van Brakel, Pieter A.

Twenty Years of Training in Online Searching: Integrating the Internet with the Teaching Programme.


Reports - Evaluative (142) -- Speeches/Meeting Papers (150)

Access to Information; Foreign Countries; Higher Education; *Instructional Design; Instructional Innovation; *Internet; *Library Instruction; *Online Searching; *Search Strategies; Technological Advancement; User Needs (Information)

Historical Background; HTML; South Africa

The Internet and its features such as e-mail; listservs; FTP and Telnet; navigators such as Archie, gophers, and the World Wide Web; and Web search engines have had a profound effect on the nature of online systems. This paper discusses the effect the Internet might have in the way online searching should be taught. The paper provides background on the history of online search instruction since the mid-1970s. Guidelines are provided to academic staff for ways to integrate these changes into their online search courses, for example: the basics of Internet searching, studying and evaluating World Wide Web Search engines, essentials of hypertext markup language (HTML), principles and practices of Web home page design, and rendering information services via the Internet. This paper represents an effort to isolate aspects relevant to current online searching techniques and to integrate these with established instructional methods. This same approach has been followed when discussing teaching online searching procedures: that is, to select and integrate relevant Internet procedures with the training program. The instructional objectives, learning content, and practical lessons are provided as examples, and serve as a frame of reference for further development by a specific training unit. (Contains 19 references.) (Author/SWC)
Twenty Years of Training in Online Searching: Integrating the Internet with the Teaching Programme

By:

Pieter A. Van Brakel
Twenty years of training in online searching: integrating the Internet with the teaching programme

Pieter A. van Brakel
Department of Information Studies, Rand Afrikaans University, South Africa

Abstract: The late seventies saw the disappearance of the dump terminal, an essential element in the online search process. Consequently, at the beginning of the eighties the popularity of personal computers was putting more intelligence into the online search process. Automatic logon procedures, offline search strategy design, and easier downloading and reformatting of search results were features that shaped the way online searching courses were offered at tertiary level. In the meantime, the high telecommunication costs were replaced by cheaper rates provided by the development of packet-switching and value-added networks. The popularity of online simulations and CAI courses consequently declined because live or hands-on searching by students became much more economically viable. Cheaper communication rates were not the only reason for this phenomenon. Special ‘classroom contracts’ with very low fees were being made available to tertiary institutions, as well as the creation of small subsets of popular databases, accessible at similar special tariffs. The publishing of databases in CD-ROM format during the second part of the eighties created yet another innovation which influenced the way in which training in online searching was done. Now hands-on practicals could take place on large CD-ROM databases, using a specific host’s command language but without being actually connected to the host. The early nineties provided the possibility of networking these CD-ROM databases, thus creating opportunities to run training programmes where multiple students can access the same database to solve the same or different search queries. The Internet and its facilities have brought about yet another need for change. But in contrast with the previous developments, facilities such as e-mail, listservs, FTP and telnet, navigators such as Archie, gophers and the World Wide Web, and Web search engines all contributed to having a profound effect on the whole online scene. This paper will discuss the effect the Internet might have in the way online searching should be taught. Guidelines will be provided to academic staff as to ways and means to integrate these changes into their online search courses: for example the basics of Internet searching, studying and evaluating of World Wide Web search engines, essentials of the HTML language, principles and practices of Web home page design, and rendering information services via the Internet.

Keywords: online searching, Internet, World Wide Web, teaching, home pages.

1. Introduction

Training in online searching techniques inevitably started soon after the advent of large commercial database host systems such as NLM’s Medline, Lockheed’s Dialog, or SDC. It is common knowledge that when these systems were launched during the mid-seventies, training was not only offered by the hosts themselves but also by the producers of databases. Training at tertiary level followed suit, more or less towards the end of the seventies. In fact, it was at the very First International Online Meeting in December 1977 that Vickery (1977) reported on a BL R&D project meant to develop the use of online services in schools of librarianship and information science in the UK. In those days time was spent on such questions as where to fit online searching techniques into the curriculum (for example, should it be part of the core curriculum?), the exact nature of the instructional objectives, the most effective teaching methods and media, and so forth.

The fact of the matter is that today tertiary training units are still seeking answers to some of the queries raised at that time. The main reason for the attention to the relevancy of such a course was probably that, for the first time, direct hardware and software costs were brought into the teaching process. Training units had, for the first time, to motivate funding for electronic equipment such as terminals, telecommunication, database fees (for hands-on exposure), system manuals, and also for training media developed soon after training started, such as work books, as well as tape/slide presentations and later videotapes portraying the essentials of online searching. Further problems of teaching online searching had to be solved by tertiary training units in the years to come. A decade after the above-mentioned report by Vickery, Large (1987, p. 288) confirmed that although ‘...online
searching is now an accepted part of the library school curriculum . . . [and] . . . has now entered the mainstream of information studies, reflecting the widespread availability of online systems in the outside world . . . a number of problems continue to plague educators.' His list is representative of the typical challenges that training units have to address:

- **Extend of hands-on experience in searching databases on remote online hosts.** This includes decisions such as the ideal duration of experience, host(s) to subscribe to, types of databases to include and the best database titles to search;
- **Content of a course.** Various topics could be included to strengthen the theoretical part of an online course, for example management and marketing of online services, the economics of the online industry, the role of the pre-search interview and so forth;
- **Which computerised teaching packages to choose?** Several approaches exist to search offline on subsets of databases, for example simulations and emulations. However, even the best examples can never replace online experience on real databases;
- **Cost effectiveness in terms of staff resources.** Because of the costs involved, extensive supervision is necessary when students search live;
- **Location of online searching within the curriculum.** Decisions in this regard proved far from easy: should it be taught as part of a current module, or rather offered as a separate course?
- **Scope of the standard course is steadily growing.** 'It is no longer sufficient to deal with bibliographic databases: full-text and numerical ones must also be covered.'
- **Keeping up to date with developments in a rapid advancing field is impossible.** There is no lack of published sources such as books, journal articles, conference proceedings and newsletters from database vendors which discuss developments such as new search commands, new databases, the latest in electronic equipment and new search approaches.

Surely, apart from faster and more reliable datacommunication facilities, two major technological breakthroughs did contribute to improve the quality of teaching online techniques. The first was the publication of large bibliographical and full-text databases on CD-ROM, and the second the availability of cost-effective local area networks. Now hands-on practicals could take place on large CD-ROM databases published, for example, by H.W. Wilson or KR Dialog, so that practising their command languages online but without being actually connected to the external host systems was possible. The early nineties provided the possibility of networking these CD-ROM databases, thus creating opportunities to run training programmes where multiple students can access the same database during the same practical.

But has the situation changed since the Internet came to the fore? Surely, the various technological and other developments which have taken place since those early days of dumb terminals and slow communication lines put their mark on the nature of online searching, but none of these innovations could surpass the impact the Internet might have on information retrieval techniques, and specifically, ways and means of integrating the Internet (with its facilities and navigators) into the online teaching programme at tertiary level.

A good starting point would be to investigate, as a frame of reference, what online searching constitutes, in an effort to establish which elements of the Internet can be categorised under the term 'online searching'. This will form the basis for the further discussion on how the Internet could be involved in the teaching process.

### 2. When can ‘Internet searching’ be regarded as ‘online searching’?

An investigation of the publications about the essentials of online searching indicated that not many differences exist between the different discussions on what exactly is meant by the term 'online'. In one of the earlier writings about online searching, Barraclough (1977) explained some of the more important concepts of online:

- **Online processing:** 'Literally it means that the user is in contact with the computer through some direct linkage . . . now [it] is accepted as being some kind of individual terminal attached to the computer either by a direct line or over the telephone network. The next prerequisite . . . is a large . . . store so that the programs for many users can be held within the computer and can be accessed rapidly. The other essential requirement is a means of sharing the resources of the machine among all the users . . . ’ (Barraclough 1977, pp. 5-6);
- **Interaction:** 'Essentially one is concerned with the reaction of the user to the output produced by the computer and vice versa. It could be thought that it is only the speed at which this occurs that makes the process interactive . . . but another important factor is the amount of skill and intelligence employed by the user during this process' (Barraclough 1977, p. 6);
• **Interface:** ‘... provide one of the easiest ways of using a computer and indeed ... allow a user knowing very little about computing to be able to access a large computer...’ (Barraclough 1977, p. 7);

• **Information retrieval problem:** ‘The first problem in setting up a retrieval system is to produce a good definition of the documents to be entered into the database. Clearly all the bibliographical information has to be held. Most systems attempt to reduce the size of the text by either using abstracts or key words...’ (Barraclough 1977, p. 7-8).

More recently, in the introduction to a volume on training for online, as part of Taylor Graham's series *Foundations of Information Science*, Jackson (1987, p. 1) provided the following, more recent, description — the emphasis is however still on bibliographic databases, as was referred to in Barraclough's 'retrieval problem':

In online searching for the retrieval of information, the searcher at a terminal is connected to the computer which holds a collection(s) of data in machine-readable form know as a database(s). The searcher interacts with the computer and poses a series of queries, in the form of search terms or indexing terms which will correspond with those in the collection of data in the computer store. The computer responds by listing items identified by these search terms. The searcher can then adapt the search according to the response from the computer.

Apart from more sophisticated machine processing, more versatile search software which reflects more powerful search commands, and friendlier user interfaces (for example Windows and Windows-based software such as DialogLink), the modern meaning of 'online searching' did not change one bit from those earlier writings. Full-text databases have been added to the list of sources being made available online, and these and other sources being published in CD-ROM format were distinguished by the term 'local online'. But even 'local online' still comprises the same idea, that is, interactive manipulation of data, bibliographic references and/or the full-text of information sources via a terminal and by an end-user.

The discussion so far can be classified as addressing the 'traditional' view of online searching. But how does this differ from information retrieval activities when using the Internet? The first point to remember is that the Internet provides the facilities for the development of a complete information system (or systems) in its own right. One cannot compare the Internet's navigation facilities with traditional online activities, as described above: that is, manipulating the content of large bibliographic and full-text databases interactively. Wittig & Wolfram's (1994, pp. 636-627) contribution provides a good starting point in establishing how the Internet could fit into the online teaching programme. They argue that where the traditional online systems permitted timely retrieval of information from large centrally located bibliographic databases, the latest Internet-based storage and retrieval tools permit access to decentralised resources.

Koenig (1994, p. 365), in an earlier writing, emphasised the trend that the location of information becomes less and less relevant, ‘... as we move information around more and more readily.’ Basili (1995, p. 459) is more specific in reminding us that traditional online database systems have a star-shaped architecture where the local PC links up with the service-supplying computer (the host). The local computer is completely subordinate to the central computer in terms of both query language and type of service. In a single session only one host is accessed and an online search is thus limited to the databases of that host alone. In the case of the Internet the connection is reticular, because it (the Internet) consists of a constellation of interconnected nodes in which in one network session (for example a World Wide Web's) it is possible to execute a search for information which can move from one node (system) to another.

These differences still conform to the 'online' idea: the Internet's navigator searching — for example Gopher, Archie and the Web — can be regarded as 'online searching', because as indicated in the above definitions, processing is executed in an online mode (albeit via TCP/IP and a client/server architecture). There is also interaction via an interface (browsers, that is), and with the main purpose of solving an information retrieval problem. The Internet should be seen as a fast growing online system, consisting of a wide variety of so-called navigators and search services. It is being accepted that the average tertiary training unit is already incorporating the Internet in its courses (see for example Wittig & Wolfram 1994, p. 630), but the argument here is that any online-based sections of the Internet should be fully integrated into those areas of curricula which were originally designed for the training of online searchers.

3. **Online via the Internet: suggestions for course work**

This section identifies and discusses those elements which could be integrated with current online courses. From a didactic point of view, these elements represent the results of the so-called situation analysis, which is always the first phase of curriculum development. The identification of these curriculum elements was based upon the author's experience in teaching online searching techniques to information science students. Articles traced on Internet training mainly covered in-service training programmes.

3.1. **Positioning of a module for Internet searching**

It is strongly recommended that teaching Internet searching techniques should commence only after the student has a good working knowledge of traditional search techniques on commercial online hosts. This well-structured
environment of database producers, hosts and clients (the players in the online industry) represents the platform necessary to acquire a solid understanding of and the skills for manipulating commercial databases cost-effec-
tively. Apart from a few exceptions the Internet's online environment cannot yet be described as a good example of how an online system should function. Lack of security, copyright protection, lack of privacy and consistency among data resources, and lack of standards and reliability of information sources still prevail. However, Notess (1995b, p. 87) recorded recently that the development of databases of Web pages and sophisticated search engines for Internet resources exemplifies the growth of the Internet as a significant player in the online industry.

3.2. Instructional objectives, learning content and practicals

In 1994 Williams (1994, p. 155) wrote that the Internet has provided an atmosphere or environment for greatly improved online services, that is 'the next generation of online'. This module's overall teaching goal should be formulated in such a way that it conveys not only the idea that although this 'next generation' represents completely new online frameworks, sources and potential clients — and accordingly, new online search techniques — the new generation still supports traditional online services. As a starting point, the first (intro-
ductive) objective could be formulated and developed into content and methods as follows:

**Instructional objective 1**

The student can demonstrate why the Internet can be viewed as a new approach to online searching.

Suggested course content and practical experience:

1. Origins of the Internet (short history of recent developments; how it is managed and funded).
2. Basic concepts, such as TCP/IP, client.server architecture (explain differences between typical inter-
active and client/server computer communication).
3. Additional facilities which can be applied to supplement the online search process:
   - **a**) telnetting to database hosts (implications of remote login; cost comparisons; procedures to follow; selection of important addresses).
     \[\text{Practical:}\]
     1. Hands-on experience of using telnet to manipulate external databases such as OPACs, as
        orientation, hence practising the phases of accessing those hosts with telnet addresses.
        See Keys (1993) for a good overview.
     2. Establishing cost-effectiveness of datacommunication of telnet versus packet switching
        and/or value-added networks; datacommunication speeds should also be compared.
   - **b**) How to use e-mail to receive downloaded search results, or ways and means to communicate
      value-added information products to clients of information services.
     \[\text{Practical:}\]
     1. Learning the use of a specific e-mail package, such as Pegasus emphasising the Windows
        version, by sending and receiving messages and file attachments.
     2. Study case studies set up by the tutor to analyse a specific enterprise's information flow
        patterns; suggesting how e-mail can improve these patterns to improve the decision-making
        process.
   - **c**) Listservs which are available to be kept updated on new online developments (distinguishing
      listservs from the popular news groups; important sources of listserv addresses; advantages of
      creating a listserv in an organisation; functions of the listserv coordinator).
     \[\text{Practical:}\]
     1. Search for listserv addresses on specific subjects or scientific fields, by using directories
        from gophers or Web sites. Use, for example, Kovac's online Directory of Scholarly and
        Professional E-Conferences (1996).
     2. Practise the procedures and commands of specific listservs, for example by subscribing and
        contributing to and cancelling from a listservs.
     3. Be involved in creating and coordinating a local listserv to keep clients updated with devel-
        opments.
   - **d**) Current awareness services (SDI) available via telnet (nature of this new approach, namely
      electronic document delivery) (such as UnCover) and SDI services combined; different systems
      currently available; criteria for establishing the best system for a specific organisation (for criteria,
      see for example Mancini (1996)).
     \[\text{Practical:}\]
     1. Searching EDD systems for SDI purposes; comparing search results between EDD and
        those of online hosts (for example KR Dialog's SAVE SDI command).
     2. Researching digital EDD systems (such as Ariel) versus the fax-based approach.
Managing Internet search activities (similar as well as alternative steps were identified by Makulowich (1995): define the area and level of expertise; require clients to complete a request form; categorise the Internet in lay terms for casual users; establish a realistic time for retrieving results; adopt a disciplined, systematic approach in your search efforts; understand your operating platform and the major tools available; maintain a file of important servers and addresses; develop your own bookmarks and home pages; learn the shortcuts).

When comparing the Internet's navigators and the amount of reporting being done in the literature, it is obvious that the World Wide Web is fast becoming the most sought after online system. The reasons are self-explanatory: standards and protocols such as HTTP and HTML have been established, friendly browsers have been developed and improved in the commercial market, inclusion of multimedia sources is supported, versatile hypertext-based organisation of information can take place, and so forth. Furthermore, the quality of the Web's online tools (search engines such as robots, worms and spiders) does provide for reasonable retrieval quality. It is hence suggested that one objective (Objective 2) is devoted to accomplishing a working knowledge of the different navigators, and thereafter an additional objective be devoted to an in-depth study of the Web's search engines (Objective 3).

Instructional objective 2

A student can describe the features of the different Internet navigators and use them to search for relevant information.

Suggested course content and practical experience:

(2.1) Working knowledge of the techniques for searching the different navigators, namely gophers, Archie (including Veronica and Jughead), Wide Area Information Server and the World Wide Web (overview of the history of each navigator; unique features; their scope, strong and weak points; application possibilities; hardware and software needed for setting up a local navigator site).

(2.2) Variables to evaluate/compare the search and retrieval features of the different navigators.

Practical:

This practical for 2.1 and 2.2 is analogous to learning to use the complexities of traditional online host systems such as command language and search procedures. Search queries are being set up and a student is expected to find, download and compare relevant information retrieved from different navigators.

(2.2) Variables to compare navigators and establishing when to search which one to use for specific type of information

Practical:

After the necessary theory has been dealt with, the student should get the chance to compare the features of the different navigators. Variables could be applied by using specific types of queries, for example date and/or subject coverage, retrieval and precision ratio, user-friendliness, reaction speed, error handling, reliability, etc.

When studying the evolution of the Internet, clearly demarcated levels of development emerge. Notess (1995b, p. 87) explained these developments as follows:

- First, the Internet was seen as a telecommunications network, with hosts such as STN, KR Dialog and OCLC making their databases available via telnet connections;
- with the growth of Internet-based information, for example from the commercial sector, these players in the online industry began to participate by developing gopher servers or home pages (see for example Notess 1995a, pp. 75-78). Although these pages contained information about their databases and services, direct searching could only take place via a suitable Web interface to a host system (interfaces should not be confused with Web browsers such as Netscape and Mosaic).

As we know, the standard interface to online systems has been an interactive, continuous connection, using a powerful command language and search facilities for effective interactive manipulation of data and/or information in these databases. Traditionally, these systems lack the graphical user interface so common in basic microcomputer packages of the day. With the development of the Web, the standard interface could be in form format (Notess 1995b, p. 88). The user can now fill out a Web form and submit a query in this user-friendly way. These developments, representing a new model of online interaction, are reflected in the following two objectives (3 and 4).

Instructional objective 3

A student can demonstrate the various levels of Web search tools for online searching and can find relevant information on different categories of search queries.

Suggested course content and practical experience:

(3.1) Study of the home pages of different online search services (interpretation of specific links from these pages, such as FAQs, publications such as manuals newsletters, pocket guides, document delivery
options, categories of databases, commands and search features, help desks).

Practical:
Given a list of the URLs of representative database host sites, the student is expected to study and compare their content for orientation, for example the EasyNet 2.0 Web site (http://www.telebase.com). A link in its home page can be used to activate a telnet connection to the real EasyNet gateway.

(3.2) Status of current Web interfaces to host systems (defining a Web interface; advantages when compared to the line mode approach; specific examples such as OCLC's or Medline's Infotrieve, which offers access to Medline using natural language as well as Boolean searching. The list of bibliographic references is ranked by relevancy and also included is a point-and-click ordering system for full-text copies (Medline over the Web 1996, p. 3)).

Practical:
Investigating the existence and features of this type of interface according to a set of guidelines; hands-on experience of using a Web interface in comparison with the standard interface (via commands and search strings). A good example is NlightN (http://www.nlightn.com), a recently developed system using a Web interface to link to a composite of hundreds of databases, including sources as diverse as Medline on the one hand and Lycos (a search engine) on the other (Notess 1995b, p. 88).

(3.3) Studying of and comparing a wide range of currently available search engines such as Lycos, Yahoo, WebCrawler, AltWeb and Harvest (terminology; different categories of search engines; typical search features; level of friendliness and ease of use; criteria for evaluation; relevancy and precision ratio; coverage of home page sites included in a search; search formats, for example forms). A number of comparative articles have appeared since December 1995, for example Courtois et al. (1995), Kimmel (1996). Jian Liu’s Understanding WWW Search Tools (1996) is probably the best directory to search engines and is regularly updated through http://www.indiana.edu/~librcsd/search/.

Practical:
It is essential to activate an extensive hands-on programme to enable students to grasp fully the extent of difficulties involved when using search engines. An appreciation for the extent of these ‘robot-generated databases’ should be developed, especially problems which occur because of the high retrieval and low precision ratios. Searches with different difficulty levels and different types of search queries should be arranged. The importance of search result evaluation should be demonstrated.

The above three instructional objectives form together a solid basis for understanding the complexities of searching online via the Internet. Depending on the structure, overall goal and student level of this module, further objectives could be formulated to teach students more advanced topics, taking them a step further than the traditional parameters of online searching. Examples are Intranet features and how an information service could utilise these to improve services to their clients; network security (building firewalls to protect databases from intruders); network management; task analysis of the Web master or Web coordinator; outsourcing Internet/WWW maintenance; Java applications; and so forth. Deliberation of these advanced topics falls outside this discussion.

However, now that the Web has become almost a standard to publish in electronic format, it is obvious that the information specialist should be well-trained in using this environment to communicate with and render services to clients. This implies demonstrating a working knowledge of such important features as HTML, home page design and rendering information services via the Web. Instructional objective 4 is an effort to convey the essentials of ‘Web admin’.

Instruction objective 4
A student can effectively apply the facilities of a Web site to support information services to clients.

Suggested course content and practical experience
(4.1.) Basics of HTML, as well as HTML editors and converters (studying the basic HTML tags as well as different editors (HTML conversion software) and converters which are available to assist in the markup process).

Practical:
Hands-on sessions are essential to grasp the technicalities of using HTML effectively. Practical to convert plain text should be arranged. One or more HTML editors should be available in the training laboratory (see, for example Carl’s list of authoring tools (1995)).

(4.2) Home page design principles (hypertext as a method to organise large full-text databases; principles of planning a hypertext environment (for example story boarding); the home page as departure point for a whole range of information services; criteria for designing good home page files).

Practical:
Time should be allocated to the methods to design hypertext documents, if the theory and
practice of hypertext was not covered in any of the other Information Science courses. Planning approaches such as storyboarding should be experienced. Group work can be strongly recommended when working on specific projects. Case studies is another approach that works well in letting students design a home page and subsequent links for a specific environment. Prototypes are then developed under the guidance of the tutor, for example to render a current awareness service via the Web, create promotional material to market the information service, develop study guides for client guidance, handle an electronic document delivery service, and so forth.

4. Conclusion

The impression one gets from articles on teaching Internet skills is that one should teach this extensive area as a large entity in its own right. This might be a good idea from a didactic point of view, because such an approach keeps the potential learning content together, with the result that a student or professional (during in-service courses) is exposed to the theory and practice in a neatly packed unit. The disadvantage of this approach is that the learner might see the Internet as something separate from the other professional information handling skills. Gone is the holistic view of information and information sources, with the result that the same professional might manage and use the Internet separately from conventional electronic and paper-based information.

This paper's approach is different from the view discussed in the previous paragraph. Instead, it represents an effort to isolate entities relevant to a specific topic — online searching in this case — and to integrate these with established methods. This same approach has been followed when discussing teaching online searching procedures: that is, to select and integrate relevant Internet procedures with the training programme. The instructional objectives, learning content and practicals were provided as examples, and therefore as a frame of reference for further development by a specific training unit.

Professor Pieter A. van Brakel
Department of Information Studies
Rand Afrikaans University
PO Box 524
Auckland Park 2006
Johannesburg
South Africa
Fax: +27 11 489 2797
E-mail: pavb@info.rau.ac.za

References


NOTICE

REPRODUCTION BASIS

This document is covered by a signed “Reproduction Release (Blanket)” form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a “Specific Document” Release form.

This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either “Specific Document” or “Blanket”).