WESTLAW is Natural v. Boolean Searching: A Performance Study

Sheilla E. Desert

Available at: https://works.bepress.com/aallcallforpapers/45/
1993 Call for Papers

WESTLAW Is Natural v. Boolean Searching: A Performance Study*

Sheilla E. Désert**

Ms. Désert reports on WESTLAW's new natural language search method and compares it to the traditional terms-and-connectors search method, using her own queries and those of law librarians and other legal information specialists.

I. Introduction

On September 16, 1992, at a New York press conference, West Publishing Company unveiled its new "WESTLAW Is Natural" (WIN) technology. It became available to the top 500 law firms on October 1, 1992, and to educational law school programs and other users in late November.

WESTLAW launched the WIN software for its case law databases. According to WIN creators Dr. Howard Turtle and Jim Olson, WIN is currently available on over 3,700 databases. It is in all full-text databases except DIALOG, Directories, Name Fields, and Prentice-Hall Abstracts and Filing. A user can check the "Scope" of any database to identify whether WIN is available. A hard copy of all the databases using WIN can be obtained from West Publishing.

WIN is first in the legal research market in allowing searchers to enter natural English language questions, which the system then translates into correct search strategies. WIN's innovative and unique feature is that "[u]nlike Boolean logic, which performs a literal match of search terms and logic embodied in a query, WIN performs a statistical analysis on each term or phrase against each document in the database."1

* © Sheilla Désert, 1993. This is a revised version of a winning entry in the 1993 AALL Call for Papers competition.

** Graduate Student, Post-Masters, Certificate of Advanced Studies: Library Automation Systems Specialist, University of Pittsburgh Graduate School of Library and Information Science; Student Law Librarian, University of Pittsburgh Barco Law Library, Pittsburgh, Pennsylvania.

This article first discusses how WIN works and places it within the context of other information retrieval systems. It then compares the results of a number of searches run using both WIN and Boolean searching, and reports follow-up comments on the search results by the designers of WIN via e-mail and in a September 1993 interview with Turtle and Olson at West Headquarters.

II. WIN in Comparison to Other Models

A. Modern Information Retrieval Models

An information retrieval model indicates the representations used for documents or objects and how they are compared during the retrieval process. Teresa Pritchard-Schoch provides an excellent description of three modern retrieval models: exact match, vector space, and probabilistic models.

The Boolean logic model falls within the definition of an exact match model. A document is retrieved using Boolean logic by matching defined criteria with the variables associated with a document. Each criterion has been assigned a truth variable as a correct description of the document. . . .

The vector space retrieval methods use weighting methods. The most common are weights . . . based on the frequency of a term in a single document, and its frequency in the entire collection. Probabilistic retrieval is based on the premise that the best overall retrieval effectiveness will be achieved when documents are ranked in decreasing order of probability of relevance.  

The traditional method used to evaluate information retrieval systems has been to measure precision and recall. "Precision is the proportion of a retrieved set of documents . . . relevant to a query, while recall is the proportion of documents in the collection . . . relevant to a query that is actually retrieved." One problem with Boolean searching is that the user has no explicit control over the number of documents retrieved. The system might retrieve 500 documents or none, while the user might want to see between ten and twenty relevant documents.

The vector and probabilistic models offer better means of retrieving relevant documents. Both use recall to broaden the scope of the search and retrieve more documents. Recall is aimed at replacing infrequent and very specific words with more general words. Precision, on the other hand, is intended to make the search more precise and, thus, retrieve fewer.

2. Teresa Pritchard-Schoch, Natural Language Comes of Age, ONLINE, May 1993, at 33, 34.
3. Id. at 34.
documents. Precision can be used to replace general words with more specific words, as well as to restrict the search to specific areas of the database. WIN represents a method used in attempting to retrieve the most relevant documents.

B. How WIN Works

WIN uses many algorithms to translate a natural language query. Four main steps are involved: (1) identification of key concepts, such as a phrase, citation, or word; (2) removal of "stop words," such as "a," "the," "what," and "when"; (3) application of stemming programs and linguistic techniques to determine root words, derivatives, and expansion of the roots; and (4) performance of statistical analysis and statistical comparison of retrieved concepts after determining the relevant cases in the WESTLAW database.

The first step that WIN performs is to identify key concepts. The query is passed through a machine-readable legal dictionary, which allows the system to identify two or more words (for example, "insurance policy") as a legal concept and search them as a unit.

The system does not automatically supply synonyms, nor does it prompt the user to ask for them. West did not make the expansion of thesaurus terms automatic, due to the possible multiple meanings of English language words (for example, homonyms). The search can be broadened, however, by using West's online thesaurus. Before or after the search has been run, the user can enter the "THES" command, which prompts the system to supply additional terms. The user can then insert new terms to supplement the query or create phrases by enclosing the words in quotation marks. Thus, the user is not limited to system-induced phrases, but can identify any group of words as phrases.

WIN's second step is to run the query through a "stop list," which removes all words too common to be of any true value in the search. If the stop word is part of a phrase, such as "at will," or "in personam," however, it will not be removed.

The third step, the "stemming program," strips each word down to its most basic, constituent parts. WIN then applies a program to add the derivatives of the roots. For example, the word "incorporate" would find the words "incorporated," "incorporating," "incorporation," and so on. This is similar to truncation in terms-and-connectors searching.

The fourth step, statistical comparison analysis based on a series of complex algorithms, is the most complicated. Instead of looking for terms, the system searches the database for natural language hits or concepts. WIN identifies relevant cases in the WESTLAW database and performs a
statistical comparison of concepts (that is, a word, phrase, key number, or
citation) to come up with the best statistical matches for the query.⁴

Sara Jankiewicz describes this statistical analysis: WIN "(1) [e]valuates
the concepts in the query, assigning frequency weights to them; (2)
[e]valuates documents in the database for these concepts and overall
occurrence; and (3) ranks the results for 'relevancy' based on measured
concept frequency."⁵ To accomplish this frequency weighting, WIN uses
a method of concept weighting that emphasizes certain terms from the
query against terms contained in the database being searched.

A high weight is apportioned for terms appearing in fewer documents
and low weights are given to terms appearing generally in many
documents. "It is considered that the fewer [the] documents in which
the word appears the higher the discrimination power of the term. The
similarity measure between a document and a query involves compar-
ing the sum of the weights of the query terms in the document." This
explains how documents can be ranked in order of their approximate
relevance to the query. Retrieval is based on perceived word impor-
tance and word occurrence.⁶

WIN's creators point out, however, that "the approach described by
Sara Jankiewicz is related, but that current retrieval models are quite a bit
more sophisticated than this description would suggest. In particular, WIN
is based on a substantially more advanced theoretical model."⁷

Two main problems arise when constructing a thesaurus: first, a
decision must be made about what terms to include, and second, the terms
specified for inclusion must be suitably grouped.⁸ Salton and McGill offer
four principles for constructing a thesaurus:

1. The thesaurus should include only those terms likely to be of
interest for content identification in a subject area (for example, a term
such as "hand" might be used in a thesaurus dealing with biology, but
it should not be included if its frequency of occurrence is due largely
to expressions such as "on the other hand").
2. Ambiguous terms should be coded only for those senses likely to be
important in the document collection (at least two thesaurus categories
should thus be used for a term such as "field," corresponding on the
one hand to the notion of subject area and on the other hand to its

⁴ Daniel E. Harmon, The New WESTLAW: English-Language Queries and Automatic Analysis of Case
Relevancy, LAWYER'S PC, vol. 9, no. 25, 1992, at 1, 2 (paraphrasing Jim Olson).
⁵ Sara Jankiewicz, Natural Language in Plain English, HALL NEWSL., Dec. 1992, at 9, 11.
⁶ Id. (quoting a telephone conversation with Julie Tate of West Publishing Co. (Nov. 20, 1992)).
⁷ Electronic-mail communication from Howard Turtle and Jim Olson, developers of Westlaw I
(WIN), West Publishing Co. (Apr. 29, 1993) [hereinafter E-mail from Turtle and Olson].
⁸ Gerard Salton & Michael J. McGill, INTRODUCTION TO MODERN INFORMATION RETRIEVAL 76
(1983).
technical sense in algebra; no provision need be made to cover the
notions of "a patch of land" if the thesaurus deals with the mathemati-
cal sciences or related technical fields).
3. In order to obtain good matching characteristics between query and
document terms, each thesaurus class should include terms of roughly
equal frequency; furthermore, the total frequency of occurrence should
be as close to equal for each class as possible, thus ensuring that the
probability of producing a match between queries and documents is
approximately equal for all thesaurus classes. (If these frequency
characteristics are grossly violated—for example, if a high frequency
term such as "computer" is entered into the same class as a more
specific term such as "minicomputer"—queries about specific topics
will produce general responses, thereby depressing the precision of the
search.)
4. Whenever possible, terms with negative discrimination values should
be eliminated; even if the size restrictions that control the thesaurus
construction do not immediately lead to the elimination of all high-
frequency nondiscriminators, the latter are best relegated to thesaurus
classes of their own (their classification together with lower-frequency
terms would produce low-precision output).9

Another concern, especially in a legal database where new cases create
new terms of art, is the question of updating. As discussed below, a term
such as "e-mail" cannot be found by a natural language search, because it
was not entered into the thesaurus as one term.10 WIN does not add new
terms to the database at the time new cases are entered. This will need to
be remedied if the thesaurus is to be a fully integrated part of the system.

Natural language programs that allow user-friendly front-end queries
have existed since the 1960s,11 and artificial intelligence software that
applies statistical analyses is now on the market. How does WIN differ
from these systems? According to Turtle, four factors make WIN unique:
"(1) the scale of the collections it will search (virtually the entire
WESTLAW database), (2) the use of 'relatively deep linguistic analysis',
(3) "a very extensive set of inference techniques",12 and (4) "a sound
theoretical model."13

9. Id. at 77-78.
10. See infra pages 736, 737-38.
11. An early system was designed in 1967 by Cherie B. Weil, a librarian, to aid in accessing biographical
12. Quoted in Harmon, supra note 4, at 2.
13. E-mail from Turtle and Olson, supra note 7. In clarifying WIN's uniqueness, Turtle states:
A "sound theoretical model" is where WIN departs from other Natural Language programs. WIN is based
on a formal retrieval model that is intended to allow multiple sources of evidence about document and
query content to be combined in order to assess the probability that a document matches a query. The
WIN retrieval model is based on Bayesian inference networks which provide a probabilistic framework
within which multiple evidence sources can be represented and used in the inference process. WIN makes
Pritchard-Schoch identifies two other programs, Personal Librarian and Dow Quest, which provide more powerful natural language capacity. An important aspect of Personal Librarian's "relevance ranking," according to its creator, Matthew Koll, is that it bridges the gap between the "Boolean logic, inverted file traditionalists and the Salton vector space probabilistics." The relevance ranking is based on five factors:

- The frequency of each query word within a document is an indicator of relevance.
- A match on a word that is rare within a given database is more indicative of relevance than a match on a word that is common in that database.
- Since long documents have a greater chance of containing query words, the weighting needs to be adjusted so that they do not rank unduly high.
- Sometimes the fact that query words appear near each other within a document is an additional indicator of relevance.
- Regardless of theory, users have an intuitive sense as to how well the relevance ranking is working and no algorithm should rank a document containing only two query words from an eight-word query higher than a document containing all eight words.

WIN does not provide these added features. As the discussion of questions 9 and 10 below makes clear, the system does not include Koll's fifth requirement, and Turtle and Olson refute its validity. They point out that there can be no hard-and-fast rule about the ranking of an eight-word query versus a two-word query, and that, at times, a two-word query may be ranked higher than an eight-word query, if the two-word query is in fact more relevant.

Stephen Weyer describes DowQuest in a 1989 *Online* article:

Matching is performed by tallying word occurrences, combining scores for different words (basically using an implicit Boolean "or" between query words) and then normalizing the score to take into account the size of documents and length of query. DowQuest then sorts the documents by their scores and displays headlines in this

use of several sources of evidence, including traditional retrieval measures (e.g., concept frequency in a document and in the collection) as well as several sources based on linguistic analysis (e.g., morphology, phrase structure) and on domain knowledge (e.g., citation structure, West Key numbers, legal thesaurus).

E-mail from Turtle and Olson (Sept. 15, 1993).
14. Pritchard-Schoch, supra note 2, at 38-42.
15. Id. at 40 (quoting Matthew Koll).
16. Id.
17. See infra pages 736-37.
18. Interview with Howard Turtle and Jim Olson, West Publishing Co., in Eagan, Minn. (Sept. 10, 1993) [hereinafter Interview with Howard and Turtle].
order. Documents containing more of the query words and more of each particular word will be generally higher in significance. Also, documents with phrases . . . where query words occur closer together are ranked more highly than documents in which the query words are scattered.\textsuperscript{19}

**C. Pitfalls of Natural Language**

Although WIN and other natural language software are bound to dominate the market in the coming years, there are a number of problems with the effectiveness of document retrieval from full-text systems. Dan Dabney, in a 1986 article, identifies three problem areas: synonymous words, ambiguous words, and complex expressions.\textsuperscript{20}

Synonymous words complicate retrieval because computers search literally, not conceptually. The computer cannot tell whether the search term and the synonym are the same or are relatively close.

Ambiguous words are a problem because one word can represent more than one concept. For example, consider that a computer search for the drug DES will also retrieve documents that contain references to the city of Des Moines.

The obstacle encountered with complex expressions is the implicit meaning contained in the expression itself. Consider the following example: If a party fails to plead a cause of action, can evidence of that cause of action be introduced at trial? The computer is required in this instance to conceptualize the relationship between the query and overall contents of the database. This creates a problem because, before WIN, computer searching was literal, not conceptual.

The power of algorithms in syntactic and semantic analyses allows natural language processing systems to address these problems effectively. WIN uses a number of models to process the search terms through statistical analysis.\textsuperscript{21}

**D. Boolean and Natural Language**

Boolean logic connects query terms. The database then performs a literal match of search terms and logic embodied in a query. WIN, on the other hand, performs a statistical analysis on each term or phrase against the database. In a comparison between the two search systems, Richard

\begin{itemize}
\item \textsuperscript{19} Stephen Weyer, *Questing for the "Dao": DowQuest and Intelligent Text Retrieval*, ONLINE, Sept. 1989, at 39, 44.
\item \textsuperscript{21} See supra § II.B.
\end{itemize}
Black, Vice President of Marketing and Sales for Personal Library Software, stated, "A Boolean searcher creates a formula and sticks with the formula throughout the search. [The researcher is] bound by the formula, which puts blinders on the search."22

Griffith describes these "blinders"--the exclusive terms used in the search phrase--as both strengths and weaknesses of Boolean research.

While the search phrase can be modified, adding synonyms or changing the proximity in which the terms must appear, the documents retrieved must include the terms of the search.

Natural language research is different. It uses the terms of the search phrase as a starting point for retrieving documents. If none of the documents in the database contain the terms of the natural language query, it will retrieve and rank results in order by relevancy.23

Griffith uses the following example: in a natural language search, the query "wheat sales in Russia" may also retrieve grain sales in the Soviet Union or the Ukraine, whereas the Boolean search, bound by specificity, would not.24

**E. Features Offered by WIN**

By typing "nat" at the prompt, the researcher enters the brave new frontier of natural language searching in the legal universe. WIN offers six important features:

1. **Twenty-document default limit** (can be reset for up to one hundred documents). Based on the assumption that users typically are searching for the most relevant cases, WIN assumes that twenty documents will most often satisfy the users' needs.

2. **Display.** WIN results are displayed not in reverse chronological order, but in order of relevancy. Thus, the first relevant case may be a 1952 case, while the tenth relevant case may be from 1992.

3. **Thesaurus.** Before or after running an initial search, users are allowed to identify concepts in the thesaurus related to the issue and add them to the description as alternative terms. Users can add alternative terms by enclosing them in parentheses immediately following the search term.

---

23. Id. at 46.
24. Id.
4. **Automatic phrasing of terms.** WIN has a phrase dictionary that automatically processes common phrases and terms of art. Users can also create their own phrases by enclosing the phrases in quotation marks.

5. **Browse.** This feature is similar to the terms-and-controllers "term mode." Users can browse all cases for relevant terms. Cases do not necessarily contain all the key terms included in the query because only the most relevant documents are selected and retrieved by WESTLAW.

6. **Best.** This feature displays the portion of the viewed document most likely to match the terms in the query.\(^{25}\)

### III. Experimental Queries and West Responses

To test West's natural language software, I ran thirty experimental WIN and terms-and- connectors queries on a variety of issues on several databases, noting patterns and problems that began to emerge. I compared these with queries run by "power searchers" (computer specialists) in published studies, and discovered that they had identified similar, as well as additional, results. To provide a broader perspective and not duplicate results, I then ran test queries to see how the system would adapt and react to possibly confusing terms. Through the LAWLIB e-mail conference, I gathered additional queries from law librarians on their experiences and comparisons with WIN and with the traditional terms-and-connectors search method.

---

25. In my on-site interview with Turtle and Olson on September 10, 1993, at West Publishing Headquarters in Eagan, Minnesota, several new enhancements were demonstrated to me:
1. WIN searches in case law databases have been enhanced as a result of the enhancements made to all case law documents. Improved headnotes and West Topic and Key information has been added to all case law documents. WIN, which makes extensive use of West's editorial material, uses this new information to improve WIN searching.
2. Changes have been made to recognize WIN descriptions that begin with "find." For example, WIN descriptions that begin with "find all cases that..." are no longer intercepted as possible FIND commands, and are properly processed as WIN descriptions.
3. Many additional thesaurus entries, phrases, and introductory clauses have been added. Customers are encouraged to call WESTLAW with suggestions of new phrases or words, or terms of art that they felt should be added to the thesaurus. The "e-mail" query (see infra page 736) is an example of a term that has been added to the thesaurus. It will no longer be split.
4. Because the BEST command proved to be popular, it was expanded into a browsing mode.
5. The THES (Thesaurus) and RES (Restrictions) commands can now be entered while browsing a WIN result.
6. If a user has entered a citation or key number as part of a WIN description, these concepts will be displayed following a "Special Concepts" label on the "Search Is Proceeding" screens.
7. When transferring a query from terms-and- connectors to WIN mode, any date, judge, and attorney restrictions are extracted from the terms-and- connectors query and placed in the appropriate field(s) for WIN searching.
After analyzing results from these groups of queries, I chose representative samples from each group to illustrate and highlight my findings. With this data, I prepared a set of questions for Turtle and Olson, the creators of WIN. This section presents the first three sets of queries (my own, those of power searches, and those of law librarians), followed by the responses of Turtle and Olson. The following section presents additional searches run after speaking with Turtle and Olson and their subsequent responses.

A. My Own Queries

1. WIN: What constitutes handicaps under state legislation forbidding job discrimination on account of handicap?
   Boolean: Legisl! or statut! or section or code w/50 employment or job w/20 discriminat! or disqualif! w/10 handicap!

   This query was performed in the Law Review database. In WIN, the first two relevant articles dealt with "AIDS" as a handicap, but did not include job discrimination. The most relevant current article dealing with the Disabilities Act was ranked eighth in relevancy. The same article was seventh in Boolean. The article that answered the question on point was ranked ninth using Boolean, but was not found using WIN. The phenomenon of finding the most relevant and important item in Boolean but not in WIN also occurred in a number of my other searches. WIN did not appear to live up to its claim that the most relevant articles would be included with the first twenty documents.

2. WIN: What is an employer's liability to an employee for failure to provide a work environment free from tobacco smoke?
   Boolean: Tobacco or cigarette or smoke /p workplace or smoke-free /p environment or workplace.

   I purposely ran this in the Supreme Court database, knowing no case would be on point. Instead of indicating "no documents satisfy your search," however, WIN searched the database and produced a gamut of irrelevant answers. It appeared that WIN was designed to provide twenty cases, even if they are irrelevant. Forcing the system to fill the twenty-case requirement creates skepticism as to the true reliability of the system. Boolean, on the other hand, confirmed that no cases could be found to match my query. 26

---

26. In a later discussion, Turtle and Olson pointed out:
When I ran the same query in the ALLSTATES database, I discovered that on WIN, the first- and seventh-ranked cases were the same case (the first was a Supreme Court case, and the seventh a Court of Appeals case). The appearance of the same case twice in the list of the twenty most relevant cases excluded another case that could have been of greater or equal importance. Another critical problem was that the most recent case was from 1991. Boolean searching, on the other hand, found a 1993 case. This, again, points out a weak link in the issue of relevance. There must be a way of including at least one up-to-date case in the list of twenty responses. In all the searches I performed, no relevant 1993 document was included in any of the WIN results.

3. WIN: The public policy exception to employment at will doctrine.  
   Boolean: "public policy" /s exception /s employment /s "at-will".

   I ran these searches exclusively in the ALLSTATES database to see if the 1959 landmark case\(^\text{27}\) would be included in WIN's top twenty, but it was not found. As pointed out below, I attempted similar searches in other databases and never found the landmark case. Novices using WIN will assume, as did I, that landmark cases will be found by WIN. Only because of my previous knowledge of this area of law did I know that a critical case was not among the top twenty. If WIN is to be truly effective, it should become mandatory practice at WESTLAW to include within the top twenty cases the landmark case covering the topic or area of law.

4. WIN: Long arm statute and in personam jurisdiction.  
   Boolean: "long arm" and "in personam" and juridict!

   The landmark case *International Shoe\(^\text{28}\)* was decided in 1945. I searched both old and new U.S. Supreme Court databases using WIN, and

---

WIN is designed to retrieve the cases that are most likely to match your query, but there is no guarantee that they are, in fact, relevant. This is particularly true when there are no relevant cases in the collection being searched. In general, relevance assessments are difficult for humans to make even when the problem statement is perfect. WIN does not attempt to make direct assessment of relevance; it attempts to rank cases in an order that will present those cases that are most likely to be judged relevant by a human first.

With respect to your statement that the Boolean query "confirmed that no cases . . . match [your] query," there is a substantial body of evidence that suggests that failure of a Boolean query to find any documents is fairly weak evidence that no relevant documents exist. The "Curse of Thamus" article you cite and the research of Blair and Maron on which it is based both point out that recall with Boolean searching is much lower than expert searchers believe it to be.

E-Mail from Turtle and Olson, supra note 7.

the case did not appear in either. However, in the current Supreme Court
database, the WIN search located the other leading cases, including Burger
King\textsuperscript{29} (ranked sixth), World-Wide Volkswagen\textsuperscript{30} (eighth), and Shaffer\textsuperscript{31}
(fifteenth). The top five cases found were less relevant. It was a surprise
to discover that the three leading cases ranked below five less relevant
documents. In Boolean, the three cases (ranked according to date) were
Volkswagen (third), Shaffer (fourth), and Burger King (tenth). Thus,
relevant cases would be easier for a researcher to find in the Boolean
search.

Another problem was the splitting of "in personam" by the system,
even though I put the required quotes around the term as instructed by
WESTLAW.\textsuperscript{32} This greatly diminishes the effectiveness of WIN. In
Boolean searching, a phrase created within parentheses does not split the
phrase. Unless WIN can assure users that phrases in quotation marks will
not be split, users will become highly skeptical with results that do not
follow the system's existing standards.

5. WIN: Race as a factor in adoption or custody proceedings.
Boolean: Race or racial /p factor /p adopt! or custody w/50 child /p
proceeding.

I ran this search to see what WIN would do with the term "race." The
first four cases were on point. The eighth and ninth ranked cases, however,
had to do with horse racing. I discovered the same pattern in a number of
other searches. To look for clues for why some cases about race and child
custody or adoption were ranked below horse racing, I turned to WEST-
LAW's thesaurus. "Ethnic," the only term that was applicable, was of little
use in solving this dilemma.

In running the same search on later occasions, I discovered that results
differed. At first, I thought this was because I had run so many different
queries that I was beginning to blur the results. However, the phenomenon
was confirmed by other searchers. Duane Strojny, Head of Reference at
Marquette University Law Library, stated, "I agree that the promotion
behind WIN is misleading. We have found that when two people conduct
the same exact search, they may get different results. Certainly, such
results do not promote effective CALR."\textsuperscript{33}

\textsuperscript{29} Burger King Corp. v. Rudzewicz, 471 U.S. 462 (1985).
\textsuperscript{30} World-Wide Volkswagen Corp. v. Woodson, 444 U.S. 286 (1980).
\textsuperscript{32} A similar problem occurred with the word "co-ed." See infra pages 727, 730.
\textsuperscript{33} LAWLiB e-mail conference communication from Duane Strojny, Head of Reference, Marquette
I concluded that WESTLAW's loading of new cases into the databases every day changes the statistical weighing. If new cases have some degree of relevance to a particular search query, they will impact and shift which cases are included in the top twenty. When asked about my conclusion, Turtle and Olson stated:

It would help if we knew when these searches were run and the user i.d. of the searcher so that we could look at the transaction logs to see what was going on. Frankly, in most of the cases that we've tracked down, the difference in the ranking resulted from a difference in the query (e.g., a misspelling, a change in word order that prevented phrase recognition). A couple of cases did lead us to bugs, but we are not aware of any outstanding bugs that would significantly alter the ranking.

The addition of new documents to a collection does alter the underlying probability distribution which can alter the ranking, but for collections of the size present on WESTLAW these changes are generally too subtle to be seen by users. Certainly, two users performing the same search on the same database at roughly the same time should get identical results. The same search run before and after a load should differ noticeably only if some of the newly added documents were good enough to get into the top ranking. If you can provide us with examples of searches that produce different rankings, we'd be very interested—it's possible that you've uncovered a problem, but the only way to find out would be to look at specific queries.34

**B. Power Searchers' Queries**

1. WIN: What waiting periods are legal for an abortion?35

Boolean: "waiting period" /s abortion

This query was run in the ALLSTATES database. WIN produced twenty very relevant cases ranging from 1992 to 1952. Boolean produced twenty-nine cases from 1992 to 1915. Only five cases were common to both. The most relevant WIN case was a 1981 Illinois case (ranked tenth in Boolean). All the WIN cases were relevant, whereas some of the Boolean cases had the terms but were not necessarily relevant. Boolean lists cases only in reverse chronological order and does not rank according

---

34. E-mail from Turtle and Olson, supra note 7. Olson later added, "In all the research we have done pertaining to this particular problem we have found that the ranking change was due to a human error in retyping the query." Interview with Turtle and Olson, supra note 18.

35. Queries 1, 2, 3, and subsequent discussion are from Richard A. Leiter, WIN: "It's the Natural Way," LEGAL INFO. ALERT, Nov./Dec. 1992, at 1, 1-4.
to the cases' relevance. Leiter noticed that each of the WIN cases actually discussed the issue he was interested in, rather than simply mentioning it.\footnote{36. Id. at 3.}

2. WIN: When is intoxication a defense to assault and battery?
   Boolean: Drunk! or intox! w/15 defense w/10 assault! [or] battery

   This query was apparently run in the ALLSTATES database. There were more pronounced differences in this search. Without exception, the twenty cases in WIN were relevant. The dates ranged from 1991 to 1972. The Boolean search had 304 cases, with dates ranging from 1992 to 1924. Five of the cases in WIN were not listed in Boolean, but the top three WIN cases were in Boolean.

   All of WIN's cases were highly relevant, compared to perhaps a five to ten percent (or more) irrelevancy in Boolean's cases. It is also more manageable and less daunting to view 20 cases rather than 304. As Leiter points out, a search like this usually will not be conducted in such a broad database; in a specialized database, fewer cases would be pulled up.\footnote{37. Id.}

   WIN's most relevant case appears sixty-ninth on the Boolean list. A researcher would have to be extremely diligent to pore through that many cases to find the best one.

   On the other hand, WIN's number one case was from 1982. No matter how relevant the twenty cases are, Leiter notes, if they do not include some from the last few years, it is hard to put one's faith in the result. He suggests that "date of decision" should be factored in as an element in ranking cases.\footnote{38. Id. at 3-4.}

   The system also presents cases as if they are equally valuable, regardless of jurisdiction. This causes concern because an older, lower court case from another jurisdiction is of little value in presenting a case at the state supreme court. It is therefore important to select your database carefully.

3. WIN: When is a college or university liable for personal injuries of a student?
   Boolean: College [or] university w/10 liab! [or] neglig! w/10 injur! [or] assault!

   Of the 20 WIN cases, 10 appeared in the list of 327 Boolean cases. Only one of WIN's ten most relevant list was found in Boolean. This search pointed out some important quirks and oddities of the system. Leiter
added the synonym "pupil" to the WIN search and found that the list of retrieved cases varied only slightly: four additional cases were retrieved, and two were in the reverse order.\footnote{39} This points out two important features of WIN that users need to understand to use WIN effectively.

First, adding synonyms appears not to affect results dramatically. Because WIN does not search only proximity relationships of terms, adding synonyms only affects the search in subtle ways. The legal terms identified in a query will continue to have the same weight and touch the same cases, no matter what words are used to identify parties. Adding synonyms broadens the search in a purer sense than it would in the Boolean environment.

Second, the addition of synonyms changes the rank of the cases. Rank is determined by relative linguistic and statistical weightings and frequency of occurrence. Adding "pupil" to the WIN search showed that it had a different weight—perhaps it appeared more frequently in the case. Leiter then added another synonym, "co-ed." Surprisingly, WIN dropped "co" and left the term "ed," and the result was a very different list. When Leiter asked Turtle and Olson about this, they were initially surprised at the results, but after considering the question, they decided that WIN probably interpreted "ed" as a personal name and gave it top relevance.\footnote{40}

I was disturbed by Turtle and Olson's response and wondered why Leiter did not question the fact that "co-ed" was put in as one term. How and why did the system split the term into two words when it was within the required parentheses?\footnote{41}

4. WIN: What rights do I have against a real estate agent who gave me wrong information about the size of my lot and the location of my property boundaries?\footnote{42}

Boolean: "real estate" /p misrepresent! or incorrect /s boundary or description (lot w/3 size)

Of the twenty cases retrieved by WIN, two were relevant and two were helpful analogous cases. Boolean retrieved two cases, one that was relevant and one that was not.

\footnote{39. Id. at 4.}
\footnote{40. Id.}
\footnote{41. Turtle and Olson's response appears infra page 730.}
\footnote{42. Queries 4, 5, 6, and the subsequent discussion are by Teresa Pritchard-Schoch. I discuss them together because they present other interesting and unexpected complications with the system. Teresa Pritchard-Schoch, \textit{WIN—WESTLAW Goes Natural}, \texttt{ONLINE}, Jan. 1993, at 101, 101-03.}
5. WIN: Is there a cause of action against a former customer who makes false statements about a product?

   Boolean: False or untrue or misleading or incorrect or derogatory /p
            (former or previous or prior w/3 customer) /s product

   Here WIN focused on production as an extension of product, and found
   numerous irrelevant cases with only two on point. Boolean retrieved
   twenty cases, but all were irrelevant.

6. WIN: Can adverse possession apply when the property is in a subdivision?

   Boolean: "adverse possession" /p subdivision

   WIN did include the most critical cases, but found too many cases
   dealing with a subdivision of the government and got sidetracked. The
   proximity of the concepts were important here because Boolean found the
   greatest number of relevant cases.

   After running twenty similar searches, Pritchard-Schoch realized that
   creating complex sentences for WIN often confused and sidetracked the
   search results or produced few relevant cases. In a complex question, WIN
   could not determine which words were important, and the query was
   thrown off. She believed that Boolean's ability to use the "but not"
   connector (not available on WIN) would have been helpful.

   C. Law Librarians' Queries

1. WIN: Does a real estate agent have the duty to disclose if a murder
   or suicide occurred in a home?\(^{43}\)

   Boolean: real estate /s agent broker /p disclos! /p murder suicide

   The search was done in the Ohio case database. Every case that WIN
   pulled up was a criminal case dealing with murder. By doing a Boolean
   search, the searcher managed to find the one relevant case in the jurisdic-
   tion: an unreported decision that has limited use within the state. Janto
   believes that, unless you are searching an incredibly narrow issue, a WIN
   search is no different than doing a long Boolean search using only the "or"
   connector. West does not point out in its literature that the cases may be
   considered statistically significant merely because at least one search term
   appeared a number of times: no context is used.

\(^{43}\) LAWLIB e-mail conference communication from Joyce Manna Janto, Deputy Director, University
   of Richmond Law Library (Apr. 1993).
2. WIN: Can an employee be fired for refusing to take a drug test?  

Most of the retrieved cases were about firefighters and drug tests. The trick is that WIN looks for the frequency of the word, and "fire" shows up a lot in those cases--firefighter, fire station, fire department. If an employee is fired, the court might say "fire" only once (or might say "discharge" or "terminate"). Boolean might give you some of the same false drops--but maybe not as many, since it is not counting how many times "fire" appears.

The searcher, Mary Wisner, suggests that WIN prompt the researcher to try the thesaurus. Her greatest concern with WIN is that its ease of use can mislead unsophisticated users, who may not understand how important it is to choose an appropriate database. Inexperienced users may not understand that WIN will always report twenty documents, whether or not all twenty are relevant, or whether there are more than twenty relevant documents. Because the researcher types in a regular sentence, it is tempting to anthropomorphize and think the computer "understands" the question. Using terms-and-connectors reminds the researcher that the computer simply looks for combinations of letters and numbers that he or she has thought of. With Wisner's search, a naïve user might think that firefighter drug cases were the leading cases on drug testing (and may wonder why the issue is litigated so much more by firefighters than by any other group of employees). With terms-and-connectors, learners accept that they need training. With WIN, they might not bother to listen to such suggestions as choosing search terms, trying alternative formulations, and being aware of database coverage.

D. Discussion with WIN's Creators

Turtle and Olson spoke with me about a number of concerns and problems that arose in my research. Following are the results of our question-and-answer sessions.

1. Question: Richard Leiter in his Legal Information article pointed out the problem with the system splitting "co-ed." I encountered the same problem with "in personam," with the system returning clearly irrelevant documents. What is being done to resolve this problem?

Answer: You are describing two different problems here. The first has to do with hyphenated terms, which are treated specially. Many systems,
including WESTLAW (terms-and-connectors and WIN) treat the hyphen as a potential word break character and treat "co" and "ed" as separate terms (terms-and-connectors also generates the run-together form "co-ed" as well as the original hyphenated form). This is generally the safest strategy since the use of hyphens is nonstandard (e.g., "long arm," "long-arm," and "longarm" generally mean the same thing), but it occasionally leads to the kinds of anomalies encountered with the "co-ed" query. The handling of hyphens is under review.

The "in personam" example you cite will be corrected for external users on May 18. WIN will recognize this as a phrase and will not retrieve cases that match only on "in."

2. Question: When I put in the query "Race as a factor in adoption or custody proceedings," the system included cases with horse and car races. The thesaurus was not helpful. What can be done to limit getting false hits with words of this nature?

Answer: The term "race" is ambiguous and, unfortunately, a thesaurus is not useful for ambiguity resolution. A thesaurus is useful when you want to broaden your search by expanding the forms that will match a term--essentially, you want to create a set of synonyms using the thesaurus as a guide. Since "race" is already too broad, you need ways to restrict its meaning (e.g., using a phrase), or you need another term that is less ambiguous but will appear in the same context (e.g., discrimination).

Ambiguity is not as large a problem as it might appear because natural language statements tend to have enough redundancy built in to establish a context in which the ambiguous term can be interpreted. For example, if "horses" or "cars" or "tracks" occur in the same query with "race," then cases with the appropriate sense will be found. If "equal opportunity," "employment," or "discrimination" are included in the query, then they will tend to disambiguate a different sense of "race."

In some cases, changing the word form can help disambiguate a term. For example, your Boolean query uses the term "racial," which is not nearly as ambiguous as "race." In general, when you see that an ambiguous term is retrieving extraneous cases, you should either find an alternative term or add terms that will give more context.

Follow-up Question: Why were car and horse races ranked higher than some custody and adoption cases?

Answer: I'd be interested in the database you ran this query in. I ran it in ALLSTATES, and all of the top twenty documents dealt with the appropriate sense of "race." In smaller databases, inappropriate rankings could arise if the frequency of "race" in a case was sufficiently high to
convince WIN that the case did deal with race, whereas the frequency of "custody" or "adoption" was not high enough to convince WIN that the case was about those topics.

3. Question: In all of my searches, there were many older cases and very few recent cases in the twenty cases retrieved. Boolean searching has often resulted in retrieving 1993 cases. How can this problem be resolved?
   Answer: The date restriction can be used to ensure that only new materials are retrieved.

4. Question: Having the cases ranked by relevancy can be confusing as to dates when the cases were decided, especially since we are used to looking at a reverse chronological display. What can I do to see them in order?
   Answer: The "sort" function is being considered for addition to WESTLAW and would be usable for both WIN and terms-and-connectors searching, but it is not available at present. Date can be a very useful presentation order, especially when there are ties (in terms-and-connectors searching, all retrieved cases are presumed to be equally relevant and ties are broken by date). It is very difficult, however, to decide when date is more important than probability of relevance: when should WIN decide to present a newer case ahead of one that it believes is more likely to be relevant and how should we explain all this to users? In the end, we decided that probability of relevance was the most reasonable presentation order. Note that in either WIN or terms-and-connectors searching, relevant cases should be checked using Insta-Cite and Shepard's to ensure that you've got a good case: the fact that a case is recent does not guarantee that it is good law.

5. Question: I was very concerned in a number of results that the landmark cases did not appear in the list of twenty. I only knew this because I was familiar with the area of law. I think it is important to consider finding a way to include landmark cases in the list of twenty. They are extremely relevant and yet there is no guarantee that they will appear.
   Answer: WIN may not rank landmark cases highly if the language used in those cases does not match that used in the query. There may be subtle reasons for this. If, for example, a landmark case gave rise to a legal concept but did not describe that concept in the same terms as it was subsequently referred to, that case probably won't be found.
   WIN has no notion of a landmark case built in, although this concept was discussed during development. In many ways the landmark problem
is the dual of the traditional retrieval problem: in traditional retrieval we have a query in search of documents, but we might also want certain documents to search for queries. The WIN model allows this alternate search mechanism, but it requires identification of those cases that are to be considered to be "landmark" and a specification of the kind of queries that each favors. This is a very labor-intensive activity, and the feature was not included in the initial implementation.

6. Question: I ran a number of "trick" queries in databases that I knew would find no relevant documents. I was surprised that several times twenty irrelevant documents were displayed anyway. Is the system forced to give you twenty cases, even if there are no relevant cases in that database?

Answer: WIN is designed to retrieve the cases that are most likely to match your query, but there is no guarantee that they are, in fact, relevant. This is particularly true when there are no relevant cases in the collection being searched. In general, if WIN can find any evidence that a document may be related to the query, it will rank that document; even documents with a single term in common with the query will be ranked. These poorly matched documents are not a problem when the collection contains a reasonable number of relevant documents, since they are too far down in the ranking to be seen by the user. The problem here is that WIN generally produces a reasonable ranking, but at present, provides no indication of how good the ranked documents are. Techniques for indicating "quality" of the cases retrieved are being tested in the internal prototypes and will be added to WIN when we are satisfied with their reliability.

7. Question: From all the examples I've gathered, it appears that WIN can only effectively do simple queries. Is that true?

Answer: WIN was developed using a wide variety of test queries, including test sets that are substantially more complex than those used in your examples or than used by a "typical" user. In general, WIN performance on complex queries is very good.

8. Question: How do you explain the problem encountered by Pritchard-Schoch with the query, "Can adverse possession apply when the property is in a subdivision?"

Answer: The problem arises because "subdivision" is ambiguous, and WIN is unable to distinguish between the various senses present in documents. This is essentially the same problem as the "race" query, and the same strategies for improving the query can be used--either make the term more specific (e.g., a phrase) or add terms that will help establish the appropriate context (e.g., using "real property" or "lots").
This query is an interesting example of when a bit too much knowledge of the legal domain can hurt. WIN knows that subdivisions are important parts of laws and that they may be abbreviated to "subd." Many of the extraneous cases found in response to this query are retrieved because "adverse possession" appears near the reference to its definition, which cites a subdivision. If WIN didn't know about legal abbreviations for "subdivisions," it would have retrieved a more plausible set of documents.

IV. Additional Queries Run After Discussion with WIN Creators

A. Queries

1. WIN: Does a defendant have a second opportunity to respond on the merits after making a motion to dismiss in an action brought as a motion for summary judgment in lieu of complaint pursuant to NY Civil Rule of Procedure 3213? (abbreviated in NY as: CPLR Section 3213).45

   Boolean: "Summary judgment" /p "in lieu of complaint" /p CPLR 3213

Here WIN was a lifesaver and on point. The first case was the relevant case. A great number of different Boolean strategies were tried without results. The correct query came from having the results of the WIN search. Even then, in the Boolean search the case was document 267 of 286. The case was retrieved without looking at every document because the date was known from the WIN search. Few, if any, users would have plowed through that many documents before finding the correct one.

2. WIN: Lesbian

This search was run to see if, in fact, the database would produce fewer than twenty documents, as had been indicated by WIN's creators. Assuming that there were probably only a small number of cases decided on the topic, the query "lesbian" was put into the Washington State database. Only five cases were found, yet when an additional term (lesbian parenting, lesbian couple, lesbian movies) was added, twenty cases came up. In each instance, cases six through twenty were irrelevant, as if the system went strictly by the second term. Even though the five cases filled the top five slots, the ranking changed for no apparent reason. In each

45. This query was provided by the University of Washington Law School WESTLAW Representative, Deanna Loy.
search with an additional term, I assumed that the document in only the first slot would change because it would be the only one on point. This is not what happened. The other four slots seemed to be randomly filled. The examples of this randomness are demonstrated in the following charts.

A. Lesbian-parent, custody.
B. Lesbian-freedom socialist party.
C. Lesbian-freedom socialist party.
D. Lesbian-marriage license.
E. Lesbian-Jr distributor movies pornographic materials.*

**Query 1: Lesbian**

<table>
<thead>
<tr>
<th>Slot/Rank</th>
<th>Query 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
</tr>
<tr>
<td>4</td>
<td>D</td>
</tr>
<tr>
<td>5</td>
<td>E</td>
</tr>
</tbody>
</table>

**Query 2: Lesbian Parenting**

<table>
<thead>
<tr>
<th>Slot/Rank</th>
<th>Query 1</th>
<th>Query 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>D</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>E</td>
</tr>
<tr>
<td>4</td>
<td>D</td>
<td>B</td>
</tr>
<tr>
<td>5</td>
<td>E</td>
<td>C</td>
</tr>
</tbody>
</table>

* The letters in the query columns correspond to the document assigned. The slot/rank columns indicate the ranking order of each query run.
There seemed to be no clear pattern to the changes of queries two through four. The randomness of documents ranked second to fourth was unexpected. The big surprise came with the "lesbian movies" query. The only case with the words "lesbian and movies" (therefore the most relevant) was ranked second. The case in the first rank, which had to do with lesbian and parental custody, did not include the second term, "movies," at all. How the weighting determined relevancy is not clear.

3. WIN: "e-mail"
   Boolean: "e-mail"
This query was run in the ALLSTATES and ALLFEDS databases. In WIN, seven cases were retrieved in the ALLSTATES database, none of which was relevant, because "e-mail" was split into two words. The cases retrieved had such terms as "i.e. mail," "(e) mail," and "Louisville & E. Mail Co." The Boolean search produced some "e-mail" cases, but also included the problem encountered by WIN. In the ALLFED database, WIN was able to produce relevant cases dealing with "e-mail," but the problem of splitting the term was also present.

**B. Additional Discussion with WIN's Creators**

Having run the additional queries to follow up my discussion with Turtle and Olson, I had the following new questions.

9. **Question:** After our first conversation, I ran a few additional queries to explore new insights you provided. For example, I was trying to see if I could put in a query that would produce fewer than twenty documents. Assuming that there were probably only a small number of cases on the topic, I put the word "lesbian" into the Washington State database. Only five cases were found. This proved your point. However, how do you explain the change of ranking of the same five cases when I put in an additional term where clearly the other four cases did not address the question equally?

   **Answer:** When you run two-term queries like these, WIN assumes that both terms are potentially important, and the ranking is based on both terms. In general, adding terms to a query will change the ranking because the top-ranked documents will be those that deal with both concepts. The fact that the same five cases remained at the top of the ranking simply indicates that WIN judged the term "lesbian" to be much more important than the added terms. Had the added terms been nearly as rare as "lesbian," then the original five documents would have moved down in the ranking and new documents would have moved up.

10. **Question:** When I ran the query "lesbian movies," why was the case that contained the words "lesbian" and "movies" slated second to a case that did not contain both terms?

   **Answer:** WIN bases its ranking on a number of factors, but the two main factors at work with a simple query like this are (1) the discrimination power or importance of the individual terms, and (2) the degree to which WIN believes that these terms are correct descriptors of individual documents. In this case, the probability that document 1 dealt with lesbian issues was quite high, and there was no evidence linking document 1 with
"movies." The probability that document 2 dealt with lesbian issues was judged to be lower than document 1, and the probability that document 2 dealt with movies was judged to be moderately high. The evidence supporting "movies" was not judged to be strong enough to offset the uncertainty with regard to "lesbian," and document 1 was ranked higher.46

11. Question: I attempted to do the same thing with "e-mail." In the Washington database, it said no documents. However, in the ALLSTATES database and the ALLFEDS database, I encountered a number of problems. In the ALLSTATES database only seven cases were given, none of which were on point. All seven were of this variety: "i.e. mail, e)mail," and so on. At the federal level, I got "e-mail" cases, but intermixed with the same responses as in the ALLSTATES. What could have been done so that the ALLSTATES would have returned no cases?

Answer: This is another example of a hyphenated problem. The real solution for this query is to add the term "e-mail" to WIN's knowledge base, with specific instructions to match only the hyphenated form. While it is difficult to know exactly what you are after, the query "electronic mail" retrieves cases that appear to be relevant.

As a general comment, WIN does not eliminate the need for a carefully worded and complete issue statement or description of the cases desired. The main advantage of WIN is that it allows searchers to focus on the issue statement itself rather than on the mechanical aspects of Boolean

46. In a follow-up conversation, Turtle gave further explanation on how WIN dealt with the issue:

Given a two-term query like "lesbian movies," the document ranking will be based on the relative importance of each term and on the level of certainty that each term correctly describes a particular document. Since "lesbian" is a relatively rare term in the database in question, it will be judged to be a more important term than "movies." Ignoring, for a moment, differences in the level of certainty that each term describes a particular document, documents containing both terms will be ranked highest, documents that contain only "lesbian" will be ranked next, followed by documents that contain only "movies."

When we factor in information about the level of certainty that each term correctly describes a document, the ranking gets a bit more complicated. Now, for example, we must decide how to rank the following two documents relative to each other. For document A, we are reasonably certain that "lesbian" is a good descriptor of document content are very certain that "movies" is a good descriptor. For Document B, we are very certain that "lesbian" is a good descriptor and very uncertain about the term "movie(s)." Which document is more likely to match the query?

Whether we will judge a document that is "probably about both" concepts to be superior to a document that is "certainly about one" concept will depend on our precise judgments about certainty that a concept describes a document and about relative concept importance. These problems, quantifying concept importance, quantifying the certainty that a concept describes a document, and combining this information to produce a ranking that matches user judgments, are exactly the problems that the WIN algorithms address.

E-mail from Turtle and Olson (Sept. 15, 1993). In our conversation on September 10, 1993, Turtle and Olson pointed out that in fact the word "movies" had been found in the retrieved document for the simple query of "lesbian." This was another factor in why "lesbian" had been ranked higher than "lesbian movies." Interview with Turtle and Olson, supra note 18.
query formulation. The searcher must still be concerned with ambiguous terms and with problems associated with concepts that can be described in several ways. These are exactly the same kinds of problems that humans must deal with in describing their information needs to each other or to an expert search intermediary, and it is unlikely that any natural language system will do away with the need for a carefully developed issue statement.

When compared to conventional technology, WIN's understanding of natural language is a significant advance: WIN has moved the very best natural language retrieval technology available out of the laboratories into commercial practice. When compared to a human's ability to understand natural language, however, all natural language systems are primitive. Natural language understanding systems can be expected to improve over time, but we are a long way from approaching human levels, and even humans misinterpret natural language statements with remarkable facility.

V. Critique

WIN is a revolutionary and exciting new addition to the online searching of full-text databases. This program will be the stepping stone to future advancement of natural language searching in legal databases.

WIN's major problems seem to be not in program or design, but in marketing. Marketing any product is a tricky endeavor. A great deal of how a product will be accepted is based on consumer perception. The manner in which WESTLAW chose to introduce its new product has caused a great deal of skepticism among users.

Although WIN is revolutionary, its introduction should have been tempered by including the realistic limitations of the system. In its early promotional materials, WESTLAW set the level of expectations too high. The following remarks are taken from these materials:

Finding the precise case law you need has never been faster, easier, or more natural. . . . Simply enter your issue in plain English and WESTLAW does the rest with results that are remarkably thorough and dependable. . . . With WIN, there's no computer language to learn. No need to enter root expanders, no proximity connectors, or other items required by other computer research services. Until now this is the only language you could use to compose a case law search request on-line. It's called Boolean. This is the language you now have the choice of using for the same research on WESTLAW. It's called English. . . . WIN is the perfect choice for the new or occasional researcher. 47

47. West Publishing Company, Westlaw Takes the Wraps Off Another Breakthrough, Promotional
This search strategy frees researchers from learning computer language and allows them to enter their request in plain English. The power of the computer translates the search and retrieves relevant documents. 48

Such materials do not properly educate the naïve consumer and, perhaps, underestimate the level of sophistication among regular users, who have come to expect a certain level of quality. WIN promised, but then failed to "take care of everything." If regular users had been properly warned about the limitations of the system, they would have been more prepared for the shortcomings and more willing to understand the problems that have developed. I have found that there is a high level of skepticism among users in my sample. Some have chosen not to use WIN at all because they do not trust the results.

The following excerpts are representative of the comments that I received:

I guess my concern with WIN is that it gives the illusion of simplicity. For example, if one enters a query and fails to take synonyms into account, the search is often not as effective as it could be. . . . I'm somewhat amazed that West Reps would state that WIN and Boolean are equivalent approaches. Our West Rep has described WIN as another research alternative. I believe this is a more accurate description. 49

At this point in its development, WIN is a quick and dirty and not very powerful way of finding a handful of cases. It is not yet a reliable tool for serious legal research. I have run numerous WIN searches on my standard teaching examples, and I have found that WIN misses important cases. Because it is deceptively simple, and because West is presenting it as a substitute for Boolean searching, there is a danger that the next generation of law students will learn only WIN. This is a good example of why academic law libraries should not allow the database vendors to take control of teaching of the use of their products. 50

I agree there is no way [WIN] can substitute for Boolean searching at the present. Someday it will be sophisticated enough that it will get everything, I'm sure, but it is nowhere near that point now. Students who learn to rely on it, who believe it is unnecessary to learn "terms-and-connectors" searching are going to have serious problems down the line. 51

49. LAWLIB e-mail conference communication from Dan J. Frechling, Director, Pappas Law Library (Apr. 1993) [hereinafter Communication from Frechling].
50. LAWLIB e-mail conference communication from David Lowe, Computer Services Librarian, University of Alabama Law Library (Apr. 1993).
51. LAWLIB e-mail conference communication from James Quinn, Gonzaga University Law Library
Some comments pointed out the positive aspects of WIN:

It's not perfect, but it will get better, and it adds an important tool that can be useful. Every 'search' need not be exhaustive, particularly in many practice situations. The issue is not, it seems to me, that WIN may or may not get the same results as Boolean searching but rather that, for some applications, it may be a cost-effective and time-effective method of doing a search.  

The greatest harm of WIN's marketing strategy will be on novice law students. The early promotion clearly identifies WIN as the only choice for this target population to use.  

Given the results of the queries presented and the concerns raised by law librarians across the country, students will be under the false impression that all "relevant" and "important" cases have been found by WIN. The marketing strategy's impact on new users will be a disservice to the legal community and clients. This is pointed out in a number of comments from the LAWLIB Bulletin Board:

I am . . . sympathetic to the issue of students thinking that WIN is the 'only' or 'best' way of searching— that could indeed be a problem, much as students sometimes think that the full-text law review databases on LEXIS/WESTLAW have all the journals.  

I agree that we must be very careful in the way we present WIN to law students. Personally, I incorporated the teaching of WIN into the regular introductory WESTLAW class (about 10 minutes at most). I think that it is probably important to discuss Boolean searching before touching on WIN. Otherwise, students may ignore the Boolean searching as being irrelevant.

West seems to feel that WIN should be the method of choice for most attorneys and law students. I believe that it must be emphasized to students that they need to think about the pros and cons of WIN vs. Boolean when using WESTLAW. I tell students that ideally they can try both approaches and compare the results.  

Some respondents, however, highly recommended WIN as a first choice for students:

---

52. LAWLIB e-mail conference communication from Gary J. Bravy, Georgetown University Law Library (Apr. 1993) [hereinafter Communication from Bravy].

53. Turtle advocates using "both WIN and Term and Connectors (Boolean) search methods for complete research." E-mail from Turtle and Olson, supra note 7. In an interview with this author, Turtle and Olson later pointed out two additional matters that students and law librarians should keep in mind. One, WIN does not do a "legal" analysis, it does a "statistical" analysis, which is an important distinction to remember. Two, a WIN search is not "legal research." It is only part of the process, not the end point. Interview with Turtle and Olson, supra note 18.

54. Communication from Bravy, supra note 52.

55. LAWLIB e-mail conference communication from Scott L. Schaffer, Villanova Law School Library (Apr. 1993).
WIN v. Boolean Searching

I have been using WIN regularly since it became available, and I recommend it to students who are intimidated by Boolean terms-and-connectors. "Term and connectors" can be a daunting new language to learn, and WIN is much simpler to understand. Initially, there were problems with some missed cases and imprecise language in queries, but as students became more familiar with the process, these problems lessened. WESTLAW has also made other refinements to the system that make WIN more effective. Hyperlinking key numbers and headnotes together means that you can jump from one case to another using the same key number. Even if your WIN search only turns up one relevant case, you can use the key number and headnotes to find other cases. In many respects, WIN and hypertext moves us closer to the day when online systems will replace books.\(^56\)

Dan J. Freehling, Director of the Pappas Law Library at Boston University, had this to say about headnotes and WIN:

The fact that WIN may have an easy connection to headnotes is helpful and important, but does not, in my mind, cause it to replace Boolean completely. For one thing, the digest (and headnotes) is hardly a perfect tool, even electronically. Second, to ignore Boolean connectors one would have to stop using field searching, and this is a major step backwards if one is interested in efficient searching.\(^57\)

At the other end of the market, legal professionals who have been trained to provide the most current cases will be in for a shock. Imagine their reluctance to trust a system that displays no cases from the last several years in the cases considered most relevant! Similarly, since WIN does not necessarily display landmark cases as part of its twenty most relevant, sophisticated users are bound to be skeptical.

VI. Conclusion

Although WIN will play an ever-increasing role in the future of legal research, the queries show that there are a number of serious problems in the system that must be addressed if the software is to live up to its potential. As demonstrated in the complex query provided by WESTLAW, WIN can be the easier and more efficient manner of retrieving a case. As indicated by other examples, however, it does not always provide the most relevant cases in complex queries. For the best results, users should combine the WIN search with a Boolean search. The two systems working together offer the user a better chance of retrieving the appropriate

\(^56\). LAWLIB e-mail conference communication from Elmer Masters, Reference Librarian, Barclay Law Library, Syracuse University (Apr. 1993).
\(^57\). Communication from Freehling, supra note 49.
documents. One of WIN's most effective tools is the thesaurus, which allows the novice access to terms not so readily evident.

As the system is used and WESTLAW receives appropriate feedback, changes will be made regularly to provide the user with a better product. For 120 years, West has delivered an excellent product, growing and changing with each new breakthrough. I expect the tradition will continue with WIN.