INSIDE A SEARCHER’S MIND: THE SEVEN STAGES OF AN ONLINE SEARCH — Part 1

by Barbara Quint
Editor, Database Searcher

Why me? Why would ONLINE call me to write this article? Sure it’s a great idea: what every searcher should ask themselves and their clients before they go ask a machine.

But why me? Everyone who knows me knows that I am a “grasshopper” searcher. And this sort of thing — going down a list checking off items — is strictly for “ants.” Remember Aesop’s story of the Ant and the Grasshopper where the ant worked hard all summer, planning in advance and storing up for the winter, while the grasshopper played and sang only to suffer and shiver in the cold until the ant showed mercy and shared its provender? Well, searchers are like that. Some searchers plot their searches out well in advance, conducting careful, detailed interviews, anticipating all the databases they may need, checking documentation, writing out all the terms, graphing the Boolean connections, and then, only then, go online.

Other searchers just grab requestors by their ties and drag them down to the machine for a merry session of searching. They conduct interviews on the way to the micro, select databases while dialing, and whistle up the first set of search terms while putting in passwords. They consider using LOGOFF HOLD equivalent to cheating at solitaire. It’s a matter of style. By the way, you will notice that the grasshopper ended up living off the ant with no more than a promise to change its lifestyle. If “grasshopper” searchers do anything diligently, it’s usually reading magazines like this one where they can pick up tips in articles written by “ants.”

So why ask a “grasshopper” searcher to write the quintessential ‘ant’ article? Scenario time . . . Possible Answer #1: Brevity. Only a “grasshopper” could keep the article short enough to prevent print overruns. Possible Answer #2: Serendipitous synergy. A “grasshopper” trying to think like an “ant” might, if you were lucky, reach a central core of mutual insight from the two searching styles. Possible Answer #3: Dastardly dealing. If I fall on my face from a great height, a competitor publication bites the dust. (Naah. They’re my friends. Besides, if I do it right . . .)

FIRST RULES OF ONLINE SEARCHING

Suddenly, the scales fall from my eyes. These three scenarios illustrate the first rules of online searching and how to use a checklist.

Rule One: Remember the bottom line when you search. Do what works. The searcher’s style is not as important as the answer. Some questions require one approach. Others require another. Whatever style answers the question is the right style for the moment. If you need a long checklist, use it. If you lack the time or resources or simply do not need extended pre-planning, then skip it. Or better, skip through the list lightly and save the careful perusal for the questions that require it.

Rule Two: Stay open and aware and flexible while you search. Prepare for good fortune and bad. Get the best feedback you can throughout the online search experience. Keep requestors with you during the search, unless they make you too nervous. Or stop the search, review and regroup, and go back on a little later with a clearer strategy. Use different approaches for different elements of a problem. Accessing one database may require a detailed, itemized plan while a quick and dirty search of another file more than suffices.

Rule Three: Look for hidden agendas. They can be benign or less than saintly, but failing to spot them will always prove malign to the search process. You can find them everywhere — your clients and the clients they serve (customer, boss, boss’s boss, professional colleagues, the next job, etc.). You have hidden agendas of your own. While you serve your immediate client, you also serve your employer, your supervisor, your profession, etc. Most important, you serve the next client. Today’s search should improve tomorrow’s — both your tomorrow and the client’s.

Rule Four: Quint’s Law of Research (plagiarized from a former client who probably adapted it from an old motto of a typesetter’s union): There are three kinds of research: good, fast, and cheap; but you only get two out of three. Sometimes you get lucky and fall into the perfect search, but that is not the way to bet. Professionals play percentages. Don’t promise your client more than you can probably deliver. Run through different tactics before you pick your first approach. Use the rejects as fallback strategies.

SEVEN STAGES OF ONLINE SEARCHING

Traditionally, the search process involves at least six stages, but this list uses seven:

- Reference Interview
- Tactical Overview
- Database Selection
- Search Strategy Formulation
- The Online Search
- Feedback or Reviewing Results
- Presenting Final Search Results
The untraditional stage is the "Tactical Overview": a general planning phase between the "Reference Interview" and "Database Selection."

The penultimate step on the list may require repeating the process until you reach a solution or run out of resources. Actually, the feedback phase should permeate the entire search process. As you plan and conduct a search, the process should educate you continuously. Looking at database sources and thesauri can change the nature of the request, initial approaches, and evaluation of results. Good online searching is interactive, frequently displaying terms and results to provide the searcher with new terminology and fresh approaches and to cut off unproductive lines of inquiry. Even after final search results have gone to the requestor, a good searcher should follow up on how the results were used and any new inquiries that may have emerged.

THE REFERENCE INTERVIEW

Without doubt, the most critical area of the entire search process, the most professionally demanding, and the most personally rewarding for all parties is the reference interview. End-user menu interfaces to online or CD-ROM databases pay intermediary searchers the ultimate compliment when it comes to this component of the search process: silence. They don't even attempt to duplicate what searchers do for clients in this phase. (And thereby hangs a tale — and perhaps another article.) Successful reference interviews are often research design on the fly, a problemsolver's quickie psychoanalysis.

The questions you ask the requestor and the way you ask them will vary from situation to situation. Much of the interaction will be nonverbal. Body language will tell you a lot about the requestor's comfort with elements of the discussion. Remember that you send off signals as well. Be careful that the requestor does not misinterpret them. As a rule, you want clients to receive a general positive impression of your ability and willingness to solve their problems. On the other hand, if you cannot meet their desires in every respect (e.g., a totally comprehensive review of a broad subject for less than $20), the client should know it before the interview is over.

The Requestor

Who are they? How do they spell their name? Where can you reach them? Who will pay for the search, if your institution bills back? How will they pay (credit card numbers, project numbers, departmental budgets, signed authorizations, etc.)? Will there be any special delivery problems: off-site, out-of-state, etc.? Did they sign the contracts or fill out the forms your organization requires?

What are they? Are they expert in the subject area of the request? Are they expert in related areas? Do they know how to handle bibliographic information effectively? How capable are they of absorbing the information you expect to find for them?

What format do they prefer for the data? Can they handle machine-readable, print, fax, etc.? Are there any machine specifications you should know about (type of microcomputer, fax number, file transfer protocols, hours and conditions of operation for their equipment, etc.)?

Will they use the information alone or with other people? What other people? Who are they? How can you reach them? What kinds of data are the other people expert enough to handle? Do they need their own copies of the results? In what formats?

The Search Request

What is the question? What do they know they need to know? What do they really need to know? Does talking through the whole information problem expand or redefine the request? What do they need to know most? What do they need to know first? What will they need to know in the future on this subject? How far is the future?

What do they know already? What can they tell you that will help you construct your search? Key terminology? Key sources — personal (e.g., author names), institutional (e.g., government agencies, companies, research facilities), bibliographic (e.g., key journals, conference proceedings, books)? Do you have the correct spelling, the meaning of abbreviations, acronyms or other jargon, specific sources or types of literature the requestor prefers, etc.? Do you have a timetable for when certain elements became important in the development of the subject (e.g., historical players, recently developed concepts and terms, etc.)? What or who have they already checked? Do they have some "perfect" references or sources? What are they and can you see them? Do they have contact information for experts in the area? Would it be alright if you contacted the experts directly?

How soon do they need the results? Does the timetable allow additional time to retrieve documents or contact identified sources? Would a full online answer change the timetable? Truth to tell, how real are the deadlines?

How much will they pay? How sure must you be that you can find the answers before you can commit the money? Any problems with payment (billing requirements, payment delays, currency, credit checks, etc.)?

What will they do with the information? What specific formats or delivery media would affect the utility of search results? Will they need documents identified by the search? Will they need supplementary information to round out answers retrieved online?

The Requestor's Expectations

Do they understand enough about the search process to form a sound judgment about the prospects for success with this and future requests? Have you used the interview to educate them on your services? Do they understand why you've been asking all those questions? Do they really know how much this will cost? Did you warn them about dangers or weaknesses in the search process (e.g., a lack of resources or skills in certain areas, flawed data sources, negative impact of cost constraints, etc.)? Did you sum up your view of the final search request, checking the relevance of the keywords you expect to use, discussing the relationships of general concepts, explaining likely retrieval limitations, etc.?

Do they need more data than you will find today? Can you deliver results piecemeal? What do they require when? Does the question involve an ongoing flow of information (breaking stories, research project results, future scheduled events, etc.)? Will other people sharing some or all of
the search results need more information? What kind? How soon?

**How does the requestor plan to handle the search results?** Will they initiate their own document retrieval activities, and under what conditions (library request forms, shelf-checking by requestor or their staff, external paid retrieval services, etc.)? What are their specific document delivery sources (office journal collection, local library, etc.)? Have or should any document delivery conditions put restraints on the nature of the search request? Can you identify any alternatives to improve the final result within their budget and time constraints (full-text online sources, information broker delivery, bibliographic results limited to deliverables, etc.)? Do such recommendations change the scope of the search request?

**What does the requestor plan to do with the data?** Will it need substantial manipulation or evaluation? Of what kind? Do they know how critical such processing would be to the use of the data? Would failure or inability to conduct such processing affect the search request? Do they have the resources to handle the processing? Can you identify any alternatives to accommodate any or all of these processing requirements within their budget and time constraints (e.g., multiple sources for cross-checking data, extensive reformatting of search results, delivery of numeric data in spreadsheet modes)? Would such recommendations change the scope of the search request? Do they understand the legal and institutional limits on use of the data in certain formats or for and beyond certain purposes?

**TACTICAL OVERVIEW**

Between the Reference Interview and Database Selection stages, you need a Tactical Overview. Wyatt Earp, the famous marshall of Tombstone, Arizona, told a biographer once that the secret of his success as a gunfighter lay more in his swiftness than his speed: he always took a split second to aim. Once you start diving into manuals and thesauri, and begin drafting long lists of synonyms, code numbers, and cross-referenced terms, your concentration may block off the "big picture." Take time to aim.

---

**Five Rules Of The Interviewer's Art**

1. **Beware of requestors who ask the answer, not the question.** When requestors start asking for checks of specific sources (databases, journals, authors, etc.), they reveal the presence of a set of assumptions, many of which may relate to their notions of what you do and how you do it. These assumptions are frequently wrong. Some requestors come down gripping, physically or mentally, a list of keywords like a roll of nickels for a jukebox. Speaking of nickels, if I had a nickel for every time a requestor sent me paper-chasing, tracking some will-of-the-wisp specific reference, only to find that they wanted some general piece of information that this material might not even contain and which I could have found in five minutes if only... well. Insist on broadening the inquiry. Try the key questions: "What's the information for? What are you going to do with it after we find it?"

2. **Never trust requestors when they tell you what they don't need to know.** Always listen, but never trust. Most requestors have no idea of the breadth of information available and/or your power to pinpoint needed data. They cannot assess what they do not always understand. Find out why they don't want you to pursue a line of inquiry — time, expense, fear of overload, etc. They may be right in their assessment. But you need to make your own independent judgment to guarantee the client receives the best you can get. Ask them what getting lucky would look like. If the search went like a dream pouring out everything they wanted at no loss of time or money, what would "everything" look like? After they tell you, don't forget to return them to the real world with warnings about the laws of probability as applied to searching.

3. **Get on the team. Get interested or, at least, look interested in the complete problem the search process will assist.** Start using first person plural words ("we," "us," "our") as the interview progresses. Maintain direct eye contact. Use the appropriate emotional sound signals on the phone ("Hmmmhh...", "pause for effect, "Veerrry interesting"). Ask intelligent questions about the topic or situation itself. Contribute a smattering of personal experience or insights. Keep in touch with requestors throughout lengthy search processes.

4. **Talk to requestors.** I hate printed search request forms. If you use them, fill them out yourself as you conduct the interview. If you absolutely must give them to clients, e.g., for after-hours requests when no searcher is available, then follow up in person before you conduct the search. Whatever the demands on searcher staff, if possible, no search should be conducted without the searcher talking to the requestor directly.

5. **Take good notes.** Jot down even remotely relevant background material. Work out a personal code for indicating the subtleties you pick up in an interview, e.g., question marks next to dubious data, plus marks next to concepts that need more terms, asterisks next to key sources or cut-off points. Make the note-taking very natural. Don't let it interfere with the interaction between you and requestors. Before you close the interview with a summation to the requestor, don't forget the spelling checks.

---

1991 May ONLINE 15
Broad Controlling Factors
What are your institutional controls? Must you search only databases on search services you already have? Does management prefer you to search offline resources first? If your institution or department pays for the search, is there a de facto or de jure limit on how much you can spend for a particular search? Do you need to get approval before conducting a search? Could these institutional controls interfere with the success of this particular search? If so, can you get around them or should you warn your requestor?

What are your personal limits as a searcher? Which services do you search effectively? What subject areas or databases disqualify you from searching? At what level of difficulty? Can you draw on other searchers to help with a search? Which search services do which other searchers search effectively? Do special arrangements need to be made to acquire external searchers skills (e.g., contacts with information brokers, formal or informal search swapping arrangements with colleagues)?

What do you think of the requestor? How sharp are they? When you asked them for ideas on sources or where they had already looked, did the answers indicate a good knowledge tracker? Do they seem like a good risk for paying back the costs of the search? Can they handle the data — technically (e.g., right equipment for format chosen), intellectually (e.g., sufficient expertise for the necessary manipulation or evaluation), psychologically (e.g., reluctance to be "shown up" by search results or special problems such as the impact of negative medical searches on patients)? If the requestor represents a group of users, do the results need special treatment? Do the group’s needs, the requestor’s limitations, and your responsibilities require transmitting results to the group directly?

Broad Approaches
What kind of search will this be? Answer or Reference? Precision (high relevance) or Recall (comprehensive)? Big or Little? Can you form an image in your mind of possible search results ranging from the ideal to the acceptable? Can you map possible approaches using the three key combinations: Good and Fast, Good and Cheap, Cheap and Fast? If money were no object (Good and Fast), what approach would you pursue? Can you acquire currently unavailable resources to answer this question (e.g., subscribe to new search services, hire a consultant or information broker, etc.)? If you had all the time in the world (Good and Cheap), what approach would follow? What offline alternatives are available — formal or informal, documents or personal contacts? If you had to live within rigid constraints (Fast and Cheap), what approach would give you the best you could afford? Would the requirements limit the search to specific kinds of data (e.g., deliverable documents)? Can you see any approach that would achieve all three qualities — good, fast, and cheap — simultaneously?

What portions of the search results need the highest quality controls? What the least? Are there any special ways of insuring that quality? Will you need to perform any special tests or cross-check more than one file for the same information? Do you have the knowledge or resources to conduct the tests yourself or will you need to arrange it? Can you rely on the requestor to judge the data? Will they need any assistance from you? What areas will you need to watch and warn them about?

What are your cut-off points? What are your time deadlines? At what time or date would even good search results cease to matter? Would any particular result stand up better to the test of time? Would any particular result delivered faster than others eliminate the need to conduct the rest of the search? What expenditures are realistic from desirable to absolutely top dollar? Do the expenditures depend on the likelihood of particular outcomes? Will the results you expect justify the expenditures? How do you plan to recognize when the results do not justify further expense? How much error in judgment will your organization or client support? What information must your search results contain? What results would be useful but not imperative? What results would stop the search? What results would expand it? When should you return to the requestor for further guidance?

Where should you start? Which approach will produce the best results? Do you need to conduct test runs (e.g., small samples from a series of databases or from a megabase in the field)? Where is the bull’s-eye on a precision search? How much can you spend trying to hit the bull’s-eye? What factors would determine the first strategy has failed and why? What types of failure would dictate what sorts of secondary strategies?

What do you hope to get out of this search yourself? Better client relations with the customer type represented by the requestor? More search experience in the field or on specific software interfaces? Guidance as to new services or sources needed by your client community? Does this particular request hold any potential for serving other ends — short-term or long-term — than the immediate search request itself?

DATABASE SELECTION
With a clear view of our goals, our needs, and our responsibilities, we begin the selection of the proper sources. In the real world, database selection involves balancing relevance against accessibility. The perfect source may demand and get extra effort from searchers, ranging from brushing up on a forgotten software interface to acquiring a whole new service or a whole new searcher. On the other hand, accessibility is a powerful attraction in itself, not simply because it makes the searcher’s life easier, but because it can also speed completion of the search, reduce costs through discount rates from heavily used services, and allow expansion of a search into important, related areas. Just to keep the dealer honest, however, every searcher should have access to a good database directory and use it often. How else will you know when you need to add a service?

Relevance
What subjects does the search cover? What databases focus on those subject areas specifically? What databases subsume the topic in broader coverage?

What type of literature carries material on the subject (e.g., books, scholarly journal articles, government reports, research reports, biographies, encyclopedias, etc.)? What databases
It is surprising how often the last rule is breached, usually with the victim's cooperation. Experienced searchers know that the toughest search to conduct is a "negative search", one where the failure to locate information constitutes information of value in itself. Yet clients often confuse simple failure with informative failure. "No directory listing found" or "No bibliographic citation verified" could mean mis-spelled, mis-spaced, mis-abbreviated, or several other errors in searcher entry of terms. It could mean the database producer missed or mis-entered a listing or a search service hasn't loaded a tape yet. But requestors often interpret a failure as meaning the data are available in any and all sources, or even that the information they used to start the search was false. End-user searches are notorious for overestimating the effectiveness of their searching power. When that happens, the Searcher's Code has indeed been breached.

Question: Does CD-ROM live by the same code? Should you post notices by the machine to warn people of what the shiny disks do not cover? What about a poster carrying in large bold-faced letters the statement, "OH, NO, YOU HAVEN'T" followed by a list of all the possibilities failure to find a reference on the CD-ROM do not indicate? How about a stack of business cards from your professional searchers, referring unsatisfied CD-ROM clients back to the full-power information services of dial-up online? —B Q

cover the format required? Do any of the format-oriented databases provide the only coverage for material eschewed by the subject-oriented files?

How comprehensive are the databases? What rules govern inclusion of material? Do the databases cover the dates required (e.g., starting dates, lag time or currency)? Do the databases incorporate the geographic areas needed (e.g., foreign bibliographic coverage, regional statistics, etc.)?

Do they cover the languages desired or required? Do the databases include document delivery (full text, online ordering, etc.)? Have rules for inclusion changed over the years? Will those changes affect database effectiveness for the specific search?

Which databases contain all the needed information (e.g., full bibliographic citation, abstract or description, full text, etc.)? Do some databases have better records while others have broader coverage? Does the search service allow you to search on needed fields and sub-fields (e.g., dates, language, geographic areas, etc.)? If you cannot search for specific information in a specific field, can you retrieve a larger set and reprocess the results using online or local computer power (e.g., download records and search through them with microcomputer software)? Which databases have better access points (e.g., good descriptors, codes, etc.)?

Which services offer better searching features for relevant databases (e.g., displaying search terms in full-text records and moving to full-text display from any point in the text, full phrase operators, field-specific displays of the inverted index, full-record free-text access, etc.)? For which databases would which access improvements be critical? For which, desirable? How serious are the differences in terms of specific retrieval for the specific search (hopeless, wasteful, fixable, inconsequential)? Can you repair deficiencies (e.g., by accessing one or more databases with related information, by supplementing the search with offline approaches, etc.)? How easily?

Which databases did the requestor suggest? Which databases will carry the citations the requestor already retrieved? Are you sure? Did you check the database directories online or in print? Did you review any specialized manuals or documentation reflecting files within files (e.g., full-text source lists for multiple-source files, collections of material from specific sources, etc.)? Did you check the cross-file index to database contents that many search services maintain online? Did you look for the unexpected source?

Accessibility

What relevant databases are available through search services you currently access? Which services do you search most effectively? If other services offer significant discounts for searchers expert in their protocols, are the savings and circumstances appropriate for acquiring such skill directly or indirectly (e.g., local searcher training or external searcher access)? How quickly can you gain access to currently unavailable databases? How quickly can you become an effective searcher of a new database or service? Do occasionally or never used search services with relevant databases offer simplified access through menus or other end-user styles?

Which databases on which services are cheapest to search? Which services carrying databases you need offer special cost savings (e.g., discount programs, cost-saving features, etc.)? What are the cost elements you include in judging the expense? Do they include overhead elements (e.g., staff training time and expenses, special equipment purchases, subscription fees, etc.)? Do they include
requestor expenses (e.g., revising or reentering machine-readable data, document delivery, expert evaluation, etc.)? What are the true costs? What are the billable costs? Which databases are cheapest under both definitions?

Assuming “Good and Fast” drives the search, how would unimpeded expenditures provide solid value? Can you acquire access to a currently unavailable search service fast enough to meet requestor requirements? Should you locate an outside searcher who already accesses the service and negotiate a deal? Are there any currently accessible databases on which your current skills may not suffice? Are there special features or commands to improve product output on specific services (e.g., extensive statistical processing on a numeric search service, mailing list labels for a directory file, etc.)? Can you find / negotiate another search route for those files? Which service or source includes document delivery or desired auxiliary material?

Assuming “Good and Cheap” drives the search, are there any alternatives for reducing the search costs (e.g., offline manual searches, acquiring access to cheaper services, CD-ROM versions of the file, transferring search costs to demonstration programs, etc.)? How much longer will the slow approach take? How will the cost reduction play affect the quality of the information retrieved? If quality is seriously affected, should you warn the requestor before conducting the search at all? Are there any services with special features that could reduce total search costs (e.g., duplicate detection, weighted search results, etc.)?

Assuming “Fast and Cheap” drives the search, to what extent will quality be affected? What warnings will be needed to assure the requestor makes no false assumptions about the results? Will an alternative approach be required (e.g., locating expert contacts rather than documents or the information itself)? What defines speed to the requestor? Does it include the whole process of reaching an answer (e.g., document delivery, phone calls to experts, etc.)? How does the requestor’s definition affect source selection (e.g., restrict to immediately available material through online catalogs or source-specific searches of format-oriented files)?

“Batting Order” Selection
Which approach appears the best now that you have taken stock of your resources? (Review the Tactical Overview checklist.) Which is “Good and Fast?” Which is “Good and Cheap?” Which is “Fast and Cheap?” Which databases best support which strategy?

Do you want to “shoot for the bull’s-eye,” i.e., try for a relatively small set of search results with high relevance using precision search techniques? Which databases contain all or most of the information you need? Which search services possess sufficient retrieval software features for accurate retrieval of high-relevance material? Do your searcher skills suffice to conduct a high-precision search? If not, do you have time or resources to acquire them (e.g., coaching, alternative searchers, etc.)? If you must operate under cost constraints, which databases have the most chance of producing the greatest relevance for the least expense?

Do you need to perform a critically comprehensive search? What array of databases cover the field? Do you need to conduct a series of “sampler” searches to test different search strategies? Or do you need to move around different areas of the search through different types of databases to improve your search strategy formulation? If so, should you start with relevant databases that will prove representative of other databases to be searched later? Should you arrange the databases by the search services they appear on and start with the ones with the closest terminology to facilitate group file searches or transporting saved searches (e.g., hoping to speed retrieval and reduce costs)? Should you move from subjects or databases or services with which you feel comfortable to those that will need more effort (e.g., hoping to have found what the requestor wants before you reach the tough ones)? Should you start with the cheapest relevant database on the most familiar search protocols (e.g., hoping to get lucky before the money runs out)?

STAY TUNED!
In the next issue, we will complete the journey to the terminal and circle back to the requestor with a small detour down “Hidden Agenda Lane” to seek guidance from today’s successful and unsuccessful searches on how to improve tomorrow’s searches.

Communications to the author should be addressed to Barbara Quint, 932 11th Street, Suite 9, Santa Monica, CA 90403; 213/451-0252; DIALMAIL — 13675; ALANET — ALA0717; CompuServe — 72557,23.