

INTERNET for HACKS

by ***MARTIN HUCKERBY***

This is an introduction to using the Internet as a journalist: how to make the most of a rich resource which could have been invented just to make our jobs easier. It is essentially an Anglo-American guide; English may not be your first language, but it can be worth knowing what is available worldwide on English-language sites — these still dominate the Net, even if other languages are growing fast.

This manual operates on the assumption that you know a bit about the Net, and have used it occasionally, so it does not include things like how to start out on the Net or how to use emails.

It is written by a British journalist, who's not a computer specialist but has used the Net a fair amount, and has listened to, talked to, and read the suggestions of, experts in the USA and the UK. Much of what follows is a digest of their advice, and the text includes links to several of their sites, which provide more detailed information.

The guide is concerned with how journalists utilise the Internet, and the individual sites listed are normally those with information or tools that the author has found useful for journalists.

If you're using the Net as a journalist, **DON'T SURF, SEARCH!**

Don't go where the waves take you; decide what you want to find and then be methodical in your search. Surfing is fun, but casual browsing can waste vast amounts of time.

When you first try out the Internet, it makes sense to explore, to wander through a host of sites, to get an idea of what is out there. But once you are working on the Net, avoid distractions.

You may not need to be so precise when you get the results of a search. Scan these fast. It's hard to do a search without getting lots of irrelevant information. Even if you miss something important on a quick scan, you'll probably pick it up through a link from another site.

A basic rule: don't believe everything you read on the Net. There's a cartoon showing a large dog sitting at a computer, typing away merrily — with a thought bubble from its head saying:

"They don't know I'm a dog."

When you are using the Net, or receiving an email, you have no certain way of knowing with whom you are dealing. If you are collecting information, or conducting interviews by email, and you have even the slightest doubt about the authenticity of the people involved, then check — try to telephone them (using a phone number obtained from a different source.)

And if you doubt the Internet is relevant to you, remember that in 1998 — aeons ago in Internet terms — 98% of American journalists had access to the Internet. The Internet was even then a leading source of information in 92% of American newsrooms.

Internet for Hacks was originally produced in November 1999. It soon grew out of date and a second version was completed in October 2000, but since then, as dot.com companies crashed, services have been merged, closed or curtailed. So while the structure of the manual has changed little, much of the detail (in March 2002) is different. By next year, doubtless it will all again need updating — that's how quickly the Internet changes.

NOTE ABOUT THE TITLE: Hack is a rather derogatory term for a journalist. It means a drudge, an uninspired writer, someone who will turn out anything. Journalists in Britain often use the term about themselves, if only to get in the insult before others use it against them.

INDEX

• Searching	Page 3
• Search Engines	Page 3
• How to Search	Page 5
• Search Engines - Specific	Page 6
• The Deep Web	Page 7
• Web Site Evaluation	Page 8
• Mailing Lists and Discussion Groups	Page 9
• Journalism Resources	Page 10
• Reference Tools	Page 10
• Definitions	Page 11
• Final Thoughts	Page 12

SEARCHING

THE BASICS:

Learn how to use search tools:

There are essentially two different approaches: using lists (a classification or directory system) or links (searching not just for sites, but for individual documents which could have the information you want).

Yahoo likens a directory to a book's table of contents, a search engine to a book's index.

Plan a search:

If you know the area of information relevant to your question, use the directory in a service like Yahoo, www.yahoo.com — search a subject area for good sources, rather than searching for all references to a particular (key) word, which can produce overwhelming amounts of data.

If you are trying to find a specific fact, or name, use something like Google, www.google.com

Use logic/common sense about where to go.

Evaluate the sources.

Learn when to give up.

Journalists always seem to rush at any task given them, without much planning, and use of the Internet may be little different. At some point, find time to study what a particular search tool can do for you. It's probably sensible to concentrate on using just one or two, so you become familiar with their capabilities and with the commands to make them operate quickly and efficiently.

Look for 'text only' buttons, to speed up the loading of information sites.

If you are not getting the results you want, ask yourself if you should choose other keywords, or a different search tool.

SEARCH ENGINES

If you want to stick with only one search engine, use Google, www.google.com. It's bigger, better and faster than almost all the competition. But if you want to understand the many possibilities for searching, here are some of the most useful:

Directories are not strictly search engines: they are rather like encyclopedias, cataloguing information which is gathered and selected by humans (of varying levels of expertise), and they allow you to search for particular subjects. Apart from Yahoo, see

dmoz.org, the Open Directory, with over 3m sites, listed by 46,000 voluntary editors, and

home.about.com/ About recently cut back its team of full-time guides, but it still offers an index to 50,000 subjects, with a million links and much useful advice.

Robot-Indexed Search Engines: These compile their indexes automatically, using software robots, or 'spiders', which regularly check the whole Internet for the latest information.

A current leader is the Norwegian-based engine, FAST, alltheweb.com, which aims to crawl the Web at least once a fortnight. It can present results about particular aspects of a topic grouped in different folders. www.altavista.com, once a star performer, is now rather unreliable.

Meta-Search Engines: These compile answers from a host of engines into a single set of results, usually sorted for relevancy, and take a little longer than an ordinary search. The concept sounds great; the reality is not always so impressive: many are weak at complex searches, and

can present poorer results than a single search engine. Most select only a limited number of hits from each engine, so they may not go as deep as a single search.

One of the best is www.ixquick.com, which proves quite adept at tailoring its search to the ten different engines it interrogates each time. A quick alternative is Pandia's Fast Metasearch, at www.pandia.com/metasearch/, which checks a useful mix of just four engines: FAST, Yahoo (including Google), AltaVista and the sizeable database of the new Wisenut.

www.copernic.com is a different type, with a programme you download to your computer, which can then access up to 80 engines; it sorts by relevancy, removing duplicates and dead links.

Newer Types: www.google.com (used on Yahoo) is a links engine, presenting results based on the importance of the sites, calculated by the number of links to them from other sites. It's not the best for the very latest information (it takes time for people to link to a new site) but it wins plaudits all over for delivering really relevant information. Its simple design has been much copied.

It is also good for finding companies and organisations if you don't have the exact web address — type in the name, hit the "I'm feeling lucky" button, and it usually finds the right place.

www.teoma.com, developed from a Rutgers University project in the US, is already gaining success with a sophisticated analysis of links, to focus on the most important sites.

There are plenty of more esoteric search tools being launched: a recent type seeks just one meaning of a particular word (www.oingo.com is an example). But you may want to wait for a while to discover which ones prove useful enough to win a wide public following.

Much research is going on into 'intelligent agents', software you can programme to go out and seek information for you (rather than sitting online while you try all the different sites). But the agents are still hard to programme, and it is likely to be a while before they become common.

Natural Language: www.ask.com (Ask Jeeves) answers questions phrased in ordinary English, but if you ask what a postillion is, it doesn't normally tell you directly — it simply gives you documents or sites which include that word. www.albert.com is an alternative which can be better at the job; Ixquick also allows such questioning.

Images: More than 330 million images (photos, graphics, maps, etc) are now accessible via images.google.com. A detailed US study by the Research Libraries Group put Google top in this field, but if you need more, try the Ithaki images metasearch engine, www.todalanet.com/images/ which consults such services as AltaVista Images, FAST and Ditto.

Multimedia: Go to multimedia.lycos.com for wide coverage of video, audio and image sources.

WHAT THEY COVER

At the start of 2002, Google was claiming to cover two billion pages (1.5b indexed, and links to 500m more). On top of that, it gives access to 700m Usenet postings and over 330m images. FAST was promising to reach two billion by early 2002 (up from 625m in the autumn of 2001).

Size is not everything: with meta-search engines in particular, bigger means slower. In the past, bigger also did not mean better: many search engines were full of out-of-date links. But technological improvements have brought much fresher databases, with engines like Inktomi and Google aiming for complete updating at least monthly.

Relevancy may be more important. For instance, it may be better to use a search engine designed for a particular country, rather than an international one.

ETHICS

A lot of well-known search engines are full of concealed advertising: links are put at the top of search lists because companies pay to be there. A business opening an account with Overture

(formerly GoTo) can reportedly expect guaranteed placement on every popular search engine except Google. This guide has tried to avoid the worst offenders. Most search lists these days seem to have links that are 'sponsored' (as on Yahoo), 'featured' or otherwise specially promoted — even if it does not sound like it, presume most of these links have been paid for.

It can be a particular problem for metasearch engines: if they only take the top listings from popular search engines, they sometimes get lists where more than half are paid links. vivisimo.com can be good at keeping out these commercial links.

Remember that online operations do not necessarily have the standards you might expect from an established newspaper or broadcasting company. The new head of editorial content at AOL UK said it was his job to get readers to go to the areas promoted by the company's paying partners; the boss of LookSmart said: "We cannot afford to have ideological debates any more."

MORE INFORMATION

www.searchenginewatch.com, Danny Sullivan's comprehensive site is probably the best around, but if you want straightforward advice, www.pandia.com, run from Norway by Per and Susanne Koch, is a fine place to start. There are good clear tutorials from New York University on library.albany.edu/internet/index.html

HOW TO SEARCH (using Ixquick)

This type of search is based on guessing which words will appear in the pages you want to find. Search engines offer many options for narrowing down the search, using WORDS, not SUBJECTS. (For the technically minded, they can use Boolean operators, words like AND and NOT, to create relationships between the keywords in your query.)

Search engines have their own specific rules: Ixquick interrogates a mix of engines, so its system allows you to try a wide variety of search techniques — these work only on engines which accommodate that particular approach. AltaVista is one with a notably good search tool.

Say you are seeking information on Arab people in Africa:

arabs and africa means you may get all the pages where the two words are somewhere in that document: many, many pages, and often there will be absolutely no connection at all between the two words in question.

arabs near africa means you can get all pages where the two words occur within ten words of each other (or perhaps in the same paragraph): many fewer pages.

"arabs in africa" means the words must appear in that exact order: very few pages — but can you be sure the documents you want will use the words in precisely that order?

You may be missing some pages which don't say 'arabs', but talk about arabia, arabian or arabians. Try truncation: write **arab***, and on some search engines it will cover those extra words; similarly **afric*** will cover african/s. But a problem with truncation is that you can get unexpected results: write **brit*** to get Britain, Britons, British etc, and you will also get Britney — and there are an awful lot of Internet references to Britney Spears. (Another valuable use of an asterisk can be as a wildcard: if you are unsure of spelling, use **lab*r** to get both labour and labor.)

Perhaps you are not interested in history; you only want material about Arabs in Africa in 1998: write **arabs and africa and 1998** (though it is possible a page written in 1998 won't actually mention the current year).

Say you want material on the whole of the 1990s: one or two search engines allow you to truncate down to just a couple of digits, so you write **arabs and africa and 199*** (the asterisk will cover the mention of any year from 1990 to 1999, plus the 1990s). If, on the other hand, you are

interested in history, you could write **arabs and africa and 18*** to get everything mentioning any year in the 19th century (though you could pick up extraneous figures starting with 18).

You can specify particular hosts, if you want to search, say, only government or education sites, or you could demand the keywords appear in the site title. You can specify a particular domain, to get, say, just South African sites, with addresses ending **.za**. But there's a downside: this would cut out sites registered overseas (eg **.com** sites) or foreign sites about South Africa.

Most searches allow you to put 'NOT' before a word. This can be really useful: if you search a British news site for stories about the city of Manchester, you can save a lot of time by specifying NOT United, and cutting out tales of the soccer team, David Beckham etc.

Rather than using AND and NOT, search engines may use plus and minus signs before words to ensure their inclusion, or exclusion. Each search engine normally has its own user's guide of search tips, available on-screen when you access the site. More sophisticated searches may be carried out in a separate section called Advanced Search.

If you use all lower-case words, they normally match everything; if you use capital letters at the start of words, they will often match only the words written exactly like that.

Once you get to a page you are seeking, you can usually find a word or phrase within it by clicking Control + F. Typing in the word will bring you to the point where it is on the page (very useful with a lengthy document). Some engines, like Google, enable you to search the whole site for the words you want.

If you are looking for an author, don't just search for Mark Twain; also try just Twain, or Twain, M (the latter is the way it will appear in library sources). Writing **Jane near Smith** can get you Jane Smith; Smith, Jane; Jane P. Smith; Jane Penelope Smith, and Jane Brown, formerly Smith.

If, after all this, you decide you just want to stick to something simple, then turn to Google and follow their search instructions, in plain English: you should still do okay.

Common words make for bad searches — choose something unusual.

If you can use a phrase you are certain will appear in the page/s you want, it is a very efficient way to search.

SEARCH ENGINES - SPECIFIC

Regional: Search engines, and directories, for different countries and continents, are listed on searchenginewatch.com/links/regional_search_engines/ and websearch.about.com/cs/regionalsearch/. For news searches, many of the big operations are overloaded with US sites, so a regional engine may be best for finding out about events elsewhere.

Individual subjects: There are many varieties, on subjects like science, women, and the law. An example is www.hon.ch/MedHunt, the Swiss-based Health on the Net foundation, with both human appraisal of health sites and robot collection of information.

Guides to news organisations: News Directory, www.newsd.com, offers 8,500 web sites for newspapers, magazines and broadcast media, all with English-language content and many with accessible archives. Try also www.newslink.org, the *American Journalism Review* site; on many countries, though, their listings are skimpy — better to use a country-specific search engine.

Breaking news: Rather than directly accessing agencies like Reuters and AP, use a portal such as Yahoo for current news: fullcoverage.yahoo.com/ Among many specialist news sites: www.msnbc.com/news/ is a multi-media combination of Microsoft and NBC; www.bbc.co.uk/worldservice/ lets you hear the last broadcast bulletin. You can get email alerts for big news events (try www.nytimes.com/services/email/email.html or www.cnn.com/email/) Fully customised versions, for detailed subjects, can be quite costly.

Current news search: Google's news search is still under development, but already looks a winner: try it at news.google.com. For obscurer subjects or places, try several different facilities, because sources, and the results, vary a lot. Worth a look: www.alltheweb.com/?cat=news, the newer www.rocketnews.com and, with specialist business wires, www.northernlight.com/news.html. www.daypop.com also searches weblogs — individual journals, with comment and information, which are often very idiosyncratic, but may offer useful insights, or eye-witness accounts of events. blogdex.media.mit.edu is a good survey of weblogs (covering over 15,000 sites).

People: www.teldir.com has links to online telephone directories across the world. (Go direct to www.bt.com/directory-enquiries/ for UK phone numbers.) people.yahoo.com has a useful phone and email search system; www.worldemail.com offers email searches of 18 million addresses, but these can be out of date — indeed, many big email providers now bar outside access to their lists.

THE DEEP WEB

These sources, also known as the Invisible Web, contain much more information than is openly available on the Net. Some of this material is in fact becoming more visible, as search engines become more advanced. But most searches still do not normally reach:

- Information in picture and formatted files
- Contents of sites requiring you to log-on first (like many newspapers)
- Contents in frames and image maps
- Intranet pages
- Sites that bar robots from accessing them
- The latest information added to a web site
- Databases within a web site — such as many news archives
- Commercial (paid for) information services
- Library catalogues and non-web resources

www.invisible-web.net is based on the recent book by Chris Sherman (of Search Engine Watch) and Gary Price (librarian at George Washington University) and has a good directory of sites. See also Price's own Direct Search site: gwis2.circ.gwu.edu/~gprice/direct.htm.

Operations which will dig down for you include www.completeplanet.com, searching 103,000 databases and speciality search engines, and www.invisibleweb.com, with a directory of 10,000 specialised databases.

More specific approaches include:

Information recently added: Update agents track changes. www.spyonit.com is an easy-to-use popular site; www.karnak.com, with on- and offline searches, storage of results online, and automatic email alerts of new findings, aims at serious researchers — ones with money to spend.

Commercial information services: Services like Dialog, www.dialog.com, and Lexis-Nexis, www.nexis.com, provide loads of data, newspaper archives etc, and usually charge plenty to do so.

But they may permit free searching, so even if you do not want to pay to read a whole article, you can quickly discover what material is available (and possibly find the article another way for free).

Library catalogues: You may not be able to get the actual information, but you can find out if a particular book is available. Most libraries can be accessed, but it may be an old-style telnet connection to a simple catalogue (the mouse will probably prove useless — you may have to type in commands). This can tell you if a copy of a book or journal is available at a library nearby.

To check what books exist, try www.amazon.com for what is in print, and library catalogues for older works.

WEB SITE EVALUATION

You've found a web site. Now you need to work out whether it is worth utilising.

The basic premise of web site evaluation is simple: use your common sense, and normal journalistic caution, about any source you don't know.

Steve Miller, assistant to the technology editor of *The New York Times*, provided a simple scale for determining the integrity of web sites:

1) Government information: It may not be correct, but it is official — you can quote a government source with a clear idea of what you are getting. (Yahoo's government page is one place to find such sites; the UN (www.un.org) has links to the missions of member countries, and thence usually to their government home pages.)

2) Universities: Most studies by recognised experts are still reviewed by their peers, so the information is likely to be good quality. You now get a lot of doctoral and masters' theses; the data here may be less valid, but there are normally still links or references to supporting data.

3) Special interest groups: Often such non-governmental organisations and pressure groups are pushing a particular line, but if they are recognised bodies, you, and your readers, have some idea of what is being provided — it could be Amnesty International, Sydney Chamber of Commerce, or whatever. Companies and commercial sites could be regarded similarly.

4) Everyone else: personal sites/hobbies/obsessions etc. If it looks as if the person responsible is an expert, try to check them out independently.

Web Site Evaluation Checklist

(Based on the system developed by the Poynter Institute in the USA)

AUTHORITY: Is this a recognised expert? A body with a known reputation?

AFFILIATION: Who is it connected with? A university? Another reputable body?

ACCURACY: If you spot mistakes while reading the site, then start worrying.

APPEARANCE: Is the site carefully put together? A lot of reliable sites are old-fashioned looking, rather than modern or flashy, but a sloppy or amateur-looking production may indicate the site is the work of an individual rather than the large operation it purports to be.

INTENT: Why does the site exist? Does it do the job it claims to be doing?

CURRENCY: Is it up-to-date? Look for recent dates, or information you know to be new.

RECOMMENDATIONS: Is it recommended by other people or organisations, by reliable experts, by people you know?

COMPLETENESS: Has it done a thorough job in covering a subject or issue?

COMPREHENSIBILITY: Does it make sense? (But bear in mind that people writing in what for them is a foreign language may be less clear than native speakers.)

OBJECTIVITY: Are there signs of bias?

CREDIBILITY: A simple test: do you believe it? Does common sense tell you it is true?

Do not feel you must use all these criteria every time you check out a site, but it is worth being aware of the pitfalls.

Remember also: it is vital that the information you get from a web site is attributed to the source (just as in other journalism).

Be aware that there can be serious attempts to mislead you: during the conflict in the former Yugoslavia, the different sides produced web sites masquerading as what they were not. And look at www.martinlutherking.org; it might seem neutral at first, but it is full of right wing ravings.

Who runs a particular site? For the USA, use www.networksolutions.com/cgi-bin/whois/whois to see to whom a site is registered. www.apnic.net (for the Asia-Pacific region), www.arin.net (for the USA) and www.ripe.net (for Europe) have Who Is/Search facilities which should lead to a phone number for the administrator of a web site.

MAILING LISTS AND DISCUSSION GROUPS

A British IT journalist reckons what you get from a mailing list has the same value as what you overhear in a bar; what you get from a news group might be a pretty drunken bar.

MAILING LISTS, LISTSERVS, DISCUSSION LISTS, EMAIL FORUMS

You sign up and receive all messages posted to the list via email; some have the messages screened by the owner, or the organisation running the list. There are one-way lists which just send out announcements; some are private or on a limited subscriber basis. Those taking part may be experts in some field; more likely they are simply enthusiasts.

www.liszt.com is a searchable directory of over 100,000 lists; tile.net/lists/ is an alternative. paml.net/ is a catalogue of nearly 7,000 publicly mailing lists.

Mailing lists often have digests stored on the Internet, so you can see what they are like before signing up. If they have a section marked FAQs (frequently asked questions), do look at these before you start launching inquiries of your own.

(You may want to use a separate email address of your own before signing up to a busy list — you might receive an awful lot of messages.)

You may be able to get a list of subscribers once you have signed up to a list — which could be very useful for contacting interesting people individually.

NEWSGROUPS, USEGROUPS, DISCUSSION GROUPS

These post messages, from people with a similar interest, so that you can read them via the Internet. Look at a newsgroup over a period, not only to see what people are talking about, but also to measure the quality of the material.

Use groups.google.com to search for news group messages by content (Google has now taken over Deja.com and offers 700m+ Usenet postings). tile.net also accesses newsgroups. You need a news server to make postings to a newsgroup; www.newsreaders.com has info.

Do read the instructions carefully, otherwise you may pester people with questions asked many times before, or send individual queries to all and sundry. (If subscribing, consider using a pseudonym and a separate email address; you can get lots of mail and SPAM — unwanted ads.)

Lists and groups can be a good way to reach individual people caught up in causes and controversies, but remember that an awful lot of these discussions are dedicated to such subjects as whether the rock singer Elvis Presley is still alive — or, indeed, much less intelligent matters.

JOURNALISM RESOURCES

www.journalismnet.com is a comprehensive site from Canadian TV journalist Julian Sher, with lots of useful services, worldwide lists, links etc. Many leading sites are US-focussed, but Bill Dedman's resources for journalists, www.powerreporting.com, still has a lot of interest to all.

Working online? www.cyberjournalist.net, from MSNBC.com's technology correspondent Jonathan Dube, is bulging with useful material, including instruction on how to write for The Net, with plenty of examples. ojr.usc.edu is the University of Southern California's review of online journalism, with some wise comment, notably from columnist J.D.Lasica, but also a fair few rants.

When a really major news story breaks, such as the September 11 attacks, many US sites now swiftly produce detailed lists of Net resources for journalists working on the subject. www.journalisttoolbox.com and www.mediachannel.org were among those that did so for Sept 11.

Journalistic training: www.poynter.org The Poynter Institute, in St Petersburg, Florida, is one of the best, concentrating on mid-career training for journalists; its site offers over 200 practical tip-sheets on various topics, resources for different types of journalist, training materials, and more.

REFERENCE TOOLS

There are several compendia of research tools: refdesk.com is a popular reference site; www.onelook.com offers 300 online dictionaries.

www.itools.com/research-it can also translate web pages between common European languages, plus Japanese, using the Canadian Gist-in-Time system; in addition, Itools has the Italian Logos system (also at www.logos.it/) to translate individual words in dozens of languages, everything from Albanian to Zulu. The Babelfish system, from the French-based Systran (go to babelfish.altavista.com/) translates texts of up to 150 words at a time between English and French, German, Spanish, Italian, Portuguese, Russian, Japanese, Chinese and Korean. Remember, though, that such machine translation systems are imperfect — they provide the gist of a document, rather than the precise meaning.

www.bell-labs.com/project/tts/voices.html If you have speakers on your computer, you can type in words and it will speak the phrases back to you; it can handle such languages as German, Spanish and Mandarin (Putonghua) Chinese. There are also specialised tools for broadcasters, such as www.voa.gov/pronunciations/, which tells you how to pronounce the names of newsmakers.

www.megaconverter.com converts amounts (e.g. metric into other systems).

www.hilink.com.au/times/ provides current times around the world (though it may be simpler just to look at a site in the country concerned and check the time shown).

www.xe.com/ucc/ offers foreign currency conversions (many sites do this, but watch when lesser-known currencies were last updated).

www.travelhealth.com If you're travelling to out of the way places (as a journalist or anything else), this US site, Medicine Planet, offers much health advice. Steve Kropla's help for world travelers, www.kropla.com, offers details of phone lines/plugs etc around the world (for plugging in laptops), as well as useful links to ATM locators, search engines for internet cafes, US government travel advisories and official tourism sites. www.worldclimate.com tells you what the weather may be like when you go.

Computer help: www.whatis.com is a place to turn to if you cannot understand a computer term or process; email.about.com offers Heinz Tschabitscher's helpful guide to using email.

DEFINITIONS

There are many different names and sets of initials, and it's hard to keep up with them all. Even if you don't remember that URL means Uniform (or Universal) Resource Locator, you may need to know this is the proper title of a computer file you can access via the Internet. HTML (HyperText Markup Language) and HTTP (HyperText Transfer Protocol) could be less important to understand. Quite a few things on the Net have a variety of names — confusing, but there is, after all, no central authority to decide what to call different features. Custom and practice mean that, in time, one name usually becomes the standard.

Thus you get Directory = Catalogue = Human-Indexed Omni-Search Engine. Portals offer entry to the Net, but can also be operations like Yahoo, with a multiplicity of facilities.

The Internet is the whole network linking computers around the globe; the World Wide Web is the system of hypertext links that enables us to jump swiftly from one computer to another. But these days the names are often used interchangeably (as they are in *Internet for Hacks*).

Domains: originally American, the name supposedly defines the sort of organisation behind a web site, with the following abbreviations appearing in addresses:

- .com is for commercial companies (but .co in the United Kingdom, for instance)
- .edu is for universities (.ac in the UK, and other countries)
- .gov is for official government sources
- .mil is for military sites
- .org is for organisations
- .net for operations linked to the Internet

But the last two in particular are very vaguely defined, and in practice you have some choice of which domain name you wish to pay to use; so the domain does not provide any certain information about the site operator.

New domain names are being introduced (some with much tighter controls over who can use them), but, with the recession in the dot com world, few examples of these have yet been spotted on the electronic highways.

The following have already been launched:

- .biz, for businesses
- .coop, for cooperatives
- .info, for general use
- .name, for individuals

The following are expected to become fully operative in the course of 2002:

- .museum, for museums
- .aero, for the aviation industry
- .pro, for professionals

After the domain name in a URL comes the country designator (except in the case of the United States, which invented the system, and thus has no designator at all).

So Britain is .uk, Germany is .de, Australia is .au, China is .cn and India is .in. A new type of regional designator, .eu, is under discussion for Europe.

FINAL THOUGHTS

This brief manual aims at encouraging journalists to take advantage of the Internet — too many are still unaware of its benefits, or unable to enjoy them. A lot of the operations involved in using it are very simple, so long as you do them regularly; the Net changes a lot, so the more you use it, the easier it is to keep your knowledge up to date.

There can be a temptation, though, to rely on the Internet for everything. You don't have to use it all the time, and you don't have to communicate solely by email. There can be advantages to email interviews: people get a chance to think about your questions and to frame their answers carefully. But it may be faster, and frequently more effective, simply to pick up the phone. Or you could even go — in the West, a bit of an old-fashioned concept this — and talk to someone in person. You never know what they might tell you face to face!

The Internet is a tremendous resource, but there is an additional danger of growing too dependent upon it — you may suddenly find that the Net is inexplicably slow, that key sites are totally unavailable or overloaded by other users. So if you are working to a really tight deadline, don't rely on the Net as the sole source for the answers you need.

Even if the truth is out there, you may not find it in time!

And always remember: the person at the other end of the computer might just be that dog.

Martin Huckerby
London
March 2002

A copy of this guide is available online at www.thomsonfoundation.co.uk

Particular thanks for ideas and information incorporated in this guide go to:
in the US, Nora Paul, now director of The Institute for New Media Studies, at the University of Minnesota, and Janet Dombrowski, senior librarian at the National Geographic Society;
in the UK, Mike Holderness, technology journalist.

Complaints, corrections and suggestions should be sent to huckerby1@aol.com