



3-1-1993

Ten Loose Guidelines for Online Searchers

Carol Tenopir
University of Tennessee - Knoxville

Follow this and additional works at: https://trace.tennessee.edu/utk_infosciepubs



Part of the [Library and Information Science Commons](#)

Recommended Citation

Tenopir, Carol, "Ten Loose Guidelines for Online Searchers" (1993). *School of Information Sciences -- Faculty Publications and Other Works*.
https://trace.tennessee.edu/utk_infosciepubs/93

This Article is brought to you for free and open access by the School of Information Sciences at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in School of Information Sciences -- Faculty Publications and Other Works by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

Disclaimer: This is a machine generated PDF of selected content from our databases. This functionality is provided solely for your convenience and is in no way intended to replace original scanned PDF. Neither Cengage Learning nor its licensors make any representations or warranties with respect to the machine generated PDF. The PDF is automatically generated "AS IS" and "AS AVAILABLE" and are not retained in our systems. CENGAGE LEARNING AND ITS LICENSORS SPECIFICALLY DISCLAIM ANY AND ALL EXPRESS OR IMPLIED WARRANTIES, INCLUDING WITHOUT LIMITATION, ANY WARRANTIES FOR AVAILABILITY, ACCURACY, TIMELINESS, COMPLETENESS, NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Your use of the machine generated PDF is subject to all use restrictions contained in The Cengage Learning Subscription and License Agreement and/or the Gale Academic OneFile Terms and Conditions and by using the machine generated PDF functionality you agree to forgo any and all claims against Cengage Learning or its licensors for your use of the machine generated PDF functionality and any output derived therefrom.

Ten loose guidelines for online searchers

Author: Carol Tenopir

Date: Mar. 1, 1993

From: Online(Vol. 17, Issue 2.)

Publisher: Information Today, Inc.

Document Type: Article

Length: 4,104 words

Abstract:

Online database searchers can benefit from ten basic rules. They should pre-plan, but not be excessive about it; have a purpose at the outset; select the best database; use cost-effective search strategies; bear in mind that corporate and personal file entries are inconsistent; understand the differences between databases; tailor search strategies to suit databases; understand the differences between online systems; employ post-search routines; and develop a search style.

Full Text:

THE FIRST WORD

The original title of this article was to be "The Ten Commandments of Online Searching." The more I thought about it, the more uncomfortable I became. Online searching is not so structured nor so exact that any one searcher can command another how to best search. One searcher's style may not be another's, but each may be as effective as the other. Certainly there are general hints to remember that will improve a search, but these are guidelines, not commandments. So, not wanting to impose an artificial sense of correctness (and not wanting to offend any religion), I offer my "Ten Loose Guidelines for Online Searchers."

These guidelines are not new--they have developed over the past 20 years as databases and online systems developed, and as searchers learned the best ways to interact with them. Most are things that experienced searchers do almost automatically. Many are a part of an "online searcher's mind" as described by supersearcher Barbara Quint in a two-part article in ONLINE in 1991 [1]. This discussion is aimed at the person who is relatively new to online searching, especially someone who is new to searching as an intermediary. Experienced searchers may want to use the guidelines as a "reality check" against their established searching style.

Guideline #1: Pre-search planning is essential, but overplanning is excessive.

We've all been told to plan an online search before logging on and I echo that advice. Take time to make sure you understand the question and to select likely databases. Check printed thesauri, database manuals, online system documentation (such as DIALOG Bluesheets or BRS AidPages), and other appropriate search aids. Specialized or general dictionaries and a general "Roget's-type" thesaurus will help you to identify alternative spellings, synonyms, antonyms, and potential false drops from truncation. The rule used to be that intermediaries should spend one hour offline for each half hour online; few of us have the luxury of so much time, but plan to spend at least an equal amount of time pre-planning and online.

If you are new to searching or new to a system, you should write out the initial commands and your "best guess" initial search strategy. This preliminary strategy can be recorded into your telecommunications software and automatically uploaded as part of the logon process. It is at this point that pre-search planning can change from essential to excessive if you never go beyond this initial uploaded strategy. Even in a \$300 per hour database, online searching should be an interactive process. Don't fall into the "fast batch" trap--the trap of losing all interactivity by uploading a prepared strategy and printing out results without viewing interim results, and without correcting or tinkering with the strategy. Batch searching (usually with 24-hour delays) was hot news in 1965, but today's searchers shouldn't cripple the interactive power of online by too much pre-planning.

Online tinkering lets you restrict the size of a set if you retrieve too many records (restricting terms to the title, descriptor, or identifier fields is one good way). It also helps you enlarge the set if you retrieve too few (try adding synonyms or eliminating one concept); or allows you to eliminate problem words or concepts that may be causing false drops. If your brain freezes when the online meter is running, run your search, print out titles of at least 10 to 20 items, logoff in a way that saves your strategy, and think offline. Commands like DIALOG's DISPLAY SETS or BRS and Data-Star's DISPLAY will help you keep track of where you are at any time.

This first guideline could really be called "keep your brain online at all times!"

Guideline #2: Know the purpose and motivation for the search.

Whether you are searching as an intermediary or helping someone plan their own search, you must understand why the requester needs the information and how it will be used. Only then can you choose appropriate databases or develop a search strategy.

If, for example, a scientist needs a search to lay the groundwork for a new research project or a doctoral student is trying to find out if anyone has written previously on his or her dissertation topic, a comprehensive or exhaustive (high recall) search is in order. A high school student with the assignment to find ten articles on a general topic or a manager trying to prepare a half hour speech may need just a superficial or high precision search. A public library patron asking about information on migraine headaches probably would benefit from a search of Magazine Index, Readers' Guide, or Health Periodicals Index; the same topic for a health professional should be searched in MEDLINE or EMBASE.

Twenty-five years ago Robert Taylor described the reference interview process as the negotiation between a library patron and a reference librarian [2]. Although he was speaking about a manual searching process, much of what he described holds true for the reference interview needed in intermediated online searching. Taylor spoke of "five filters through which a question passes and from which the librarian selects significant data to aid in his search." The five filters that allow an intermediary to develop online search strategy are:

1. Determination of subject
2. Objective and motivation
3. Personal characteristics of inquirer
4. Relationship of inquiry description to file organization
5. Anticipated acceptable answers

Asking questions that use these filters helps the online intermediary extract all the information that is needed to begin to formulate strategy.

Guideline #3: Choose the best database or databases for each search.

Twenty years ago online searchers could be intimately familiar with all the databases on all the major online systems. Each year that becomes more difficult, and is impossible today with over 5,000 databases and nearly 700 online services.

So many choices mean the chance of finding useful databases is great, but it also means that online searchers (and online systems) must spend more time evaluating and selecting databases. There are both printed and online aids to help the database selection process.

Printed database directories are somewhat overwhelming--facing a fat volume of thousands of databases, many on systems you will never search, can make the database selection task look daunting. Still, these directories are indispensable to advise you about new databases, tell you what other systems a familiar database may be on and what it costs, and give a general idea of how much coverage there is online for a subject. Every online searcher should have access to a database directory such as the new Gale Directory of Databases [3]. (Database directories are also available online.)

Don't use database directories just to find databases on your favorite online system (use the online service catalog for that). By broadening your online system horizons, you may discover additional (or different) systems that will better meet your information needs. Many novice searchers rely too much on a single system. Learn commands of several and use the system that offers the most appropriate databases at the best price for any particular search.

A database directory is only the first step in the database identification process. It provides broad subject categories, but doesn't tell you which databases actually have information on your specific topic. Online database selection aids such as DIALOG's DIALINDEX, BRS and Data-Star's CROS, or ORBIT's Database Selector, fill that need.

Many experienced searchers tell me they use DIALINDEX more than any other DIALOG database. It tells them which of the more than 400 DIALOG databases have postings on their search terms. Ranking DIALINDEX results puts the databases with the most postings at the top of the list. Strategies can be saved in DIALINDEX, so the next step is to begin a OneSearch in the top few databases and execute the strategy.

What a database selection aid does not tell you is how good the information is or how appropriate it is for the user. Paying close attention to Guideline #2 helps with the latter--experience and studying database reviews are the only surefire method to help with the former. Reviews in journals such as DATABASE, ONLINE, RQ, and Online Review help identify the strengths and weaknesses of specific databases.

Guideline #4: Use cost-effective search techniques.

Unfortunately, high costs are a fact of online searching life. Most systems still charge by the amount of time you spend online plus the number of records you print. Other systems, such as LEXIS/NEXIS, charge less for connect time, and charge a flat fee for each

search on each file.

Either way, good online searching depends on formulating a strategy that gives the best results for the least amount of money. Good, cost-effective searchers pay close attention to cost-cutting techniques when they formulate strategies.

Cost-cutting tips have been covered in recent issues of Barbara Quint's Wilson Library Bulletin column and in my Library Journal column [4,5]. Both recommend switching to 9600 bps searching (or the fastest speed available to you), and increased reliance on uploading strategies and downloading results. Since so many databases are on more than one online system, both advise shopping around for the best price. Quint even advises "sometimes you can save money by using one system to search and another to print data from the same source." The same goes for telecommunications costs--access over the Internet is now possible for many services including DIALOG, OCLC EPIC, LEXIS/NEXIS, BRS, and ORBIT, often at lower costs than SprintNet or TYMNET. If you don't have access to the Internet or have to pay a surcharge, DIALNET is cheaper for DIALOG access than other networks.

Other cost-cutting tips include: use multifile searching whenever possible; use print documentation rather than online versions; take advantage of free files, free documentation, alternative databases on the Internet and any other bargains you can find; and use LOGOFF HOLD or similar commands when you want to pause to think instead of PAUSE or PARK commands that still charge up to \$1 per minute.

Guideline #5: Use special care when searching for corporate or personal names.

Libraries traditionally have spent time and energy on maintaining authority files to ensure consistency in their catalog entries. It seems that most database producers have never heard of authority files! Personal author names, corporate sources, and company names all suffer from inconsistent entry practices. Even Library and Information Science Abstracts (LISA) has at least a dozen different forms of entry for the prolific author, Ching-chih Chen. Truncating doesn't even retrieve all variations since some are entered Chen, Ching-chih and others are entered Ching-chih Chen.

A safe solution for the online searcher is to view and select all possible forms of entry when searching for any kind of name. On DIALOG this means the EXPAND command, on Data-Star and BRS it means ROOT, and on ORBIT and NLM it is NEIGHBOR. (On LEXIS/NEXIS it can't be done at all.) Choosing terms from an index display reduces the amount of typing and ensures that you find most variations of an author's name.

To be fair, some databases do go to a lot of trouble maintaining authority files. The H.W. Wilson Company databases in particular, verify all names using a system-wide authority feature. UMI/Data Courier's Newspaper and Periodical Index also spends much time and effort on authority work. But consistency is the exception, not the rule.

Guideline #6: Know and understand the differences between databases.

It goes without saying that the articles you will locate in a search on MAGAZINE INDEX will differ in level and treatment from the articles you locate on MEDLINE. Bibliographic databases differ in their structure and record content as well.

Even with databases on the same online system, you cannot assume that field labels are consistent across databases. Most online systems are trying to build in some consistency, such as always having an author (AU) field and a title (T1) field. Recently, previously inconsistent fields on DIALOG (such as publication year) have been standardized as PY.

Other fields may vary widely from database to database. Source, corporate source, journal name, document type, etc. may or may not be present. If present, they may have different field tags in different databases. In some databases, a field such as corporate source is word indexed only (meaning you must search it with proximity operators); in some it is phrase indexed only (meaning you must view the index display to find the exact form of entry); and in some it is both word and phrase indexed.

Always check system database documentation such as DIALOG Bluesheets or BRS AidPages before searching an unfamiliar database. Checking the printed sheets before you go online and keeping them next to you as you search is less costly than fumbling online or than relying on the online versions of the documentation.

Databases also differ widely in the terms used as descriptors or identifiers. Descriptors usually are assigned from a subject heading list or thesaurus, typically created and maintained individually by each database producer. If you want to search with descriptors (and research shows that expert searchers do so more than 50% of the time [6], you must purchase thesauri from the database producers or display the terms online.

All these differences among databases make multifile searching tricky. When using Data-Star's StarSearch or DIALOG's OneSearch, database differences are seemingly masked, but the differences still exist. Don't forget to check every database sheet, formulate generic strategies, avoid database-specific field tags. Remember to OR together descriptor terms from each database or use free-text search strategies.

Guideline #7: Plan your search strategy to match the type of database.

Major online systems today contain a variety of database types, including bibliographic, full-text, directories, numeric, and hybrid (combination) databases. The commands to access the different types are the same, but good search strategy often varies with the type of database.

For example, in a bibliographic database search it makes sense to use the Boolean AND to link concepts, such as health care AND costs. In a full-text file that strategy would probably retrieve an unacceptable number of false drops. The costs of a phone call could be mentioned in passing in an article about poor health care for the elderly. In full-text searching, substitute a proximity operator for the AND between concepts to force a closer relationship. On LEXIS/NEXIS, for example, you might search health care w/10 cost; on BRS health adj care WITH cost (requests the terms appear in the same sentence).

Numeric and directory database searching almost always require precise searching--specifying fields, numeric ranging, sorting results, and perhaps even printing out in a REPORT format. DIALOG's numeric and directory databases may have dozens, or even hundreds, of special fields such as address, phone number, gross profits, square footage, etc., each of which must be individually specified if it is to be used in a search.

Beware the OneSearch trap if you use DIALOG's preset database categories instead of specifying your own selected database combinations. Sometimes these categories mix types of databases. This not only fouls up your best search strategy, it may result in weird results. When a full-text database is searched in a category that contains mostly bibliographic databases you may wonder why most of your hits come from one, relatively small, database. Or, a full-text file, with more words to search, may contribute a disproportionate number of hits (mostly false drops.) Build your own categories or use NOT in the BEGIN statement to exclude unmatched types (i.e., begin infosci not 170).

Today you may encounter mixed record types even within a single database. Hybrid databases like the newly renovated ABI/INFORM or PTS PROMT are a mix of bibliographic and full-text records. To get best results you can restrict search terms to the title or descriptor fields (excluding full text) or always use full-text strategies, such as replacing ANDs with proximity operators.

Guideline #8: Know and understand the differences between online systems, and between the same database on different systems.

Although the major online systems developed from the same traditions and are fundamentally the same, there are some subtle differences that will affect your search results. All major online systems create dictionary (a.k.a. inverted) indexes from complete records to increase searching speed. It is actually the indexes you are searching when you enter a term on any system, not the records themselves. The rules of index creation may differ subtly among systems.

DIALOG creates one "Basic Index" that includes words and phrases only from subject-related fields such as titles, abstracts, descriptors, and text. All other fields, such as authors, dates, journal name, etc. are put into separate "additional indexes" that are only searched when specified by a two-letter code followed by an <=> sign. (e.g., AU=). LEXIS/NEXIS, on the other hand, puts words from every field into a single master index. When you search carpenter on DIALOG you will not retrieve authors with that name, but you will on LEXIS/NEXIS.

In most systems every word is searchable, but a "word" is defined differently. DIALOG says a word is any character string surrounded by blanks or by punctuation. That means that CHILDREN'S in a title on DIALOG is two words: CHILDREN and S. BRS does text editing for specified punctuation marks, so CHILDREN'S is just one word (CHILDRENS).

Stopwords also vary system by system. Designated stopwords are not searchable--they are not put into the dictionary indexes. DIALOG takes a conservative approach, designating only nine stopwords. These nine words (an, and, by, for, from, of, the, to, and with) are consistently applied across all databases. BRS fills a page with its stopword list, LEXIS/NEXIS "noise" words vary from file-to-file.

Most systems use Boolean operators, but the default order of execution varies. On DIALOG, BRS, ORBIT and most others AND is processed before OR. That means parentheses are required to give you the desired results of (cholesterol or fat) and (heart or cardiac). On LEXIS/NEXIS, ORs are processed first, so no parentheses are required on the same search.

You have to have a good memory from beginning online training classes or be a hard-core reader of system manuals to find this information because it often is not included in those handy pocket sized "cheat-sheets" we end up using when searching. (You can also refer to the table in the REFRESHER COURSE on Boolean operators by Doug Morton that appeared in the January ONLINE. --NG)

Other system differences that affect searching are more obvious and widely documented. These include which proximity operators are available for use, update schedules for each database, and loading of special features such as online thesauri.

Guideline #9: When you logoff, your search is not over--post-search routines complete a good search.

In the old days (the 1970s) online searchers used dumb terminals that produced just a paper copy of the search instead of microcomputer work stations. Many searchers got away with just handing the raw, unedited thermal paper scroll to the paying customer, who, not knowing any better, was usually satisfied. In this day of high-capacity disk drives, high-speed modems, laser printers, and good software for word processing or database building, every online search can be an attractive, edited product.

Post-processing should be as much a part of the search as the online search itself. Downloaded search results can be transferred to a word processor for cleaning up and adding annotations. Transferred to a text retrieval software package, citations can be reformatted and made into a customized database (within the restrictions of the copyright law and your system contract.) Citations can also be reformatted according to a variety of citation style standards, and imported into a special multipurpose program such as Pro-Cite. Numeric results can be transferred to a spreadsheet program for manipulation, enhancement, and report generation.

Evaluation of search results is another part of post-logoff strategy. Intermediary searchers especially should verify that each client got what he or she expected, that the results were presented in a way that was most useful, and that they were satisfied with the total product. No good business person would dream of neglecting this step, but a surprising number of searchers in libraries don't take the time.

One word of warning: when you ask for feedback, be prepared to take action if it is negative. That means having a budget set aside for redoing unsatisfactory searches at no extra cost to the customer, honing system skills if problems resulted from your errors, or referring the client to another search service that may better meet their needs.

Guideline #10: Develop a style, but know when to try something different.

In this art of online searching, everyone has his or her own style. There are no right or wrong ways to do a search--just different ones that work best for individuals. No one, not experienced searchers, online educators, nor expert system designers, has yet figured out an optimal way to search. Find a style that is best for you and develop it.

Individual searching styles have been studied and described for years. In ONLINE in 1982, Hawkins and Wagers summarized styles from the strategy point of view, describing four common approaches to search strategies [7]. Whether you prefer "building blocks," "successive fractions," "citation pearl growing," or "most specific/least posted facet first" in part depends on your own stylistic approach. (Three of these styles were also discussed by Bob Buntrock in ONLINE in 1979 [8].--NG)

It would take an entire article to cite and describe all the studies that look at style from the human point of view. Work over the last decade by Fidel, Bates, Saracevic, and many others [9,10,11] show that how many steps we take online, how many changes we make in strategy, how many terms we add (or delete), and what overall approach we prefer is a very individualistic thing. For the most part, individual search styles don't seem to affect our overall success as searchers.

The most important thing about searching style is to realize the variety of possibilities and to be flexible, changing techniques as required in different search situations. When you have no luck with your highly structured and pre-planned "building blocks" style, perhaps finding one good citation online with a "quick and dirty" approach will let you interactively build a search using "pearl growing." Or, instead of always starting with the narrowest "least posted facet" sometimes try building a huge set (for example, LIMITALL/HUMAN in MEDLINE on DIALOG) and successively chipping away fractions of the monster set.

THE LAST WORD

Today's online searchers are blessed with faster modems, powerful online searching machines, better searching and post-processing software, hundreds of online systems, thousands of databases, and more nontraditional alternatives such as the Internet. We are cursed with higher online costs, hidden online costs, inconsistent databases, inconsistent systems, and poor quality control.

Top-notch online searchers are the ones with style, flexibility, and a willingness to try new approaches or new systems (always with one eye on the cost).

REFERENCES

- [1] Quint, Barbara. "Inside a Searcher's Mind: The Seven Stages of an Online Search: Part 1." ONLINE 15, No. 3 (May 1991): pp. 13-18, and Part 2 ONLINE 15, No. 4 (July 1991): pp. 28-35.
- [2] Taylor, Robert S. "Question-Negotiation and Information Seeking in Libraries." College & Research Libraries (May 1968): pp. 178-194.
- [3] Gale Directory of Databases. Detroit, MI: Gale Research, 1993. 3 vols.
- [4] Quint, Barbara. "Online Self-Defense, Punch One." Wilson Library Bulletin (January 1992): pp. 89-91 and "Online Self-Defense, Punch Two." Wilson Library Bulletin (February 1992): pp. 70-72.
- [5] Tenopir, Carol. "Cost Effective Online Searching." Library Journal (October 1, 1992): pp. 65-66.
- [6] Fidel, Raya. "Searcher's Selection of Search Keys: II Controlled Vocabulary or Free-Text Searching." Journal of the American Society for Information Science 42 (July 1991): pp. 501-514.
- [7] Hawkins, Donald T. and Robert Wagers. "Online Bibliographic Search Strategy Development." ONLINE 6, No. 3 (May 1982): pp. 12-19.
- [8] Buntrock, Robert E. "The Effect of the Searching Environment on Search Performance." ONLINE 3, No. 4 (October 1979): pp. 10-13.
- [9] Fidel, Raya. "Searchers' Selection of Search Keys: 3. Searching Styles." Journal of the American Society for Information Science 42 (August 1991): pp. 515-527.
- [10] Bates, Marcia. "Information Search Tactics." Journal of the American Society for Information Science 30 (July 1979): pp. 205-214.

[11] Saracevic, Tefko. and Paul Kantor. "Online Searching: Still An Imprecise Art." *Library Journal* 116 (October 1991): pp. 47-51.

Copyright: COPYRIGHT 1993 Information Today, Inc.

<http://www.infotoday.com/default.asp>

Source Citation (MLA 8th Edition)

Tenopir, Carol. "Ten loose guidelines for online searchers." *Online*, Mar. 1993, p. 27+. *Gale Academic Onefile*,
https://link.gale.com/apps/doc/A13567241/AONE?u=tel_a_utl&sid=AONE&xid=404a8b95. Accessed 9 Nov. 2019.

Gale Document Number: GALE|A13567241