

An email discussion: search engines, quality, and consumer information on the internet

Fiona McLean, Health Care Information Officer, British Library
96 Euston Road, London NW1 2DB
fiona.mclean@bl.uk

Introduction

The world wide web is a vast and rapidly expanding resource. Publicly indexable webpages have increased from at least 320 million in December 1997 to about

800 million by July 1999 (1). A report in October 1998 estimated that nearly 7.5 million people, representing 16 % of UK adults, had used the Internet in the last six months. This compared with 12.4% or 5.4 million in the previous six months (2). A recent market research report estimates that the number of UK households with internet access will double in the next four years (3). There is not equal access: UK households with internet connections in 1998-99 range from 1-4% in the lowest income groups to 32% in the top income group (4).

There has been concern about the content of some health websites, and various quality criteria have been suggested for evaluating them (eg 5,6,7). It is these criteria I am evaluating for my Information Science dissertation. I need to use a search engine which will retrieve websites of all levels of quality, so quality checking gateways such as OMNI, Health on the Net, and Healthfinder (8) will not be an option.

I sent an email to three mailbase discussion lists likely to include relevant information professionals to find out what they would use (9,10,11). The question I sent was unusual for these lists as it was hypothetical and excluded the subject gateways:

'I am about to do some research about quality criteria for assessing medical/ health websites. The scenario will be an LIS professional looking for a website with information on a condition. Non technical level ie suitable for patient/ public/ consumer/ junior PAM student. I need to use a search engine to find some sites to assess, and want some idea of what YOU would do. So if you were looking for such information on condition X (and there wasn't anything on a quality checked gateway such as OMNI, Healthfinder, Health on the Net or Medical Matrix):

1. Which search engine would you use?
2. Would you just search for 'condition X' or would you add more search terms?'

Results

29 members replied, several in their own time from home. This cannot be

*"The world wide
web is a vast and
rapidly expanding
resource"*

seen as a representative sample, but does give some useful information. All quotes are from the email replies.

Question 1: Which search engine would you use?

The number in brackets after the resource name indicates number of times mentioned (some people listed more than one).

Search engines and URL

- Alta Vista (13) www.altavista.com/
including Alta Vista advanced (1)
Alta Vista was placed at the top and bottom of lists when searching several search engines (see below), but others gave it as sole source. '...being the largest search engine and usually yields good results'. 'It seems to produce good results though there is the usual headache of wading through lots of sites, though I often use domain:uk to reduce the quantity found if a UK slant is appropriate.'
- Yahoo (7) www.yahoo.com/
including Yahoo health (1)
- Northern Light (4) www.northernlight.com/
including Northern Light power/advanced search (2) '...sorts hits into folders.'
'I like Northern Light's power search capability. It allows you to limit your search to particular types of Web sites (government, academic, etc.), dates, etc.'
- Google (4) www.google.com
'I have had a lot of luck with - even with single word searches. Nice uncluttered front end too, although I have no idea who is responsible or the way it works.'
- Hotbot (3) www.hotbot.com/
- Ask Jeeves (3) www.askjeeves.com/
'...natural language searching an advantage.'
'If I got nothing I might try Ask Jeeves- because I find that 'he' sometimes succeeds where others fail.'
- FAST/ All the web (2) www.alltheweb.com/
'It has had good reviews so I feel I ought to use it, but I'm a bit slow to change my habits.' (Another person also mentioned not having checked other sites in depth yet).
- Metacrawler (1) www.go2net.com/search.html
- Oingo (1) <http://oingo.com>

*"Question 1:
Which search
engine would you
use?"*

- Infoseek/Go (1) www.go.com/

Comments:

- 'Yahoo and Hot Bot (because these two attracted the most stars* in an article about search engines & health information in: John Scott Cree and Ron LeBruin "The search for the best search engine" Library technology 3 (3) June 1998, 45-47. (This is the insert in the LAR). They checked for target home pages in the health subject area on given days in September and November 1997.'
- '...search a minimum of 3.'
- 'My favourite is HotBot, but I use AltaVista and FAST as well if I get nowhere. Perhaps for a general enquiry (i.e. not needing to restrict results by domain or file type etc.) I would choose AltaVista first.'
- 'I would use more than one search engine. I think there are more than 1 billion pages on the internet and the largest search engine (Fastsearch) only covers about 200 million. Alta Vista covers around 140-160 million. There is some overlap however a good deal of unique material can be found on different search engines.'
- 'If I were looking for sites in general, I'd go and use Yahoo, because most of the public (I guess) are familiar with Yahoo and my guess is they'd start there. Personally, I'm an Alta Vista Fan.'
- 'I suggest that my physicians categorise their inquiry into one of the following before searching:
 1. General medical information or consumer information
 2. Support groups or other resources
 3. Medical literature or evidence
 4. Someone with medical expertiseBy doing this, they can select a site or search engine that will best meet their needs.'

'I would use more than one search engine'

Order of searching multiple sites (where specified):

- Northern Light / Alta Vista
- Northern Light / HotBot / Alta Vista
- HotBot / Alta Vista and Fast; but Alta Vista first if general enquiry
- Alta Vista / Metacrawler / Google / Oingo / Yahoo
- Alta Vista / WWW Virtual Library / Yahoo Health
- Yahoo and HotBot / Ask Jeeves

Meta-search engines

'Search engines are dead. Long live meta-search engines!'

- Searches.com (1) www.searches.com
- Copernic (2) www.copernic.com 'I have the trial version, ... it is brilliant in that it searches so many other engines.' 'There is a piece of software called copernic that searches using your terms entered into the top 10 www search engines and collates the results from all of the search engines. It is an excellent way of condensing a search.' [Unfortunately, use of Copernic now requires downloading software from the internet to a PC with Windows NT or similar]
- Dogpile (2) www.dogpile.com
- Simplifynet (1) www.simplify.net/
'One physician likes Simplifynet. You input your request, and Simplifynet sends your search to a number of commercial search engines (GoTo, Yahoo, Altavista, etc.). You can then check your results on each very simply.'

Other sites mentioned

- WWW Virtual Library for Health (1) <http://vlib.org/Medicine.html>
- Medscape (1) www.medscape.com/
- Achoo (1) www.achoo.com/
- Patient UK(1) www.patient.co.uk/
- Health Education Authority (1) www.hea.org.uk
- MedLINEplus (1) www.nlm.nih.gov/medlineplus/
'My first suggestion for a lay-person would be to visit MedLINEplus.'
- Office of Rare Diseases (1) <http://rarediseases.info.nih.gov/ord/>
- DISCERN www.discern.org.uk
'..DISCERN is a project in progress to provide a tool to assess consumer health information sites.'
- Internet Detective www.sosig.ac.uk/desire/internet-detective.html

General comments

- 'I must admit that disease info isn't something I search for often, but if it were then I would probably have a handful of favourite sites (not search engines as such, but sites with good collections of documents) that I would look in. A bit like going to a specialist cheese shop instead of just to the cheese counter at Sainsbury's.'
- 'I would also direct the patron to print resources such as medical

dictionaries and encyclopedias as well as Merck (both print and online) and regular encyclopedias.'

- 'A very good and recently updated site which looks at the capabilities of various search engines is available at <http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/FindInfo.html#Outline>

This compares a selection of good search engines.'

- 'I suggest that my physicians categorise their inquiry into one of the following before searching.

1. General medical information or consumer information
2. Support groups or other resources
3. Medical literature or evidence
4. Someone with medical expertise

By doing this, they can select a site or search engine that will best meet their needs.'

- 'After searching the gateways, I find no particular search engine my best option. Unless the question is particularly Canadian, in which case there are a couple of engines particular to that instance.'

- 'My approach is to go first to our own library's compendium of sites, which is partially homemade and partially links to the gateways. This can be found at: www.library.dal.ca/kellogg/internet/internet.htm Then I go through the list of search engines at our university library site: www.library.dal.ca/general/searchtools/searchtools.html Sometimes, with particularly difficult topics that yield very little, I move pretty quickly to engines like Dogpile and Metacrawler. Most of the time, I find there are too many peripheral at best hits and the restrictions available through the various "ADVANCED" engines do little to bring things back into context.'

- 'Yahoo and Hot Bot (because these two attracted the most stars* in an article about search engines & health information in: John Scott Cree and Ron LeBruin "The search for the best search engine" Library technology 3 (3) June 1998, 45-47. (This is the insert in the LAR). They checked for target home pages in the health subject area on given days in September and November 1997.'

- '1. Simple enquiry using Altavista (or any other competent search engine); if necessary, adding words to refine search if there are lots of 'hits';

- 2. Gain an overview of the differing perceptions adopted by a range of sites that profess knowledge in the subject matter e.g. from 'establishment' sites to those preferring 'alternative medicine', observing the extent of reliance on evidence in each case;

- 3. Simple enquiry on Medscape's Whole Text Articles/News/Patient

Information;

4. The act of undertaking 1 - 3 above usually suggests links to the most currently useful global or national websites dedicated in whole or in part to the subject in hand, be they recognised organisations, or dedicated 'amateur' sites (the latter often run by people who have the condition, or who have a relative who has the condition; either way, their dedicated interest invariably adds useful perspectives and 'wisdom' to our response).

Clients are then presented with a manageable spread of information, whose depth and technicality varies according to their need and capability. This includes available and meaningful data (and including the absence of data!) on pros, cons, success rates, etc. In my experience, virtually all clients, regardless of education or perceived 'intelligence' - express strong identification at this point with one or more specific perspectives. Invariably, their reasons are sound and straightforward to them, even if they are not what would suit me or the next person! Usually, the client expresses interest in obtaining more information on a small number of specific aspects to the condition, which can be addressed by:

5. Detailed search on the specific aspects using Medline, and Medscape whole text articles;
6. Ad hoc web pages (articles, user groups, etc.) from any source on the specific aspects, using Altavista, or any preferred search engine'.

Question 2: Would you just search for 'condition X' or would you add more search terms?

Of those who replied to this question, 7 said they would search for 'condition x' only, at least initially. One would search for 'condition' and 'intervention'.

- 'It depends on what I was looking for - I would start by just searching for condition X, but would then expand it, so if I was looking for say, autism, I would also search under mental health, autism, children, autism etc.'

Search strategies to increase relevance of results if too many hits were retrieved:

- adding the search terms 'patient' or 'consumer'
- limiting results to UK (2)
- adding more terms if term has other non-medical meanings
- 'Within other sites I may find terms for this condition that are also used and try them.'

Search strategies if few relevant results retrieved:

- search with synonyms/ related terms
- check spelling in medical dictionary
- 'Yahoo/Alta Vista: '. I would use the condition itself only. Depending

"Clients are then presented with a manageable spread of information..."

on how much appeared and what I was doing it for I would limit to UK material.'

- 'I think to get some 'bad' sites as well as poor, you'd have to do a simple search on just the condition, without extra search terms, after all, not many public use extra search skills.'
- 'If nothing appeared anywhere I would go away and think about it, possibly look in a medical dictionary to check the terminology.'

General comments

- 'Surprisingly difficult to answer since your scenario specifies a less routine condition than others I guess condition X to start with. Then I'd either look to narrow it or to try out as many lay synonyms / 'hot' aspects as possible. Depending on what it is, I might have checked the NORD site before starting to 'surf'. This is not what you asked for but-fact is that by this point I'd be looking for what I could find, and I wouldn't be looking for quality filters. I'd expect to screen the findings and then discuss issues around authority, bias, currency etc. with the client - just as I would if handing out the address of a self-help group or printed information. I'd do that even if I found it through one of your "quality checked gateways." In practice, I can't seem myself working methodically through any appraisal tool (Discern, or any of the listings cropping up in The Lancet, CHiQ [13] or whatever).'
- 'You will note I reject the notion that health sites should in effect be simply categorised as "good" or "bad". There are no doubt some "bad" ones around, but are they truly "bad", or just saying things that some clinicians do not personally agree with?'
- 'I would delve deeply into the DISCERN web site for guidance: I would also consult the printed version of Internet Detective. These are both tools to help assess quality of information available on the Internet.'

Other information

- Mark Duman sent me an extract from a forthcoming book he is involved in that looks really promising: 'The POPPi Guide- Practicalities of producing patient information' (Kings Fund, early 2000).
- A useful overview site is Search Engine Watch (12).
- The Centre for Health Information Quality is doing a lot of work in this field (13).

I also received some very useful information on the topic of my research, including a list of guidelines developed by someone whose job is searching for health info for the public, and a brief, clear summary of the literature. The author of that comments 'as for assessing health sites, you are walking into a minefield!'

References

1. Lawrence, S et al (1999) Accessibility of information on the web. *Nature*, vol. 400, pages 107-109
2. BMRB International's Internet Monitor (1/10/98) U.K. home Net traffic surging. Quoted in Reuters press release at <http://news.cnet.com/news/0-1007-200-333823.html?tag=>
3. Griffiths, J (1999) Consumer internet usage; 1999 market report. Key Note Ltd, Middlesex
4. Government Statistical Service (25/11/99) Family spending in the United Kingdom 1998-99. Press release, Office for National Statistics at www.ons.gov.uk/pr_f.htm
5. Shepperd et al (99) Helping patients access high quality health information. *BMJ*, 319, 764-766 at www.bmjjournals.com/cgi/content/full/319/7212/764
6. Kim et al (99) Published criteria for evaluating health related web sites: review. *BMJ*, 318, 647-649 at www.bmjjournals.com/cgi/content/full/318/7184/647#T2
7. Murray, P et al (97) Web site reviews and evaluations. *Nursing Standard Online*, JULY 30/VOLUME 11/NUMBER 45/1997 at www.nursing-standard.co.uk/vol11-45/ol-art.htm
8. Brief information and links for these quality checked gateways is at www.bl.uk/services/stb/weblink.html
9. Lis-medical at www.mailbase.ac.uk/lists/lis-medical/
10. Info-allied-health at www.mailbase.ac.uk/lists/info-allied-health/
11. Quality-consumer-health-info at www.mailbase.ac.uk/lists/quality-consumer-health-info/
12. Search engine watch at <http://searchenginewatch.com/>
13. Centre for Health Information Quality at www.chiq.org.uk/

All urls in this article were correct at 25/11/99

Acknowledgement

I am very grateful to the members of lis-medical, info-allied-health and quality-consumer-health for their replies.