Borgeson Paper Archive   Law Librarianship Program

2018

Are Congressional Resources Created Equal: An Evaluation

Matthew P. Neely

Follow this and additional works at: https://digitalcommons.law.uw.edu/law-lib_borgeson

Recommended Citation

Are Congressional Resources Created Equal: An Evaluation.

MATTHEW P. NEELY

Gallagher Law Library, University of Washington School of Law, Seattle, Washington, USA

This paper examines the four major online databases for government documents. The databases include ProQuest Congressional, HeinOnline Congressional Documents, govinfo.gov, and Congress.gov. This paper evaluates the four databases by using standardized searches to compare the search results from the different databases given a standardized data set. This paper primarily compares the item type, frequency, order, and duplication of the results. Conclusions are drawn from this data and suggestions for future research are offered. Finally, concluding lessons and practical suggestions are offered.

Table of Contents

I.   Introduction

The inspiration for this paper came from Susan Nevelow Mart's article *The Algorithm as a Human Artifact: Implications for Legal [Re]Search*.[1] Mart gave a presentation on this material at the 2017 American Association of Law Librarians in Austin, Texas, which the researcher attended.[2] In her study, Director Mart discusses the algorithms used in Westlaw, Lexis Advance, Fastcase, Google Scholar, and Casetext and attempted to reveal something of the hidden algorithms of the various databases.[3] Mart focused on this area in part because of how clouded these databases were in providing their algorithms to researchers because algorithms are considered Trade Secrets.[4] Mart hoped to enlighten the public about the worldviews of the different databases, noting that, "[t]he 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."[5]

This study will examine Proquest Congressional, HeinOnline's Congressional Documents Module, govinfo.gov, and Congress.gov to determine their worldviews and how these worldviews impact first year law students. This study will analyze what types of results are

---

[1] Susan Nevelow Mart, *The Algorithm as a Human Artifact: Implications for Legal [Re]Search*, 109 Law Libr. J. 387 (2017).

[2] American Association of Law Libraries, *UNDERSTANDING THE HUMAN ELEMENT IN SEARCH ALGORITHMS (A7)*, https://www.aallnet.org/recording/understanding-the-human-element-in-search-algorithms-a7/ (last accessed July 10, 2018).

[3] Mart, *supra* note 1, at 387.

[4] Mart, *supra* note 1, at 389 n.11.

[5] Mart, *supra* note 1, at 390.

returned from standardized searches in an effort to help legal research instructors and practitioners better understand the distinctive of each of these databases.

These databases have unique quirks, strengths, and weaknesses. These distinctions are highlighted in the age of the natural language search. This study presumes that the average first year law student will perform a simple natural language search about a legal research topic. Therefore, this study will run a similar search to examine the characteristics of each database and the implications for first year students.

II. Literature Review

a. Susan Nevelow Mart's Study

Mart's study focused on the fact that "human creators made choices about how the algorithm would work that have implications for the search results returned to the researcher."[6] Those choices resulted in significantly different search results when queried with an identical search.[7] These differences ultimately produced trends that impacted the way researchers should search for cases.[8]

b. Natural Language Searching

Most of the literature regarding student searching covers undergraduate students. This literature is relevant to this study. Most first year law students are coming straight from undergraduate studies. Therefore, their searching behavior mimics undergraduate research behavior. Undergraduate students do not always understand that Natural Language searching

---

[6] Mart, *supra* note 1, at 388.
[7] Mart, *supra* note 1, at 412.
[8] Mart, *supra* note 1, at 416.

is less effective than an advanced search, like a Boolean search.[9] Students often prefer

convenience to efficiency.[10] Most undergraduates are part of the Millennial Generation and

often start their research on the internet.[11] They tend to choose popular search engines that

use natural language searching.[12] Millennials have a general set of characteristics "based on

their generational, cultural, and social environments."[13] These seven traits are, "special,

sheltered, confident, team-oriented, conventional, pressured, and achieving."[14] Millennials

have not experienced a time without the internet; but, this does not make these individuals

tech savvy.[15] Nor do Millennials "follow organized, hierarchical structures of information

gathering or processing."[16]

Successful search strategies includes "two important pieces: identifying the right terms for

the search and developing an effective strategy to search for the needed information."[17] "[The

attributes of the system being used and the search behaviors of users" impact the effectiveness

of search results.[18] This can be particularly important when "many companies do not publicize

their retrieval process for fear of competition and loss of revenue."[19] Millennials believe that

the first link in a search is the best link.[20]

---

[9] M. Sara Lowe, Bronwen K. Maxson, Sean M. Stone, Willie Miller, Eric Snajdr, & Kathleen Hanna, The *Boolean Is Dead, Long Live the Boolean! Natural Language versus Boolean Searching in Introductory Undergraduate Instruction*, 79 College and research Libraries 517, 518 (2018).

[10] Lowe, *supra* note 6, at 518.

[11] Brandi Porter, *Millennial Undergraduate Research Strategies in Web and Library Information Retrieval Systems*, 5 J. Web Librarianship, 267, 268 (2011).

[12] Porter, *supra* note 8, at 268.

[13] Porter, *supra* note 8, at 269.

[14] Porter, *supra* note 8, at 269.

[15] Porter, *supra* note 8, at 269.

[16] Porter, *supra* note 8, at 270.

[17] Porter, *supra* note 8, at 270.

[18] Porter, *supra* note 8, at 270.

[19] Porter, *supra* note 8, at 270.

[20] Porter, *supra* note 8, at 281.

c. **History of the Databases**

Electronic legal research is hampered in many ways by the ontologies and classification systems, especially in terms of automation.[21] This is true to the point that several of the major legal database providers note how much human work goes into the systems.[22] The general histories of the database systems help demonstrate their purposes and illuminate why databases produce certain results.

i. **ProQuest**

In 1938 Eugene Power founded University Microfilms to preserve works from the British Museum on microfilm.[23] ProQuest has a long history of merging or purchasing other research companies in order to increase its service base.[24] Among these purchases was the Congressional Information Service from LexisNexis in 2010.[25] This content included "congressional publications, bills, laws and other research materials, data produced by U.S. Federal agencies, States, private organizations, and major intergovernmental organizations, and the microfilm vault of government documents encompasses text of congressional hearings dating from 1789."[26]

---

[21] Mart, *supra* note 1, at 382 n. 24.

[22] Shepard's Citations Service, LEXISNEXIS https://www.lexisnexis.com/en-us/products/lexis-advance/shepards.page (last visited July 10, 2018) (commenting on how "our attorney-editors follow rigorous quality controls for case law and a strict 29-step editorial process"); Topic and Key Number Overview, WESTLAW, https://lawschool.westlaw.com/marketing/display/RE/24 (noting that "attorney editors read the case and pick out the points of law addressed in the case").

[23] *History & Milestones*, ProQuest, https://www.proquest.com/about/history-milestones/ (last visited July 10, 2018).

[24] *Id.*

[25] *ProQuest Acquires Acclaimed Congressional Information Service and University Publications of America from LexisNexis*, ProQuest, https://www.proquest.com/about/news/2010/ProQuest-Acquires-Congressional-Information-Service-LexisNexis.html (last visited July 10, 2018).

[26] *Id.*

ii. **HeinOnline.com**

Kevin Marmion drafted a three-page memo titled "Electronic Law Review Project" in December of 1995.[27] That memo outlined potential benefits of an online law review library. The memo described monthly deliveries of CDs.[28] The memo also speculated "If on-line services were used [delivery] could be done daily."[29] Hein went online in May 2000.[30] Xerox Developed Docutech in the early 1990's that allowed the creation of electronic facsimiles.[31] Hein obtained Docutech in 1994. Thus, enabling Hein to produce small batches of print journals and books.[32]

HeinOnline added the Federal Register in September of 2002, covering up to 1974.[33] The full printing was soon available.[34] The Historical Code of Federal Regulations was added in 2005. The U.S. Statutes at[35] large was added in 2006. HeinOnline began scanning in the Congressional Record in June 2007.[36] This was eventually compiled into the "U.S. Congressional Documents" Module. From this history, one can derive that HeinOnline's collections have started with the historic data and moved into providing current law.

iii. **Govinfo.gov**

---

[27] Joe Gerken, *Invention of HeinOnline: The Story of Hein's Dramatic Transformation in Response to the Coming of the Electronic Age*, 18 A.A.L.L. Spectrum, February 2014, at 17.
[28] Gerken, *supra* not 27, at 17.
[29] Gerken, *supra* not 27, at 17.
[30] Gerken, *supra* not 27, at 17.
[31] Gerken, *supra* not 27, at 18.
[32] Gerken, *supra* not 27, at 18.
[33] Gerken, *supra* not 27, at 20.
[34] Gerken, *supra* not 27, at 20.
[35] Gerken, *supra* not 27, at 20.
[36] Gerken, *supra* not 27, at 20.

Congress passed the GPO Electronic Information Access Enhancement Act on June 8, 1993.[37] This act adapted the GPO's mission to include access to Federal Government information electronically.[38] This lead to GPO Access.[39] On June 8, of 1994, the GPO Accesses site went online.[40] This was the public source for government information for 15 years until FDsys was launched on January 15, 2009.[41] FDsys was available along with the existing GPO Access site until December of 2010 when FDsys replaced GPO Access.[42] FDsys initially included "Congressional Bills, Calendars, Committee Prints, Hearings, Reports, Documents, Congressional Directory, Congressional Record, Public Laws, and United States Code."[43] The GPO launched govinfo on February 3, 2016, which updated the searching features and redesigned the look of the website.[44]

iv. Congress.gov

The website THOMAS was launched in January 1995, concurrently with the convening of the 104th.[45] THOMAS originated when Congress instructed the Library of Congress to provide free access to federal legislative information. Congress.gov replaced THOMAS, which was retired on July 5, 2016.[46] The Library of Congress released Congress.gov in September 2012.[47] Congress.gov describes itself in the following way:

---

[37] *History*, govinfo, https://www.govinfo.gov/about/history (last accessed July 10, 2018).
[38] *Id*.
[39] *Id*.
[40] *Id*.
[41] *Id*.
[42] *Id*.
[43] *Id*.
[44] *Id*.
[45] Library of Congresss, About Congress.gov, https://www.congress.gov/about (last visited July 10, 2018).
[46] *Id.*
[47] *Id.*

Congress.gov is the official website for U.S. federal legislative information. The site provides access to accurate, timely, and complete legislative information for Members of Congress, legislative agencies, and the public. It is presented by the Library of Congress (LOC) using data from the Office of the Clerk of the U.S. House of Representatives, the Office of the Secretary of the Senate, the Government Publishing Office, Congressional Budget Office, and the LOC's Congressional Research Service.[48]

In 2013 the URL for THOMAS was officially redirected to Congress.gov.[49]

III.   The Empirical Study

a.   Methodology

Given the potential challenges for first year law students conducting natural language searches in the government documents arena, the purpose of this research study was to investigate the utility of four highly utilized congressional document database and the search results that are returned by simple, natural language searches. Original research was conducted in the form of an empirical study. From the data obtained, several trends arose

One of the greatest challenges to this study was the creation of a set of database modules that would result in a uniform data set from which to search. This study chose to search four commonly used databases to obtain congressional materials: ProQuest Congressional, HeinOnline's Congressional Documents Module, govinfo.gov, and Congress.gov. Table 1 summarizes the availability of the different kinds of the congressional documents and the date ranges for each item type. This coverage was derived from each database's website.[50] As table

---

[48] *Id.*

[49] *Id.*

[50] The coverage for each database was obtained from their own websites:
- ProQuest, Content coverage Chart, ProQuest.com, https://congressional-proquest-com.offcampus.lib.washington.edu/congressional/saleable/contentcoverage?accountid=14784&groupid=95339 (last visited July 9, 2018).
- HeinOnline, *U.S. Congressional Documents*, Heinnline.org, https://heinonline.org/HeinDocs/USCongressionalDocumentsCollection3.pdf (last accessed July 9, 2018).

1 demonstrates, database coverage varied widely from database to database.[51] This study

combined the available coverage ranges for each database and item type and found that there

was overlap in all databases from 1994 to 2014.[52] Congress.gov was an exception in that the

only item types that this database covered were Congressional Bills, the Congressional Record,

and Congressional Reports.[53] Once the available coverage was determined, the study then

conducted the same search in the four different databases.

Table 1

Database Coverage of Select Congressional Materials*

Resource Type	HeinOnline	ProQuest Congressional	govinfo.gov	Congress.gov
Congressional Bills	1909 (61st Congress) to 2009* Congressional Record Only	1989 to Current 1789 to Current in PDF	1994-1995 (103rd Congress) to Current	1989 (101st Congress) to Current
Congressional Hearings	1905 (59th Congress) to Current	1988 to Current	1995-1996 (104th Congress) to Current	N/A
Congressional Reports	1909 (61st Congress)- Current	1990 to Current	1995-1996 (104th Congress) to Current	2005-2006 (108th Congress) to Current.
House Journals	1980-2014	1817 to Current	1992-2015	N/A

- U.S. Government Publishing Office, *What's Available*, govinfo.gov, https://www.govinfo.gov/help/whats-available (last visited July 9, 2018).
- Library of Congress, *Coverage Dates for Legislative Information*, Congress.gov, https://www.congress.gov/about/coverage-dates (last visited July 9, 2018).

[51] *Id.*
[52] *Id.*
[53] Library of Congress, *supra* note 50.

Senate Journals	1981-2015	1817 to Current	2005-2006 (108th Congress) to Current.	N/A
Congressional Record	1980 to Current	1985 to Current	1995-1996 (104th Congress) to Current	1995-1996 (104th Congress) to Current
Code of Federal Regulations	1938 to Current	1981 to Current	1996 to Current	N/A
Federal Register	1936 t0 Present	1980 to Current	1936 to Current	N/A

*Coverage dates vary widely. The date where the service begins to provide general coverage is listed.

This study conducted an advanced search limiting the database coverage to the 1994 to 2014 for the phrase "do not call registry" without quotation marks. The study was revised to also include a search in each database for the phrase "do not call registry" within quotation marks when the initial search in govinfo.gov returned twenty irrelevant and unrelated CFR. sections when the search was conducted without quotations. Additionally, govinfo.gov had such a wide assortment of collection of resource types available that the study limited the collections to the following:

- Code of Federal Regulations

- Congressional Bills

- Congressional Committee Prints

- Congressional Documents

- Congressional Hearing

- Congressional Record

- Congressional Record Bound

- Congressional Record Index

- Congressional Reports

- Federal Register

- Journal of the House of Representatives

- Public and Private Laws

- Statutes at Large.

The researcher recorded several pieces of data for each of the top twenty results for each of the two searches in all the databases. The researcher recorded the title of individual result as the database presented it.[54] Then, the researcher manually assigned a resource type to account for differences in naming conventions within each database.[55] Next, the year of the item was recorded. Finally, the item was determined to be relevant or not relevant. Relevancy was given a very broad definition. If the item was in any way related to the topic, then the result was scored as relevant.[56]

This study did not look at advance search, filtering, or sorting features other than the default result, except for setting a date of results range or limiting govinfo.gov's collections.

---

[54] This study did not standardize the title types.

[55] Names for congressional hearings and congressional reports varied slightly between databases.

[56] The researcher wrestled with the idea of relevancy. A Congressional Record Daily where an associated bill was discussed could be very relevant for a faculty member doing a very deep dive on a topic. However, the same result may not be relevant for a student looking for more general information on a specific statute. This study decided to err on the side of overinclusion of relevant data.

This was done to mimic a first-year law student's information seeking behavior as closely as possible and remove the element of skill from the search results. This allowed the researcher to compare the databases initial results without introducing the searcher's own bias into the search. Each searcher develops their own searching style than often influences the results of search, even if in a minor way. Similarly, Boolean connectors were omitted in an attempt to evaluate each database's search algorithms. This search question was limited by the number of searches completed. The limitations of this study are discussed in section V. Future Research *infa*.

Each database had a total of forty results, twenty from each search. This resulted in a total of 160 item results. The researcher then compiled the data into a spreadsheet and analyzed the data. The study examined the item types, the number of items within each item type, the difference in item types between searched with and without quotation marks, and compared the results for each database.

b. **Hypotheses**

This study hypothesized that the paid databases, ProQuest Congressional and HeinOnline, would provide the best relevancy filtering in their results. This is because the researcher believes that the paid databases have the resources available to develop complex algorithms to filter and sort the data. The researcher believed that because of the advanced algorithms, the paid databases would have the widest variety of item type in the top twenty results, with the database pulling the most relevant items. The study also hypothesized that

the free databases would have comparable results to each other and just a little less variety in item type than their paid peers.

c. **Results**

    i. **General**

This study confirmed its hypothesis that the paid databases would have the most diverse results in item type. Table 2 identifies the total number of item types in each database across both searches.

Table 2

Total Item Type Results by Database

	ProQuest Congressional	HeinOnline	govinfo.gov	congress.gov	Total
Bill Profile	4				4
Bill Text	6		4	18	28
Congressional Report	6	6			12
Congressional Record	12		15		27
Hearings	2		1		3
Legislative History	6				6
Public Law	4	4		22	30
Congressional Hearing		2			2
CRS Report		8			8
Senate Journal		13			13
House Journal		7			7
CFR			20		20

	40	40	40	40	160

Item types varied significantly between databases with the paid databases having the most

diverse results. The single most common item type was public laws. The second most common

item was bill texts. At least sixty-four-percent of each of these two item types were results

from Congress.gov.[57] These results are mainly due to Congress.gov's item type restrictions.

Several item types were found in only one database. The second-to-least reported item

type was a bill profile found exclusively in ProQuest.[58] One important piece of information to

note is that ProQuest's data set for each search was exactly the same. The implications of this

result are discussed in section III.C.ii. ProQuest, *infra*.   This number was reasonable because

bill profiles are produced only by ProQuest. Similarly, Congressional Research Service (CRS),

Reports were only found in HeinOnline. HeinOnline was the only database queried that

provides access to the CRS Reports.

Other single database results were not as easily understood. Congressional Hearings

and Senate and House Journals were only found in HeinOnline's results when every database,

with the exception of Congress.gov, has those item types in their databases. This phenomenon

will be discussed in section III.C.iii. HeinOnline *infra*. Similarly, all twenty Code of Federal

Regulations (CFR), results were from a single govinfo.gov search when every other database

---

[57] Congress.gov had eighteen out of twenty-eight Bill Text Results or sixty-four-percent of the results. Congress.gov had twenty two out of thirty Public law results or seventy-three-percent of the results.
[58] Two unique item types are discussed in this study: ProQuest's Bill Profile and HeinOnline's Congressional Research Service Reports (Hereinafter "CRS Reports). These resources were not excluded for several reasons. First, CRS Reports were very difficult to exclude without searching or filtering by individual item types, defeating the point of the searches to analyze the databases' interpretation of relevant data based on a first-year law student's hypothetical search.

also contained the CFR. These results demonstrate significant distinctions in algorithms, which will be discussed in section IV. Discussion.

The databases did not return a result in 18 of the possible item types across all searches in all databases. Individual results will be discussed in database's section. Each item type was represented in nor more than two databases, except for public laws and bill texts. No item type was represented in every search. This indicates a preference item type in each database. The number of item types in each search results varies between databases.

ii. ProQuest

ProQuest had some of the most surprising results. The searches with and without quotation marks returned the same results. There was no difference in the order of results. Table 3 shows the search results sorted by the item type. Table 4 shows the search results in their original order.

Table 3

ProQuest Data without Quotation Marks and sorted by item type

Title	Resource Type	Year	Relevant
110 Bill Profile S. 781 (2007-2008)	Bill Profile	2008	Yes
110 Bill Profile H.R. 2601 (2007-2008)	Bill Profile	2008	Yes
110 H.R. 2601 Engrossed in House	Bill Text	2007	Yes
110 S. 781 Engrossed in Senate	Bill Text	2007	Yes
110 S. 781 Reported in Senate	Bill Text	2007	Yes
TRIBUTE TO EILEEN HARRINGTON AND THE DO NOT CALL REGISTRY TEAM	Congressional Record	2005	No
DO NOT CALL REGISTRY	Congressional Record	2003	Yes

DO-NOT-CALL REGISTRY FEE EXTENSION ACT OF 2007	Congressional Record	2007	Yes
MAKING IN ORDER AT ANY TIME CONSIDERATION OF H.R. 3161, RATIFYING AUTHORITY OF FTC TO ESTABLISH A DO-NOT-CALL REGISTRY	Congressional Record	2003	Yes
DO-NOT-CALL REGISTRY FEE EXTENSION ACT OF 2007	Congressional Record	2007	Yes
DO-NOT-CALL REGISTRY FEE EXTENSION ACT OF 2007	Congressional Record	2008	Yes
Telemarketing: Implementation of the National Do-Not-Call Registry	Congressional Report	2005	Yes
Do-Not-Call Registry Fee Collection Extension	Congressional Report	2007	Yes
Do-Not-Call Registry Fee Extension Act of 2007	Congressional Report	2007	Yes
The Do-Not-Call Registry	Hearings	2003	Yes
Do-Not-Call Improvement Act of 2007	Legislative History	2008	Yes
FTC Do-Not-Call Registry Implementation Authority	Legislative History	2003	Yes
Do-Not-Call Implementation Act	Legislative History	2003	Yes
[FTC Do-Not-Call Registry Implementation Authority]	Public Law	2003	Yes
Do-Not-Call Registry Fee Extension Act of 2007	Public Law	2008	Yes

Table 4

ProQuest Search Results with Quotations in the Original Order

Title	Resource Type	Year	Relevant
Do-Not-Call Improvement Act of 2007	Legislative History	2008	Yes
Telemarketing: Implementation of the National Do-Not-Call Registry	Committee Report	2005	Yes
The Do-Not-Call Registry	Hearings	2003	Yes

TRIBUTE TO EILEEN HARRINGTON AND THE DO NOT CALL REGISTRY TEAM	Congressional Record	2005	No
DO NOT CALL REGISTRY	Congressional Record	2003	Yes
[FTC Do-Not-Call Registry Implementation Authority]	Pub. Law	2003	Yes
Do-Not-Call Registry Fee Collection Extension	Congressional Report	2007	Yes
Do-Not-Call Registry Fee Extension Act of 2007	Congressional Report	2007	Yes
FTC Do-Not-Call Registry Implementation Authority	Legislative History	2003	Yes
DO-NOT-CALL REGISTRY FEE EXTENSION ACT OF 2007	Congressional Record	2007	Yes
Do-Not-Call Implementation Act	Legislative History	2003	Yes
MAKING IN ORDER AT ANY TIME CONSIDERATION OF H.R. 3161, RATIFYING AUTHORITY OF FTC TO ESTABLISH A DO-NOT-CALL REGISTRY	Congressional Record	2003	Yes
110 Bill Profile S. 781 (2007-2008)	Bill Profile	2008	Yes
110 Bill Profile H.R. 2601 (2007-2008)	Bill Profile	2008	Yes
Do-Not-Call Registry Fee Extension Act of 2007	Public Law	2008	Yes
DO-NOT-CALL REGISTRY FEE EXTENSION ACT OF 2007	Congressional Record	2007	Yes
DO-NOT-CALL REGISTRY FEE EXTENSION ACT OF 2007	Congressional Record	2008	Yes
110 H.R. 2601 Engrossed in House	Bill Text	2007	Yes
110 S. 781 Engrossed in Senate	Bill Text	2007	Yes
110 S. 781 Reported in Senate	Bill Text	2007	Yes

ProQuest had the most diverse search results with seven different item types. Table 5 summarizes the different item types. Table 5 only lists twenty results because of the duplication of the search results between both searches. Two of these types, bill profiles and Legislative Histories[59], were unique to ProQuest. These results were unique because they are a proprietary product of ProQuest.

---

[59] HeinOnline also has legislative histories, but of a different type.

Table 5

ProQuest Search Results by Item Types

Item Type	Number of Redults	Percent of Total
Bill Profile	2	10%
Bill Text	3	15%
Congressional Report	3	15%
Congressional Record	6	30%
Hearings	1	5%
Legislative History	3	15%
Public Law	2	10%

The item types were otherwise evenly distributed with about two to three results, except for

the Congressional Record. This was the most uniform numbers of individual item types. The

next database for balance of item type and number of items within each type was HeinOnline.

ProQuest returned no Congressional Hearings, CFRs or House or Senate Journals. ProQuest

favored the Congressional Record.

iii. HeinOnline

HeinOnline presented a balanced set of results with few surprises. Table 6 presents the

HeinOnline's search results without quotation marks. Table 7 presents the HeinOnline's search

results with quotation marks. These two table illustrate the difference that quotation marks

make when using HeinOnline. All but two of HeinOnline results were relevant.

Table 6

HeinOnline without Quotation Marks

S. Rept. 110-244	Congressional Report	2007	Yes
H. Rept. 110-485	Congressional Report	2007	Yes
122 Stat. 635	Pub Law	2007	Yes
Do-Not-Call Registry: Hearing before the Committee on Commerce, Science, and Transportation, United States Senate, One Hundred Eighth Congress, First Session	Congressional Hearing	2003	Yes
117 Stat. 1006	Pub Law	2003	Yes
H. Rept. 108-8	Congressional Report	2003	Yes
H. Rept. 110-486	Congressional Report	2007	Yes
S. Rept. 110-246	Congressional Report	2007	Yes
Telemarketing Regulation: National and State Do Not Call Registries	CRS Report	2016	Yes
Telemarketing Regulation: National and State Do Not Call Registries	CRS Report	2014	Yes
122 Stat. 633	Pub. Law	2007	Yes
Regulation of the Telemarketing Industry: State and National Do-Not-Call Registries	CRS Report	2003	Yes
Regulation of the Telemarketing Industry: State and National Do-Not-Call Registries	CRS Report	2003	Yes

Regulation of the Telemarketing Industry: State and National Do Not Call Registries (RL31642)	CRS Report	2004	Yes
Regulation of the Telemarketing Industry: State and National Do-Not-Call Registries	CRS Report	2004	Yes
117 Stat. 557	Pub. Law	203	Yes
Regulation of the Telemarketing Industry: State and National Do Not Call Registries	CRS Report	2003	Yes
Regulation of the Telemarketing Industry: State and National Do Not Call Registries	CRS Report	2004	Yes
H. Rept. 115-301	Congressional Report	2017	No
Admission of vessels to American registry : hearings before the Merchant Marine and Fisheries Subcommittee of the Committee on Commerce, United States Senate, Eighty-Seventh Congress, Second session.	Congressional Hearing	1962	No

Table 6 illustrates how HeinOnline's results were spread throughout its search results, unlike ProQuest's search results, which tended to be clumped together at the same location within the search results. HeinOnline's results, however, were spread throughout the search results in clumps. The search without question marks yielded CRS Reports bunched in the middle of the search, whereas the search with quotation marks returned no CRS Report results.

Table 7

HeinOnline with Quotation Marks

Title	Resource Type	Year	Relevant
2003 J.S. U.S. H.B. 81 (2003)	Senate Journal	2003	Yes
2007 J.S. U.S. 1171 (2007)	Senate Journal	2007	Yes
2003 J.S. U.S. I-1 (2003)	Senate Journal	2003	Yes
2007 J.S. U.S. 1104 (2007)	Senate Journal	2007	Yes
2008 J.S. U.S. 60 (2008)	Senate Journal	2008	Yes
2007 J.S. U.S. H.B. 103 (2007)	Senate Journal	2007	Yes
2003 J.S. U.S. H.B. 71 (2003)	Senate Journal	2003	Yes
2003 J.S. U.S. 782 (2003)	Senate Journal	2003	Yes
2007 J.S. U.S. I-1 (2007)	Senate Journal	2007	Yes
2003 Part 2 J. H.R. U.S. 1857 (2003)	House Journal	2003	Yes
2005 J.S. U.S. 351 (2005)	Senate Journal	2005	Yes
2007 Part 3 J. H.R. U.S. 4029 (2007)	House Journal	2007	Yes
2006 Part 1 J. H.R. U.S. 1105 (2006)	House Journal	2006	Yes
2008 Part 2 Journal of the House of Representatives of the United States 3615 (2008)	House Journal	2008	Yes
2003 Part 2 J. H.R. U.S. 3641 (2003)	House Journal	2003	Yes
2007 J.S. U.S. 819 (2007)	Senate Journal	2007	Yes
2008 J.S. U.S. H.B. 75 (2008)	Senate Journal	2008	Yes
2003 J.S. U.S. 115 (2003)	Senate Journal	2003	Yes
2006 Part 2 J. H.R. U.S. 1941 (2006)	House Journal	2006	Yes
2005 Part 2 J. H.R. U.S. 3485 (2005)	House Journal	2005	Yes

Table 8

Hein Online Item Types

Without Quotation Marks				With Quotation Marks		
Item Type	Number	Percent of Total		Item Type	Number	Percent of Total
Congressional Hearing	2	0.1		Senate Journal	13	0.65
Congressional Report	6	0.3		House Journal	7	0.35
CRS Report	8	0.4				
Public Law	4	0.2				

Table 7 and Table 8 illustrate the importance of the use of quotation marks in searching. When quotation marks were used to search for an exact phrase, HeinOnline returned only two item types compared with no quotation mark's four results. Additionally, the item types that quotation marks returned were distinct.

HeinOnline had the second most diverse set of item types with six item types. CRS Reports were unique to HeinOnline. HeinOnline was the only database to return the item types of Senate Journals or House Journals, even when ProQuest and govinfo.gov has these items in their databases. HeinOnline did not return any CFR results, legislative histories, or congressional records. Hein showed the second most significant changes between searches

with and without quotation marks. There were not duplicate results with HeinOnline's search results.

iv. Govinfo.gov

Govinfo.gov had the widest disparity between the two searches. Govinfo.gov returned the fewest relevant results. The results in Table 9 show how govinfo.gov returned twenty irrelevant results in the search without quotation marks. This likely occurred due to how govinfo.gov algorithms process search results. Most of govinfo.gov's results in this section returned the word "do" in the title or somewhere within the subtitles. This implies that govinfo.gov's algorithms run searches word by word in the order that the search terms were entered.

Govinfo.gov also returned only CFR. results in the search without quotation marks. This may be due to govinfo.gov's algorithms. The Code of Federal Regulation Collection was the first selected collection in the list of collections selected to be searched. If these results are standard across other searches, then govinfo.gov's algorithms search by search term order and by collection. This means that natural language searches are the least curated of any of the databases in govinfo.gov. The results from the search with quotation marks, Table 10, produced a more robust selection of results.

Table 9

Govinfo.gov without Quotation Marks

Title	Resource Type	Year	Relevant

42 CFR 137.220 Do section 314 of Public Law 101-512 - [25 U.S.C. 450f note] and section 102(d) of the Act [25 U.S.C. 450f(d)] (regarding, in part, FTCA coverage) apply to compacts, funding agreements and construction project agreements?	CFR	2002	No
40 CFR 279.30 - it-yourselfer used oil collection centers.	CFR	1996	No
40 CFR 414.100 - of-pipe biological treatment.	CFR	1996	No
40 CFR 414.101 - of-pipe biological treatment.	CFR	1996	No
45 CFR 2516.850 - learning program?	CFR	1996	No
14 CFR Appendix F to Part 25	CFR	2002	No
14 CFR Appendix F to Part 25	CFR	1998	No
14 CFR Appendix F to Part 25	CFR	1997	No
14 CFR Appendix F to Part 25	CFR	2001	No
14 CFR Appendix F to Part 25	CFR	2003	No
14 CFR Appendix F to Part 25	CFR	2000	No
14 CFR Appendix F to Part 25	CFR	1999	No
13 CFR 102.32 - What do Systems Managers do?	CFR	2005	No
13 CFR 102.32 - What do Systems Managers do?	CFR	2000	No
13 CFR 102.32 - What do Systems Managers do?	CFR	2006	No
13 CFR 102.32 - What do Systems Managers do?	CFR	2007	No
13 CFR 102.32 - What do Systems Managers do?	CFR	2002	No

13 CFR 102.32 - What do Systems Managers do?	CFR	1998	No
13 CFR 102.32 - What do Systems Managers do?	CFR	1997	No
13 CFR 102.32 - What do Systems Managers do?	CFR	2004	No

Quotation marks proved to be significant in this search. Govinfo.gov's search with quotation marks, Table 10, produced twenty relevant results, none of which were CFR.s. Govinfo.gov's search with quotations lends support to the theory that govinfo.gov's algorithms search by collection order. In the search with quotation marks, govinfo.gov returned results in blocks of item types: Bill Texts, Congressional Hearings, Congressional Records, and Congressional Reports. These item types mirror the order of collections available for searching at govinfo.gov.[60] More searching is needed to confirm this result, however there is a significant correlation in these results.

This result was outside of the expected result possibilities. The researcher did not consider that a database would search its datasets in this way. This indirectly confirms the need for 1) database providers to disclose their algorithms, and 2) the need to educate searchers about the searching process generally and advanced search methods to ensure that searchers have sufficient data and training to use these databases effectively. This study did not determine how this result held outside of the top 20 results.

---

[60] *Advanced Search*, govinfo.gov, https://www.govinfo.gov/#advanced (last accessed July 10, 2018); bill texts are found in the collection "Congressional Bills", Congressional Hearings, Congressional Records, and Congressional Reports item types collections have the same name as their respective item types. *Id*.

Table 10

Govinfo.gov with Quotation Marks

Title	Resource Type	Year	Relevant
S. Res. 242 (IS) - To express the sense of the Senate concerning the do-not-call registry.	Bill Text	2003	Yes
S. 1655 (IS) - To ratify the authority of the Federal Trade Commission to establish the do-not-call registry.	Bill Text	2003	Yes
H.R. 3161 (IH) - To ratify the authority of the Federal Trade Commission to establish a do-not-call registry.	Bill Text	2003	Yes
S. 1654 (IS) - To ratify the authority of the Federal Trade Commission to establish a do-not-call registry.	Bill Text	2003	Yes
S. Hrg. 108-982 - The Do-Not-Call Registry	Congressional Hearing	2003	Yes
149 Cong. Rec. S11957 - DO-NOT-CALL REGISTRY	Congressional Record	2003	Yes
149 Cong. Rec. S11890 - DO NOT CALL REGISTRY	Congressional Record	2003	Yes
149 Cong. Rec. S11903 - DO NOT CALL REGISTRY	Congressional Record	2003	Yes
149 Cong. Rec. S11900 - RATIFYING THE DO-NOT-CALL REGISTRY	Congressional Record	2003	Yes
153 Cong. Rec. S15788 - DO-NOT-CALL REGISTRY FEE EXTENSION ACT OF 2007	Congressional Record	2007	Yes
CRI 2008 - DO-NOT-CALL REGISTRY FEE EXTENSION ACT	Congressional Record	2008	Yes

CRI 2007 - DO-NOT-CALL REGISTRY FEE EXTENSION ACT	Congressional Record	2007	Yes
154 Cong. Rec. H588 - DO-NOT-CALL REGISTRY FEE EXTENSION ACT OF 2007	Congressional Record	2008	Yes
153 Cong. Rec. H15265 - DO-NOT-CALL REGISTRY FEE EXTENSION ACT OF 2007	Congressional Record	2007	Yes
151 Cong. Rec. E906 - TRIBUTE TO EILEEN HARRINGTON AND THE DO NOT CALL REGISTRY TEAM	Congressional Record	2005	Yes
149 Cong. Rec. H8916 - RATIFYING AUTHORITY OF FTC TO ESTABLISH A DO-NOT-CALL REGISTRY	Congressional Record	2003	Yes
153 Cong. Rec. E2553 - DO-NOT-CALL REGISTRY FEE EXTENSION ACT OF 2007	Congressional Record	2007	Yes
153 Cong. Rec. (Bound) 34618 - DO-NOT-CALL REGISTRY FEE EXTENSION ACT OF 2007	Congressional Record	2007	Yes
S. Rept. 110-244 - DO-NOT-CALL REGISTRY FEE COLLECTION EXTENSION	Congressional Report	2007	Yes
H. Rept. 110-485 - DO-NOT-CALL REGISTRY FEE EXTENSION ACT OF 2007	Congressional Report	2007	Yes

Govinfo.gov returned the second fewest item type results with five item types, as shown in

Table 11. No CFR. was relevant to the search. All other items were relevant to the search at

hand. Congressional Records were the most frequent result and accounted for sixty-five-

percent of the results in the search with quotation marks. This more than triples the next most

common item type, Bill Texts, which accounted for twenty-percent of the search results. Govinfo.gov did not duplicate any of its search results between the two searches.

Table 11

Govinfo.gov Item Types

Without Quotation Marks				With Quotation Marks		
Item Type	Number of Results	Percent of Total		Item Type	Number of Results	Percent of Total
CFR	20	100%		Bill Text	4	20%
				Congressional Hearing	1	5%
				Congressional Record	13	65%
				Congressional Reports	2	10%

v. **Congress.gov**

Congress.gov is the least useful of the four. Congress.gov produced the second fewest relevant results and nine irrelevant results. This is the only database to return a significant mix of relevant and irrelevant results. Govnifo.gov produced more irrelevant results; however, those irrelevant results were due to difficulties in its basic search algorithms. Congress.gov's

irrelevant results are scattered among the relevant results, consist of different item types, and

occur only in the search without quotation marks. Govinfo.gov returned results in only six out

of the ten possible item types.

Table 12

Congress.gov Without Quotation Marks*

Title	Resource Type	Year	Relevant
S.Amdt.3867	Bill type	2007	Yes
H.R.1585	Bill Text	2007	No
H.R.2454	Bill Text	2009	No
H.R.2768	Bill Text	2007	No
H.R.5136	Bill Text	2010	No
S.1390	Bill Text	2009	No
H.R.116	Bill Text	2009	Yes
H.R.1391	Bill Text	2009	Yes
H.R.248	Bill Text	2007	Yes
H.R.2601	Bill Text	2007	Yes
H.R.372	Bill Text	2008	Yes
H.R.4298	Bill Text	2007	Yes
H.R.5769	Bill Text	2008	Yes
S.2096	Bill Text	2007	Yes
H.R.1105	Pub. Law	2009	No
H.R.2647	Pub. Law	2009	No
H.R.4173	Pub. Law	2010	No
H.R.6523	Pub. Law	2010	No
H.R.3541	Pub. Law	2007	Yes
S.781	Pub. Law	2007	Yes

* Shaded items indicate duplicated results in Congress.gov searches with and without quotation
marks.

Table 13

Congress.gov With Quotation Marks*

Title	Resource Type	Year	Relevant
S.Amdt.3867	Bill Text	2007	Yes
H.R.116	Bill Text	2009	Yes
H.R.1391	Bill Text	2009	Yes
H.R.1950	Bill Text	2005	Yes
H.R.248	Bill Text	2007	Yes
H.R.2601	Bill Text	2007	Yes
H.R.372	Bill Text	2008	Yes
H.R.4072	Bill Text	2005	Yes
H.R.4298	Bill Text	2007	Yes
H.R.4931	Bill Text	2006	Yes
H.R.5769	Bill Text	2008	Yes
S.1652	Bill Text	2003	Yes
S.1654	Bill Text	2003	Yes
S.1655	Bill Text	2003	Yes
S.2096	Bill Text	2007	Yes
S.Res.242	Bill Text	2003	Yes
H.R.3161	Pub. Law	2003	Yes
H.R.3541	Pub. Law	2007	Yes
H.R.395	Pub. Law	2003	Yes
S.781	Pub. Law	2007	Yes

* Shaded items indicate duplicated results in Congress.gov searches with and without quotation marks.

Tables 12 and 13 also illustrate how many overlapping results were in the two searches. This was the only database other than ProQuest to duplicate results. All relevant results from the search without quotation marks were returned in the search with quotation marks. The nine irrelevant results were not repeated, and nine relevant results replaced them. This was a substantial increase in relevant results. Congress.gov did not return all possible item types.

Table 14

Congress.gov Item Types

Without Quotation Marks				With Quotation Marks		
Item Type	Number of Results	Percentage of Total		Item Type	Number of Results	Percentage of Total
Bill Text	14	70%		Bill Text	4	25%
Public Law	6	30%		Public Law	16	75%

Table 14 displays the number of item types per search. One interesting result was that the number of each item types almost exactly inverted. The first search retuned seventy-percent Bill Texts and thirty-percent Public Laws. The second search returned twenty-five-percent Bill Texts and seventy-five percent Public Laws. No explanation was found for this change.

IV. **Discussion**

    a. **ProQuest**

ProQuest had the most diverse search results. It had the most item types and the most similar number of results across the item types. This likely reflects complex algorithms that are designed to provide balanced results to users. Likewise, the lack of change between the search with quotation marks and the one without indicate that the algorithms disregard quotation marks or treat natural language searches as a phrase. ProQuest also had two unique item types: Legislative Histories and Bill Profiles. Both item types are proprietary to ProQuest. Post search filtering would be very useful here to narrow the search by item type. Searching by individual item types would also narrow the search results in a useful way. As with all databases, ProQuest did not return results from every item type possible.

    b. **HeinOnline**

HeinOnline has a good balance of search results overall. It had the second most item types and a good spread of results. HeinOnline favored CRS Reports, which are only found on HeinOnline in this study. Quotation marks impacted the search results in HeinOnline more than expected. The search without quotation marks yielded four item types, where the search without quotes yielded two different item types. Searching by phrases seemed to limit HeinOnline's algorithms. Future searchers should be aware of this phenomenon and run both kinds of searches to ensure a variety of results. As with all databases, HeinOnline did not return results in each of the possible item types. CRS Reports are a research powerhouse, containing hordes of valuable information. This item type alone is a significant reason to search in HeinOnline. As with all databases, searchers should utilize advanced search features. Here, quotation marks seem to greatly influence search results. HeinOnline is particularly useful in historical searches as it is nearly the only database to have historic data.

c. Govinfo.gov

Govinof.gov had the most mixed results of any database. This is a free database. Thus, the algorithms may suffer due to the lack of financial incentive. There is a possibility that the public may view results from govinfo.gov as more official or better than other databases because it is provided by the government and the search material is government documents. This assumption would be unfounded.

Govinfo.gov's search results varied the most widely. Since the searches without quotation marks produced twenty irrelevant results than the search with quotation marks, searchers should always use quotation marks or Boolean operators when searching in

govinfo.gov. Govinfo.gov's algorithms appear to return data based primarily on the order of the search terms and then the order of the databases selected. This indicates the use of advanced filtering or searching by item type to narrow the results and ensure that each item type is returned.

Since this database is free and has a vast collection of documents, user education is of primary concern. Disclosing the algorithms so legal research professionals can produce better guides on searching would be helpful. Teaching users to utilize advanced searching is critical to the successful use of this database because members of the public are the least trained and most dependent upon this database. This database can be a powerful tool, if the searcher knows how to use it properly.

d. Congress.gov

Congress.gov had the fewest item types. These two types of data represented two out of the three possible item types that Congress.gov offers. Overall Congress.gov returned the least helpful results. Congress.gov also produced the only overlap between results in each of the searches, excluding ProQuest's mirrored search results. This database is useful for finding bills or public laws. However, govinfo.gov will produce the same data from a slightly wider timeframe. The overall utility of this database makes this the least useful tool, save for the cost.

e. General

All government document databases are not created equal. Each database was created

for a different purpose.[61] The searcher needs to choose the best tool for the correct

information need. A searcher looking for legislative histories or CRS Reports in Congress.gov

will never find any of these items because Congress.gov does not currently contain these types

of items. Similarly, a search for the Journals of the House or Senate will probably not yield

results in the first search results of ProQuest. These fundamental differences in database

design necessitate legal research instruction. This is true of legal professionals, but especially

true of the public.

The mixed results between databases underline the need to utilize complex search

strategies. All the databases preferred certain item types to others, if only slightly. The

breadth of government documents available also contribute an abundance of search results.

The number of results and preferences within databases make more advanced search features

more useful. Filtering by item type will be very useful to a search in an area with so many types

of items. Similar to case searches that have several item types, limiting searches to certain

collections or item types within the databases will produce more manageable loads of search

results. This complicates and lengthens the search, however, possibly leading to fewer

searchers availing themselves of this method.

A major problem within the government document world is the lack of disclosure from

database providers. Database providers are under no duty to disclose any changes in their

search results, so a default "AND" database may become a default "OR" database without

---

[61] IV.C History of the Databases, *infra*.

warning. The lack of disclosure combined with the underlying differences produces an and every shifting search landscape where novice searchers can easily lose their footing. The shifting sands of algorithms and search features provide unique challenges to searchers that require constant vigilance to overcome.

This Study confirmed that paid databases are the ones that are the easiest to search. Paid databases provide the most balanced search results. The range of item type aids the searcher by ensuring that the searcher is exposed to a wide breadth of information. This breadth can hinder searchers, however, when the search results become too numerous to wade through. Paid databases also provide the only unique item types; some of these item types are limited to one database, as in the databases searched in this study. These unique item types also aggregate other information, thereby magnifying the utility of the database to a searcher. The free databases provided the least variety in their item types returned. The free databases also tended to return the most irrelevant results.

The expected discrepancy between the free and paid government document database imply access to justice issues. The free government document databases are created to disclose government information to the public and to educate and inform the electorate. The difficulty of searching the free databases frustrate these twin purposes. The difficulties in searching the free databases are compounded by the lack of training of the public. Legal professionals often undergo formal and sophisticated training in research strategies and still have difficulty in locating information. Members of the public often have little training in formal research, especially in legal research. This creates an inverted world where the trained researches have access to the resources that provide the best and simplest searches. The

36

public have access to the databases with the least diverse search results, the least sophisticated search algorithms, and the most limited data sets. This underscores the need for research training.

Two primary areas of research training are needed. First, a training regimen ensuring that students are trained to search well. Teaching students the fundamental principles of research is more important than teaching them how to search. Search methods have changed over the years. Print materials have become digital but the principles are still the same. Search tools limit and sort the amount of data through which a researcher must comb whether they are print or digital. This principle will remain the same whenever a database provider updates the layout or algorithms of their search platform. The specifics of how this filtering is accomplished will change. Similarly, other principles of research remain the same when the methods change. Choosing the correct tool is just as important when choosing the right print resource or electronic database. *Corbin on Contracts* is about as likely to give the research information on intellectual property as govinfo.gov is to produce a CRS Report.

The second important research training area is public outreach by law librarians. Just as training legal professionals to search well is vital, so too is it important to train the public to search well. All the lesson from legal professional are true for the public. The difference is that the public often cannot afford specialized databases or legal professionals and must resort to free research databases. Specific recommendations on how this public training should occur is beyond the scope of this study.

V.   Future Research

Significantly more research is needed in the government document database area. First, this study was limited by the number of searches. This limited data set makes drawing certain results about the algorithms of the databases impossible. The best this study can do is to draw general inferences and point future researchers in the right direction. This study was focused more on learning what could be drawn from basic searches for instructional purposes, rather than an in depth statistical analysis.

Searches should also be conducted in different areas of the country and using private browsing without enabling location searching location services. Google uses a searcher's location to determine relevant results.[62] Google also uses past search results to determine relevant search results. Blocking cookies, location tracking, and utilizing a variety of searchers in locations in different states should ensure a diverse data set that could provide insight on how databases, especially paid services, utilize data. This is important to consider with the challenges to privacy that social media has brought to the public's attention.[63]

Another limitation of the study was the use of no filters or other post-search sorting. Filtering and post-search sorting can be very useful to a searcher by narrowing the search results to exclude irrelevant results. This study did not utilize these tools to avoid skewing the results toward certain item types. This study presumed a basic natural language search and one or two pages of results was the data set of a generic first year law student. This will not necessarily be representative of the diverse populations of first year law students and their experiences.

---

[62] Find source.
[63] Age of data privacy

Searches by individual item type, especially in govinfo.gov, would yield a better comparison of each database's algorithms because this data set was too large to return significant results on individual item type within each data set. The nature of govinfo.gov's algorithms also meant that data was often clustered into groups based on the originating collection. Govinfo.gov likely searches by collection in alphabetical order of the collections selected. Many more searchers are necessary to confirm these findings and to discover additional quirks.

More searches with and without quotations within ProQuest should be run to determine whether quotation marks matter. The limitations of this study did not provide enough data to test this theory, nor was this phenomenon expected. Further searches would demonstrate whether the breadth of search results held over time. Further searches should also be completed in HeinOnline to confirm that the variety of search results holds true.

VI.   Conclusion

Databases exhibit characteristics that should be taught to students. Quotation marks matter in most databases. Unless the searcher is using ProQuest Congressional, then the search results vary significantly from database to database. Quotation marks are only a first step in the process of searching. Other relevant steps, like filters, post-search sorting, likely have as strong of an impact to the search results as the use of question marks. Other Boolean Operators may impact the search results as well. Legal research instructors should ensure that they teach students more advanced search features. The idea that natural language searches alone are sufficient for an adequate search is false. Certain databases have preferences for

specific item types, especially if it is their own content or content that they license. Students should be instructed in these preferences.

This study revealed several latent access to justice concerns within the government document searching arena. Free search databases are more difficult to use and produce fewer item types and more irrelevant results than their paid peers. Two steps would help to remedy this issue. First, ensuring that legal professionals are well trained to maximize the utility of databases helps promote access to justice by providing quality attorneys who can use their skills to serve those who don't have access to the courts through pro bono legal work. Second, legal research professionals can provide outreach programs to provide some training to the public.

Legal research professionals can assist the public by offering free legal research instruction. Govinfo.gov has a detailed help page that includes tutorials.[64] These tutorials are not sufficient, however, and further instruction is needed to train the public in the use of this database. Legal research professionals can partner with their local libraries, local bar associations, legal aid societies, and other public service organizations to provide free instructional programs to interested individuals. These classes should focus on teaching search strategies, including limiting source material, post-search sorting, Boolean operators, and the use of quotation marks. These courses will only reach a limited number of persons. However, these courses would provide an opportunity for legal research professionals to give back to their communities in a way that is often difficult to do outside of their daily work. This is

---

[64] History, govinfo. https://www.govinfo.gov/help (last visited July 24, 2018).

especially true if the legal research professional's institution is not open to the public. Legal

research professionals who do not routinely interact with the public should review their state's

rules regarding the unauthorized practice of law prior to teaching any course to ensure that

they maintain their ethical obligations.

Bibliography

"Advanced Search." *govinfo.gov*, https://www.govinfo.gov/#advanced.

American Association of Law Libraries. "UNDERSTANDING THE HUMAN ELEMENT IN SEARCH ALGORITHMS (A7)." https://www.aallnet.org/recording/understanding-the-human-element-in-search-algorithms-a7/.

"Content coverage Chart." *ProQuest*, https://congressional-proquest-com.offcampus.lib.washington.edu/congressional/saleable/contentcoverage?accountid=14784&groupid=95339.

Gerken, Joe. "Invention of HeinOnline: The Story of Hein's Dramatic Transformation in Response to the Coming of the Electronic Age," *18 A.A.L.L. Spectrum*, February 2014, at 17.

HeinOnline. "U.S. Congressional Documents." https://heinonline.org/HeinDocs/USCongressionalDocumentsCollection3.pdf.

"History." *Govinfo*, https://www.govinfo.gov/help

"History & Milestones" *ProQuest*, https://www.proquest.com/about/history-milestones/.

Library of Congress. "Coverage Dates for Legislative Information." *Congress.gov*, https://www.congress.gov/about/coverage-dates.

Lowe, M. Sara, Bronwen K. Maxson, Sean M. Stone, Willie Miller, Eric Snajdr, & Kathleen Hanna, "The Boolean Is Dead, Long Live the Boolean! Natural Language versus Boolean Searching in Introductory Undergraduate Instruction." College and Research Libraries, vol. 79, no. 4, 2018, pp. 517-534.

Mart, Susan Nevelow. "The Algorithm as a Human Artifact: Implications for Legal [Re]Search, *Law Library Journal.* vol. 109, 2017, pp. 387-422.

Porter, Brandi. "Millennial Undergraduate Research Strategies in Web and Library Information Retrieval Systems," Journal of Web Librarianship, vol. 5, no. 4, 2011, pp. 267-285.

"ProQuest Acquires Acclaimed Congressional Information Service and University Publications of America from LexisNexis." *ProQuest,* https://www.proquest.com/about/news/2010/ProQuest-Acquires-Congressional-Information-Service-LexisNexis.html.

"Shepard's Citations Service." *LEXISNEXIS* https://www.lexisnexis.com/en-us/products/lexis-advance/shepards.page.

"Topic and Key Number Overview" *WESTLAW*, https://lawschool.westlaw.com/marketing/display/RE/24.

U.S. Government Publishing Office. "What's Available." *govinfo.gov*, https://www.govinfo.gov/help/whats-available.