Cite as: Annalee Hickman, *How to Teach Algorithms to Legal Research Students*, 28 Persp. 73 (2020).

How to Teach Algorithms to Legal Research Students

By Annalee Hickman

Annalee Hickman is the Faculty Services Librarian and a Legal Research Professor at the Howard W. Hunter Law Library, Brigham Young University.

I. Introduction

Recently, the role of algorithms in legal research has become a hot topic among legal research professors—namely how various legal research databases return different results from searching because of their unique algorithms. Because of how algorithms affect (and limit) electronic legal research, legal research students should be informed about their effects. Although many legal research professors have called for the inclusion of algorithms in instruction,[1] a specific and ideal curriculum for teaching algorithms to legal research students should be developed. While it is not necessary for law students, or even legal research professors, to be experts on how algorithms work and how they are created, a basic understanding of the role of algorithms in electronic legal research will serve any attorney well. This Article calls for legal research professors to include in their curriculum the role of algorithms in electronic legal research and the warning that different algorithms on different databases cause results to appear in varying relevancy orders. To assist, I include recommendations for readings, lecture content, and assignments. Each recommendation is based on my firsthand experience teaching algorithms to first-year law students over the past three years.

II. Readings

The leading expert in this area of algorithms in legal research is Susan Nevelow Mart. Mart has published a plethora of relevant pieces which can assist legal research students in understanding the role of algorithms in legal research.[2] The one I (still[3]) recommend is *Every Algorithm Has a POV*.[4] This article is shorter and feels more appropriate for a law student reader, as opposed to Mart's other articles, which are more thorough and better fits for a law librarian or attorney reader. E*very Algorithm Has a POV* provides a useful overview on the importance of understanding the issue of algorithms, and it offers "Mart's candid analysis of various search engine algorithms and users' ignorance about how these providers write their algorithms."[5] It "serves as an eye-opening warning to law students," emphasizing the importance of understanding algorithms and making

"This Article calls for legal research professors to include in their curriculum the role of algorithms in electronic legal research and the warning that different algorithms on different databases cause results to appear in varying relevancy orders."

1 *See, e.g.*, Jamie J. Baker, *2018: A Legal Research Odyssey: Artificial Intelligence as Disruptor*, 110 Law Libr. J. 5, 28 (2018) [hereinafter Baker, *2018: A Legal Research Odyssey*] ("It behooves law librarians to bring . . . issues surrounding the use of algorithms to light during legal research instruction."); Jamie J. Baker, *Beyond the Information Age: The Duty of Technology Competence in the Algorithmic Society*, 69 S.C. L. Rev. 557, 575–77 (2018) (sharing "practical tips for teaching competent use of algorithms"); Sherry Xin Chen & Mary Ann Neary, *Artificial Intelligence Legal Research and Law Librarians*, 21 AALL Spectrum, May/June 2017, at 16, 20 ("The time is ripe for law librarians to incorporate background knowledge of . . . database algorithms . . . into the legal research curriculum."); Iantha M. Haight, *Digital Natives, Techno-Transplants: Framing Minimum Technology Standards for Law School Graduates*, 44 J. Legal Pro. 175, 208 (2020) ("Law schools should cover the basics of . . . [a]lgorithms."); Nicholas Mignanelli, *Critical Legal Research: Who Needs It?*, 112 Law Libr. J. 327, 342 (2020) (Legal research professors "should use [their] pedagogy to instill in [their] students a healthy dose of skepticism about claims of objectivity and neutrality . . . in the context of technology."); Paul D. Callister, *Law, Artificial Intelligence, and Natural Language Processing: A Funny Thing Happened on the Way to My Search Results*, 112 Law Libr. J. 161, 204–05 (2020) ("Helping students see . . . categories of potential bias [of search algorithms] is something worthy of classroom discussion.").

2 *See, e.g.*, Susan Nevelow Mart, *The Algorithm as a Human Artifact: Implications for Legal Research*, 109 Law Libr. J. 387 (2017) [hereinafter Mart, *The Algorithm as a Human Artifact*]; Susan Nevelow Mart, *Every Algorithm Has a POV*, 22 AALL Spectrum, Sept./Oct. 2017, at 40 [hereinafter Mart, *Every Algorithm*]; Susan Nevelow Mart et al., *Inside the Black Box of Search Algorithms*, 24 AALL Spectrum, Nov./Dec. 2019, at 10; Susan Nevelow Mart, *Research Algorithms Have a Point of View*, 46 Colo. Law. 10 (2017); Susan Nevelow Mart, *Results May Vary in Legal Research Databases*, 104 A.B.A. J. 55 (2018).

3 *See* Annalee Hickman, *Engaging Legal Research Students Through Supplemental Readings from the Last Decade*, 26 Persps.: Teaching Legal Rsch. & Writing 65, 70 (2018).

4 Mart, *Every Algorithm*, *supra* note 2.

5 Hickman, *supra* note 3.

informed research decisions, since algorithms may impede their finding of relevant law.[6]

Some legal research textbooks are beginning to include references to algorithms.[7] They vary, though, in their explanations and treatment of algorithms. Depending on the robustness of the discussion of algorithms in the textbook you use, you may consider using the supplemental article discussed above,[8] especially if your textbook does not specifically warn students about algorithms or adequately explain how algorithms work.

"When discussing how algorithms affect attorneys, legal research professors should emphasize . . . that legal research databases have different algorithms, resulting in different search results."

III. Lectures

For algorithms, I recommend providing students with a fifteen-minute lecture about how algorithms affect and limit electronic legal research. The lecture's structure should be twofold. First, it should answer: why should I, as a law student, care about algorithms in legal research? Second, it should answer: how do algorithms affect me when I do legal research?

Some important points about why law students should care about algorithms can include the following:

- "For lawyers, 2018 was the year of the algorithm—the year that sophisticated computer intelligence emerged both as a legitimate aid to legal decision-making and as a potential source of discrimination, bias, and error."[9]
- "[T]he future of legal research lie[s] with algorithms."[10]
- "Algorithms lull us into believing our searches will always give us the most relevant results."[11]
- "Without understanding how the algorithms generate results, it is difficult, if not impossible, for attorneys to vet the information."[12]
- Law students should be "disillusion[ed]" that the "computer intelligence" and algorithms can "do[] the [legal research] work for them" because of all "the limitations of these [legal research databases]."[13]

These quotes show that algorithms are now an integral part of legal research and will affect the results and research law students and attorneys do; thus, students and attorneys should understand algorithms and make adjustments in their research because of them.

When discussing how algorithms affect attorneys, legal research professors should emphasize the crux of Mart's scholarship—that legal research databases have different algorithms, resulting in different search results.[14] To be thorough while researching, law students should run searches in multiple legal research databases when they have the opportunity to do so. Further, they should not necessarily give deference to the top ten search results presented in a particular legal research database, nor to the order in which the results are listed as relevant.

Moreover, the legal research class lecture should explain to law students that it is currently impossible to fully comprehend the algorithms because legal research databases do not reveal to us all the factors involved in producing results through algorithms. This is a challenge to teaching algorithms. While we do not fully understand how the algorithms work, the information we do have is nonetheless beneficial to legal research students, as it will help them to better critically assess search results and see how algorithms affect their legal research.

6 *Id.*

7 *See, e.g.*, Kent C. Olson, Aaron S. Kirschenfeld & Ingrid Mattson, Principles of Legal Research 21–24, 312–16 (3d ed. 2020); Amy E. Sloan, Basic Legal Research: Tools and Strategies 230–34 (8th ed. 2021); Beau Steenken & Tina M. Brooks, Sources of American Law: An Introduction to Legal Research 24, 28–29 (4th ed. 2019).

8 *See* Mart, *Every Algorithm*, *supra* note 2.

9 Tad Simons, *8 Things a Good Lawyer Does That a Computer Algorithm Can't*, Thomson Reuters: Legal Executive Inst. (Feb. 19, 2019), http://www.legalexecutiveinstitute.com/lawyers-can-do-algorithm-cannot.

10 Yolanda P. Jones, *Expansive Legal Research*, 44 Int'l J. Legal Info. 241, 267 (2016).

11 Haight, *supra* note 1, at 208 (footnote omitted).

12 Baker, *2018: A Legal Research Odyssey*, *supra* note 1, at 22–23.

13 Callister, *supra* note 1, at 210.

14 *See supra* note 2.

At one point, we knew that Westlaw had the following factors that affected its algorithms:

- West key number system,
- statutory indexes,
- KeyCite,
- secondary sources, and
- document usage patterns.[15]

And that Lexis had these factors:

- phrase recognition,
- case names and citations,
- concentration of terms,
- coverage of terms,
- prominence,
- recentness,
- document segment the search term appears in, and
- number of hits within document.[16]

The factors of Westlaw and Lexis are significant because (1) these may not even still be the factors as they are not publicized on either's database; (2) the factors on Westlaw and Lexis differ from one other, and (3) some of the factors may be ones that students may not necessarily want to influence their searches. For example, "document usage patterns" under Westlaw means that if a document is opened and/or saved a lot then it will show up higher (signifying more relevance) in the search results list. Perhaps that case is popular for a particular issue and so it is higher up on "relevancy," but the attorney is using it for a less common reason and, therefore, the prioritization of the case in the search results is misleading as to its importance with this particular research question. For the West key number system, if relevant cases with the search terms appear in the results, and those cases have a common West key number among them, then the algorithm may include other cases that do not have the search term in the results because they have the same key number attached to their headnotes. These results may occur because they may have synonyms of the search term (or the term searched for shows up in the editor-written headnotes and not in the opinion of the case).[17] If a legal researcher wants a specific term to appear in the opinion, then irrelevant results may be in the search results list because of this factor.

For additional lecture content on algorithms, consider including in your legal research curriculum the list by law librarian Iantha Haight of minimum technology standards of law school graduates for the topic of algorithms. She argues that law students "must understand:

a. What an algorithm is
b. The impact of data quality on an algorithm
c. How experts analyze algorithms
d. Critiquing an algorithm
e. Values choices and biases in algorithms."[18]

Using these technology standards as an outline, legal research professors can make sure that their students are able to intelligently apply algorithms with caution when they are running searches in legal research databases. For example, they can encourage students to think about the data quality that an algorithm on a legal research database may have or the values choices and biases in the algorithm. This means that the students may want to use terms and connectors searches and

"The factors of Westlaw and Lexis are significant because (1) these may not even still be the factors as they are not publicized on either's database; (2) the factors on Westlaw and Lexis differ from one other, and (3) some of the factors may be ones that students may not necessarily want to influence their searches."

15 Mart, *The Algorithm as a Human Artifact*, *supra* note 2, at 400 & n.75 (citing Thomson Reuters, *WestSearch: WestlawNext Search Technology*, https://docmh.com/embed/wlnsearch [https://perma.cc/S8GV-R5SL]). Mart's article was published in 2017, and the Westlaw article cited in it is not still published on Westlaw's website.

16 *Id.* at 401–02 & n.83 (citing LexisNexis, *Lexis Advance Faculty FAQs*, http://www.lexisnexis.com/documents/pdf/20111216091630_1arge.pdf [https://web.archive.org/web/20160620151641/http://www.lexisnexis.com/documents/pdf/20111216091630_large.pdf]). Mart's article was published in 2017, and the LexisNexis source it cited could not be retrieved in January 2021. A search to see if more recent literature had been published by LexisNexis on this topic yielded no additional source.

17 *See* Thomson Reuters, *supra* note 15.

18 Haight, *supra* note 1, at 212.

search within specific secondary sources in order to avoid a bias against treatises, for instance.[19]

IV. Assignments

An example of an assignment that I do during class time is to have my law students do their own version of the exercise that Mart expands on in *The Algorithm as a Human Artifact*.[20] Using the facts from one of the memos they have researched and written that semester, they craft a natural language search, filtering to cases in a specific state jurisdiction. They then compare the top ten results in Westlaw, Lexis, and Google Scholar[21] and chart them similar to Figure 1 below.[22]

"By doing a verification like running the same search on more than one database to see how the algorithm affects a specific research question, the students can tangibly see how much they should (or should not) trust the search results from the legal research databases."

Figure 1[23]

Top Ten	Westlaw	Lexis	Google Scholar
1.	Scott v. Scott (Utah Ct. App. 2016)	Scott v. Scott (Utah Ct. App. 2016)	Haddow v. Haddow (Utah 1985)
2.	Myers v. Myers (Utah Ct. App. 2010)	Myers v. Myers (Utah 2011)	Garcia v. Garcia (Utah Ct. App. 2002)
3.	Myers v. Myers (Utah 2011)	Scott v. Scott (Utah 2017)	Myers v. Myers (Utah 2011)
4.	Pendleton v. Pendleton (Utah Ct. App. 1996)	Black v. Black (Utah Ct. App. 2008)	Bagshaw v. Bagshaw (Utah Ct. App. 1990)
5.	Black v. Black (Utah Ct. App. 2008)	Myers v. Myers (Utah Ct. App. 2010)	Bridenbaugh v. Bridenbaugh (Utah Ct. App. 1990)
6.	Scott v. Scott (Utah 2017)	Bagshaw v. Bagshaw (Utah Ct. App. 1990)	Knuteson v. Knuteson (Utah 1980)
7.	Haddow v. Haddow (Utah 1985)	Williamson v. Williamson (Utah Ct. App. 1999)	Jeppson v. Jeppson (Utah 1984)
8.	Wacker v. Wacker (Utah 1983)	Jeppson v. Jeppson (Utah 1984)	Black v. Black (Utah Ct. App. 2008)
9.	Levin v. Carlton-Levin (Utah Ct. App. 2014)	Garcia v. Garcia (Utah Ct. App. 2002)	Barber v. Barber (Utah Ct. App. 1990)
10.	Garcia v. Garcia (Utah Ct. App. 2002)	Andersen v. Andersen (Utah Ct. App. 1988)	Munns v. Munns (Utah Ct. App. 1990)

Generally, students are shocked when they see these differing results about a memo they have been researching for most of the semester. However, they are also comforted to see that the databases do agree to a certain extent. By doing a verification like running the same search on more than one database to see how the algorithm affects a specific research question, the students can tangibly see how much they should (or should not) trust the search results from the legal research databases.

Assignments such as this one develop awareness in students and tend to send the message more clearly than just reading an article on algorithms. This exercise is valuable because it is important for students to realize that they should neither rely on just one legal research database, nor look at only the top ten results. With Google Scholar

19 Callister, *supra* note 1, at 203. Additional biases include prioritization, classification, association, and filtering. Mart, *The Algorithm as a Human Artifact*, *supra* note 2, at 394–95 (quoting Nicholas Diakopoulos, *Algorithmic Accountability: Journalistic Investigation of Computational Power Structures*, 3 Dig. Journalism 398, 399 (2015)).

20 Mart, *The Algorithm as a Human Artifact*, *supra* note 2, at 406–16.

21 While Mart also compares search results across Fastcase, Ravel, and Casetext, I have found I do not have sufficient time to include more than three research databases. Additionally, I teach the class period on algorithms before the law students are familiar with Fastcase, Ravel, and Casetext.

22 The blank version of this figure is adapted from Mart, *The Algorithm as a Human Artifact*, *supra* note 2, at 421.

23 This search was for "termination of alimony" without quotes in the Westlaw, Lexis, and Google Scholar legal research databases. The search was limited to cases in the state of Utah. The cases are highlighted light blue if they appear in the top ten results of two databases and are highlighted in a darker blue if they appear in the top ten results of all three of the databases.

being free and all state bars currently giving licensed attorneys free access to either Casemaker or Fastcase,[24] most law students, once they graduate, should be able to research in more than one legal research database, even if they do not have a subscription to Westlaw or Lexis. Students should also consider various search terms and understand that a result may appear simply because it is viewed often by other researchers on the database or is cited often by other sources, even if the result is not important or relevant for their needs. This assignment drives these points home.

V. Conclusion

Overall, the role of algorithms in legal research is a topic that should be integrated into legal research courses when possible. With readings, lectures, and assignments, legal research professors can adequately prepare law students to handle algorithms in their future legal practice. If legal research professors teach law students how to search in legal research databases but neglect to warn them about algorithms and their varying effects on search results, legal research professors are not adequately preparing their students to become successful attorneys.

"With readings, lectures, and assignments, legal research professors can adequately prepare law students to handle algorithms in their future legal practice."

24 For a list of which state bar associations have subscriptions to which legal research database, see *Legal Research via State Bar Associations*, Goodson Law Library, https://law.duke.edu/lib/statebarassociations/ (last visited Feb. 19, 2021). Note that while Casemaker and Fastcase announced their merger on January 5, 2021, current free access to either Casemaker or Fastcase remains unchanged. *Id.* (citing *Casemaker* and *Fastcase Merge to Become the Leader in Legal Research and Analytics*, Fastcase (Jan. 5, 2021), https://www.fastcase.com/blog/fastcase-casemaker-merge/).