database provider offers, I am using this query from Appendix B, available online at <http://scholar.law.colorado.edu/research-data/5/>. The search is: **criminal sentence enhancement findings by jury required**; entered as key words; the jurisdiction is the Sixth Circuit). The legal framework for evaluating results is: You are looking for cases discussing the constitutionality of increasing the penalty for a crime when the jury did not make a factual determination about facts that enhanced the penalty. The enhancements are what show up on the screen after limiting the jurisdiction and entering Query 1. For a full discussion of the search protocols, see *infra* ¶¶ 30–44.

reported status, judge, attorney, law firm, key number, party, and docket number. In addition, Westlaw's results page recommends secondary sources on the right, as "Related Documents." These documents may or may not be relevant to the information need of the researcher, depending on the level of the search's success in retrieving relevant results.⁷⁹

Lexis Advance

¶22 Lexis Advance's stated worldview is that its search results should be as inclusive as possible but the same whether jurisdictional filtering is performed before or after entering the search terms.⁸⁰ Lexis Advance engineers produced a video on the search algorithm, and they state that their customized search algorithm uses automatic phrase recognition, case name recognition, proximity search between terms, and activity score boosting in the ranking algorithm.⁸¹ The video calls search without connectors "natural language" searching.⁸² Relevance ranking purportedly gives priority to legal phrases and moves documents up based on concentration of terms (how close the terms are to each other), coverage of terms (whether all the terms appear in the document), prominence of the opinion, and how recent the opinion

79. The relevance of the secondary sources recommended is related to the relevance of the results generated by the search. See *infra* ¶ 54. In this *one* instance, the secondary sources were not helpful.

80. LexisNexis, *Differences That Deliver: The Power of Lexis Advance Search*, http://www.lexisnexis.com/documents/pdf/20160803090042_large.pdf [https://web.archive.org/save/_embed/http://www.lexisnexis.com/documents/pdf/20160803090042_large.pdf]. This information was provided by LexisNexis after the author discussed the concept of algorithmic accountability with a Lexis representative, so asking for help is a useful thing to do. Interview with David Dilenschneider, Senior Dir. of Client Relations for LexisNexis (Mar. 4, 2015) (notes on file with author). In Dilenschneider's view, this is what differentiates Lexis Advance from Westlaw. As part of this philosophy, the algorithms have been adjusted so that the search returns the same number of results whether the researcher pre-filters the search results by limiting the search to a particular jurisdiction, or performs it post-search.

Here were the search instructions for limiting a search to a specific court database in Lexis Advance:

- Enter your search (e.g., agency follows clear congressional intent).
- Click the filters arrow next to the search box.
- The top limiter is jurisdiction. Check federal district courts.
- Run the search (e.g., agency follows clear congressional intent).
- Scroll down on the left to Jurisdiction and click on your circuit. Note that you have to know which circuit you are in (e.g., the Sixth).
- Now you can see, under Court, the Eastern District Michigan. Click on that.
- ALSO use the filters on the left to limit your searches to reported cases.

Note that in Lexis Advance, prefiltering and postfiltering should give you the same number of cases (e.g., 338 cases). I have not figured out another way to limit the search to a specific district court.

81. LexisNexisLawSchools, *Understanding the Technology and Search Algorithm Behind Lexis Advance*, YouTube (Oct. 8, 2013), <https://www.youtube.com/watch?v=bxJzfYLwXYQ&feature=youtu.be> (discussed at 13:06 and following).

82. *Id.*

is.⁸³ Lexis Advance also states that changing word order or stop words will generally not alter the search results.⁸⁴

¶23 When searching in a case database, Lexis Advance offers further filtering by court, date, publication status, practice area, attorney, law firm, most cited, keyword, and judge. Lexis does not offer any secondary sources when searching in a case database until the researcher clicks on a case; then the researcher can view a “Legal Issue Trail” that highlights important passages of the case being viewed, or the researcher can Shepardize the case and view secondary sources.⁸⁵

Fastcase

¶24 Fastcase supports using a natural language algorithm.⁸⁶ According to Fastcase’s promotional materials, “natural language searches are much less precise” than Boolean searches, “but they are a good place to start if you are new to legal research, or if you are delving into a new area of the law.”⁸⁷ Ed Walters, CEO of Fastcase, says that natural language is good if you are totally at sea, but only then.⁸⁸ The natural language algorithm returns cases with “the highest relevance scores based on [the researcher’s] overall mix of search terms.” The results may include cases that do not have all of the search terms.⁸⁹ In Fastcase, the researcher chooses to use natural language searching, and the same number of results are returned whether he or she filters before or after the search.⁹⁰

83. See LexisNexis, *Lexis Advance Faculty FAQs*, http://www.lexisnexis.com/documents/pdf/20111216091630_large.pdf [https://web.archive.org/web/20160620151641/http://www.lexisnexis.com/documents/pdf/20111216091630_large.pdf]:

Lexis Advance uses a variety of proprietary methods in producing relevant results for our users; “relevance” as we have defined it means that the document a user would expect to find in their results appears as one of the first five documents in a user’s results set. To ensure this result, Lexis Advance includes, but does not limit, the following:

- a. Automatic phrase recognition
- b. Case name recognition . . .
- c. Proximity search between the terms
- d. Activity score boosting in the ranking algorithm (i.e. “landmarkness” of the case).

84. LexisNexis, *supra* note 80.

85. For this *one* instance, there were no relevant further materials to be found either following the relevant Legal Research Trail or reviewing secondary sources in the *Shepard’s* report.

86. *Frequently Asked Questions: What Are Boolean Searches, Natural Language Searches and Citation Searches?*, FASTCASE, <http://www.fastcase.com/faq/> [<https://perma.cc/K77N-HB9W>]; E-mail from Ed Walters, Chief Exec. Officer, Fastcase, to author (May 4, 2015, 15:01 MST) (on file with author).

87. *Documentation and Resources*, FASTCASE, <https://www.fastcase.com/support/> [<https://perma.cc/KQ5X-UEZR>].

88. E-mail from Ed Walters, *supra* note 86.

89. Fastcase finds that “[l]awyers, law professors, and law students will always get better results searching with Keyword (aka Boolean) searches.” *Id.*

90. Here are the instructions given to research assistants to limit results to a specific jurisdiction:

- On the Start page, click Advanced Case Law search on the left, or the link at the top of the page.
- Enter your search (e.g., agency follows clear congressional intent).
- Leave the Keyword Search button on.
- Click on the Individual Jurisdictions radio button on the left, under Select Jurisdictions select US District Courts, and click on your specific court (e.g., Michigan Eastern District). Make sure that the start date (under Search Options) is January and Before 1925.
- Make sure the results are sorted by relevance.

¶25 Once case search results appear, Fastcase supplements the results with Forecite, an algorithm that suggests relevant cases that do not include the words in the researcher's search, and results from HeinOnline's law reviews and journals database. Again, the relevance of the law review articles is related to the relevance of the search results.⁹¹ Fastcase recently released "Customize Relevance Algorithm," which allows the researcher to see which factors Fastcase is using in its ranking algorithm and allows researchers to adjust the weights to suit particular research strategies, which, of course, may change with the context of the problem being solved.⁹²

Ravel

¶26 Ravel Law's Quick Start Guide describes how its search algorithm works when key words are entered:

Ravel finds all cases that contain those keywords and then ranks them based on a combination of how those keywords appear in the case, and how important that case is more broadly. Ravel ranks the importance of each case by looking at the citation network—assessing how many and which other cases cite to a given case.⁹³

Ravel therefore appears to be using term inclusion, term proximity, term frequency, and citation analysis to determine results, and then, on the main results page, providing a unique visualization tool to help lawyers find more relevant cases. Ravel's visualization shows the top seventy-five results based on the keywords entered.⁹⁴ The visualization map shows circles for cases; the larger the circle, the more important the case is in terms of the number of times it has been cited.⁹⁵ Of course, the relevance of the cases shown in the visualization is related to the relevance of the cases returned in the results set as a whole.⁹⁶ Once a researcher clicks on a case in the list to read, "case analytics" on the left direct the researcher to information

-
- Run the search.
 - Filter out unpublished cases.

Note that following this path leads to an exact number of results (e.g., 68 results); and filtering to the specific court after the search is done yields the same results (e.g., 68 cases) versus filtering to the specific court after the search is done (e.g., 927,000+ cases).

91. In this *one* instance, the secondary sources were helpful.

92. This is available on Fastcase 7, once a researcher clicks on "Advanced Search."

93. RAVEL, RAVEL LAW QUICK START GUIDE 16, https://d2xkqp20fm9wy8.cloudfront.net/downloads/Ravel_QuickStart_Guide.pdf [https://web.archive.org/save/_embed/https://d2xkqp20fm9wy8.cloudfront.net/downloads/Ravel_QuickStart_Guide.pdf].

94. *Id.* at 5. Ravel returns a list of cases on the right, ranked by relevance, and a visual map of the seventy-five most relevant cases shows on the left, which can be filtered by court and relevance.

95. *Id.*; see also Ravel Law, *Overview of Ravel's Data Visualization*, VIMEO (May 11, 2015), <https://vimeo.com/127559698> [<https://perma.cc/GK5M-NXAH>].

96. Daniel Lewis, the CEO of Ravel, explains, "What we try to communicate about visualization is that it's tailored to the way that *lawyers* do research, which is about fitting together 20–30+ cases, which is a very different kind of research than doing a Google search." E-mail from Daniel Lewis, Chief Exec. Officer, Ravel Law, to author (Mar. 31, 2016, 15:18 MST) (on file with author). For this *one* query, the visualization map was not that helpful without further refinement.

about how each page has been cited.⁹⁷ (Note: In June 2017, LexisNexis announced that it had acquired Ravel Law.⁹⁸)

Casetext

¶27 Casetext's relevance algorithm is a function of key word frequency, citation count, date, and jurisdiction.⁹⁹ The user interface lists the results by relevance, although the researcher can re-sort by date or citation count. There are tabs that lead the researcher to other texts, regardless of jurisdiction, and to organizations, communities, and posts.¹⁰⁰ Once the researcher opens a relevant case, Casetext provides "Summaries from Subsequent Cases" and "Key Passages from this Case." The summaries are parentheticals, showing how the case has been paraphrased by later judicial opinions. Key passages are extracted important language from the case, together with the number of times the extract has been cited by later courts.¹⁰¹ Casetext does not provide a filter for jurisdiction.¹⁰²

97. RAVEL, *supra* note 93, at 7. Ravel Law Case analytics show how pages within a case have been cited: the left column within a case shows how each page in an opinion has been cited by later cases, and the citations are grouped when they discuss a similar principle of law. The more citations a page has, the more stars appear next to the page number. Ravel results are the same whether you filter before or after running the search. Here are the instructions given to research assistants to limit results:

- At the home page, enter your search (e.g., agency follows clear congressional intent).
- In Jurisdictions, start typing your court (e.g., Eastern District of Michigan).
- Do not collect unreported cases or Supreme Court cases.

Note that the results are the same whether you prefilter or postfilter. If you limit jurisdiction first, you get 216 cases. If you search before filtering, the search returns 26,483 cases, but after postfiltering for jurisdiction, the results are the same: 216 cases.

98. News Release, LexisNexis Announces Acquisition of Ravel Law (June 8, 2017), <https://www.lexisnexis.com/en-us/about-us/media/press-release.page?id=1496247082681222> [<https://perma.cc/8PND-N2VK>].

99. E-mail from Pablo Arredondo, Vice President, Legal Research, Casetext, to author (June 12, 2016, 17:33 MST) (on file with author).

100. In this *one* query, the top post results were not relevant. The posts are crowd-sourced by lawyers, students, and librarians, and as such, are a form of heteromation, where the labor of a group is free and is used to enhance an automated experience. See Hamid Ekbia & Bonnie Nardi, *Heteromation and Its (Dis)contents: The Invisible Division of Labor Between Humans and Machines*, FIRST MONDAY, June 2, 2014, <http://firstmonday.org/ojs/index.php/fm/article/view/5331> [<https://perma.cc/DT45-GLS6>] (discussing an experiment using a free online community to solve folded protein puzzles, called FoldIt).

101. Hannah Doherty, *Master Case Law in Just 5 Steps with Casetext Pro*, CASETEXT (Mar. 23, 2016), <https://casetext.com/orgs/casetext/posts/mastering-case-law-in-just-5-steps> [<https://perma.cc/Q6KY-VLLW>]. Step 1 is understanding the law, which is supported by the summaries; step 2 is finding the important key passages, which is supported by the passages and a heatmap that shows how frequently other court opinions and articles cite to each case in the opinion. For a detailed discussion of the utility of parentheticals to understand the evolution of a case over time, see Pablo D. Arredondo, *Harvesting and Utilizing Explanatory Parentheticals*, 1 LEGAL INFO. REV. 31 (2015–2016).

102. Here are the search instructions given to research assistants to limit results by jurisdiction in Casetext:

- At the home page, enter your search (e.g., agency follows clear congressional intent).
- Under Jurisdictional Filters, open the plus sign, select US District Courts, and then select your court (e.g., Eastern District of Michigan).
- There is only postfiltering in Casetext. Filter out unreported cases.
- Filter out any Supreme Court cases.

Google Scholar

¶28 Google Scholar's case law database is meant to provide the general public with access to the law.¹⁰³ According to its "About" page, "Google Scholar aims to rank documents the way researchers do, weighing the full text of each document, where it was published, who it was written by, as well as how often and how recently it has been cited in other scholarly literature."¹⁰⁴ A study was done of Google Scholar's ranking algorithm, and the authors believe that Google Scholar gives the most weight to citation counts.¹⁰⁵ The occurrence of a search term in the title is also important, but search term frequency in the full text does not seem to impact the ranking.¹⁰⁶ Google Scholar does not search for synonyms.¹⁰⁷ When limiting a search by jurisdiction, the results are the same whether filtering before or after searching.¹⁰⁸ When entering a query into the search box for Google Scholar, the page with the results of the search is just a list of cases. A researcher has to look at an individual case to get any links to other resources, such as the "how cited" function, which will take the researcher to other cases and related documents.¹⁰⁹

Cost

¶29 The cost of these six databases varies widely. Without discussing actual cost, it is safe to say Westlaw and Lexis Advance are the most expensive options; Fastcase is a low-priced option; the cost of Ravel's¹¹⁰ publicly accessible resources is currently unknown; Casetext¹¹¹ is free, although new additions to the database, like CARA, require payment; and Google Scholar is free.

103. Dr. Anurag Acharya, Distinguished Engineer, Google, Presentation at AALL Annual Meeting & Conference: Searching Legal Opinions: The Google Scholar Approach (July 22, 2012) (session recording on file with author).

104. *About Google Scholar*, GOOGLE SCHOLAR, <https://scholar.google.com/intl/en/scholar/about.html> [<https://perma.cc/UN94-R8TP>].

105. Joeran Beel & Bela Gipp, *Google Scholar's Ranking Algorithm: An Introductory Overview*, in 1 PROCEEDINGS OF THE 12TH INTERNATIONAL CONFERENCE ON SCIENTOMETRICS AND INFORMETRICS (ISSI'09) 230, 230–32, 241 (Birger Larsen & Jacqueline Leta eds., 2009), <http://docear.org/papers/Google%20Scholar's%20Ranking%20Algorithm%20--%20An%20Introductory%20Overview%20--%20preprint.pdf>.

106. *Id.*

107. *Id.*

108. Here were the search instructions given to the research assistants to limit results to a specific jurisdiction:

- On the home page, enter your search (e.g. agency follows clear congressional intent).
- Click on the case law radio button and then click on the "Select courts" link.
- Under the Sixth Circuit, click the ED Michigan box and then scroll up or down and click the DONE button (652 results).

The number of results is the same if you limit by court first, click DONE, and then go back and enter your search terms. Note that Google Scholar does not include unreported cases.

109. For the results of the *one* query, there were no secondary sources in the first few pages of results that listed in "related documents."

110. Lexis acquired Ravel on June 8, 2017. Robert Ambrogi, *Harvard's Statement On Ravel Law's Acquisition by LexisNexis Confirms Continued Public Access to Cases*, LAW SITES (June 9, 2017), <https://www.lawsitesblog.com/2017/06/harvards-statement-ravel-laws-acquisition-lexisnexis-confirms-continued-public-access-cases.html> [<https://perma.cc/W86U-ZD8V>]. Ravel and Harvard state that they are committed to keeping the case law free. *Id.* It is no longer clear whether free access will include visualizations and annotations, as was the case before acquisition.

111. Casetext's case search and annotations are free, and some other enhancements are currently free. CARA, Casetext's brief analyzer, is a subscription product. *CARA Research Suite*, CASE-TEXT, <http://casetext.com/pricing> [<https://perma.cc/6DXM-U3KL>].

The Empirical Study

Methodology

¶30 One of the unique things about legal databases is that you can set up a sandbox of nearly identical sets of information. By limiting searches to the subset of reported cases within a specific jurisdiction and using the same search terms in each database, it is possible to compare search results in a nearly identical group of documents.¹¹² If, for example, a search is executed in a database of all reported cases from the Northern District of California, there should be, in theory, an identical set of documents regardless of which legal database is searched. In actuality, a very small margin of potential difference occurs because each database may have a slightly different start date for coverage.¹¹³

¶31 The current study arose out of a single search prepared for a presentation on algorithms in 2013.¹¹⁴ The same search entered into different databases produced starkly varied results. For the presentation, the author looked at a single keyword search—the right to receive information¹¹⁵—across four databases. The results are shown in table 1.

¶32 Cases marked in light grey are unique and relevant,¹¹⁶ while cases that are not colored are unique and not relevant. Of the forty cases shown, seventy percent

112. The coverage in each of the databases studied is as follows:

- Casetext has published federal circuit and district court cases from 1925 to present. *Search Queries*, CASETEXT, <https://casetext.com/search-queries> [<https://perma.cc/KD8T-96SQ>].
- Ravel has published circuit court cases from 1925 and published district court cases from 1933. *What Does Ravel's Case Coverage Include?*, RAVELAW, <https://ravellaw.zendesk.com/hc/en-us/articles/212634578-What-does-Ravel-s-case-coverage-include-> [<https://perma.cc/JPG8-BRQS>].
- Fastcase has published circuit court cases from 1924 and published district court cases from 1932. *Scope of Coverage*, FASTCASE, <http://www.fastcase.com/coverage/> [<https://perma.cc/RP4Q-WSXU>].
- Google Scholar has published circuit and district court opinions from 1923 to current. *Google Scholar Search Tips*, GOOGLE SCHOLAR, <https://scholar.google.com/intl/en/scholar/help.html#coverage> [<https://perma.cc/G8CF-39WG>].
- Lexis Advance has published circuit and district court cases from 1791 forward, as jurisdictions are added. *LEXISNEXIS, LEXIS ADVANCE CONTENT LISTING 3–5* <https://web.lexis.com/help/research/ContentListing.pdf> [<https://web.archive.org/save/embed/https://web.lexis.com/help/research/ContentListing.pdf>].
- Westlaw has published circuit courts cases from 1891 to current. *U.S. Courts of Appeals Cases*, WESTLAW, <https://1.next.westlaw.com/Browse/Home/Cases/USCourtsOfAppealsCases> (click on “i” button) (last visited July 4, 2017). Westlaw’s coverage for district court cases is 1779 to current, as jurisdictions have been added. *Federal District Court Cases*, WESTLAW, <https://1.next.westlaw.com/Browse/Home/Cases/FederalDistrictCourtCases> (click on “i” button) (last visited July 4, 2017).

113. *Id.* All database providers have coverage for all federal cases since 1933; since only two cases in the three thousand cases reviewed in the project were earlier than 1933, minor differences in the earliest years of coverage may be deemed a matter of little importance.

114. Susan Nevelow Mart et al., Panel discussion at the Association of American Law Schools 2013 Annual Meeting: Understanding Search Engine Algorithms: Can We Effectively Teach Legal Research Without Them? (Jan. 6, 2013) (session notes on file with author).

115. The right to receive information is a small but well-defined concept in constitutional law. See Susan Nevelow Mart, *The Right to Receive Information*, 95 LAW LIBR. J. 175, 2003 LAW LIBR. J. 11.

116. “Relevance” is a highly disputed term. Relevant for the purpose of the searches discussed in this article means that, measured against a statement of relevance given for each search, the

Table 12013 Comparison *right to receive information*

Lexis Advance	Fastcase	WestlawNext	Google Scholar
Celotex v. Catrett	Alderson v. US	Bd. of Educ. v. Pico	Virginia State Bd. of Pharmacy
Anderson v. Liberty Lobby	Fuller v. Dep't of Navy	Virginia St. Pharm. Bd. v. Virginia Consumers	Stanley v. Georgia
Branzburg v. Comm'r	Aitken v. Comm'r	Stanley v. Georgia	Martin v. Struthers
Houchins v. KQED	Gal-Or v. US	Martin v. EPA	Griswold v. Connecticut
US v. Polizzi	Warren Browne v. IRS	Neinast v. Bd. of Trustees	Kleindienst v. Mandel
Kreimer v. Bd. of Police	US v. Whitnery	Kreimer v. Bd. of Police	Red Lion v. FCC
Virginia St. Pharm. Bd. v. Virginia Consumers	Doe v. City of Albuquerque	Doe v. City of Albuquerque	Bd. of Educ. v. Pico
Santa Monica Pictures v. Comm'r	Texas Medical Providers v. Lakey	Sund v. City of Wichita Falls	Procurier v. Martinez
San Antonio Sch. v. Rodriquez	US v. Morriss	Cline v. Fox	Lamont v. Postmaster Gen.
Columbia Broadcasting System	US v. Holmes	Kleindienst v. Mandel	Pell v. Procurier

(twenty-eight cases) are unique, while forty-two percent are both relevant and unique. Cases colored dark grey appear in more than one database; seven of the cases are in two results sets. One case is in three databases. No case is in all four databases.¹¹⁷ Three takeaways result from this effort:

1. All four databases included irrelevant results in the top ten results.
2. Seventy percent (28 of 40) of the cases were unique to one database.
3. Of those unique cases, slightly more than half (16 of 28) were both relevant and unique.

¶33 At least for one search, every algorithm was offering unique and relevant cases not returned in the top ten results by the other databases. So every algorithm had something interesting to add to the legal construct created by the searches being entered into the database. Knowing that algorithms do not remain the same and that later-decided cases will be added to the database, the identical search was run in the same four databases in 2016. The results are shown in table 2.

case would or most likely would be put in a review pile for more detailed examination; that is, the case has the potential to be helpful to a research project on this legal subject. The statement of relevance for this search was “You are looking for cases that discuss factual situations involving the right to receive information as a listener or recipient of information.” For a more detailed discussion of relevance, see *infra* ¶¶ 47–48.

117. In 2013, eighty percent of the LexisNexis results were relevant by my standards. But only ten percent of the Fastcase results were relevant. The natural language search information in Fastcase says if you type in words, “You will get cases that best match the words and phrases in the query,” but many of the cases in the results did not have the phrase “right to receive information.” The algorithm was not privileging proximity in a way that would return cases first that had those four words right next to each other.

Table 2*2016 Comparison right to receive information*

Lexis Advance	Fastcase	WestlawNext	Google Scholar
Illinois v. Gates	Morris v. Equifax	Stanley v. Georgia	Stanley v. Georgia
Branzburg v. Hayes	Wis. Rt. to Life v. Barland	Bd. of Educ. v. Pico	Virginia St. Pharm Bd. v. Virginia Consumers
Virginia St. Pharm. Bd. v. Virginia Consumers	Hull v. IRS	Virginia St. Pharm. Bd. v. Virginia Consumers	Martin v. Struthers
Kreimer v. Bd. of Police	Sorrell v. IMS Health, Inc.	Doe v. Governor of N.J.	Griswold v. Connecticut
Bd. of Educ. v. Pico	Wood v. Ryan	Martin v. EPA	Kleindienst v. Mandel
Gregg v. Barrett	Essential Info. v. USIA	Doe v. City of Albuquerque	Bd. of Educ. v. Pico
Family Trust v. Wolnitzek	Info. Resources v. US	Neinast v. Bd. of Trustees	Red Lion v. FCC
Stanley v. Georgia	McKesson v. Bridge Med.	N.Y. Times v. Sullivan	Procunier v. Martinez
Paris Adult Theater v. Slaton	Sense v. Shinseki	Kreimer v. Bd. of Police	Pell v. Procunier
Kleindienst v. Mandel	Weaver v. USIA	Near v. Minnesota	Lamont v. Postmaster Gen.

¶34 LexisNexis now has just one irrelevant result, but only three of the cases from 2013 show up here. Seven cases are new, even though all of the new cases existed in the database when the first search was performed.¹¹⁸ Fastcase now has two relevant results, but all ten results are new, and only two of those new results were not available in the database when the first search was done.¹¹⁹ Westlaw now has two irrelevant cases, and three cases are new.¹²⁰ All of Google Scholar's cases are the same, but the order in which they were returned changed. The Google Scholar algorithm seems to have been worked on the least. This chart illustrates very clearly that search results change over time by more than the mere addition of new cases to the database. The percentages of unique cases—sixty-seven percent (27 of 40)—remained about the same over time, while the percentage of cases that were both relevant and unique went up slightly—sixty-six percent (18 of 27). But the results from just one search, while provocative, are not statistically significant.

¶35 The study was expanded to fifty different searches. Many of the searches were taken from the author's previous study of digests and citators,¹²¹ and new searches were generated by the author's random reviews of current law review articles for legal concepts that might make a good search. Each search had to turn up at least ten

118. None of the new cases are more recent than 2013, when the first search was performed.

119. Of those ten new results, two are dated 2014, after the first search was performed.

120. One of the three new cases was decided in 2015, after the first search was performed.

121. 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).

results in each of the six legal databases, so that the research assistants could compare the top ten results from each database. Not every search worked in every database, and not every jurisdiction returned sufficient results in some of the databases.¹²² It is possible that focusing on cases with a robust search history, as was done for this study, introduces some bias of its own, but this was unavoidable.¹²³

¶36 The searches were all simple keyword searches, and each was crafted to include multiple words and at least one legal phrase or legal concept and to work in each legal database.¹²⁴ Because not all databases seem to recognize legal phrases with the same consistency, this may have introduced some bias in favor of algorithms that recognize more legal phrases without quotes. But researchers intuitively expect that cases that have the words next to each other will, because of proximity, be returned in the top results. Algorithms that do not privilege proximity to the extent that a case with the four words from the search right next to each other are returned before results with those same four words scattered through the text are not meeting researchers' expectations.¹²⁵ The searches were designed to be medium-good¹²⁶ starting searches. For an actual research problem, of course, if the results from the first search were disappointing, the researcher would refine the search, adding or changing words and word order, or trying different resources. One hopes no researcher would stop with one search or one resource.

¶37 Here are three of the searches used in the study:

- special relationship constitutional duty protect public from crime (N.D. Cal.)
- job performance racial classification constitutional (D.D.C.)
- administrative search 4th amendment warrant requirement (S.D.N.Y.)¹²⁷

122. Research assistants were asked to run the searches and find the number of results in each database prior to entering any data into the spreadsheet or reading cases to determine relevancy. To make the comparison of results consistent, each search had to return at least ten results in the chosen jurisdiction. If the original query did not return at least ten results in each of the six legal databases, first the search was tried in other, larger jurisdictions. If that did not work, a new variation of the query was crafted. If the query did not work in any jurisdiction, the query was discarded and a new query crafted. Of the fifty queries originally created, the jurisdictional case database was changed fifteen times, three search queries were revised, and two queries were discarded.

123. The statistical analysis performed on the fifty queries in six different legal databases for the top ten queries (3000 cases) is descriptive statistics. See *Analysis: Descriptive Statistics*, RES. METHODS KNOWLEDGE BASE, <http://www.socialresearchmethods.net/kb/statdesc.php> [<https://perma.cc/WJ7B-6U22>]. Regarding the selection of cases, see also Mark A. Hall & Ronald F. Wright, *Systematic Content Analysis of Judicial Opinions*, 96 CALIF. L. REV. 63, 105 (2008) ("The goal in selecting cases is not a perfect match between sample frame and research conclusions, but only a reasonable connection between the two. Inevitable imperfections in case selection methods often will not seriously threaten the entire validity of the study's findings. It usually suffices to acknowledge limitations fairly briefly."). The data sets for this study are available at <http://scholar.law.colorado.edu/research-data/6/> (data files).

124. For example, just adding quotes around a legal phrase will improve most searches. However, at the time the study was designed, one could not add quotes in Westlaw without adding **adv.**, and that changes the search to a Boolean search. Since the searches had to be identical in each database, a search without quotation marks worked in all databases.

125. This expectation is based on years of discussions the author has had with legal research students about searching. This expectation may not be correct, but then researchers need to know that. This is another instance where algorithmic accountability would be useful. It is also possible that databases that return cases where the words are scattered are not using citation analysis as one of their forms of relevancy ranking.

126. Medium-good means a little thought has gone into creating a search, but not a lot.

127. See Appendix B, *supra* note 78, for the text of all fifty searches.

Where the search was “special relationship constitutional duty protect public from crime,” the student coder was told: You are looking for cases where, despite the fact that state officials normally have no constitutional duty to protect the public at large from crime, the duty is (or is not) imposed by virtue of a special relationship between state officials and a particular member of the public (you are looking for the factual contours of a special relationship).

¶38 For the search “job performance racial classification constitutional,” the research assistant was told: You are looking for cases that discuss situations where job performance is or is not related to race (parameters of acceptable racial classifications for work). And for the search “administrative search 4th amendment warrant requirement,” the coder was told: You are looking for cases about administrative searches and whether or not the search requires a 4th amendment warrant.

¶39 These instructions set the stage for relevance determinations. If a case could be helpful to the legal construct in the statement of relevance in any way, it would be coded as very likely or likely relevant enough to go into the pile for later, more thorough review. Cases that seemed not to be relevant were either very likely or likely to go into the discard files; these were the cases that would not need later review. This expansive view of relevance was meant to ensure that cases that might work by analogy would be included as relevant. While this is certainly a subjective view of relevance, it is the way lawyers actually do a quick review of cases, mentally sorting results into helpful or not helpful to the legal problem. Stuart Sutton sees this as the creation of mental models of an area of law and as a basic determination lawyers make; in his view, “[a] relevant case is one that plays some cognitive role in the structuring of a legal argument.”¹²⁸ Even though the study has tried to define relevance in the most expansive way, human coding of relevance has its own biases. Additionally, percentage of relevant documents may not be the best measure of relevance. For example, if the top ten results from one database have only two relevant cases, but those two are the most relevant in that area of the law, that might be a better result for the research than a search that returns eight relevant documents but misses those top two most relevant cases. To determine relevance at that level of granularity requires subject expertise in each specific legal domain related to each search, which was not possible for this study.

¶40 The decision to limit the review to the top ten results was based on several factors. One factor was the sheer amount of time it takes to review cases. Fifty queries in six legal databases and ten results per database is three thousand cases to review. The ten-result limit also fits in with actual user studies of the primacy of the top ten results in searchers’ behavior: the default page view on Google, for example, is ten results, and studies have found that researchers who are used to online searching will usually stop reading after the top ten results¹²⁹ and try

128. Stuart A. Sutton, *The Role of Attorney Mental Models of Law in Case Relevance Determinations: An Exploratory Analysis*, 45 J. AM. SOC’Y INFO. SCI. 186, 187 (1994).

129. Jessica Lee, *No. 1 Position in Google Gets 33% of Search Traffic [Study]*, SEARCH ENGINE WATCH (June 20, 2013), <https://searchenginewatch.com/sew/study/2276184/no-1-position-in-google-gets-33-of-search-traffic-study> [<https://perma.cc/A6TC-L7SF>] (showing that page 1 results (top ten) received ninety-two percent of all traffic); Daniel E. Rose & Danny Levinson, *Understanding User Goals in Web Search*, in PROCEEDINGS OF THE 13TH INTERNATIONAL CONFERENCE ON WORLD WIDE WEB 13 (2004), <http://dl.acm.org/citation.cfm?id=988675> (documenting a decrease in willingness to look at more than one page of search results). In academic research, researchers are noticing

another search¹³⁰ or another legal resource.¹³¹ One would expect in any event that the goal for each legal database provider would be to present the most potentially relevant results in the top ten, and Lexis Advance explicitly defines “relevance” as a “document a user would expect to find in their results appears as one of the first five documents in a user’s results set.”¹³²

Hypotheses

¶41 There were three hypotheses for the study. The study was framed in the usual way, by stating a null hypothesis and then testing to see whether it was proved or disproved.¹³³ One null hypothesis was that, because the search algorithm for each legal database was trying to achieve the same result in the same pool of information by finding relevant cases, each algorithm would find the same cases. Another null hypothesis was that, because the algorithms all rank relevance, and the goal is to return relevant cases, the top ten cases would all be relevant. The last null hypothesis was that the research assistants coding the results for the statistical study would not agree on relevance. To test the last hypothesis first, in addition to the ten queries each research assistant reviewed across all six legal databases, each of the five research assistants reviewed five of the same queries, chosen with a random number generator. Using standard tests, the research assistants had moderate concordance.¹³⁴ More than moderate concordance can probably not be expected of legal reviewers, as ranking relevance is a highly subjective task even when constrained by the parameters set out for the research assistants in this study.¹³⁵

“horizontal information . . . skimming . . . where people view just one or two pages from an academic site and then ‘bounce’ out.” Ian Rowlands et al., *The Google Generation: The Information Behaviour of the Researcher of the Future*, 60 ASLIB PROC. 290, 294 (2008), <http://www.emeraldinsight.com/doi/pdfplus/10.1108/00012530810887953>. But see Gerard de Melo & Katja Hose, *Searching the Web of Data*, in PROCEEDINGS OF THE 35TH EUROPEAN CONFERENCE ON ADVANCES IN INFORMATION RETRIEVAL 869, 869 (Pavel Serdyukov et al. eds., 2013), <http://dl.acm.org/citation.cfm?id=2458300> (positing that searchers don’t want “10 blue links,” they want “actionable information”).

130. In fact, a recent study showed that researchers using online databases to solve an ill-structured legal research problem used multiple searches; one researcher entered ten search strings in seven minutes in an attempt to get to a case on a legal principal she thought of. Stefan H. Krieger & Katrina Fischer Kuh, *Accessing Law: An Empirical Study Exploring the Influence of Legal Research Medium*, 16 VAND. J. ENT. & TECH. L. 757, 775, 778 (2014).

131. On the principle of least effort in information theory, see Marcia J. Bates, *An Introduction to Metatheories, Theories, and Models*, in THEORIES OF INFORMATION BEHAVIOR 1, 4 (Karen E. Fisher et al. eds., 2005). For an example of the principle of least effort at work, see Scott A. Moss, *Bad Briefs, Bad Law, Bad Markets: Documenting the Poor Quality of Plaintiffs’ Briefs, Its Impact on the Law, and the Market Failure It Reflects*, 63 EMORY L.J. 59 (2013).

132. LexisNexis, *supra* note 83.

133. Regarding the null hypothesis, see CHARLES WHEELAN, *NAKED STATISTICS: STRIPPING THE DREAD FROM THE DATA* 148–50 (2013). To put the issue of null hypotheses in a simple legal context, in a criminal trial, the null hypothesis is that the defendant is presumed innocent; to reject the null hypothesis, the jury must find that the defendant is guilty “beyond a reasonable doubt.”

134. The tests used were Krippendorff’s alpha and an intraclass correlation. There was moderate concordance for all five raters based on Krippendorff’s alpha (.50) and an intraclass correlation of .55.

135. See also Jeffrey T. Luftig, *Statistical Analysis of the Data, Susan Nevelow Mart Study of Search Functions in Lexis and Westlaw*, <https://dspace.library.colostate.edu/handle/10974/12902?show=full> [https://web.archive.org/save/_embed/https://dspace.library.colostate.edu/bitstream/handle/10974/12902/StatisticalAnalysisDataUploadVersion.pdf?sequence=6&isAllowed=y], for a positive concordance finding for the research assistants making legal document relevance determinations, where the concordance results were similar to the findings in this study.

¶42 The importance and immutability of the relevance rankings in this study must not be overstated. The relevance determinations were subjective, were constrained by the state of each algorithm at the exact time the research was performed, and were further limited by the precise legal problem posed by the statement of relevance for each query. Different researchers, using the same search terms, could be trying to solve a slightly different legal problem and could reach much different determinations of relevance. What can be concluded from the data the study used is that at the time the queries were run, with the human, algorithmic, and legal constraints that then existed, the relevance rankings of the raters give an accurate although subjective snapshot of relevance for the six databases.

¶43 In addition to a relevance ranking for each case, the research assistants also noted the name of the cases returned, the jurisdiction that was being searched, the date of the case, the number of results returned by the search for each database, and the number of databases that case citation was found in (ranging from unique and only in one database to found in all six databases). The determinations that were made from this raw data are discussed in the next section.

¶44 While each search within a jurisdiction had to return at least ten results to qualify for the study, the range of cases returned by query in each of the databases was large. As an example, the range of the number of cases returned in the results for Query 1, using the same search terms in the same jurisdictional case database, was 123, 909, 1730, 1197, 677, and 25.¹³⁶

Results

Uniqueness of Cases

¶45 In the study of one query, when the search was performed, there was very little overlap¹³⁷ in cases in the results for the four databases tested.¹³⁸ The first null hypothesis was that when there was a large number of searches, that result would not hold true. Because the search algorithm for each legal database was trying to achieve the same result in the same pool of information by finding relevant cases, the algorithms would find the same cases. As chart 1 illustrates, the null hypothesis was disproved by the study, and each algorithm returns an average of forty percent unique cases in its search results.

¶46 The percentage of cases in each category is very consistent across all of the searches. An average of 25 percent of the cases are in only two databases. An average of 15 percent of the cases appeared in three databases, while an average of 9.5 percent of the cases appeared in four databases. Slightly less than 7 percent of the cases appear in five databases and in six databases. So each group of human engineers is solving the search problem in very different ways and illustrating that each algorithm has something interesting to say about what a searcher is looking for, but not the same interesting things. A fair percentage of cases from each database provider will give a researcher a unique set of cases to look at.

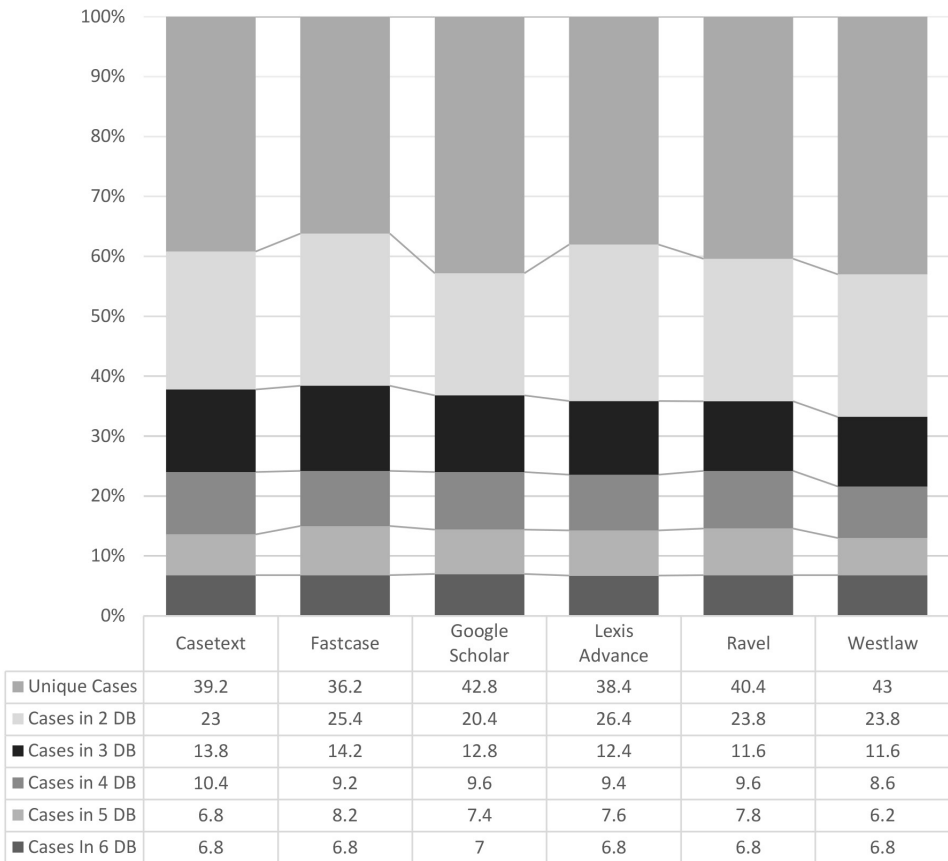
136. See chart 1 for a full representation of the data on the number of cases returned. Query 1 can be found in Appendix B, *supra* note 78.

137. Overlap in this comparison is the occurrence of a specific case in two or more databases.

138. See *supra* ¶ 33.

Chart 1

The Uniqueness of Cases in Each Database



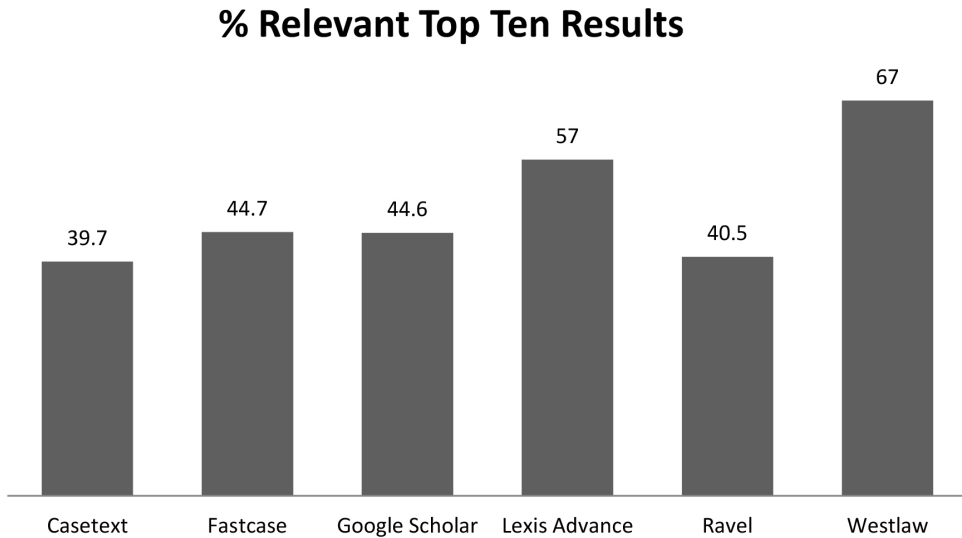
Relevance

¶47 The next null hypothesis was that because the algorithms all rank relevance, and the goal of each algorithm is to return relevant cases, the top ten cases would all be relevant. That hypothesis was disproved. Chart 2 illustrates how many of the results found in the top ten results for each database were relevant by the standards set out in my hypotheses.¹³⁹ Recall that the standard for relevance was subjective but expansive. There were more relevant results for the venerable legal database providers LexisNexis and Westlaw, at fifty-seven percent relevance for Lexis Advance and sixty-seven percent for Westlaw; the newcomers Casetext, Fastcase, Google Scholar, and Ravel were clustered together near forty percent relevance.

139. See *supra* ¶¶ 41–44.

Chart 2

The “Relevance” of the Top Ten Cases in Each Database

**Table 3**

The “Relevance” of the Top Ten and Top Five Cases Compared

Database	Top Ten	Top Five
Casetext	39.7	45.0
Fastcase	44.7	49.2
Google Scholar	44.6	46.0
Lexis Advance	57.0	64.4
Ravel	40.5	43.8
Westlaw	67.0	77.6

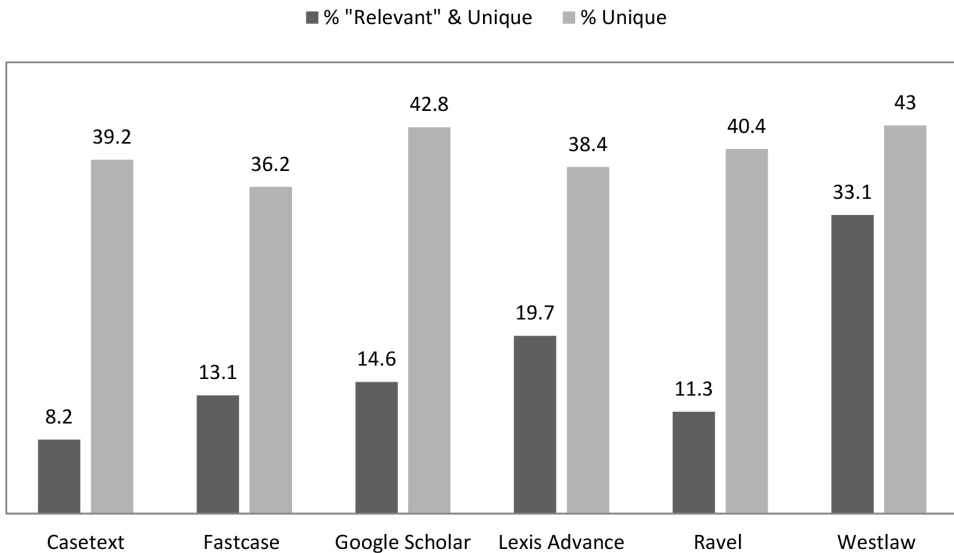
¶48 If you take a look at only the top five results, the number of cases Lexis Advance posits should all be relevant,¹⁴⁰ the percentages of cases that are relevant increased slightly for every database, but no one achieved one hundred percent relevance for those cases (see table 3).

140. LexisNexis, *supra* note 83.

Chart 3

The “Relevance” of the Top Ten Cases in Each Database

Percentage Relevant and Unique

*Relevant and Unique*

¶49 As chart 2 illustrates, each database has an average of forty percent unique cases in the top ten results, and the remaining cases have little overlap. That means that each database is providing a significant number of unique results. One question a researcher might want to know is, of those unique results that each database is providing, how many are relevant?

¶50 Chart 3 shows that Westlaw returns the most relevant results in the unique category. Although there is a diminishing payoff for the remainder of the databases, each unique relevant case is one more opportunity to find a “relevant case . . . that plays some cognitive role in the structuring of a legal argument.”¹⁴¹ The idea that every database has an individual worldview of cases, classification systems, and commentary that it mines for relevant cases, and that therefore each database’s algorithms return unique, relevant cases that may contribute to solving a legal problem that is not fully resolved by searching in only one database, is not an easy concept to communicate to novice researchers. The best way for researchers to internalize the concepts is to figure it out on their own, and Appendix A provides a sample problem for students to illustrate the work of algorithms in legal databases. Since practicing lawyers do not have access to all resources, knowing how to design multiple strategies and searches in the resources one has access to is also a lesson novices need to learn, but it is a lesson legal research professors are used to imparting.

141. Sutton, *supra* note 128, at 187.

Number of Results Returned by a Query

¶51 The study also looked for any statistical correlation between the number of results returned and relevance. For each query in a specific database, the number of results is the same, but since the relevance rank given to each of the top ten cases may not be the same, an average of the relevance rankings across the top ten results was used to see whether relevance changed as the number of results changed. The number of cases returned by each database is reported in chart 4 at the 25th, 50th, and 75th percentiles.

¶52 Lexis Advance returns the highest number of cases, which is interesting when compared to the average relevance rating across most databases. In table 4, you can see that the average relevance of the top ten results stays fairly constant even when the number of results increases. For Lexis Advance, average relevance actually increases as the number of results increases.

Age of Cases

¶53 For the age of cases returned in each database, the following results show that Google Scholar returns the highest percentage of older cases, while Westlaw, Fastcase, and Casetext return the highest number of newer results. (See chart 5.) Age of cases did not correlate in any statistically significant way with relevance.

Discussion

¶54 There is no way to account for the higher percentages for Lexis and Westlaw and the clustering effect for the newer database providers, beyond speculation. It may be that the much greater investment in classification (as in key numbers and concomitant legal phrase recognition), mining of secondary sources, and leveraging machine learning from user search history gives Westlaw the greatest edge, as its largely human-generated classification system is the oldest.¹⁴² In an earlier comparison of the classification systems in Westlaw and LexisNexis, the human-curated system in Westlaw had an advantage over LexisNexis's largely algorithmically generated classification system that despite changes in each company's algorithms and interfaces still seems to make a difference.¹⁴³ LexisNexis's results are the second most relevant, and this may be because LexisNexis's algorithms utilize a topical classification system that it has been refining with machine learning since 1999,¹⁴⁴ a large collection of secondary sources, and, as the first online legal database, the longest repository of user history.¹⁴⁵ As Frank Pasquale has noted, having a large user base that contributes to the search algorithm with every search means that the "incum-

142. The West Key Number System began in the 1890s. Ross E. Davies, *How West Law Was Made: The Company, Its Products, and Its Promotions*, 6 CHARLESTON L. REV. 231, 234–35 (2012).

143. Mart, *supra* note 121, at 13, 16, 25–29, 59. In the study, comparing the results of Key Number searches (key numbers are largely human-generated) with Lexis Topics searches (topics are largely computer-generated), Westlaw's results were 61.7% relevant, while the Lexis Topic results were 36.7% relevant. *Id.* at 37.

144. *Id.* at 16.

145. LexisNexis became publicly available in 1973. For an interesting history of online access to legal opinions, see Gary D. Spivey, *Remembering James M. Flavin: The Origins (and Unintended Consequences) of Online Legal Research*, N.Y. ST. B. ASS'N J., Feb. 2008, at 11, 18.

Chart 4

Number of Results Returned by Each Search, by Percentile

25th, 50th, and 75th Percentile for Number of Results Returned

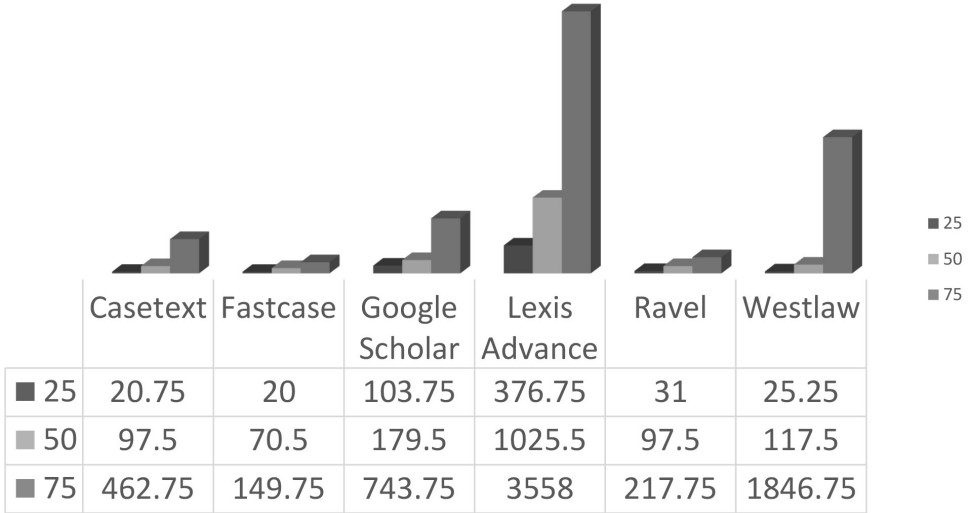


Table 4

Average Relevance Rating Within Number of Results Returned

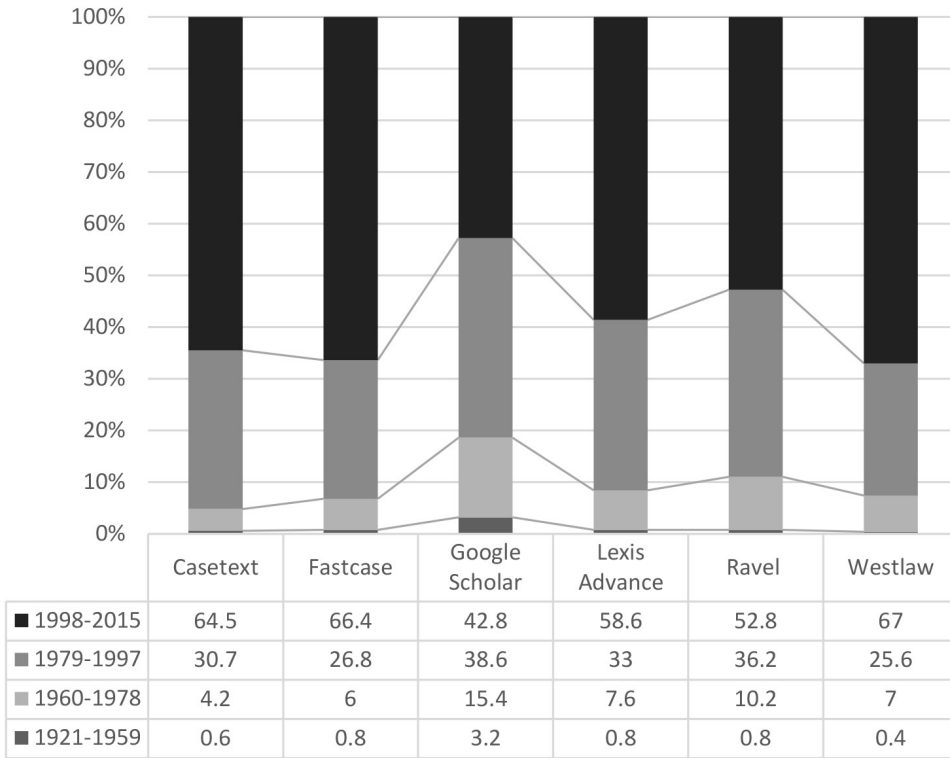
# Cases	Casetext	Fastcase	Google Scholar	Lexis Advance	Ravel	Westlaw
<100	2.2316	2.3896	2.3091	2.6	2.4692	3.1455
100-499	2.4	2.4583	2.45	2.55	2.1289	2.8857
500-999	1.95	1.9967	2.6429	2.4	1.72	3.2
1000-1999	2.04		1.86	2.8727	2.1	3.0167
>2000	1.425	2.05	2.16	2.96	1.3500	2.65

bents with large numbers of users enjoy substantial advantages over smaller entrants.¹⁴⁶ It is therefore not that surprising that the legal database providers with the largest user base and the longest search history to mine exhibit one kind of advantage over the newcomers. In terms of the general relevance of top ten results, there is an advantage to using the older providers’ algorithms.

146. Frank Pasquale, *Dominant Search Engines: An Essential Cultural & Political Facility*, in *THE NEXT DIGITAL DECADE: ESSAYS ON THE FUTURE OF THE INTERNET* 401, 411-12 (Berin Szoka & Adam Marcus eds., 2010).

Chart 5

Age of Cases in Each Database



¶55 In addition, the West classification system and the LexisNexis classification system reflect a nineteenth-century worldview.¹⁴⁷ The classification systems are not identical,¹⁴⁸ of course, and Westlaw and LexisNexis each has a unique set of sec-

147. The asserted hegemony of the West worldview is thoroughly discussed in Robert C. Berring, *Legal Research and the World of Thinkable Thoughts*, 2 J. APP. PRAC. & PROCESS 305 (2000). But see Joseph A. Custer, *The Universe of Thinkable Thoughts Versus the Facts of Empirical Research*, 102 LAW LIBR. J. 251, 2010 LAW LIBR. J. 14. Of course many lawyers did not use legal publications that incorporated the West classification system. But to the extent that the Langdellian method of teaching law recreates a similar classification decade after decade, generations of law students have parsed out the levels of classification in, for example, the formation of a contract, in ways very similar to the West system. Compare CLAUDE D. ROHWER & ANTHONY M. SKROCKI, *CONTRACTS IN A NUTSHELL*, at ix–xxvii (7th ed. 2010), with WEST’S *ANALYSIS OF AMERICAN LAW: GUIDE TO THE AMERICAN DIGEST SYSTEM* 370–82 (2015). Although the exact outlines differ, the subject matter is broken down into similar patterns of essentials for formation, interpretation, performance, and defenses or breach. So is the Topic outline for contracts in LexisNexis, where Topics to look at include formation, condition, performance, interpretation, breach, and defenses. Whether the worldview is based on the West classification system itself or the Langdellian worldview that older classification systems reflect, newer legal research databases may be freer of whatever limitations that worldview imposes.

148. Mart, *supra* note 121, at 18–21.

ondary sources for its algorithms to mine. The classification differences and the differing set of secondary sources voices that contribute to search results lead to two possible kinds of viewpoint discrimination. The first is one law librarians have long dealt with in acquiring treatises and secondary sources for law library collections. Budgets allowing, librarians want more than one authorial viewpoint in their collections because a treatise is only *one* author's view of the law; it is not the law. Since the different authorial viewpoints provided by the very different list of secondary sources in Westlaw and Lexis Advance are baked into their respective search results, it is not surprising that the results from Lexis Advance and Westlaw are different. Every database has about forty percent relevant cases. Of those, Westlaw has thirty-three percent relevant and unique cases in the search results while Lexis Advance has twenty percent.¹⁴⁹ It is possible that those different relevant results reflect the different classification systems and secondary sources. So long as researchers have to deal with databases that import viewpoints into their algorithms, it seems to be a positive that each offers results based on differing classification and authorial viewpoints.

¶56 The second kind of viewpoint discrimination is one we don't think about that much, and that is the nineteenth-century worldview of the legal system explicitly embedded in Westlaw's Key Numbers and in Lexis Advance's Topics. These classification systems, while not identical, follow a pattern that is familiar to anyone who has taken contracts in law school. It is firmly based in the Langdellian view of the world, where the subject matter is broken down into similar patterns of essentials for formation, interpretation, performance, defenses, and breach.¹⁵⁰ This view is a form of filtering, for better or worse, and the newer legal research databases may be freer of whatever limitations that worldview imposes. Over the years, legal researchers have complained that these older classification systems break down as new legal concepts emerge, and may actually impede research.¹⁵¹ The newer entrants into the legal markets may be offering, in their forty percent of cases that are unique, something outside the range of that old worldview, and with the value-added results that users see on the results page,¹⁵² they may be offering new forms of serendipity in search. Researchers looking for an alternative may gravitate toward database providers that offer search results from different worldviews of the legal universe.

149. See chart 3 *supra*.

150. See discussion ¶ 54 *supra*.

151. Hanson, *supra* note 30, at 569, ¶ 16 (noting that researchers in "developing areas of law such as civil rights sometimes have found the West digest system and other traditional research techniques to be more a hindrance than a help").

152. See *supra* ¶¶ 20–28 for a discussion of the value-added features each database provider presents to the researcher once a search has been executed.

Conclusion

¶57 This study produced several specific findings about the six legal databases studied. For results that return the largest number of more recent cases, researchers should turn to Casetext, Fastcase, and Westlaw. The highest percentages of relevant cases, as defined in this study, are found in Lexis Advance and Westlaw. Google Scholar has the most older cases in its results. And Lexis Advance returns searches with the most results. More generally, the study shows that every algorithm starts with a different set of biases and assumptions. Even for returning results from searches in a specific case database, every algorithm draws on a different set of sources and processes, whether those sources and processes are classification systems, secondary sources, citation networks, internal case analyses, mined user search history, or machine learning deployed in the unique environment each legal database provider offers. These algorithmic variations in worldview lead to substantial variations in the unique and relevant results each database provides. The knowledge of this variability expands the opportunities for researchers to find relevant cases that can play “some cognitive role in the structuring of a legal argument.”¹⁵³

¶58 Legal information literacy requires lawyers to be “self-reliant in their investigations” of the law.¹⁵⁴ Legal research professors hope to teach their students to achieve the metacognitive skills required to be self-reliant. Black-boxing the research process is not helping educators or students achieve this goal. Algorithmic accountability will help researchers understand the best way to manipulate the input into the black box and be more certain of the strengths and weaknesses of the output. Asking for that kind of accountability can be successful. It was successful for doctors using Watson in oncology departments,¹⁵⁵ and, in response to requests, some of the database providers have made some of their assumptions much clearer.¹⁵⁶ This article is, in part, a call for each of us to request more accountability from database providers and for database providers to proactively think of algorithmic accountability as a way to improve research results for their users. It is as true now as it was in 1963, when Reed Lawler wrote *What Computers Can Do*: “If you ask the wrong question, you will get the wrong answer.”¹⁵⁷ We need more detailed information to ask the right questions. The answers will allow legal researchers to be the engaged humans they need to be when working with computer algorithms.

153. Sutton, *supra* note 128, at 187.

154. Dennis Kim-Prieto, *The Road Not Yet Taken: How Law Student Information Literacy Standards Address Identified Issues in Legal Research Education and Training*, 103 LAW LIBR. J. 605, 606, 2011 LAW LIBR. J. 37, ¶ 1.

155. See Friedman, *supra* note 54.

156. LexisNexisLawSchools, *supra* note 81.

157. Reed C. Lawler, *What Computers Can Do: Analysis and Prediction of Judicial Decisions*, 49 A.B.A. J. 337, 338 (1963). Lawler was a “pioneer investigator in the application of computer technology and modern logic to the law.” *Id.* at 337. He believed that only those “trained in the law have the skill for asking good legal questions. The computer scientist’s job is to translate the question into machine language.” *Id.* at 338. The human/machine teamwork Lawler foresaw continues to evolve.

- A. Are the cases that you found in Google Scholar relevant to your legal issue? Are they about your motion, your area of law, and your defense? If they are relevant, highlight them.
- B. Are the cases that you found in Lexis Advance relevant to your legal issue? Are they about your motion, your area of law, and your defense? If they are relevant, highlight them.
- C. Are the cases that you found in Westlaw relevant to your legal issue? Are they about your motion, your area of law, and your defense? If they are relevant, highlight them.
- D. How many of the cases are in only one database? ____ In two? ____ In all three? ____