Internet search and the searchers: what they know, what they really know, how they do it and how we can help them do it better

NICK PAVLOVSKI
Reference Librarian
Swinburne University of Technology

Abstract
Information seeking has changed for the X and Y Generations, whom primarily use Internet search engines rather than more traditional resources. What they learn about search engines and how they work is often poor, but through previous successes doing simple searches, they have developed false confidence in their tool to the exclusion of all others. When these students can’t retrieve the information they need, they seek help in academic libraries where the academic librarians must adopt new ways to help and educate them in information literacy. How these users think and search is explained and then corresponding “unlearning” and solutions are offered to meet their particular educational requirements and technological skills.
Internet search

The Internet has been as big a revolution for libraries as the arrival of the desktop computer. It has allowed the freer exchange of information, whether of MARC records, journal articles from a remote database or humble text-only e-mails. It has also brought about the requirement for many librarians to teach users information-seeking and information literacy skills. School libraries and academic libraries in particular involve large numbers of users seeking information for specific purposes, usually with deadlines and other conditions adding pressure to their searching. Using the Internet has become a major part of the information-seeking process for these users. Librarians may be asked to provide help when these users fail to find what they need. This paper seeks to present information gathered that demonstrates the techniques and processes employed by older teen and adult users when they search in order to improve librarians’ understanding of these particular library users and to outline the corresponding redesigning of information literacy tuition. Brown, Murphy & Nanny (2003) have presented findings and made recommendations for students attending American colleges; this paper recognises their efforts but follows a different tangent by examining the Australian condition and exploring the psychology of Generations X and Y.

In most academic libraries, users can access information electronically. This information may come from databases subscribed to by the institution; from free-access databases or repositories; from web sites etc. One method of seeking information is to use an Internet search engine, with two of the most popular being Google and Yahoo. The literature is replete with anecdotal and formal reporting of user success (or otherwise) when using them. In particular, there is evidence that many students will use an Internet search engine first when seeking information, rather than databases or traditional library materials such as books and other print resources. Brenda Bailey-Hainer (2005) reports a finding from the Pew Internet & American Life Project study of 2002, that 94 percent of US teenagers with Internet access have used the Internet for school research, with 71 percent having used the Internet as their major information source for their most recent research (at that time). These statistics can easily be transposed to Australia, due to our swift uptake of new technologies and with greater penetration of the population than many other major countries. During
Employment at a Victorian secondary school in 2004, I repeatedly witnessed great emphasis being placed on using the Internet to research assignments by both teachers and fellow students, with one student actively stating that she could not see the need for books in the library at all! My observation is supported by similar evidence from Wade (2006). At tertiary level, this observation is confirmed and reinforced by a TAFE librarian’s brief questionnaire response: “Gen X student confident with Google and Web – and see no need to look beyond that” (Fafeita 2006, p.153). The trend is perhaps even more advanced overseas, where a primary teacher in the US commented that: “(Google) doesn’t teach (students) the basic skills they need, because they’re getting quicker access in a shorter amount of time. They hardly go to the library or encyclopaedia anymore” (Graham 2003). Information seeking has changed as the X and Y Generations have grown up – US Millennials (the next Generation) have been born into this period of change and won’t know the difference.

**Peer to peer**

How are students learning to use and search the Internet? I propose that they develop some of their knowledge from school teachers, some from parents and some from their friends. The quality of the knowledge they develop will depend on the proportions of value assigned to these three sources. Therefore, their success in finding authoritative or appropriate information will depend largely on two resulting factors – the development of that knowledge and the quality of that knowledge.

School teachers and parents, whilst being preferable sources of information about using and searching the Internet, will not get as many opportunities to help and instruct the younger student in comparison to those of the student’s friends and peers. There is a great deal of interaction and information exchange (don’t forget peer pressure!) as teenagers develop. There is also the accepted teenage phenomenon of rebellion or rejection of authority figures. Younger students will give primacy to what information they learn or hear from their friends and peers, including about search engine (SE) selection and instruction. This is summed up succinctly by Vine (2001): “Most people have learned about web searching the way they learned about sex – from their friends”. These students’ understanding of what they are doing is crude and possibly misinformed and their searches, while occasionally successful, are equally crude.
The depth of students’ searches using the Internet is worth closer examination as it can clearly demonstrate their beliefs and understanding of what a SE does and how it works. This, in turn, influences how they use them. A further study by the Pew Internet & American Life Project in 2005 found that the adults interviewed about SE use were mostly doing simple searches that resulted in a satisfactory answer, which causes them to develop false confidence in SEs and their own search abilities (Sherman 2005). This false confidence is dangerous for a number of reasons:

1. it builds up reliance on using SEs rather than other sources of information (already mentioned above);
2. a single answer obtained for a simple search may not be suitable for study or academic research, where a range of answers may be required;
3. searches during academic research will be more complicated, often requiring use of multiple terms and phrases as well as Boolean Operators - these may not be learned or employed by the searcher.

The danger of this false confidence gained from basic SE search success is further compounded by the fact that few users actually understand how SEs work (Graham & Metaxas 2003; p 73). Information Literacy carried out by librarians should therefore involve tackling misperceptions and this false confidence through some “unlearning” or “deprogramming” in order for the retention and application of correct information to be most effective.

The foundations of the false confidence in SEs

Librarians can assess a student’s grasp of searching skills when the student asks for help. Through the reference interview, the librarian should be able to find out what the student has tried, with what they have tried and the depth of their searches. The depth of their searching is, I believe, very strongly influenced by the false confidence they have developed. This false confidence’s strength is the result of the following three factors.

Firstly, there is the increasingly held belief by many that “it’s all available on the web” (Fafeita 2006, p. 151) or the perception that “It’s gotten to the point where people think that if it’s not in Google, it doesn’t exist” (Graham 2003). Success with simple searches
Internet search and the searchers

of only one or two keywords for a number of different subjects, scholarly or otherwise, will be compounded with influential feedback from peers and other societal factors such as advertising, media and teacher direction to include online resources at primary or secondary school levels. Other methods and resources are thus ignored.

Secondly, since many searchers tend to use only one or two SEs for their searches (the second often only being employed when the first fails them), the successes they enjoy when they do get results for their searches builds up confidence in that particular SE or SEs. When a scholarly search must be done, they automatically choose those same SEs because of their previous successes, even when other tools may be required. The Pew Internet Search Engine Users Report (Fallows 2005, p. 3) found that 44 percent use one SE regularly, and that a further 48 percent use just two or three.

SEs are not simple tools, though: many can search for text, images, and video; some use Boolean Operators and many offer Advanced Searching. Sadly, few searchers use these latter features. Gord Hotchkiss (2004 a) found that 26 percent of subjects never use Advanced Search and a further 48 percent use Advanced Search rarely, for a total of 74 percent of searchers surveyed. Searchers carry out simple searches, develop a strong preference for one or two SEs and rarely even use the special features available that could improve their searches. They are choosing one tool over all others but don’t even understand its full capabilities. Their trust is blind.

Thirdly, SE users will stop searching too quickly – either because they have found an answer, or because after reading one or two pages of results, they haven’t found anything appropriate. This is borne out by two studies: one from Wellesley University, where students would stop searching once they had found one answer – they would not examine multiple sources (Graham & Metaxas 2003, cited in Sherman 2005); and one by SE marketing company Enquiro, where research noted that if no relevant answer was found on the first page of results obtained, only 5 of 24 searchers would go to the second page of results while the rest started a new search; and that if a relevant answer was found on the first page of results, only one searcher of the 24 would look beyond it to also check the second and third pages (Hotchkiss 2004 b). SEs have improved over last few years, with competition between SEs being a key contributor to improvements in design, functionality and algorithms. Inversely, the iProspect Search Engine User Behaviour Study (2006) found that users have decreasingly searched
Internet search and the searchers

beyond the third page of results over the last six years (p. 10). There are theories as to why a search may be abandoned after one or two pages of negative results, which are discussed in the literature. Regardless of theories, the facts are that a user will continue using a SE even if they get no results, demonstrating that users still believe the SE is the best tool for searching with. Their false confidence may be challenged by poor or negative results but the studies indicate it is not easily broken.

This begins to form a picture of users relying on a select few online resources, and not even probing too deeply in them if they can’t find a suitable answer. They try to make their chosen tool do the work of a range of different tools, but these other tools must be handled differently – the SE only works in one way. They have possibly forgotten how to use encyclopaedias and books in printed format. Amazingly, negative results aren’t ascribed to the tool, but elsewhere. If their peers can’t help them, then they begin to look to others for help rather than reassess their methods.

The students that are seeking help are thus used to using a SE for research and usually only one or two popular ones, such as Google or Yahoo. They rarely use other repositories of information because of their past successes using their SE. What actual training or awareness of how the SE works (or more importantly, how to make it work) will have been garnered from a range of sources, with peer or social marketing playing an important role in which sources were paid most heed. They come into the library seeking help because they have not had the expected successes they are accustomed to and they are further having to perform under significant pressure – an assignment in the first semester of tertiary study, or perhaps their lecturer has banned use of non-scholarly resources. They eventually make contact with the reference librarian, who has a greater knowledge and appreciation of the many tools available with which to locate information. The task before them however, should be threefold: to help the student; teach better information seeking skills and also to reassert use of other resources for research.

*Librarians have the knowledge. The users often don’t want it.*

During the reference interview the librarian is able to gain insight into what the student really wants, what the student has tried so far and other factors such as timeliness.
Through the use of the wide variety of resources at their disposal, they are able to deal with the first part of the threefold task, helping the student. The other two parts of the threefold task may be attempted simultaneously but unless the student later engages with the Librarian, the Librarian may not ever know if their tuition was successful or even retained. Rita Vine’s challenging and provocative article “Real people don’t do Boolean” (2001) makes strong statements against the value and effort of many current Information Literacy offerings (especially those offered by academic libraries), the boldest being:

Real people want advice, not technique. They want us, the information professionals, to simplify their lives. They want us to help them identify a few really great resources quickly and help them avoid false drops and crass promotions.

Certainly, the student seeking help wants the librarian to find one or more resources. The librarian, through the reference interview, may have established that the answer or research is needed for an assignment due in the immediate future, in which case a speedy result would be eagerly sought. The librarian may be able to provide some advice and even correct some poor search techniques but it may well go “in one ear and straight out the other”. Vine may in fact be the little boy who said, “The Emperor has no clothes on” when actual retention of information literacy tuition by students is examined. While we may be patting ourselves on our backs as the student walks away at the completion of what seems to be a successful reference session, our efforts may have been for naught. One evening early in 2006 I patiently and slowly showed three First Year Sociology students how to use our OPAC to see if we held a certain journal; how to follow the link in the OPAC to the database where it was held; how to log in to the database and then search by author until we successfully found the article they were seeking. They were extremely grateful and went away to a nearby computer to search for other articles. Two minutes later, they called me over to help them – they’d got stuck at the first stage, using the OPAC to see if we held a journal. Vine’s words rang true in this instance.

Promoting and teaching information literacy is a significant part of a library’s role in society. Library associations such as ALIA, ALA and CILIP all have portions of their websites dealing with information literacy and Lifelong Learning. Authors such as Alan
Bundy have written extensively about it and plenty of research has been carried out concerning methods, content and measuring success of instruction. It is comparatively easier to gauge the success of formal instruction of information or library skills to a class than assess the success of instruction imparted during a reference session where instruction is imparted informally and outside a traditional lesson plan. Institutions such as Swinburne University of Technology and Holmesglen College of TAFE offer prizes of iPods and vouchers to try and attract students to participate in their library surveys, as too few participate without any incentive. Patting oneself on the back after helping a student is premature – an expression of appreciation from a student after the fact, although rare, could be considered a more concrete sign of success. Unless the student is explicit in stating what they’ve learned, the value of an expression of appreciation is limited. Soliciting students to give valuable feedback is difficult.

A quick, positive result is often sought by someone seeking a librarian’s help. This has already been discussed above by Vine. The reference interview may establish that a student has a rapidly approaching deadline, but a speedy reference session leading to a speedy result is often the norm regardless. Factors such as the demonstrated short attention span of Generation Y (Jennings 2006); the developed world being afflicted by information overload; impatience bred from technology’s increasing speeds of delivery may all have roles to play in the demand for faster delivery. There is a proposal of the existence of a type of mindset that incorporates facets of all these factors. Trendwatching.com have identified a type of information consumer dubbed the “Ready-to-know” consumer. The “Ready-to-know” consumer’s requirements for information is that it is:

1. accessible via both wireless and wired;
2. on their own terms;
3. pull, not push (Trendwatching.com 2005).

Their everyday needs may be met by the many technologies and services they employ, but when they cannot obtain information as they have previously been accustomed to, then they have a serious problem. These students are perhaps those for whom effective information literacy tuition would be most successful, as they are keen searchers and users of information. They have extensive hands-on knowledge of technology. Getting them to commit to surrendering enough time for a formal class is
the problem. Some other delivery style or key tip they can race away with is required – or at least an effectively baited hook with which to entice them to attend a future class. A traditional-style class simply won’t be attended.

Dr Murray Banks, while discussing psychiatry, lamented: “There are none so blind as those who will not see. Some people just don’t want to see” (1961). Helping students is one thing – trying to then educate them after giving them an answer or the assistance they needed is another. Tuition or classes may be actively or passively refused because the student honestly believes that they don’t need them. It is a brave individual who will readily profess to their ignorance. Students with their extensive hands-on knowledge of the latest technologies may be resistant to the formal classes run by academic librarians because of their own self-confidence – Fafeita reports a remark by one TAFE librarian, that: “You can always tell a teacher, but you can’t tell them much” – teachers typically are awful students!” (p. 152). In an earlier section, I highlighted the fact that many younger students have very poor and limited understandings of what a SE does and how it works. For them, it works and that’s all they need to know about it. Fox (2005) examined the reasons underlying client vagueness displayed during reference interviews and identified Fear as a cause for lack of understanding:

Fear of appearing naïve or dumb….Don’t make the assumption that your client understands the topic. You could easily be providing your client with information that is beyond their level of understanding. If the client walks away not admitting to you that they will not be able to comprehend or be able to use the information provided then you have lost that customer.

It is easy to transpose this cause and apply it to understanding resistance to information literacy tuition, either formal or informal. It also begins to point to the myriad of other causes that have been identified that prevent students from learning – cultural and racial differences, learning styles, etc. Peer marketing, web pages and online tests where a user remains anonymous partially removes this fear, yet even so, juicy carrots may still have to be dangled as use of the stick may only result in the blame for ignorance being shifted elsewhere.
Reviewing what has been presented above paints a picture of technologically-capable but arrogant students seeking help from librarians who can assist them reach their information goal but fail to impart much long-term knowledge, much less entice them to attend formal tuition which is actually to their benefit. Failing to educate is mercifully not attributable to only the librarian and their methods, as should now be evident. Still, the librarian is limited to trying to plant one or two seeds that may hopefully sprout before the student zooms away to their next appointment. This means that for many of us, information literacy will have to change again, just as our students have changed.

What we teach, what we use and how we teach it must embrace this generation with its particular attitudes, beliefs and learning styles. Existing information literacy and “library literacy” materials (UMUC 2005) will still be relevant, as there will always be mature students and those who prefer formal tuition styles. It is the repertoire of tuition types that needs to be expanded to accommodate these generation of students and greater flexibility in handling those tuition types.

A formal class offered by an academic library or an extended informal session with a student should try to include:

1. informing them of how a SE really works as well as it limitations and failings;
2. demonstrating a SE they may be unfamiliar with but that more closely matches their criteria, such as one that clusters or has some other valuable trait;
3. other ways to search.

Demonstrating Boolean Operators or Advanced Search features can wait – students should learn to crawl before they learn to walk.

**The foundations of revised information literacy**

A search engine can retrieve plenty of results, depending on the search attempted. The better a search engine’s algorithm and the more popular it is, then the further it can search and bigger a database it can build. A search by Google can yield millions of results (Wade 2006). While most students may see this as a smorgasbord of opportunity, a few will correctly identify this as information overload and they have to
spend time carefully picking a good result to suit their needs. What nearly all students don’t know is that SEs only cover a small fraction of the World Wide Web, even mighty Google (Long 2004). Information professionals are aware of this and of the vast amounts of information that cannot be retrieved by SEs, known by names such as “the Deep Web or “the Invisible Web” (Lewandowski & Mayr 2006). Regardless of what figure is proposed as representing the ratio of visible to invisible Web resources, it is only a miniscule fraction. Using 600 billion pages as an estimate (Webb 2004) - if a single word search in Google finds approximately 3 billion results, then Google only covers 0.5 percent of what is potentially available. Showing students a search that finds a billion or more results, then telling them that this great result is only 0.5 percent of what possibly exists is in my experience a very sobering “reality check” for those who may have no idea or have firm false confidence in their favourite SE. This is a great opportunity for a librarian to then promote other ways of information seeking and retrieval, once this “unlearning” has been done.

Until very recently, SEs have been in the ascendant whilst human-indexed Internet information has seemingly been in decline. Since students want information quickly and neatly, a human-indexed gateway or directory that requires consideration and time does not appeal to them. Human indexing is slowly coming back into vogue thanks to sites such as Flickr, where users can create their own terms when cataloguing or share those used by others. This is a golden opportunity for librarians to re-emphasise the value of web directories and gateways. Information professional Mary Ellen Bates (2005) wrote a concise description of how she carried out a internet search about VOIP over two hours, in which only 15 minutes were spent using SEs. Through use of web directories, web sites hand-picked by her as being useful resources, databases and a couple of SEs, she carried out a thorough search that paralleled a librarian using print dictionaries, fact books and industry journals to do the same. Many academic libraries have web gateways or directories of their own for student and staff use. A good information literacy class about the Internet should include demonstrating these gateways. Rather than emphasising the time saved using them, Bates’ example could be taught as instead demonstrating the danger of relying exclusively on a SE or two. Alternatively, suggest students only use a SE for 15 minutes when they are searching, and if nothing useful is found, to go to a library gateway or some other resource.
Through these alternatives reliance on SEs doesn’t have to be broken or denigrated, just placed back to equal footing with other search strategies and resources that should provide a range of options for students to employ in their academic searches—as academic information is well represented in the Invisible Web (Lewandowski & Mayr 2006). Without other methods to find information, students will be in increasing danger as they try to progress through their studies and then seek employment, as many industries have industry-specific tools that are similar to library-subscribed databases and print sources. Employers in Australia generally want employees whom they don’t have to train or re-train too much, especially in what should be basic skills.

Academic library gateways need to be constantly reviewed and redesigned. The parallel drawn in Bates’ article of using appropriate online resources just as a librarian in previous times used the most appropriate print resources reinforces the validity of library gateways and that they must be kept current in content and layout, if only for changing aesthetic reasons. Price (2003) states:

> Little effort as a profession has been placed on local collection development of free Web-based resources and information professionals “learning” good answer sources before (author’s emphasis) they need them….In fact, doing this allows the library the chance to maximise the free material on the Web and use directories of Web resources as collection development tools….So share some Google tricks with them, but then start demonstrating what you and your library can do – with blazing speed and clear authority – that Google can’t do well or at all.

There are plenty of SEs freely available for anyone to use. All have strengths and weaknesses. Since students rely on one or two well-branded and popular SEs, they may be missing out on using a smaller SE that may well target what they require more effectively – for example, I find Exalead (http://www.exalead.com), a French SE, excellent for finding information from European sources, rather than Google which is over-represented with American sources. New information literacy tuition should raise awareness of these SEs and encourage trialling them. Also, some students would greatly benefit from using visual interface search engines. Through graphic displays of smart organisation of search results into groups or clusters that help show their
relevance to the search terms, better selection and subject narrowing can be made (Luther, Kelly & Beagle 2005).

While visual interface SEs are still few in number and only recent arrivals to the scene, already some libraries are seriously examining these visual interfaces and their clustering technology, seeing applications that can be added to their own OPACS and commercial database interfaces which will benefit users.

The future is here

The future brings faster changes and challenges to information searching. Yahoo is working on improving their SEs and they can see the future being improved by people. Bradley Horowitz of Yahoo Labs elaborates: “Writing algorithms to understand what is in an image is extremely difficult, for example. People plus algorithms is better than algorithms alone” (Thurow 2006). With Wikis, Flickr and other repositories being developed, catalogued and moderated by people, a balance of power has been re-established between SE robots and humans, whereas until recently the robots had taken over. Librarians have plenty of opportunities to involve themselves in cataloguing, organising and tagging, setting standards and improving searches. Horowitz adds: “Subjective queries rely on domain expertise. With social search, we are handing back domain expertise to the users”. Librarians have the expertise; they have to get these positions of power. They can teach effective SE literacy – they just need to connect with their users. There opportunities are arising as the robots step back.

Both Yahoo and Microsoft have stated that they shall track SE user activity so that their SEs can suggest material they think are relevant to the searcher’s topic (Grayson 2005; Jette 2005). Microsoft’s SE labs are working on behavioural targeting, including trying to predict searcher age and gender (Thurow 2006). New search services focus on extensive personalisation, going beyond the current offerings and allowing searchers to customise both the search topics and the presentation of results (Bradley 2006). Specialist SEs have been developed to search podcasts.

Generation X and Y students have entered academia, and Millenials will soon follow. Their understanding of the tools they use may be equal to students from other generations - and as reference interview clients they may demonstrate fears and
misunderstandings common to all - but they also demonstrate lives more tightly interwoven with technology and expectations that they should be able to almost instantaneously find what they need. As they come to their University or TAFE library to seek help in finding information, librarians are able to assist and also educate them at the same time. This eduction in library literacy and information literacy skills can no longer be delivered in the ways they were – they need to be adjusted to match the students and adjusted as the online world develops and changes. Far from becoming increasingly obsolete and powerless, currently there are tremendous opportunities for librarians to seize and reassert their long-developed skills and organisational abilities, as well as give single techniques or tips to busy or short-attention-span students. Web tutorials and quizzes for those afraid of being exposed as ignorant or uneducated can be developed or improved upon. Gary Price has said: “The spoken word is now becoming as searchable as the printed word has always been” (Zetter 2005). We must raise our voices and guide our students through the many dialogues.
Internet search and the searchers

References


Internet search and the searchers


University of Maryland University College (UMUC) 2005, Information literacy and writing assessment project: tutorial for developing and evaluating assignments, Faculty
Internet search and the searchers

Resources and Services, Information and Library Services, University of Maryland University College, viewed 2 October 2006, http://www.umuc.edu/library/tutorials/information_literacy/sect2.html


**Bibliography**


Bradley, P 2006, *Your search, your way part two*, Search Engine Watch, viewed 6 October 2006,


Internet search and the searchers


University of Maryland University College (UMUC) 2005, *Information literacy and writing assessment project: tutorial for developing and evaluating assignments*, Faculty Resources and Services, Information and Library Services, University of Maryland University College, viewed 2 October 2006, http://www.umuc.edu/library/tutorials/information_literacy/sect2.html


