

# **The Net for Journalists**

*A practical guide to the  
internet for journalists  
in developing countries*

**By Martin Huckerby**



**The Thomson Foundation**



# CONTENTS

Foreword	5
1 INTRODUCTION	6
2 THE BASICS OF SEARCH	9
3 GOOGLE IT	13
4 YAHOO! & MICROSOFT	24
5 OTHER TYPES OF SEARCH ENGINE	28
6 DIRECTORIES	33
7 SPECIALISED SEARCH	37
8 DEEPER SEARCHING	40
9 NEWS	42
10 HELP FOR JOURNALISTS	51
11 REFERENCE TOOLS	53
12 CHECKING WHAT YOU FIND	62
13 MORE HAZARDS ON THE ELECTRONIC SUPERHIGHWAY	67
14 NET CODES	69
15 DESKTOP SEARCH	72
16 BROWSERS & TOOLBARS	74
17 GROUPS & LISTS	77
18 BLOGS	81
19 MULTIMEDIA - PICTURES, AUDIO AND VIDEO	85
20 INTERNET PHONES AND EMAIL COMMUNICATION	93
21 THE FUTURE	96
22 SLOW NET ACCESS, AND HOW TO GET ROUND IT	99
23 SECURITY	103
24 PRACTICAL HELP & SOFTWARE	114
25 WHAT NOW?	117
TOP TEN TIPS	118
ADDITIONAL RESOURCES	119
<i>A. Do-It-Yourself Net Training</i>	119
<i>B. Online Training Courses</i>	123
<i>C. Lesson Plans</i>	124
<i>D. Eight Ways To Save Words And Write Better English</i>	134
<i>E. Freelancing Via The Net</i>	135
<i>The Author</i>	140

**DESIGN: STEVE HOSELITZ**

[www.sha-pr.com](http://www.sha-pr.com)

© UNESCO, 2005

All rights reserved. No part of this publication may be reproduced or modified without the prior permission of the publisher. Free copies of the publication are available on the websites of UNESCO, the Thomson Foundation and the Commonwealth Broadcasting Association.

**Disclaimer**

The author is responsible for the choice and the presentation of the facts contained in these articles and the opinion expressed therein, which are not necessarily those of UNESCO and do not commit the organisation.

The designations employed and presentation of material throughout the publication don't imply the expression of any opinion whatsoever on the part of UNESCO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the frontiers or boundaries.

All illustrations are reproduced with permission (where this was required). Google Brand Features, including web pages and logos, are trade marks or distinctive brand features of Google Technology Inc.

The author has asserted his moral rights.

## **The Communication and Information Sector of UNESCO**

The Communication and Information Sector (CI) was established in its present form in 1990. Its programmes are rooted in UNESCO's Constitution, which requires the organisation to promote the "free flow of ideas by word and image." Besides its staff at UNESCO Headquarters in Paris, the CI Sector is represented in 27 UNESCO field offices.

The Sector also provides the secretariats for two intergovernmental programmes: the International Programme for the Development of Communication (IPDC) and the Information for All Programme (IFAP).

In addition to the Regular Programme, the CI Sector implements various inter-regional, regional and national projects with extra-budgetary funding mainly in Africa, the Arab States, Asia, the Pacific, Latin America and the Caribbean. The Sector collaborates with UN agencies, bilateral development agencies, international and regional non-governmental agencies.

## **The Thomson Foundation**

The Thomson Foundation, founded in 1962, exists to foster international development through assistance to the media in emerging democracies. The Foundation works in around 40 countries each year and also hosts programmes at its editorial training centre in Cardiff, Wales, for selected journalists.

Support funding comes from a variety of national and international bodies, including the UN, and the Foundation is grateful for the backing from UNESCO that has made this book possible.

## **The Commonwealth Broadcasting Association**

The CBA works for quality broadcasting round the Commonwealth. It is a membership-based organisation with over 100 corporate members. It carries out training for them, and consultancies, and offers fellowships and bursaries.

It encourages the exchange of know-how through conferences, the magazine *Commonwealth Broadcaster*, and its website, [www.cba.org.uk](http://www.cba.org.uk). The CBA also publishes various broadcasting related books such as *Guidelines for Public Service Broadcasters*, *Parliamentary Broadcasting round the Commonwealth*, and - to appear shortly - *Guidelines for Broadcasting Regulators*.

---

## Foreword

WELCOME to this guide to the world of the Internet, aimed at journalists. It is a free resource and training handbook, designed for print and broadcast journalists and journalism students in developing countries around the globe.

The manual opens a window on a virtual world which is becoming more accessible, even to journalists in the least developed countries. The net is a multi-functional tool which plays a vital role in today's reporting and editing, and this book provides a step-by-step guide to understanding and utilising its many and varied aspects.

In today's Knowledge Society there is what the handbook's author, Martin Huckerby, calls a "feast of information." Unfortunately, in many parts of the world - especially in the least developed nations - there is no feast, only a "famine of information". It is indeed incomprehensible that such a famine exists when, in our midst, this new media technology, the Internet, can provide a vast amount of information on any topic. This book aims to help address this starvation by providing journalists, especially from developing countries, with some practical skills in "exploring and exploiting" the net for day-to-day journalistic assignments.

The book teaches how to search the net more effectively and efficiently, not only for facts and figures, but also for images, audio and video.

However, we must go beyond increasing the number of stories which can be written and produced. UNESCO hopes that with this handbook, journalists from developing countries can produce stories on more relevant topics such as those related to the UN Millennium Development Goals - poverty, education and literacy, health, nutrition and HIV/AIDS, among others. We also hope viewpoints from various sectors of society can be reflected in those stories.

An important feature of the book is that it not only explains where and how to get information, but also how to evaluate and verify the information gathered, including explanations on some technicalities and warnings about hazards. In short, it keeps the tradition of journalism of verification.

UNESCO collaborated with the Thomson Foundation and Commonwealth Broadcasting Association in the production of this handbook, which definitely contributes to achieving UNESCO's mandate in promoting a free and pluralistic media, ensuring universal access to information and knowledge, and mobilising the use of communication media for human development.

UNESCO invites users of the handbook to send their comments and suggestions on how to make future versions of the handbook more useful.

Dr. Abdul Waheed Khan  
Assistant Director-General for Information and Communication  
UNESCO

---

# 1 INTRODUCTION

Imagine a giant buffet, a great array of dishes, stretching further than the eye can see. Most of it is free – all you have to do is help yourself.

This feast of information is almost too good to be true. For a journalist, it is hard to believe anyone could invent something quite so useful, and then give most of it away.

It's easy to jump in, grabbing the first tasty morsel, and then browse till you can consume no more (or get indigestion).

Okay, so there are drawbacks: some of the dishes are stale, others are downright rotten – even deliberately concocted to do you harm – and there are more which turn out to have price tags after all.

But, for most of us, the real problem is finding a way round this vast mass. The choice can be so great that it's not hard to find something adequate; it may be less easy to find something really good.

**THE AIM OF THE BOOK:** This book seeks to help you find the best information, and learn how to exploit the facilities on offer. It explains the wealth of information available: facts on people and places, statistics (local and international) about health, wealth and much more, news and material to provide background and explanation for news stories and features, and even advice on doing your job better.

The idea of the book is to help make you a more successful journalist – producing better stories, and spending less time doing it.

So, how is this going to be achieved?

The major part of the book is devoted to explaining how to search the net: looking at the tools to use, like search engines and directories, and methods to find what you want efficiently. It covers not only facts and figures, but also searching for images, audio and video.

It looks at what you can find out there, in terms of news and other areas important to journalists. Then it examines the specialised reference tools available – you can, for instance, translate material from Albanian, or into Zulu. Other aspects include evaluation of what you find, explanation of some of the technicalities, and warnings about hazards.

The internet is not only about search. It's a communication system, linking people around the world. So the book explores what's available, looking at internet telephones, email, and at groups and lists – places to find information, and also to make inquiries and build contacts.

In the course of this, you are sure to meet obstacles. This book cannot cover all eventualities, but it offers practical advice on using your computer and recommends good free software to help you do the job.

There is also guidance on dealing with spam, spyware, and other perils that stalk the net – and how to find help. A great thing about the net is that it's a co-operative place, with people willing to provide free assistance.

Finally, the guide provides some exercises, to get you used to aspects of searching and to see how the whole thing works (or, occasionally, fails to work).

**HOW THE GUIDE OPERATES:** This manual is designed on the basis that you know a bit about using a computer, and have accessed the net before. (If you want to go over the basics, there are recommendations at the end for tutorials.) The book assumes that, like a car driver, you know how to start the machine, but you may want guidance on how and where to drive it – preferably without too many accidents along the way.

Keyboard instructions, or other computer operations mentioned, relate to PCs – personal computers – with a Microsoft Windows system; the book does not cover Apple Macintosh systems, little seen in developing countries.

In such a country, reaching the net can be frustratingly slow; there is a section here on ways to overcome some of these barriers. It may be a while before a mass of local information is available on the net, but even a modest online supply of official data can beat getting facts the old way – wasting forever trying to reach someone important enough to release information.

In the past, the usual source of national statistics would be a local ministry or organisation. Now, geographical distances are irrelevant: if facts or figures are available internationally, then you can access them from, say, New York – probably quicker, and with less hassle, than from an office in the next street.

This guide concerns only the English-language net (just a third of the total). Many sources quoted here are American, because the US has the money to develop such resources and the willingness to give them away. Some of this is commercial strategy, but the generosity of American universities in freely sharing library facilities and teaching programmes, the help from foundations in sponsoring information services, provide valuable support to journalists everywhere.

The book concentrates on free services – it occasionally mentions what else is available, but presumes readers are not going to be flourishing credit cards in pursuit of information or software. Sometimes, you may find a free service mentioned here has started charging – someone has launched an idea, which has taken off in a big way, and now expects people to pay for using it. In such cases, though, imitators often start offering similar services for free.

New services are being launched every week. At present, the biggest search companies are in a fury of competition, with Yahoo! and Microsoft battling Google for supremacy, while smaller operations bid for a place in the cybersun.

Experimental ideas described here may be fully-fledged services by the time you read these pages; others may have vanished. Things change fast on the net, and the big operators can be ruthless about dropping services which attract insufficient users.

**WHAT THE GUIDE DOES NOT DO:** This book is not a list of the 'best' sites for different subjects: there are too many permutations, and they soon alter. What it does is explain where to find such sites, and teach the skills to work out which source may be suitable for you.

Where a site or service is listed, it is often one choice out of many. This book tries not to list lots of options, leaving you to investigate each and work out which might be best. It prefers to offer a single selection – usually with one alternative, in case the original site stops working.

If you want to explore a subject further, a recommended site will normally have links to other, often equally good, possibilities across the web.

(The World Wide Web is the system of hyperlinks that enables you to jump from site to site across the internet; the net is the global communications and computer network, and all the activities that take place on it. The internet grew out of decades of military and academic projects; it was not until the mid-1990s that it began to explode into the consumer network we know today.)

If you already have some net experience, scan fast through the book's initial sections, concentrating on what is new to you; if you're starting out, don't feel you must learn about every service – just make sure you understand the main systems.

A hint as you go along: there are so many different points made in the following chapters that you're unlikely to remember them all. Note down what seems especially useful – try bookmarking links that are particularly relevant.

So, there's lots on offer that could be useful in your job. How do you turn yourself into an expert at finding this?



## 2 THE BASICS OF SEARCH

Searching for information on the net can be a bit weird at first. Instead of looking in a book, or searching a library, or even asking a direct question, you most commonly start by thinking of words you expect to find on an electronic page, one which should have the information you are seeking. (This may seem an oddly indirect approach, but it soon becomes second nature.)

You go to a search engine, one of the many available free on the net. Here (*Figure 1*) is a typical search page, [Yahoo!](#) – in the fashionably minimalist style pioneered by [Google](#).



*Fig 1: A typical search page*

A search engine is a system which sends electronic crawlers, a type of software, out across the web, finding information from across the globe; it then indexes all the words from the pages found. It constantly revisits the different sites, updating its database.

When you write in a search query, you ask the system to find web pages containing those words. The engine searches its database, and produces a list of pages. It's a process which can take less than a second – you are not asking the system to scour the whole web for the information; it just looks in its files. So it's not bang up-to-date material; what you get is from a snapshot of the web in recent days, or even weeks.

Originally, such a search was fairly simple, but as the net has grown in complexity, a search can return a list of, theoretically, millions of pages. It is getting ever easier to publish information on the net: establishing websites, or starting blogs – weblog diaries – can take mere minutes and cost nothing, which means the quantity of material is expanding enormously; the quality is another matter.

**FINDING SUITABLE ANSWERS:** Out of all this, how does a search engine decide which pieces of information to show you first? The system makes a guess about the sort of pages you might want (based not just on what you ask, but its experience of other inquiries). Pages are selected on their likely relevancy to your query.

As search engines evolved in recent years, [Google](#) took the lead by developing a special measure of relevancy, based initially on links: if lots of sites had links to a particular site, then [Google](#) concluded the site in question could be worthwhile. Its method proved so successful that the biggest search engines all now operate along such lines.

In the current [Google](#) system, the page ranking it designates is of crucial importance. A page rises in rank, not due so much to the number of links or the amount of web traffic, but rather to the links made from other pages regarded as authoritative – more than a hundred factors help determine which documents are likely most relevant.

**FOCUSING ON THE RIGHT QUESTIONS:** The majority of people use only one word, or two, when they search online – not much, but enough to set them on the trail they want. Most searchers are casual, not thinking hard about what they do. If you write 'Kylie' into a search box, you can expect to get a mix of sites associated with the Australian pop singer.

On the other hand, if you need to know about new treatments for a type of malaria, you have to enter a series of precise words to define what you want.

Many users are happy to roam around the web, in the expectation of finding something interesting or entertaining. When you first try out the net, it makes sense to explore, but casual browsing really consumes time. People talk of surfing the web, but if you're using it as a journalist, don't surf. Don't go where the waves take you; decide what you want to find, and then be methodical in your pursuit.

A March 2004 [survey](#) by an American-based internet expert, Jakob Nielsen, found six out of ten people used just one word to start looking for something, and another two people used only one extra word – though a later survey suggested much larger numbers using extra words. (Treat surveys of net use with caution – the figures often vary widely.)

Nielsen's survey also indicated that only a tiny proportion of people used any sophisticated search techniques. Just three per cent used quotation marks to tie together search terms, and only one per cent used any other advanced methods. So it is not going to take much for you to become a more expert searcher than most people in the world!

There's an argument for trying a "quick and dirty" search first – can you find what you need just by sticking in a few words? But, after a couple of minutes without locating much, start getting organised.

### THREE SIMPLE WAYS OF EFFECTIVE SEARCHING:

- **Use more words**

By writing in extra words, you give a search engine more to work on, so it can home in on what matters to you

- **Use unusual words**

Common words crop up anywhere, and so pages are selected which have nothing to do with what you want. You need something less ordinary. Try to think of specific words, likely to appear only in pages related to the subject you want – people's names or the titles of organisations are obvious things to choose.

- **Use a phrase in quotation marks**

A combination of two or more words, that you are fairly sure will crop up in the pages you want, is a highly effective way to search. Putting words in double quotes is the most common way to tell a search engine to look for a phrase.

Also, put the most important word first (unless it is in a recognised phrase or title). Search engines usually list first the pages which best match that initial keyword.

This all sounds straightforward enough. So why don't more people do it?

There is an impatience that pervades what we do on the net – everything may be available speedily, but that does not stop us wanting it instantly. Rather than spend half a minute working out a good query, we would rather plunge in, and then waste effort clicking back and forth among the inadequate results we've received.

Reporters seem little different from anyone else. Perhaps by nature, we rush at tasks, eager to pursue the news. But as a professional journalist, so you should be professional about searching.

What other basic rules should you learn to aid your digging for information?

One worth learning as a priority is based on the simplest arithmetic: adding or subtracting words. You use a plus sign '+', or the word 'and', to ensure a word is included in your query; a minus sign '-', or the word 'not', aims to keep a subject out of your answers.

This approach is important because results lists are easily swamped by irrelevant subjects. If you need to find the Paris Hilton, write '+hotel' if you want to avoid results full of stuff about the American celebrity of the same name. (Since the woman is an heiress to the hotel fortune, the word hotel may crop up in stories about her, but if you put the '+hotel' instruction in [Google](#), the machine is sharp enough to understand you are inquiring about accommodation in the French capital, and not about the celebrity. Mind you, tales from the gossip columns still sneak in.)

If you are interested in an aspect of football in Senegal, you might write: Senegal football finance -Diouf, which would cut out the many stories about the local star's activities in Europe.

Be careful, though, about exclusions. Words can crop up in unanticipated places, especially in a long page; you might inadvertently bar a document you want.

**THE WRONG RESULTS?** If the results don't fit your bill, consider the following:

If your initial results – the first page or two – yield nothing fruitful, don't waste time. Start again, with different search terms.

Too much information? Narrow the search, with extra search words, or more specific ones.

Too little information? Widen your inquiry, with broader search terms and fewer specifics.

And if you are still getting nowhere, ask: am I trying the right sort of question? Or using the wrong search system? You may need a different search engine, or a completely different tool.

The simplest possible advice comes courtesy of [Microsoft](#), which found, after much research for its new search engine, that the most common problem with web searches was – by far – wrong spelling.

If you are unsure of the spelling of a word (no surprise if you are writing in a second language), check it.

If you write 'Arnold Schwarzenegger' in the search box of [Yahoo!](#), [Google](#) or [MSN](#), they'll all lead their results with a question like 'Did you mean Schwarzenegger?' (The supersize actor turned American politician has one of the most mis-spelled names on the net.)

Try copying names, long words and numbers, or phrases, and then pasting them into the search box; it should reduce typing errors.

### 3 GOOGLE IT

Until recently, the big search engines battled for supremacy by boasting how big they were, how fast they were and how up-to-date they were.

Things are changing. The big three in the field, [Google](#), [Yahoo!](#), and even [Microsoft](#), which is making a massive effort to catch up with the others, now have a fair equivalence.

- [Google](#) has been saying it searches over eight billion pages, but the others also have giant databases ([MSN Search](#) offered *only* five billion – but how many of us can tell the difference?)
- [Yahoo!](#) will tell you what fraction of a second it took to answer your query – but who cares whether the search took twice as long on one system, when the difference is between 0.08 of a second and 0.16?
- [MSN](#) will emphasize their constant crawling of the web, and thus how up-to-date their results are (but frequency figures mean little with so much variety in the level of visits to different types of site).

They still argue about these things: in August 2005, Yahoo reckoned it searched 19.2b pages (more than double Google) – a claim not immediately substantiated. Google later claimed its number of unduplicated pages was three times as big as any competitor, but said it would no longer list any specific number of pages searched.

This makes sense: increasingly, searchers take such achievements for granted. The crucial area of competition now lies in a harder task: giving you the answers you want.

To do this, the organisations offer a barrel-load of different techniques and services, vying for your custom. It is all essentially free – the companies aim to make money from advertising or selling other services.

The main search engines are now sufficiently similar that if you understand how one works, you should be able to operate the others without much trouble.

[Google](#) was selected as the system to explain in detail here because it remains the world's leading search machine (used probably more than the rest put together), and it does most things pretty well. Founded in 1998 by two graduate students at California's Stanford University, it became, in mid-2005, the most valuable media company in the US. Yet it has retained the trust of users to a notable degree.

## A. THE RESULTS IT PROVIDES

When you make a search with [Google](#), the results appear on a page like (Figure 2):

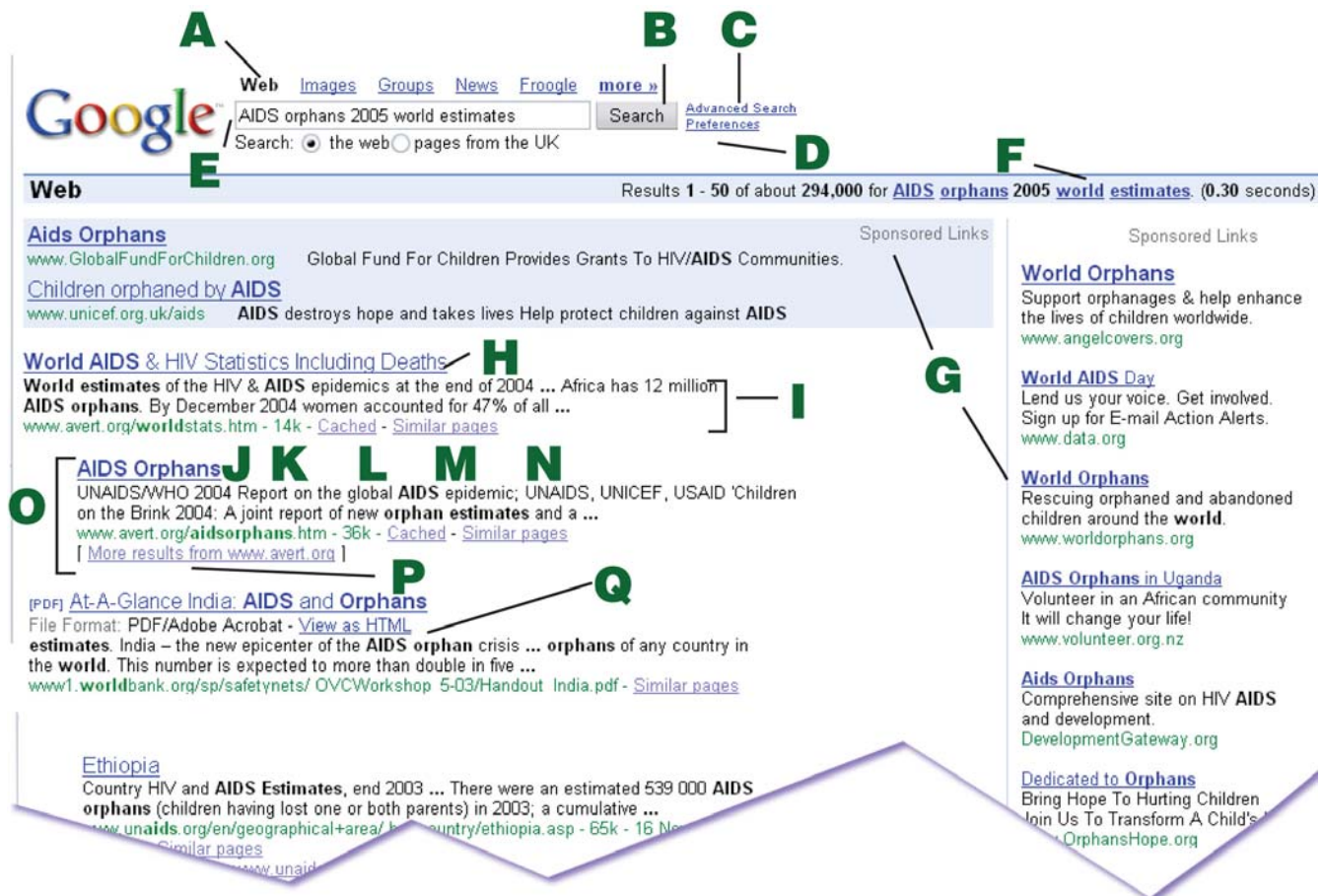


Fig 2: Google search results for the number of AIDS orphans

At the top is your original search [E], with the different areas [A] that you can search, the button [B] for searching, the links to [C] Advanced Search and [D] personal Preferences.

Then [F] is the total of what has been found, and the time it took. If words in the query are underlined here, you can click on them for a dictionary definition, or more information, from [Answers.com](#). If you are meeting problems, this is a quick way to check whether you have an incorrect spelling, or the wrong word completely.

At the top and the side, you will often find [G] adverts, called Sponsored Links, which are chosen for connections with the search topic. Ahead of the results, you may also find tips, for alternative ways to find information, or relevant news results.

The initial result has the title [H] of the web page [Google](#) thinks best matches your query.



Under it, there's an excerpt [I] from the page, showing the words you searched for, in bold type. These lines give you an idea of whether the result is what you want. You click on the blue title to open the full page: the search engine should go to the site and bring up the page.

Underneath in green is [J] the full web address of the page – the URL (the initials mean Uniform, or Universal, Resource Locator). This gives you the source, and may indicate whether it is reliable, or information to ignore.

The size of the page (K) may be important if your computer loads pages very slowly: 17k (17 kilobytes) should not be bad; 69k (below) could be a bit slow

Sometimes there are problems in getting a page to load – and not just due to its size. It may be hard to get through to the site. If so, click on *Reload/Refresh* at the top, or give up, hit the *Stop* button and then click on *Cached* [M], which provides a copy of the page from when [Google](#) last indexed it. This may date from a while back, so decide whether you need a more recent copy. If you want something which appears to have vanished or changed, the *Cached* facility may help.

Clicking on *Similar pages* [N], also in paler blue, can be a way to sites with similar content, though I have not found it especially useful. The results can seem strange because [Google](#) is finding sites similar to the one listed – these may be unconnected to the subject of your query.

If there is another result from the same website, it is indented [O]. There's often a note at the end, saying [P] *More from the same site*, in case you want to pursue the results deeper into that source.

There are other file types, like PDF or Powerpoint, which require special software. Google enables you to open such a file by clicking on the blue link [Q] *View as HTML*.

**EXAMINING THE RESULTS:** Scan them speedily. A glance will probably show which could meet your needs. Don't just start into the first and laboriously go through what it offers, before beginning the next. Search engine choices are approximate – don't presume you are missing something just because a site does not yield anything worthwhile.

But don't go to the other extreme of ignoring a site's potential. Most people visiting a website only look at a couple of pages; if they do not find something they want after two clicks, the majority are off. As a serious searcher, you need a little more patience.

People used to search for a good site – hoping to find there what they wanted – but surveys indicate they no longer tend to bother. They expect to go directly to a page with the answer they require. If, though, you are seeking hard-to-find information, or material from a wholly reliable source, spend a little time delving when you reach a promising site. With serious websites – governmental or academic – you may need to progress through several levels to find what you seek.

When you get to a page, if the item you want is not immediately obvious, use Control and F from your keyboard to search within that page for the term/s you need.

Learn when to give up: don't spend forever on a fruitless pursuit. Try a different search or another approach.

A final thought on results: what happens when you want to go back to something you glanced at hours, or weeks, ago – and you cannot remember where you found it?

[My Search History](#) is a new facility under trial, which keeps a record of everything you visit, stored on [Google's](#) servers, and accessible once you sign in. It lists things in time and date order, with links to bring up the results you viewed previously. You can search the history for the subject you were investigating, or, if you do an ordinary search on the same topic, a note at the top of the results reminds you what you saw earlier.

Be aware that it records everywhere you've been, on personal as well as professional searches. You can, if necessary, remove items from your search history, or temporarily interrupt its operation. If you share a computer and use this, log out of [Google](#) when you finish.

## **B. SHARPENING YOUR QUESTIONS**

Having seen the results from an ordinary search, let's go to the next stage: how to refine a search.

More words in a search can bring better results, so you could keep adding words to refine things further. ([Google](#) will let you enter up to 32 words, but that is usually only useful for tracing a particular passage in a document, or checking on plagiarism – seeing whether someone has illicitly copied material from another author, or tracing one of your own stories.)

[Google](#) searches as if there is an automatic 'and' between each search term, so normally leave out 'and'. If you want to exclude something in [Google](#), don't write 'not' beforehand – always use a minus sign, and leave a gap between this and the previous word (otherwise it could be treated as a hyphen).

Incidentally, when there is a word which some hyphenate and others don't, like round-table/roundtable, [Google](#) should search for both. (But redlight, red-light, and red light, used as adjectives, got varying results, so these things are not straightforward.)

Another useful 'operator' – a term or symbol used to give orders to the search engine – is 'OR' (always capitalised). This is important when you want to cover an alternative spelling, say Chad OR Tchad, or different names, Kalashnikov OR AK47.



To include words with a similar meaning, use the little wavy sign called the tilde '~'. Put this in front of a word like ~plane, and you don't have to think of all the possible synonyms; it should cover aeroplane, aircraft, jet etc.

A couple of basic rules: do not worry about writing capital letters; you can use either upper or lower case and gain the same results. Google also ignores most punctuation, but does usually include apostrophes.

Google has claimed accents on words do not, in English language search, make any difference to the results. But a check suggested that cafe gets different results from café. If in doubt, search for both. Generally, regard Google's rules as applying most of the time, but not invariably.

**STEMMING, PROXIMITY & RANGES:** It uses a version of stemming, finding variations on the word you are searching for: write 'flood', and it can also return 'floods' and 'flooding'. Stemming can be awfully useful when you are trying to track scarce information, where any variant of a word could be vital. But this does not operate universally: search for 'drive' and it may not cover 'driving' or 'driver'. Regard stemming as a bonus in Google, and consider using another search engine if this technique is important.

A similar approach uses the asterisk as a wild card, a symbol standing in for any word. If you cannot remember an exact phrase or quotation, put an asterisk instead, with a space on each side. Write "Political power \* out of the barrel of a gun", and Google supplies 'grows', the full quote from Mao Zedong. If you are seeking more missing words, you can try up to ten asterisks – a type of proximity search.

A facility that sounds beneficial, but often proves less so, is searching within a range of numbers. The idea is to cut down a mass of material by looking at, say, only what happened in a specific period, and not the whole of history.

To look into Nelson Mandela's early years at Robben Island prison, you could search within a range of 1964...1974 (you separate the numbers by two full stops). But such results often include extraneous material – for instance, search engines pick up the dates of publication as well as dates in the text. Similarly, if you check the price of computer printers, \$100...\$200, the information usually comes on pages with many different prices – you are little better off; a specialised shopping search engine is required.

**STOP WORDS:** When making precise searches, be aware of what Google calls "stop words". These are very common words, like 'the', 'is' and 'in', which are omitted to speed up searching. When the results appear, it says at the top of the page, next to the number of results, which words have been left out.

If it is strangely hard to find a particular name or phrase, check what is being ignored. A few words in languages other than English are also stop words, which explains why, when searching for De la Roche, the listings only seem concerned with Roche pharmaceuticals.

Google does not supply a complete list of such stop words, but trial and error by users suggests the following will be omitted: a

an	for	it	the	when
and	from	of	this	where
are	how	on	to	who
at	in	or	was	will
be	is	that	what	with;

from foreign languages: de, en, la, und; the figures: 1 or I; and also: www.

Rather than worry about this long list, though, just put what you want in quotes, or use a plus sign in front of a particular word you want to keep in your search.

(Google is reticent about details of how its search works because there is an industry, search engine marketing, spending large sums to exploit the Google system, trying to get sites to the top of results lists.)

**ADVANCED SEARCH:** A lot of these instructions contain hard-to-remember detail. If you are unlikely to recall the symbols and names required to target your search, then consider using the [Advanced Search](#) facility (*Figure 3*).

The screenshot displays the Google Advanced Search interface. At the top, the Google logo is on the left, and 'Advanced Search' is in the center, with links for 'Advanced Search Tips' and 'About Google' on the right. Below this is a section for finding results with different criteria: 'with all of the words', 'with the exact phrase', 'with at least one of the words', and 'without the words'. Each criterion has a corresponding text input field. To the right of these fields is a dropdown menu set to '50 results' and a 'Google Search' button. Below the search criteria are several filter sections: 'Language' (Return pages written in), 'File Format' (Only return results of the file format), 'Date' (Return web pages updated in the), 'Occurrences' (Return results where my terms occur), 'Domain' (Only return results from the site or domain), 'Usage Rights' (Return results that are), and 'SafeSearch' (No filtering or Filter using SafeSearch). Each filter section has a dropdown menu or radio button and a text input field. Below these filters is a 'Page-Specific Search' section with two options: 'Similar' (Find pages similar to the page) and 'Links' (Find pages that link to the page). Each option has a text input field and a 'Search' button. At the bottom is a 'Topic-Specific Searches' section with two links: 'Google Print - Search the full text of books' and 'Google Scholar - Search scholarly papers'.

Fig 3: Google's Advanced Search page

Calling this 'advanced search' is a bit off-putting; in fact, if you want to do any search beyond listing a few words, this is the simplest way to do it.

Click on [Advanced Search](#), on the search page, then write what you want into the appropriate slots, altering any of the controls you require. This is a good way to handle some specific searches: to find a word only in a page title or in a URL, try Occurrences.

Another control relates to the freshness of the material: *Date* enables you to select pages updated in the past three, six or twelve months. Web pages can be deceptive, often providing few clues as to whether they contain anything recent. When a page is superseded, the old one may remain, and when it appears in a search, you can think it is the latest information.

The *Domain* feature enables you to confine searches to certain classes of site, indicated by the suffix at the end of the URL. If you want to look only at US government sites, you could put 'site:gov' in the query, while 'site:es' would bring up pages from Spain. The latter is a quick way to obtain returns from a particular country, but it is not foolproof – many sites do not use their country's domain name; they prefer to use a universal domain, like .com.

If you prefer not to use the advanced facility, [Google Help](#) offers direct operators to use on the main search page, so you write 'date:' before a range of months, or 'site:' to search a particular website, as well as for domain names. Put an operator as near as possible to the end of your search words, so it is clear to which words it applies.

This guide does not list every detailed use of [Google](#); for more, try the [Cheat Sheet](#), or explore Nancy Blachman's [Google Guide](#).

**KEEPING SEX AT BAY:** One feature on [Advanced Search](#) aims to protect you from pornographic material, of which there is plenty on the web. SafeSearch is the default, the normal state, on [Google](#), so it is regularly filtering content, and, as a journalist, you do not know exactly what you are missing. They now operate a moderate filter, which seeks to exclude explicit images, but not words.

You can control this when you set Preferences, which may be the last important thing to do before setting off searching. Click on *Preferences* on the search page, and then adjust the controls to fit your needs. On *SafeSearch*, you can choose strict filtering, to keep out sexually explicit text as well as images, or the moderate filter, or switch off the control altogether.

If you regularly search for images, you may find yourself bumping into plenty of porn, so the moderate filter can make sense. And if you want to search an area like sex education, which might bring up dubious material, put 'safesearch:' before the search term, to exclude such pages.

**OTHER PREFERENCES:** You can choose how many results are listed per page. The default number of ten results is quicker to load, but if you regularly skim many results, it is easier to scroll down one long list, of up to 100 results, than keep jumping to new pages.

You can search in a specified language – currently a choice of 35. This could be a boon if you read one of those languages well, but it may cut out a lot. [Google](#) was designed for English: its system does not necessarily work so well with other languages, and cultural differences also mean results in other languages may not prove so useful.

There is, though, a valuable facility presenting the [Google](#) homepage, site messages and buttons in different languages – providing helpful clarity when you want to understand the instructions. It's already available in a hundred languages. And if you wish to fly the flag, you can use a [Google](#) homepage for your country, available in the local language, for over a hundred states – like [Tonga](#) or [Turkmenistan](#).

## C. SPECIALISED FUNCTIONS

There is lots more you can do with [Google](#), though you are unlikely to require all the proffered services. (Don't worry about remembering the addresses – just look up the name ... with a [Google](#) search.)

Skill on the internet usually comes from frequent practice, and that means regular use of a few things, not occasional use of many different systems. We'll touch only briefly on services less relevant to journalists, and others will be examined in detail later, as part of overall surveys of important facilities.

New services are usually launched in the US, and then, if they win success, introduced elsewhere. But they can be confined to markets which produce profits – it may be a long time before some reach less-developed nations.

Like other search engines, [Google](#) has been putting much effort into developing local services. Initially covering the US, Canada and the UK, [Google Local](#) (still officially in the beta, trial, stage), searches an area for a type of business or service. Then it provides details of each place, directions to get there and shows the locations on an accompanying map. It can, if you type 'movies:' and the name of a current film, show film details and performance times at local cinemas.

Another commercial operation under development, but already in use, is the online shopping service, [Froogle](#), which includes searching within price ranges. If such services are not on the main [Google Search](#) page, click *More* to see a fuller listing. Beyond these is an area called [Google Labs](#), with experimental facilities. Have a browse, to see what is coming down the road.

One very much on the way is [Google Maps](#), which we'll look at in Reference Tools, chapter 11. But to indicate just how much can be available online, one recent venture is [Google Ride Finder](#), which (for some US cities) shows on a map the location of taxis currently available for hire.

**PERSONALISED SEARCH:** One idea, tipped as the future of search, is [Personalised Search](#). It allows you to register your particular interests (such as areas of knowledge, parts of the world, types of business, sport or entertainment). Then, when you've done a search, you move a slider at the top of the page to maximise the importance given to your preferences.

Designing searches to fit your particular concerns would appear an important and logical development, but thus far [Personalised Search](#) does not seem to have created too much interest among users.

It can be helpful to differentiate between subjects with the same name. Look for something about Georgia, and you commonly get lots on the US state and little on the former member of the USSR. Look for 'Georgia election results' and its all American politics, with, down the page, one entry on Georgia the country. But with personal preferences set to include Europe and Asia but not USA, move the slider to maximum and the country listing immediately jumps to the top of the results.

**ADDITIONAL SERVICES:** Google now has a [Q&A](#) feature, so you can ask questions in ordinary language, such as: What is the capital of Niger? But the results are not always much different from a regular search.

Other tools being created include [Suggest](#), which finishes off words as you start typing them into the search box, and can propose phrases specifying more clearly a possible search. [Scholar](#) permits the search of academic research – useful if you are pursuing serious inquiries into a specific area of knowledge.

One more in the beta stage is [Google Print](#), for searching the full texts of books. It's an operation set up with publishers, but joined in 2004 by Harvard, Stanford, Michigan and Oxford Universities and the New York Public Library. Older books – especially those out of copyright - may be available complete, so you could read them online. Books for purchase may let you read a few pages, around the term you searched for; books still in copyright, screened from libraries, will only provide a few sentences next to your search term, and you can only look at three instances of each term.

Another recent innovation – swiftly becoming a norm among search engines – is desktop search (see Chapter 15, Desktop Search). An alternative way to search, looking for subjects, is the [Google Directory](#) (see Chapter 6, Directories). Also, in Chapter 9, News), we'll look at Google's fully-automated [News](#) search and at the valuable news and web alerts, like [Google Alert](#).

Meanwhile, if you want to know what subjects people are searching this month, [Google Zeitgeist](#) gives the latest American queries, with less up-to-date top ten lists from a score of other countries. Zeitgeist is a fancy expression – in reality, it's most likely to tell you about current teen obsessions.

**WORDS & NUMBERS:** Language Tools offers translations (somewhat rough) between English and eight other languages, including French, Spanish and Chinese. A more sophisticated system is on the way. (More on translation in Reference Tools, Chapter 11.)

The Google [Calculator](#) allows you to do your sums by writing figures with symbols in between: plus '+', minus '-', multiply '\*' and divide '/'. It can be used to convert units, and you can instruct it in ordinary language: write *14 gallons in litres* and it tells you 53 litres (well, actually 52.9957652). It's nice and simple for temperatures, too: put *59f in c* and you get the answer *15 degrees Celsius*. You can use common abbreviations: thus *pounds* or *lbs* will work, as will *kg* and *kilo*, though not everything – you can say *yd* rather than *yard*, but not *yds*. For more, try [Calculator help](#).

There's a built-in spell checker, which may automatically offer an alternative spelling if you are using something which occurs more often on the net in a different way. This can be useful for proper names, not always covered by dictionaries – but make sure the word they are offering is indeed the one you want.

Google has lots of number searches, usually applicable only in the USA: you can search for stock quotes, bar codes on products, aircraft registrations and even the tracking numbers of parcels being transported by Federal Express, UPS etc. US phone numbers can also be looked up – see [Web search features help](#).

**MORE THAN JUST SEARCH:** [Google](#) is fast expanding into areas unconnected with search. Thus there is online email, [Gmail](#), which debuted with a gigabyte of storage space – well ahead of most competition – subsequently increased to 2.5gb; more space will follow. Only a few users took up more than one per cent of this space, but with the spread of big files like video, demands on capacity are rising.

[Gmail](#) can organise material by email exchange, so you see the whole of an online conversation with someone, and, of course, it's searchable. Normally only available by invitation, you can now obtain it in the US by registering a mobile phone number; presumably such access will soon be more widespread.

There's [Google's Short Message Service](#) for mobile phone users, supplying answers to general knowledge queries, dictionary definitions, and more. Then there's [Orkut](#), a social networking system; [Picasa](#), software for editing and sharing photos; and [Blogger](#), where you can put a personal diary on the web. (There's more in Blogs, Chapter 18.)

And then, if you are able to spend money, there is a service where you pay experts to do the searching: [Answers](#). Anything from \$2.50 up will attract their researchers; offer enough and you should get enthusiastic help.

Google is trying to pull together these varied strands in a single offering to users, in a project called Fusion, and in the process is experimenting with a new [Personalised home page](#). Instead of the original uncluttered look, it offers news, weather and other services, and can include [Gmail](#). Still pretty basic, it does not yet compete with the pages full of information which greet you at [Yahoo!](#) and [MSN/Microsoft](#). But this may be the start of turning [Google](#) into a portal – an all-purpose entry-point to the net, and not just a search engine.

#### **D. TOO POPULAR FOR ITS OWN GOOD**

[Google](#)'s success creates problems. Getting to the top of a [Google](#) results list can, for companies selling online, be the difference between fortune and oblivion. So it's no surprise that much effort is devoted to manipulating those results – and consequently making it harder for ordinary users to find what they need. [Google](#) engineers change their system, trying to stymie the marketers; the latter then seek a way around the new barrier.

So you may get results swamped with unwanted material. If you are searching for an organisation and find it buried by sales pitches – like seeking a hotel whose website is obscured by companies offering 'big' discounts – you can try writing '+home page'. (This can work better in the [A9](#) search engine, or try [Yahoo's Mindset](#) – see the next chapter). But the battle will not be easily won: commercial operations aiming to top the search results now produce not just individual pages for each subject, but whole websites.

You could, though, try a favourite of mine from [Google](#)'s earliest days: on the search page, click on the I'm Feeling Lucky button, and it will take you to the first search result, which can still turn out to be your exact target.



## 4 YAHOO! & MICROSOFT

What are the main alternatives?

The 'search engine wars' of recent years have vanquished most contenders, leaving three search conglomerates astride the net.

While [Google](#) grew organically, and [Microsoft](#) is building a new search system from scratch, [Yahoo!](#) created its machine out of several engines, acquiring [Inktomi](#) (once used by [MSN](#)), the powerful [FAST](#) search engine, and [AltaVista](#) (an early star in the search firmament, fallen on hard times, but still offering useful techniques).

From these technologies, [Yahoo!](#) produced a strong challenger to the [Google](#) leviathan. At the start of 2005, a survey covering 100 countries, from [OneStat.com](#), put [Yahoo!](#) on 21% of world search engine usage, as against [Google](#) with 57%, and [MSN](#) at under 10%.

### YAHOO! SEARCH

So, should you choose [Yahoo!](#)?

[Yahoo's](#) results are now regarded as broadly comparable with [Google's](#). If you're already a [Google](#) user, it's probably not worth switching. But check to see if one suits you better. The [Dogpile](#) metasearch engine can compare the results from a whole variety of search engines, using a [Comparison](#) tool to show them alongside each other.

[Dogpile](#) has a nifty [Missing Pieces](#) tool to compare [Google](#), [Yahoo!](#) and [Ask Jeeves](#) schematically. You can see how many items appear on all three lists, and then you hover the mouse pointer over individual items, to see what was found. [Dogpile](#) points to a [study](#), including [MSN](#) as well, indicating that just 1% of page one results from those engines appeared on all four lists – demonstrating how varied are the selections you get from different engines. 85% of results appeared on only one list.

If the immediate results do not sway you either way, then the mix of services might.

[Yahoo!](#) search is virtually a clone of [Google](#) in its clean look, and operates similarly. Normally you'd access it via the [Yahoo!](#) portal, with lots of services you select according to your interests, and a search box at the top with such special areas as News, Images, Video and Local.

If you go to the main search page, you get further options, such as Maps, Travel, Directory and People Search – and the opportunity to edit this selection. The Directory is particularly comprehensive, one of the biggest on the web.



There's an advanced search page and preferences, both similar to [Google](#). Operators, like '+' and 'OR', are essentially the same, too, but with relatively few stop words. There are language tools and also [Yahoo!](#) home pages for some 30 countries.

You could use [Shortcuts](#), a system which jumps straight to specialised results, listing the exact information above the main results lists. Many categories are only for US users, but others offer a quick way to worldwide weather (just write in a search box 'Buenos Aires weather' to find the situation in the Argentine capital), encyclopaedia reference information, and calculations. (Don't bother with the currency converter, though – at present, it works with only a few currencies.)

One experimental idea, which can reduce annoying commercial offerings, is called [Mindset](#). After receiving the results, you move a slider to make the list more shopping oriented, or more research biased.

[Yahoo! Site Explorer](#), now under trial, offers an in-depth probe of sites, including all the links to a site or to its individual pages.

**SAVE WHAT YOU SEARCH:** Going further than [Google](#)'s My Search History, [Yahoo!](#) is trialling [My Web](#), for saving pages you look at. Previously, you could note web pages as bookmarks – links that would let you jump back to that web address. But URLs change or disappear, and it is easy to forget what a bookmark, or even the URL, was actually about; this new system retains the whole page.

With [My Web](#) operating, click on Save at the end of a results listing. A box appears where you can make accompanying notes. The saved pages may be placed in different folders, but to avoid extra work, save them unsorted and just search later within your [My Web](#) collection. It can all be accessed from any computer, not just your regular one.

[My Web](#) normally displays saved sites in date order, but you can also sort them by title, or into groups based on the searches you did. Click on Search History to see the sites you visited, but did not think to save at the time.

You can block sites from appearing in results, but unless you are regularly searching in one field, and want to exclude very common sites, it's probably not worth the bother – blocking rubbish sites could be a full-time job.

A development like [My Web](#) is possible now partly because electronic storage costs are so low; search engines are happy to keep large amounts of information for you, in the belief this will tie you to them permanently.

### **AN INDEPENDENT SEARCH MANAGER**

If you use various search engines, and want a central online store, look at a stand-alone web research manager, like [File URL](#). On a page you want to save, you click on a 'Furl It' instruction, adding notes if you prefer. You receive up to five gigabytes of space for storage – enough for tens of thousands of pages – and the system will recommend new pages that might interest you, based on what you've already saved. But set it for 'private' if you don't want your selections publicised to other users.

Another potentially important development is context search, though initial offerings have not attracted as much user enthusiasm as might be expected. [Yahoo's](#) version is [Y!Q](#), currently in beta.

It narrows your search, based on the page you are looking at. You select a passage and the search can operate simply, perhaps with only a single word, because it distinguishes what you want on the basis of what you are reading. You can also look for related material by searching on the passage you are reading, without entering another query; just highlight the words or paragraph you wish to pursue.

You can set this up on your computer, but it is also possible to highlight and copy some text, and then go to [Y!Q](#), paste it in, and look at related results.

[Yahoo!](#) has much more than just search, offering such services as email, messenger, and chat. And to see what else it has planned, go to: [Next](#).

If you want more about [Yahoo's](#) facilities, try [Search Help](#). It's not too obviously displayed, but you could use an exclamation mark after keywords, an original method for navigating round [Yahoo's](#) services. Write: 'help!' to go to the Help section; 'finance!', 'tv!' and many others work likewise.

### **MSN SEARCH**

At present, [Microsoft Online Search](#) is a weak contestant for your custom. But [Microsoft](#) is pouring resources into this sector and before long it will doubtless have a highly competitive system.

Categories with its search box are Web search, News, Images, Desktop and Encarta – the latter providing access to its online encyclopaedia. A drop-down menu from the Search button lists further options.

Settings allows you to establish preferences for languages, SafeSearch and the rest; if you specify your location, then the Near Me button may bring local results – in the US, anyway. [MSN Search](#) has more than 30 country versions, mostly in the main local languages, and these will typically search for material more relevant to that country.

The results pages can be a bit heavy with sponsored links, but are otherwise similar to the opposition. When entering search terms, generally use 'AND', rather than a plus sign; otherwise it can be 'NOT' or a minus sign, 'OR', etc. Accents on words do not seem to affect searches too much.

Instead of the standard advanced search page, MSN offers SearchBuilder, with a variety of options and a tool to adjust the ranking of results: there are three sliders you move up and down to accentuate whether the listings should be recently updated sites, popular sites, and an approximate, or exact, match to the search terms.

Reaching an Encarta item via [MSN Search](#) gives you access to the main encyclopaedia, free for two hours (and readily renewable). It will also answer general knowledge questions: [MSN](#) offers some 1.5m [Instant Answers](#) from Encarta, covering such areas as health, music, maths and dictionary definitions – just write: 'What is catarrh?' or '278 x 16?' (Helpfully, you can write simply: '26% of 1200'.) For more complex calculations, there is a detailed list of symbols.

Like [Yahoo!](#), [MSN](#) offers a mass of services. As well as [Hotmail](#), you can sign up to [MY MSN](#), which allows you to organise a customised homepage of news services, communications links and more. For projects under development, check [MSN Sandbox](#). It has already announced Windows [Live](#), which will integrate a much more sophisticated webmail service Live Mail (taking over from Hotmail) with Messenger, Spaces and other facilities.

At present, you would not choose [MSN Search](#) over the others – but come back in a year's time.

## 5 OTHER TYPES OF SEARCH ENGINE

### ORDINARY LANGUAGE

While [MSN](#) has developed its [Instant Answers](#) programme, there are search engines specialising in 'natural answers'.

The best known, [Ask Jeeves](#), with its image of a stereotypical butler, allows you to ask a question in everyday language, and receive a straightforward answer. It never seemed to deliver too well on the latter – frequently offering just a set of results containing, somewhere, the information sought. But recently it has sharpened its act, and you are more likely to get a simple answer at the top of the page.

Phrase the question as you would in ordinary conversation – don't put the most important words first. Thus: Who is president of Kiribati? Or: What causes hurricanes?

The company is trying to improve its responses with features like Smart Answers, where you can enter a name with a keyword (like 'time in Jakarta', to find what's on the clock there), and Web Answers, with opportunities like asking about quotations by writing: 'Who said religion is the opium of the people?'

The successful delivery of natural answers has been described as a holy grail of the search world, but the appeal may not be that great: only 20% of [Ask Jeeves](#) queries are now phrased in natural language (it was 70% at one point).

A Zoom feature offers suggestions, on the right of the results page, for how to narrow your search, or broaden it if necessary. Click on [Smart Search](#) for a list of all the search options.

The organisation includes such net operations as [Teoma](#), the search engine it uses to crawl the web; [Excite](#); and [IWon](#). Surveys give the group only a couple of per cent share of search engine usage, but it may be number four globally. [Ask Jeeves](#) now has an array of search features more comparable with bigger competitors, even if the number of pages indexed may be less – one quoted figure was 2.6 billion pages, although a recent study put it as high as 5.3 billion.

The advanced search page has a precise date tool, enabling you to search for pages modified in a particular month or even on a specific day. You can also select sites based on the continent, rather than the country, where they are located.

There is a [My Jeeves](#) system, still under trial, to save pages as you search. You can attach notes, and also tags to help identify the pages, and there is a search history feature as well. It will save up to 1,000 pages and more if you sign up for the service, which then enables you to access the saved material from any computer.

[Ask Jeeves](#) is changing its image shortly, abandoning the butler mascot and calling itself just Ask, or Ask.Com.

Through [Google](#), you may have met another system which tackles natural language questions: [Answers.com](#). It concentrates on providing information from reliable reference sources, and also has a modest directory.

## MULTIPLE-SEARCH ENGINES

If one search engine is good, why not use several simultaneously?

That's the idea behind Metasearch engines, which interrogate a series of search engines and present the results in a single collection. Given that results vary so much between engines, such a system sounds ideal.

My experience is that they rarely deliver as much as you hope. A [Google](#) search often produces as much, and more quickly. Metasearch results can come laden with paid insertions (i.e. adverts) from less ethical search engines, and can find it difficult to handle complex queries, because the search engines they visit respond in different ways to the instructions.

It can be sensible to use them to pursue obscure search terms, like parts numbers, but perhaps their most convenient service is when a single search engine fails to deliver on a topic. If it is an out-of-the-way matter, then try a metasearch to scan a mix of engines – it may turn up one with good coverage of that particular subject.

Among the better metasearchers is [Ixquick](#), which shows a consensus of the search engine judgements – adding stars to pages that appear in the top results of several engines; it covers a dozen, including [Yahoo!](#), [Teoma](#), [MSN](#), and [Gigablast](#), but not [Google](#).

[Vivisimo](#) also lacks [Google](#), but presents some comprehensive results, and uses a clustering technology to group these neatly under different headings, beside the main results. [Dogpile](#) includes [Google](#), [Yahoo!](#) and plenty more.

## CLUSTERS

Clustering has been around for some time, but gained new attention with the introduction of [Clusty](#), a [Vivisimo](#) project (still in beta).

When a word has various meanings, it can help separate results, and let you concentrate on what matters to you. Look for 'Dido latest research', and it (*Figure 4*) goes some way to dividing pages about a British pop singer from those on a type of nuclear reactor and on the reputed founder of Carthage.

But it is hard to differentiate material clearly. Groups entitled Articles or Reviews could apply to any of the meanings; you still have to check through them to be sure you're missing nothing. (Click on Queen and you get both the rock group and the mythical Carthaginian.)

The screenshot shows the Vivísimo metasearch engine interface. At the top, there's a navigation bar with links: company, products, solutions, customers, demos, press. Below this is a search bar with the query 'dido latest research' and a dropdown menu set to 'the Web'. A 'Search' button is next to it, along with links for 'Advanced Search' and 'Help'. A banner below the search bar says 'NEW try your query at Clusty.com'.

The main section is titled 'Clustered Results' and shows 'Top 171 results of at least 20,300 retrieved for the query **dido latest research** (Details)'. On the left, there's a sidebar with a list of clusters: **dido latest research** (171), **Performance** (13), **Release** (13), **CD** (12), **Rent** (12), **Review** (11), **Blog** (9), **Dido and Aeneas** (9), **HighBeam Research** (8), **London, England** (7), **Book** (8), and a 'More' link.

The main results area displays a list of search results. The first two are sponsored links: 'Research Articles' from The New York Times and 'Free Research info' from www.getbestinfo.com. The third result is 'MATLAB OPTIMIZATION - TOMLAB' with a snippet about optimization and a link to www.tomlab.biz. The fourth is 'DidoZone - The Unofficial Resource For The Latest Dido News And Media' with a snippet about Dido and Aeneas. The fifth is 'Latest News from the Marlowe Society' with a snippet about Dido and Aeneas. The sixth is 'Re: Love of Dido & Aeneas: Latest news on the opera, in electronic...' with a snippet about Henry Purcell. The seventh is 'DIDO.PA: Research Reports for DIDOT BOTTIN - Yahoo! Finance' with a snippet about DIDOT BOTTIN.

At the bottom left, there's a 'Find in clusters' section with a text input field and a 'Go' button.

Fig 4: A metasearch by Vivísimo, showing (left) clusters of results

Rather than treat clustering as something deserving a special search engine, it makes sense as an adjunct to your regular search. The cluster list alongside [Vivísimo](#) results gives the option of inspecting the groupings, or just flicking down the main results. [Teoma](#) works in a somewhat similar way, with Refinement suggestions, and other engines are introducing such facilities.

## THE CUTTING EDGE

There are plenty more varieties competing for your attention. No harm will come if you skip the rest of this section and concentrate on those already covered. If, though, you want to increase your capabilities, there are newer possibilities which could develop into important options:

**Exalead** is a French offering, with an ultra-minimalist search page and over a billion pages indexed. With every result, there is a thumbnail image of the page, which sounds good – you can have a quick glance to see if it is worthwhile before clicking on it. But these are usually too small to read, and, without a fast net connection, it can mean slower loading of the results.

The pages can be viewed within the search screen, and the query words are automatically highlighted. Its advanced search has some sophisticated tools, allowing phonetic search, approximate spelling, and the automatic stemming of search words, with the notable plus of 'fuzzy' queries, where an option seeks results 'preferably containing' certain words or phrases. And there's proximity search, to locate pages where two search terms occur within 16 words of each other.

You can look for material modified after a specified date – useful for getting only the latest material, but revealing that at present the **Exalead** index has little up-to-date material. Still, it is only a trial at the moment.

**A9** combines [Google's](#) databases with a system from [Amazon](#), the seller of books (and much else). It can store pages you search, plus your notes, and list that history alongside a new search around that topic. As well as web results, you can open a column of results from an allied search of images, books, an encyclopaedia or other sources.

Hover the mouse pointer over 'Site Info' on a result and it tells you the time it takes to load. You can also find out how long the site has been in operation and the traffic rank – the lower the number, the more popular it is. With, in addition, the figure for how many other sites link to it, you can judge how valuable a source it might be.

For an eyebrow-raising novelty, try **A9's** yellow pages, where, in some US city centres, you seek a business address and it provides not only a map, but a photo of the building. With Block View, on [Maps](#) you can check the look of neighbouring places, too. It could be useful for journalists, unable to reach the area of some event or disaster. (This is an expensive operation, so you may not see a photo-vehicle in your neighbourhood any time soon.).

**Blinkx** is still pretty experimental, with a limited database as yet, but it offers ideas which are attracting much attention, notably Smart Folders: you put items in a folder and then the system automatically updates it with new material from your computer or links to relevant web pages.

This could be very useful to those covering specialised areas – the latest information, delivered automatically. A folder can be configured for different types of material, or information from a particular location, and it can alert you when something new arrives.

On your computer, Smart Folders can group all files with material meeting your search criteria, while making the same files available in other folders. As we accumulate more and more files on our computers, finding them again gets harder; this can provide a variety of paths to locate them.



With downloadable software, [Blinkx](#) can continuously read the contents of your screen and search for related content in your computer files or emails, as well as on the net. Then it will provide the results whenever you request them. The aim is to work more as people do, making connections between related ideas, rather than by keywords or mathematical logic.

[Blinkx](#) has notably clear listings, with biggish thumbnail images on the first page, and then, against lower results, just figures showing how much a page matches the search terms. A straightforward refinement tool allows you to search further from a particular page, and there are clustered results on the left.

It also has a splendid looking Visualiser, where you click on gently circling items or topics from your results, and bring up suggestions of linked materials. It may not be a utilitarian way to search, but unite two of the globes and you spawn whirling new 3D worlds of related information. The system was offline in September 2005, but an updated version should return.

Several engines offer intuitive visual approaches, including the French system, [UJIKO](#) (using results from [Yahoo!](#)), with colour-coded displays which get more sophisticated as you gain experience. [Grokker](#), also using [Yahoo!](#), has an over-the-top style of visual clustering, but can deliver information usefully, with background on sites as you click on them.

In time, the above ideas may well be utilised by bigger companies, which are already making some developments along parallel lines.

There is, however, a completely different approach to searching which you might wish to consider: directories.



## 6 DIRECTORIES

Once upon a time, when the web was young – a mere decade ago – directories were what everyone wanted: somewhere all the information out there could be indexed and categorised and easily looked up.

[Yahoo!](#) built its initial success as a directory. Back in 1994, the creators (another two graduate students at Stanford University) started making a list of their favourite links on the web, and that list just grew.

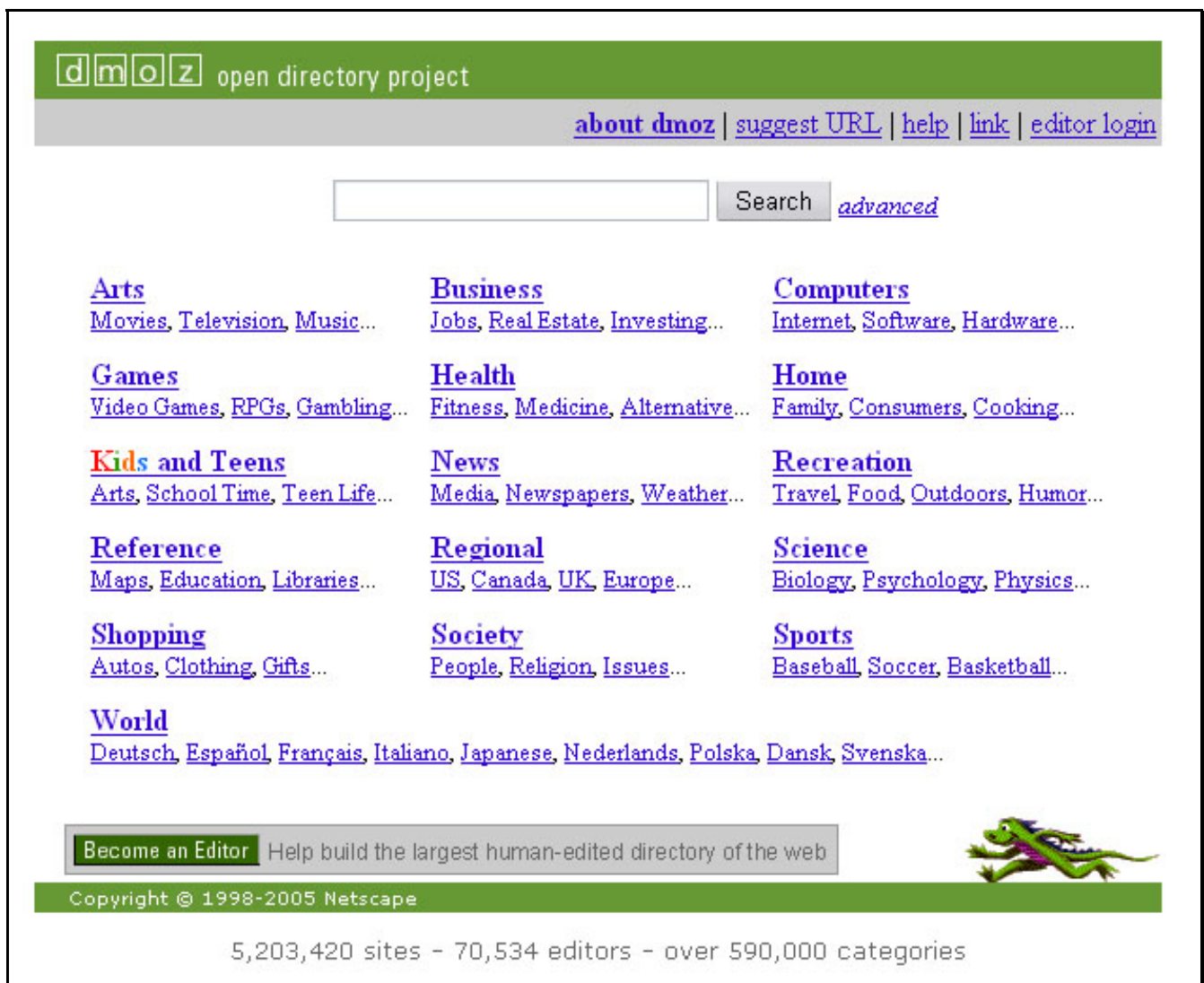
The other leader in the field, the [Open Directory](#), has been one of the wonders of the net . Run by volunteer editors, it grew into the world's largest directory, [DMOZ](#), linked to the Mozilla Open Source project (DMOZ means Directory Mozilla).

The net is a great source of information, but not necessarily so helpful when you want to understand something. A good directory helps find material to explain a subject.

Directories differ from search engines in that they divide knowledge into subject topics, and they take you to sites rather than pages. You normally reach home pages, so you can study what a site has to offer. Directories may be alphabetical lists, as in a print encyclopaedia, or indexed according to a system, like a library.

When we first ventured onto the web, familiarity with scanning reference books meant we were happy with searching as we would in a dictionary, or in the broad categories we had used in libraries. But as we sought more detailed, or more complex, material, we discovered the web enabled us to jump fast to what we wanted, without the electronic equivalent of flipping through the pages.

So even if we use directories these days, it can be quicker to search them as you would with a search engine: put in the keywords for the subject you seek and then browse the resulting sources. If it is something straightforward, you can follow the categories on the main page (*Figure 5 on the next page*), and on down through sub-sections. But for anything less obvious, you are dependent on understanding how subjects have been categorised on that particular system.



*Fig 5: The categories of information for the Open Directory*

If you want to know about forensic photography – such as pictures from crime scenes – then in [Yahoo! Directory](#), you go to Arts and Humanities, click on Photography, and then on Forensic Photography. But in [DMOZ](#), you reach that subject through: Science>Science in Society>Forensic Science>Forensic Photography. So a better bet is simply to put what you want in the search box, probably saying 'crime photography' – a more natural description of the subject.

One of the pluses of directories is that human beings are involved – someone is making a considered judgement about whether a site is worth viewing. Given the dross out there, this can be beneficial.

There is, though, a tendency, when good new sites appear, simply to add them to ever-lengthening lists – so when you look, it is hard to work out which might be best. Both [Google](#), which uses [DMOZ](#), and [Yahoo!](#), operate a ranking system, aiming to put the most useful sites first.

You have to take on trust that the editor responsible is expert and reliable, but if you know little about a subject, it's a reasonable bet the editor for a big directory will be a decent guide. The descriptions attached to entries tend, however, not to offer any opinion on the quality of the sites.

**DIFFERENT OPTIONS:** The oldest catalogue on the web is the [Virtual Library](#), established by the inventor of the web, Sir Tim Berners-Lee, and run by volunteer experts. The information sources tend to be high quality, but some of the less academic areas can be dated and limited in range.

Another approach is offered by [Pinakes](#), a project hosted at Scotland's Heriot-Watt University. It presents 40+ gateways to different areas of knowledge, using serious collections developed by academic institutions and the like, in the UK and around the world.

**PRACTICAL MATTERS:** An alternative type of directory is offered by [About](#). It's a source of practical information, on anything to do with everyday living – from writing a business plan to the health benefits of green tea.

It can be a bit simplistic if you know much about the subject, but it provides good introductions for newcomers – presented in simple language. Try it, for instance, when you want to understand some basic aspect of computing. You can browse through the topic headings, but using its alphabetical search is probably simpler.

Much of what [About](#) offers is its own material, rather than links to web pages – in this, it's rather like the reference sources we'll see later. And rather than choosing directories or search engines, an old-fashioned encyclopaedia – even if in an electronic version – may still be the best way to get to grips with a new subject.

**FUTURE PROSPECTS:** The day of the big directory may be nearly done. As the net expands, so the task of compiling a comprehensive index becomes ever harder. They are very costly to maintain, which is why the [Open Directory](#) has thrived, relatively – it has a tiny paid staff.

It's difficult to review all the new arrivals, or monitor what is happening with previous choices. 'Net rot' sets in – links become obsolete and no longer lead anywhere. One in five of Yahoo's links, according to a report, failed to work or was out-of-date.

Rich Skrenta, one of the [Open Directory](#) founders, pointed out recently that [DMOZ](#) now listed 4m. sites in nearly 600,000 categories, which was "almost too large to be useful – one can't click around, browsing through a 600,000 page directory". Yet, he noted, it was still covering only 0.1% of the net, or less.

Fewer people now use directories, and as a result they are slipping further from public view: [Google](#) has moved the [DMOZ](#) directory link off its search page (click on 'More' to find it); [Yahoo's](#) directory is still on its main home page, but can be tucked down the bottom.

So what other methods can we use? Up till now, we've been looking at systems designed to scan the globe, for any and all information. You might prefer a narrower focus – a search engine or a portal designed for your local region or special subject.

## 7 SPECIALISED SEARCH

We are concerned here with tools to seek information on a particular field or country. These include search engines, directories, portals, and more.

Let's see how it works in one particular field, medicine. You could start with a metasearch engine, [Omni Medical Search](#), which covers medical search engines, databases and other sources. But the results may be confusing, so it could be sensible to explore more suitable services which appear in its lists.

If you are after serious medical research, [PubMed](#), from the US National Library of Medicine, offers its Medline database, with 12m+ citations from scholarly journals.

But that is aimed at health professionals and scientists, so you may prefer something more suited to ordinary readers, like [WebMD](#). Obviously American in approach, some of the material will not apply in your country, but it has a tremendous amount of detail about diseases, drugs, and treatment methods, and a nicely presented symptom checker.

Another service couched in everyday language is [familydoctor.org](#), from the American Academy of Family Physicians, which is full of practical information.

For health statistics in your country, try the [World Health Organisation](#). It can usually supply the latest on the local health situation. Browse the site's topics for surveys of health and diseases worldwide. It has many links to other material and sites.

Directories play an important role here. There is a strong argument for using a catalogue rather than a search engine, because the accuracy of what you discover can be crucial. For that, you need some guidance from recognised experts. And a directory is more useful at a level where the number of subjects is not completely out of hand.

In the UK, you might go to the government health care site, [NHS Direct](#), or use a different [OMNI](#), a university-based system of directories with resources selected and evaluated by experts in health, bioresearch, veterinary matters and agriculture.

**SEARCHING CLOSER TO HOME:** You, however, will want the equivalent for your part of the world, preferably for your own country. That's when things get a little harder.

A national medical directory could be much more useful, as it would focus on the diseases, treatments and health system you deal with in daily reporting – information in the language, and using the terms, familiar to local people.

Unfortunately, reliable services can be hard to come by. In too many countries, there is not the money to pay skilled staff and provide computer capacity – other priorities for public funds tend to be too urgent.

Checking on Zambia, as an example, it was clear the [University of Zambia Medical Library](#) had an unusually comprehensive site. But it had not been updated for several years.

It was obvious when inspecting national directories and sites, and regional ones as well, that many were not being properly maintained. As a result, the information had become obsolete and many links failed to work. You will need to check what is available locally, and then test their comprehensiveness and reliability.

You may find the best on offer comes from foreign initiatives, perhaps at academic institutions, or from expatriates, focussing back on their home countries. Thus someone looking for net resources on Africa could find a good gateway in California (at Stanford University – where else? you might ask): [Africa south of the Sahara](#). A more popular style product, [WoYaa!](#), is British-based, but presents mixed results – it depends on which country you check. There's also a pretty professional directory of websites from the South African-based [All Africa Internet Guide](#)

Many services are not comprehensive, perhaps designed to concentrate on a single aspect of local life; others take a strongly commercial approach. Plenty are multi-purpose gateways to countries, so you get such services as email, dating and tourism info.

**SEARCH FOR SEARCH ENGINES:** This guide does not try to name the best search services for every country and subject – such a list would be almost endless, and individual suggestions could date quickly.

So what do you do? Use the net.

There are quite a few listings of regional search engines, but some omit a lot of countries, while others are full of dead links.

Some may not be too helpful, either: there are international search engines that list themselves as local ones, and others come laden with paid-for results.

Try the following for search engines and directories from individual countries:

- [Phil Bradley's website](#), run by a UK internet consultant, concentrates on listing just regional engines, with over 2,600. It probably has the widest selection, from 216 countries and regions, but there are no notes about individual ones – so if there is a long list, as with the bigger countries, it can be hard to find the best one for your needs.
- [Search Engine Colossus](#), based in Canada, is not quite as comprehensive as it sounds, but its international directory offers 2,100 links to 198 countries and 60 territories. It gives only minimal explanations for the individual entries, but says where each is based and the language used. There are frequent checks on links, so it should be up-to-the-minute.

For a wider spectrum of search engines, covering specialised topics, try:

- [Search Engine Guide](#), edited from the US. It offers 3500+, indexed and annotated.
- [Pandia](#), based in Norway, has a much smaller but strong selection, up to date, and all on one long page. (There are links to more, if necessary.)
- [Search EngineZ](#), run by Australian Robert Skelton, thoughtfully provides search boxes for certain engines, so you can insert your query right there, and then lists other selections, with notes on most of them. Some of the background information is getting out of date, but the links still seem reasonably good.

**SHOPPING:** One additional type of search is shopping. Mostly, it's irrelevant to journalistic activities, but note that [Froogle](#), [Google's](#) comparative shopping tool (still in beta), [Yahoo!](#), which bought the successful European shopping engine, [Kelkoo](#), and very many other web operations, all seek a share of the multi-billion dollar online shopping pie.

Services are usually designed for particular countries, though people buy from overseas when prices are much cheaper. They may be useful to check international prices and product information, and some sites also collect reviews from customers. For detailed information about a product, or instructions on using it, try the manufacturer's site.

## 8 DEEPER SEARCHING

The invisible web – also called the deep web – describes information that is not openly available on the net. Some of this material is becoming more visible, as search engines grow more advanced.

But many searches still do not reach:

- Databases within a website – such as news archives
- Sites that bar robots from accessing them
- Contents of sites requiring you to log-in first (like some news organisations)
- Intranet pages (inside organisations)
- Commercial (paid-for) information services
- Library catalogues and non-web resources

A deep web directory, [Complete Planet](#), reckons this area is approximately 500 times greater than the surface web (gulp!), with more than 200,000 searchable databases. It offers to search more than 70,000 of them.

[Complete Planet](#) argues that it should be called the deep web because it is not invisible; it is just not easily accessible to ordinary search engines. It reckons 95% of the deep web is publicly accessible, without fee. You just need to go and ask the questions in the right places.

Much of the data does not exist like a normal web page – one that links to others. When you make an enquiry, the answer is generated from a database; the page has no existence beforehand, so a search engine has nothing to find or index.

Query [Complete Planet](#) and it calls up databases, listed in order of relevancy. The sources could be months old – the operation is not checking as frequently as the big search engines. But, with luck, it will take you to databases that cover what you want. Then you query databases individually, perhaps after signing up to gain access.

Another site offering access to much of the invisible web is [Search Systems](#), which covers 32,000+ public record databases. It's slowish if you are not paying for the service, and it concentrates on the US. But on its international listings, 'legal information' often covers the main government databases in a country. It works quite well as a directory or you can make searches.

A lot of this material may be found through specialist directories, gateways and the like. For instance, a mass of aviation information is available via a hub like [Landings](#). It's US-centric, but its many databases cover such subjects as planes, airports, licensed pilots, and air crashes.



If your local library can be accessed online, then you can make an inquiry and discover if a copy of a book or journal is available at a branch nearby. (For general inquiries about books in print, try [Amazon](#).)

You don't necessarily need specialist tools. Just look for databases: type into your regular search engine a request like 'air pollution databases' and start scanning the many places to make detailed enquiries.

The subject of the invisible web has attracted less attention in the last year or two, perhaps because search engines are getting better at indexing a wide range of different file types, and other material. Probably the best introduction is the book, [The Invisible Web](#), co-authored by Chris Sherman and Gary Price. It's getting old now, but look at Price's [Direct Search](#) site – it dates from 2002, but just scrolling down the long, long page gives you some sense of what is out there.

An indication of how things may develop comes with the recent launch of a trial for Yahoo! Search [Subscriptions](#), which enables users to access subscription files for such organisations as the international financial newspapers, the [Wall Street Journal](#) and the [Financial Times](#), and the big news archive, [LexisNexis](#). You cannot see the articles from their archives without contacting the companies (to pay for the access) – but the launch shows how big players are getting interested in this field.

Commercial information services may be important even to journalists without cash. Scan news archives like [LexisNexis](#) to find out what articles have been published, and then, if you find something promising, see if you can track down a copy elsewhere on the web. So much has been copied and republished that there is some chance of finding items on alternative sites.

There is one other place you can inspect: the past. [The Wayback Machine](#) is a repository of what people thought was transitory: over 40 billion web pages, ranging back to 1996. New pages start appearing on the machine after six months or so.

It is only a small selection of what has been published – the system crawls the web every few months, so it collects just an occasional snapshot of a site. As well as historical information about how the net has developed, it can show roughly how long a site has been active, or provide evidence of what was being said by a person or an organisation in past years – particularly useful in a place where public records may not be kept.

## 9 NEWS

Thus far, we've concentrated on finding general information. Now we want to look at what matters most to journalists – news. This section will cover some of the best news services on the net, then look at news delivery systems – how the news you want can be sent direct to you – and at archives for researching older material.

Many journalists go to individual sources for their regular news: a local paper or broadcaster, or a national news organisation. But the great thing about the net is that we are no longer confined to our own country's borders – online, we can access, just as easily, a regional network or publication, or a worldwide news producer, like the [BBC](#), or [CNN](#).

The big three in the search engine business play a pretty dominant role in the news supply business, too. They're not normally originators of coverage – they bring together, or aggregate, reports from a multitude of news producers, and package them for audiences.

### YAHOO! NEWS

[Yahoo! News](#) has long been an effective news supplier. It uses a mix of human editors, content from a hundred or so leading sources, and a searchable database covering over 7,000 sources in 35 languages. And it can all be customised to meet your personal requirements. Nielsen/NetRatings calls it the most trafficked current events/global news site.

The news home page serves up headlines for top news stories, in categories such as World/Business/Science/Sports. It typically offers a selection from one leading provider, like [Reuters](#) or [Associated Press](#); then you can click on tabs to see alternative lists from, perhaps, [Agence France Presse](#), [The Washington Post](#) or the [Canadian Press](#). Hover your pointer over the link, and you can read the intro of a story; click on the link, and go to the original site for the full text.

You can search the news coverage much as you would the web – use quotation marks around a phrase and asterisks for variations of words. The advanced search page lets you look for stories from the last hour or the last day, as well as longer periods. For other search operators, look at the [News Search help](#). You can save your news searches.

As a shortcut, write 'news' into the main [Yahoo!](#) search box, followed by a topic, like 'cricket', and get the latest headlines. Click on the title, 'news results for cricket', if you want more.

There's an added wrinkle for journalists: if you want to follow the work of a particular writer or photographer, put their name in [Yahoo! News Search](#), with the word author in front – as in 'author:Sainath' – or when searching on the full name, in curly brackets, as 'author:{Sebastiao Salgado}'.

[Yahoo!](#) retains Reuters and AP stories for two weeks, and reports from other providers usually for a month, though it can be as little as seven days.

Like many news suppliers, it offers links to alternative sources and background material. Thus for an important election or major crisis, its [Full Coverage](#) section offers links to websites, feature articles, opinion pieces, audio and video reports, and a picture gallery. (The links are more to international sources than local ones).

## GOOGLE NEWS

[Google's News](#) is revolutionary in that it is wholly computerised – humans do not make the selections. The links to the news stories on one subject are clustered together, with the total number of stories indicating the scale of worldwide interest. The report selected for display is not always the most obvious one – but this should be a less of a problem for journalists, used to assessing the quality of a news source.

Depending on robot editors may be off-putting at first, but regular use indicates the service is surprisingly quick and efficient at delivering breaking news and in covering a subject widely. Topics are updated every 15 minutes.

You can customise the news, by arranging [Google's](#) subject areas, including sections from their 20+ international editions, in various languages. Really advantageous, though, is that you can choose your own subjects: select keywords required in the stories you want (with a choice of nine languages), and [Google](#) will search them out for you on a regular basis.

This can be particularly useful in the middle of reporting a story: after searching for a subject, simply click on 'Add this search as a section on your front page' and it will provide regular updates on that subject, for as long as required.

The advanced search includes the choice of anything published in the past hour, day, week or month, or between specific dates, and you can also seek reports from one news provider of your choice.

[Google](#) says it gathers stories in English from over 4,500 sources worldwide. An unofficial survey over three months in 2005 showed [Google](#) using nearly 3,000 sources on its main home page (everything from the Wodonga Border Mail to the World Socialist website, though mostly from leading US and international sources). Articles are retained for up to a month.

It taps into wider sources than some competitors. Recent US research concluded that [Google News](#) contained more items than [Yahoo! News](#) that were in some way 'biased'. But this was not bias in any particular direction – it reflected that [Google](#) includes a greater variety of sources, encompassing alternative approaches to the news. Look at a trial system, [Newsiness](#), to compare results from the two

It may be a bit misleading to mention it here, because it is early days yet, but [Google](#) has filed patents on a system which would allow it to search news on a basis of quality and veracity. Companies like [CNN](#) and the [BBC](#) were “widely regarded as high quality sources of accuracy in reporting, professionalism in writing”, and thus could be of greater interest to its users. So it was looking at technology to pre-rank such news outlets, to ensure they appeared high in the results.

The [Google News](#) founder, Krishna Bharat, spoke in the summer of 2005 of how they were seeking subtler filters for news. They want to measure how many people are employed by a news operation, and how many foreign correspondents; they want to check whether parts of a story are simply copied from elsewhere. They want to take into account the length of story, the number of bylines, as well as the amount of traffic to the site and the variety of countries accessing it. They want to see if computers can recognise clarity, conciseness and originality – can software spot a scoop?

(If they are successful in their programme, such a qualitative control could be equally important for general searches.)

## **MSNBC**

[Microsoft](#) took a different approach, establishing [MSNBC](#) jointly with the US TV network [NBC](#). The resulting package, with emphasis on the TV programmes and strong video coverage, won it much success in the US.

[MSN](#)'s sites worldwide made less fuss about news, but an automated news service is under development: [MSN Newsbot](#) operates similarly to [Google News](#) – and there are already over 20 sites in different countries or regions (often, like Latin America, in an appropriate language).

[Newsbot](#) has stayed in a lengthy beta trial, but currently searches over 4,800 sources. It has clear clustering of the reports and photographs on each story. As yet, the news sources seem variable – in, for instance, the [Italian Newsbot](#) service, the sources were not comparable to those used by [Yahoo!](#) or [Google](#). And it also has that computer habit of picking random sources for stories, so you could get the top story on Palestine from the *Grand Forks Herald* (a story actually produced by AP).

The main [MSNBC](#) pages can be personalised, with different pre-formatted modules of news, including BBC news from different world regions. There is a novel system whereby your reading is monitored: Related Stories checks the news you are reading and displays the five stories most related to it. It will also track the 16 stories you read most recently, and suggest similar stories which might interest you. (You can delete items from your history, if you prefer not to have such a record kept.)

## OTHER NEWS SUPPLIERS

Solely devoted to news, [Topix](#) currently searches more than 10,000 sources, aiming to do so every 30 minutes. It categorises them into some 300,000 topics (48,000 celebrities and musicians for a start, plus countries, sports teams, businesses, diseases and lots more).

Click on a broad category and then, within the main page that comes up, click on the 'subsections' link to see what is on offer. Or simply do a search for what interests you. [Topix](#), established by some [Open Directory](#) pioneers, provides material which can seem overwhelmingly American, but the site searches widely; sometimes the international items just get rather swamped.

A British alternative, [NewsNow](#), has perhaps a simpler system, but fewer categories. It claims over 20,000 sources, searched every five minutes. Its layout emphasizes the most recent items, with top stories running alongside. You can search the contents, but most of its services are fee-based, so it only allows searches on a single word at a time. It covers weblogs as well as more traditional sources, so you may get something different. Most of the material seems to be from the US or UK, but it displays stories with the flags of the source countries, to give you a quick idea of where things are from.

For more services, try the news [portal](#) run by Belgian Wim Wylin. It's one very long page of links to just about any specialist news search engine you can imagine.

## WORLD NEWSPAPERS AND AGENCIES

None of the listings of newspapers and agencies around the world are truly comprehensive. But there's a simple directory by country, [Newspaper Index](#), kept by Danish journalist Hans Henrik Lichtenberg, while you can search by city, country or paper in [The Paperboy](#), from Australia. If the country you want is ill-covered by them, try [ABYZ Newslinks](#) or [Kidon Media-Link](#).

## ALERTS – get the news sent to you

You may not care too much about the detail of world events; rather, what you need is better coverage of your own country. This can be achieved by going out and searching for the latest news, or setting up a system which captures the news you need.

Presuming that you know the main online news sources in your own country (and, if not, use a local or regional search engine to check them out), you will want to augment what they say with any other available sources.

To read the latest news, it can be a tedious business if, each time, you have to seek the right sites, call them up individually, and then work your way to the items that concern you. Why not reverse the system, and let the news come to you? This used to be the prerogative of those with cash to spare, but now there are free alternatives.

News alerts have been the most common system for this, usually delivered to you by email – though increasingly there are other delivery options. For instance, the [BBC](#) (see links at the foot of its news page) offers breaking news alerts not just to your PC, but for mobile phones, palmtop computers and other wireless handheld devices.

You can access BBC news round-ups for about 20 broad categories, but also news on whatever detailed subjects you choose – from 'Armenia international football' to 'Brazil Amazon cotton'.

Google's [Alerts](#) system (currently still under trial) has four alternatives: you can be contacted when a new article enters the top ten of your Google news search, or when new web pages enter the top 20 results of a web search – or both, and it also covers [Google Groups](#).

It can do this as it happens, or once a day, or once a week. If it's a busy subject, and you are not overly concerned with breaking news, you may prefer the daily option, a single email rather than a constant flood. You can create ten alerts at a time, and more subsequently – if you've the capacity to absorb so much.

[Yahoo! Alerts](#) also now allow you to set the keywords for alerts, with the useful ability to exclude terms. They can be delivered as the news stories appear, or sent once a day – and the service operates by email, mobile or instant messenger. It also offers breaking news, from AP, and emails of news bulletins in [Yahoo! News](#) categories.

[MSN](#) has much less at present – breaking news alerts, and a few newsletters offering daily roundups – but they'll probably offer more before long.

At [Topix](#), you can get daily alerts on any of its regular pages – just click on the link at the bottom.

You will quickly find you need to refine your news alerts, even more than with a web search: too broad a topic and there's an endless flow of messages. And you need to be careful to exclude extraneous subjects: it was no good asking [Yahoo!](#) for news updates on possible Italian elections at a time when Pope John Paul II was dying – stories from Rome about his health usually included mention of the 'election' of his successor.

## WEB ALERTS

A wider canvas than news can be covered with web alerts, a system telling you when a particular page is updated. This sounds wonderful, but there can be trial and error in establishing them. A site may be updated in various ways – just because a page has changed, it does not mean it has been altered to include the sort of information you want. So don't set up lots of web alerts until you've tried one or two.

As well as [Google](#), there are various specialist monitors of web page changes, some free. Since successful free services have a habit of vanishing (or starting to charge fees), several are mentioned here:

[WatchThatPage](#) provides updates daily or weekly; you can select the time of day the pages are checked – thus it can deliver an update to you in the morning, local time. It will monitor changes in key words, as well as a page as a whole; it can email the changes that have taken place, or just a note of pages which have altered. There can be delays because of heavy use – the two individuals from Norway who run it seek contributions to help with the running costs.

Two that offer to monitor up to five pages for free are:

[Track Engine](#). A Singaporean system, it involves installing a button on your computer browser; it delivers the changed pages as a link, or as a summary of the new content, with the new material highlighted. A paid package allows much more.

[ChangeDetect](#). It gives you a 30-day free trial of the advanced service (including searching sites, as well as single pages, and password protected content, plus anonymous stealth monitoring of sensitive sites, so those running the site do not know it is being tracked regularly). After the trial, you only receive info for free once a week.

## REALLY SIMPLE SYNDICATION

This system, RSS, has the potential to change much of what we do on the web – instead of going to look for information, more of it can come to us. The idea, as a Yahoo! staffer put it, is “anyone can subscribe to anything”.

Currently it is a good way to see the latest headlines, from various sites. It's a type of automated websurfing, which also covers most blogs. It can present updates, usually in a short series of headlines, perhaps with a line or two of text alongside each item – it partly depends on how much space you want to devote to it. You click on what interests you and it opens the full story at the site of the news producer.

Sophisticated 'readers' – systems for accessing RSS – can customise how the items will be presented, showing only items you have not looked at before, or storing headlines. The simpler systems have the disadvantage, in comparison with alerts, that you need to check the RSS feed regularly or you might miss something.



They are a good way to keep up to date with what is appearing on a wide mix of sites. If you only have slow dial-up access, then time is saved by not having to check lots of sites individually.

There are plenty of specialist RSS readers, but for the moment a simple way to access feeds is via [Yahoo!](#), which displays them on a My Yahoo Page. To pick up a new feed, go to the site of a service you want and look for a little orange button saying 'RSS' or 'XML'. Right click on the button and then on 'Copy Shortcut'; go to My Yahoo's Add RSS page, paste in that shortcut and hit 'Add'.

Sites like [CNN](#) and the [BBC](#) simplify things with an Add My Yahoo! button alongside each of the RSS feeds they offer for different areas of coverage. As with alerts, you can set up an RSS feed for a particular inquiry: make your search and then click on the button Add to My Yahoo!/RSS.

My [MSN](#) now enables RSS incorporation, too. And you can click on the RSS button at the foot of the first page of results – for a news or a web search – and then get the top ten results of the search delivered to you by RSS, so you can watch for any changes.

There are specialist search engines for feeds (see later section on Blogs), but the regular search engines, like [Gigablast](#), covering blogs and XML, are moving into the field. Expect the big three doing likewise before long.

For more on RSS readers, see Search Engine Watch 's new [survey](#); there's loads more on the whole subject at Peter Scott's [RSS Compendium](#).

Talking about RSS feeds makes things sound technical. Don't worry, such facilities will soon be so universal that receiving news in this manner will be a normal part of net use. (RSS adverts are already on the way.) A plan for making things more straightforward across the board is reaching fruition: [Microsoft](#) is making RSS available as part of the new version of its Windows operating system, Vista, promised for late 2006, and is incorporating RSS in the latest version of its web browser, Internet Explorer 7, and its Windows Live system.

## **TICKERS**

News tickers deliver a running feed of headlines to your PC screen. You need to download the software, which then enables a stream of news, sport, business and other headlines across part of your screen. It sounds good, though it can be distracting.

The advantage is that you get instant updates drawn to your attention; you click on a headline to get the full story. But for all this, you normally need to be online, so if you are dialling-up to the net for short periods, it has less appeal. You might as well arrange an RSS feed to give you the latest headlines, and then when you dial in, skim them all in one go.



Free tickers tend to come and go. The [My Yahoo! Ticker](#), though still in the beta stage, may be the best bet at present. It's especially useful because it offers the opportunity to build a ticker out of RSS feeds from all over the web, not just from Yahoo's own news services. To learn more, see [Ticker help](#).

[BBC News](#) offered a free desktop ticker, which, among other things, allowed you to dial in and download an update, in a tiny 4k file, and then disconnect so it could subsequently run offline. During much of 2005, they were upgrading their system, and the final shape of the new service was unclear.

Tickers can be delivered to other devices – you can even get a ticker-toting watch in the US.

## **BUILD YOURSELF A PERSONAL INFORMATION SYSTEM**

You can also organise a Personal Information System, especially if you cover a regular beat or specialism. Study the sources you are consulting and decide which ones you want check regularly; then set up RSS feeds or web alerts to deliver all the information you want.

It is easiest to do this with broad topics, but obviously you get a mass of information. The material may be too much to download daily, or too much to read when you have got it. Thus, instead of general topics, refine RSS feeds to concentrate on what you really need.

For a field like African agriculture, select a single reliable digest of news rather than several different ones; and choose very targeted news alerts – perhaps only items on your specialism which refer to your country or the surrounding region. It won't take long to discover which sources produce material that is new and, more important, relevant to your audience.

If it cannot always be delivered, at least not for free, set up on your web browser a list of 'favourites' or 'bookmarks' to consult regularly. In Dar Es Salaam, covering Tanzanian agriculture, you might check international sites like, say, AllAfrica.com on [agriculture](#), alongside East African regional sources. Some you would check on a daily basis; others would be on a weekly round, or even monthly for very specialised sites. Build up a regular routine to check the best potential sources.

Increasingly it should be possible to set up an automatic system: if you want, say, the UN-operated [IRIN](#) news, they will provide an RSS feed for an individual country and for particular topics.

All this should help you keep on top of your field – and ahead of the opposition.

## **REPLACING THE CUTTINGS LIBRARY**

The free supply of news over the net is impressive – given the cost of gathering that news in the first place – but the main news search engines only cover the last 30 days at most, and you'll want to call up older stories. Some may be found through a general web search, but not easily. So you need a good news archive.

[LexisNexis](#), covers 3.8b documents from over 20,000 sources, but, without paying, you can only glimpse what is available. Most of the big international newspapers, and magazines, guard their old stories zealously, aiming to make money from them. Likewise, the big US papers expect you to pay, though [The San Francisco Chronicle](#), is a pleasant exception, offering free search back to 1995.

Among the few international news operations that do offer their old stories freely are [CNN](#) and the [BBC](#). Both provide nearly a decade of news coverage. For the Middle East, [Al Jazeera TV](#) goes back for a couple of years.

A newspaper with good world coverage that gives you free access – to its reports going back to 1998 – is Britain's [The Guardian](#); its new search system is pretty straightforward. Another British paper, [The Daily Telegraph](#), offers archives back to 1996 (you may have to sign up, to access the main newspaper site).

To see if there is a free news archive in your country, consult the Special Libraries Association. The main international [list](#) has over 160 news archives, and there's a separate page for around a hundred [Asian](#) news archives. They normally provide the age of each electronic archive, the language and whether there is any charge; some go back to the mid-1990s.

The lists are getting a little out of date, but will lead you to such papers as Egypt's [Al-Ahram](#), the [People's Daily](#) of China, [The Times of India](#), and Kenya's [Daily Nation](#), as well as some news agencies and a few broadcasters.

## **A THOUGHT FOR THE FUTURE**

After [Google News](#) and the MSN [Newsbot](#), you might want to see just how far the computerisation of news can go. Take a look at an American project, the [Newsblaster](#), from Columbia University's Department of Computer Science. It not only collects news from various sources, but automatically summarises what it finds into a single story on each topic.

There were some weirdly-written stories when the project was launched in 2001. Today's machine-produced summaries can still seem odd, but already there are quite a few which look like standard news fare.

The software is clearly raising its game. It is not going to put our colleagues out of business (not for a while, anyway), but as a demonstration of what you can already achieve without a human editor, it's quite impressive.

## 10 HELP FOR JOURNALISTS

The best journalism sites have often come from individual journalists who want to provide the sort of facilities that they, and we, require. Individual enthusiasm is rarely enough to sustain such a service permanently, and sites have faded as their creators ran out of energy or cash. But several remain recommendable sources of help.

[JournalismNet](#), from the Canadian TV producer Julian Sher, these days bills itself as the Investigative Guide to Internet Research, but offers a fine one-stop shop for material that matters to hacks. It has links to most of what you can think of, including news operations around the world, reference tools, story beats, journalism organisations, courses, jobs and more – though other journalistic commitments mean Sher struggles to keep it all up to date.

Working online? [CyberJournalist](#), from the one-time technology editor for MSNBC.com, Jonathan Dube, has news, commentary and resources. American-oriented, it nevertheless has much of general use. Dube, currently editorial director of Canada's [CBC.ca](#), offers tips and tools like online writing guides, examples of good work online, and lists of journalism blogs.

Another US-centric site, but important for the study of journalism is the Florida-based [Poynter Institute](#), a mid-career training organisation. It's well worth exploring for its tipsheets, articles and other resources on reporting, photo-journalism, editing and design; its tools and advice on the craft of writing are excellent. It's also strong on such issues as ethics and leadership. Poynter grew out of newspapers, but caters for broadcasters, too. The site has a search engine covering 220+ journalism sites.

There are also regional sites, such as the New Delhi-based [Media Foundation](#), which covers South Asian journalism as well as Indian media.

If you want to see what the rest of the world is doing on a big story, the [Newseum](#) offers the front pages of 460+ papers from 46 countries – it's strong for the Americas and Europe, weaker elsewhere. (This US site is image-based – slow loading for those with poor net connections.)

When a massive story breaks, the bigger sites swing into action with detailed lists of net resources for journalists covering the subject. [The Journalist's Toolbox](#), at the American Press Institute, and the New York-based [Media Channel](#), were two that did so in 2001 for the 9/11 World Trade Center attack. Among many offering breaking news help on the Asian tsunami in December 2004 was the [South Asian Journalists Association](#). Academic and specialist organisations also rush to put up background briefings.

A study of 9/11 search results showed it was vital in such a situation to go to NEWS search and not general web search. Even if ordinary web sites are updated quickly, it may be days before they are seen by search engines.

There are specialist sites for different types of reporting, like business or environment, or for specific skills like copy editing/sub-editing. To find them, consult the sites above, or look at [Power Reporting](#), from Bill Dedman, AP's first director of computer-assisted reporting. His site has positively exhausting lists for a considerable range of beats.

For points to remember whenever you are writing, try a checklist of my own, *Eight ways to save words and write better English*, (Section C under [Additional Resources](#)).

For information on wider journalism and media issues, try the [International Federation of Journalists](#), including authors' rights, gender issues, and trade union matters, the [International Freedom of Expression Exchange](#), and [Reporters Without Borders](#).

## 11 REFERENCE TOOLS

Compilations of reference facilities abound on the web. [Infoplease](#) has a useful selection, including almanacs for current information. Jim Martindale's [The Reference Desk](#) has an idiosyncratic touch. The journalism sites mentioned above list plenty.

But these collections are all-embracing. Out of the wealth of possibilities, we need to choose a few tools that do a really good job in finding facts and figures.

### FACTS

**GENERAL KNOWLEDGE:** One of the innovations of the web is the great co-operative encyclopaedia, [Wikipedia](#), which now attracts 50m+ hits a day. Written by volunteers, it spans over 700,000 articles in English, with smaller quantities in dozens of other languages. Individuals can edit the contributions, but their changes are monitored by teams of other volunteers, so there are controls over what appears. A real plus is that entries can be updated in a matter of hours when something significant changes.

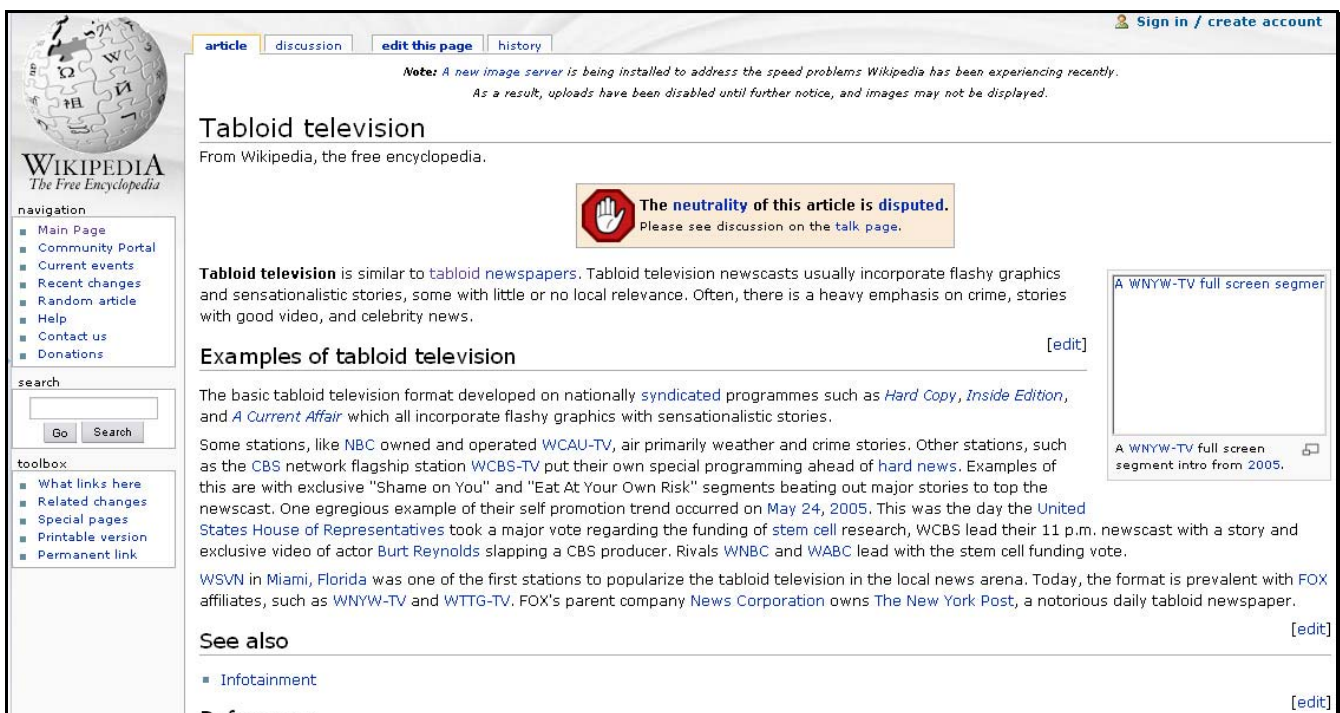


Fig 6: Wikipedia, the volunteer encyclopaedia

You may feel nervous about depending on variable volunteers rather than a system of paid experts. But after four years, it has built quite a reputation, concentrating on information from recognised, impartial sources; it avoids publishing original thought. Yahoo!, for instance, now features [Wikipedia](#) articles at the top of relevant search results. Take a look at a subject you know, and see what you think. You could be impressed.

Traditional encyclopaedias have migrated online and can be useful for general knowledge queries, especially in fields where current events are not likely to outdate entries. Free offerings tend to be smaller, or older, versions: [Encyclopaedia Britannica](#) offers only limited results unless you buy a subscription. Other possibilities include the [Columbia Electronic Encyclopaedia](#), with over 50,000 entries, and Microsoft's Encarta, through the free pass offered in [MSN Search](#) (their paid service covers some 60,000 articles).

Yearbooks or almanacs may lack changes that have occurred since publication; it's another area where the net can keep ahead. The online version of the [CIA World Factbook](#) is updated through the year. It offers a welter of facts and figures on the countries of the world, and is also downloadable.

**EXPERTS:** A reporter searching the net should know when to halt: even if the information is out there, you're not going to find it anytime soon. At such a point, you could turn to an expert.

[ProfNet](#) provides links to thousands of news and information officers in the Americas, Europe and Africa, and offers a searchable database of 16,000+ experts (mostly in the US, UK and Canada, it appears). They operate for email queries on weekdays from 9am to 11pm, Greenwich Mean Time. Profnet has probably been overused in the west, but is worth considering, so long as you are aware these can be sources with an official or commercial concern to press upon you.

You could contact [Ask Now](#), operated by the California library services. But to benefit from the knowledge of a librarian, you need to have internet chat enabled on your computer, so you can converse on screen over the net.

To explore more, [JournalismNet](#) has a good round-up of sources for experts from various countries and different disciplines. [RefDesk](#) has other possibilities.

## FIGURES

**STATISTICS:** The [UN](#) is a central starting point. Go to the [Statistics](#) division – as well as social indicators for the member countries, there are links to figures for trade, environment and much else. Go to the different world bodies, like the [World Bank](#) or the [Food and Agriculture Organisation](#), for detail on specialist areas. Via the home pages of the UN missions for each state, there are links to government information for each country.

[NationMaster](#) offers country statistics on a mass of subjects (over 4,600 stats, at last look), but you can only publish up to 20 individual figures or lines of text.

Don't ignore educational resources: the UN's [InfoNation](#) is a really straightforward site for creating charts of comparative data from groups of countries. You can produce bar charts on anything from crime to tourism. As a way of providing an instant customised illustration, for articles on subjects ranging from TV ownership to refugees, it's worth trying. (Email [cyberschoolbus@un.org](mailto:cyberschoolbus@un.org) for permission to use material in a commercial publication.)

**CALCULATIONS:** For everyday computations, use a search engine like [Google](#).

To convert all sorts of amounts, from wind chill to wire density, as well as metric measures and the like, try [MegaConverter](#). It's good for abstruse matters, and even ancient measurements – find the length of a Roman cubit.

An intriguing tool is the [inflation calculator](#) available at the [Columbia Journalism Review](#) site. It converts a US dollar figure, from 1800 onwards, into contemporary values by adjusting for inflation. It only applies specifically to the USA, but as well as historical uses, it can quickly indicate whether more is being spent on something now than a decade or two back.

[Economic History Services](#) can do this for both the US dollar and the UK pound, including figures for purchasing power going back centuries more; their permission is needed to publish the results.

**MONEY:** Plenty of search engines and portals will convert foreign currencies, but watch when little-traded currencies were last updated. A specialist site is [XE.com](#), from Canada; it aims to cover every known currency.

## WORDS

**DICTIONARIES:** There seem to be shelves of dictionary sites. But given that [Google](#) provides word definitions, why bother going elsewhere? After all, it would be another site to remember, or add to the lengthening list of bookmarks.

Reason number one is that [Google](#) may not do the job.

If you come across a word you don't know – like 'refettle' – you can do a search on it. If it comes up on top of the [Google](#) results with [\[definition\]](#) next to it, you can find the meaning. But [Google](#) did not know the word.

Neither did [Yahoo!](#), nor [MSN](#).

So the obvious answer: try some dictionary sites, like [Dictionary.com](#) and [Dictionary.net](#). They had not heard of it either.

In fact, this was a query the net could not easily answer. After much time fruitlessly seeking illumination, the nearest to a definition came from studying the references collected in an



ordinary web search. It was possible to work out from the context (mostly in reports on motor-racing) that the word means redo/repair/improve.

The dictionaries may offer a definition, but it might not be what you need. Neither [Google](#) nor [Yahoo!](#) could help with the meaning of 'haver'. Some specialist sites did better, offering meanings such as 'oats' and 'talking foolishly'. But they missed out more contemporary usages, including a software project of that name and the meaning 'to vacillate' or 'equivocate'.

[OneLook Dictionary Search](#) (which consults 993 dictionaries), produced a list of ten relevant dictionaries. Several offered what was required: the American [Merriam-Webster Online](#) defined 'haver' as "to hem and haw", which was fine, except that some people might then need a definition of the definition. [AskOxford.com](#), based on the *Compact Oxford English Dictionary*, was probably the best, with the definition: to act indecisively.

The conclusion? The search engines do well enough with ordinary words; if it is something obscure, then be ready to try a series of sites – and don't take the first suggestion, either.

If, by the way, you are struggling to find an alternative for a word, try [Thesaurus.com](#).

**STYLE GUIDES:** There are no absolutes when it comes to writing styles: it is up to an organisation to decide what is best for its readers or listeners, in terms of grammar, spelling, names, titles etc.

But if no-one is laying down the rules for you, consult a manual from an established media concern. AP style is commonly used, not only in the USA, but the AP Stylebook must be paid for – likewise the Reuters guide.

There are several UK guides offered freely over the web (downloadable in some cases):

- [The BBC](#)
- [The Times](#)
- [The Economist](#)
- [The Guardian](#)

**QUOTATIONS:** The net should be a good place to find quotations – they can be easily indexed, and it is possible to put a phrase or sentence into a search such as Google to find out who said it.

There are plenty of sites for quotations, but they are none too comprehensive, and are weak on recent quotes. Scouting across the net may be the only way to find or check a quote, but be warned that it is very easy to get wrong answers.

"If we make peaceful revolution impossible, we make violent revolution inevitable" is a shrewd and oft-repeated quotation.



Search for it in [Google](#), and you get perhaps 25,000 references, explaining it is a statement from US President John F Kennedy. But did he actually say those words? After all, it's a slightly surprising comment from a US leader.

You might think some 25,000 sources cannot be wrong, but search further, for as reliable a citation as possible. [Bartleby](#), a well-known site for quotations, offers a version from the *Columbia World of Quotations*, with a source, the official Presidential Papers, plus the place and date in 1962.

It turns out the actual statement was more general, not necessarily implying that America might be responsible. Kennedy told Latin American ambassadors: "Those who make peaceful revolution impossible will make violent revolution inevitable."

Be aware that people think they know things and write them inaccurately on the net; then others copy them. There is no safety in numbers.

### **TRANSLATORS:**

Machine translation is one of the boons of the net. While it used to be good enough to give only the gist of the original text, it is increasingly reliable. Nevertheless, human intervention is still required to ensure accuracy and make perfect sense.

Take a random example of a machine translation, by [Google](#):

*Tres asesinatos en menos de una semana han decidido al gobierno brasileño a enviar un contingente de 2.000 efectivos militares para evitar el desborde de la violencia en el estado amazónico de Pará, en el norte de Brasil.*

*Three murders in less than one week have decided the Brazilian government to send a contingent of 2,000 cash military to avoid the overflowing of the violence in the amazonian state of Pará, in the north of Brazil.*

Typically, this translated passage can just about be understood, but requires editing to be used in a news report. It also needs a check on the meaning of the word 'efectivos', translated as 'cash'; you could find out from an online dictionary – and typing 'Spanish dictionary' into a [Google](#) search is probably easier than going to a reference site. (Here 'efectivos' means 'effectives' i.e. troops.)

Some passages still turn out as gibberish, and the enthusiasm of software for translating every word means it will try to translate names into words it understands. Try an awkward passage in a variety of translators.

Both [Google](#) and [Yahoo!](#) offer a 'translate this page' link alongside search results in foreign languages. Like most web translators, they utilise the Systran system, from France.

[Google](#) has improved its facility by a huge throughput of translated documents, from the UN and elsewhere – educating its own system.

It has a 'language tools' button on its search page, offering to translate, between a variety of languages, either a text (which you paste in) or a web page (write the URL). It provides 18 pairs of languages – a 'pair' would be English into Chinese, or German to French.

Yahoo's [Language Tools](#) translator is tucked away, but will translate text or web pages between 36 language pairs.

Alternatives to consider (some can get busy and be hard to access):

- [Free-Translator](#) has quick access to a variety of translation tools and dictionaries.
  - SDL's [FreeTranslation](#) has 17 language pairs, handling up to 750 words at a time.
  - Smartlink's [Im translator](#), powered by PROMPT, offers 20 pairs, including Russian.
  - [World Lingo](#) has dozens of language pairs. It will also translate email messages before sending them.
- 
- [Your Dictionary](#) is a good source for multi-lingual dictionaries, covering around 300 languages.

## PEOPLE

Tracking people down has been aided by the net, but not necessarily simplified. Start by Googling their name – if the name is clearly distinguishable, you may find clues to their whereabouts.

[ZoomInfo](#) uses references it finds to build a summary, or brief resume, of people on the net. It is still under development, and erratic, but making progress – I was listed in spring 2005 as in London, Arkansas; by the summer, I had been moved (correctly) to the UK.

If people don't appear clearly in a simple search, then look for alternatives. If you are trying to trace someone, think first about where their details might appear. The more unusual their name, and the more specific a location, the better chance you have of success.

If a person uses the net a lot, it can be worth pursuing email addresses, or searching for users of mailing lists or discussion groups. (See Chapter 17 on Groups.)

The vast number of email users worldwide means most people can no longer be identified from their address. Email search also got harder with the advent of spam. Big email providers bar access to their lists, and individuals avoid leaving their addresses around on the web, for fear of attracting a deluge of rubbish.

[MESA](#), MetaEmailSearchAgent in Germany, will try several email searches for you. [Infospace](#) offers international coverage, of variable value.

A landline telephone is still a useful way to locate people. You may have local online directory inquiries (as, in the UK, [British Telecom](#)), but for anywhere abroad, the place to start is [Teldir](#), [www.infobel.com/teldir](http://www.infobel.com/teldir), a great Belgium-based listing of online phone directories across the world. Even if it might be taxing to find the Comoros or Vanuatu on the map, their phone directories are close at hand.

Depending on what exists in a country, you may get white pages (residential) and yellow pages (commercial) directories, plus other business directories and access to various locator sources.

An additional possibility is reverse look-up – if you have a number, you may be able to get the name and address, though this is mainly available in North America and parts of Europe. (See [JournalismNet](#) for a selection of reverse directories.)

The [Ixquick](#) metasearch engine aims to interrogate international phone directories for you. None of these systems are likely to be truly comprehensive: with the break up of monopoly phone companies, people use different suppliers; others will hide their numbers ex-directory. And, of course, people just use mobile phones, or have none at all.

A lot of the tracing tools offered on the net are not terribly useful: they only work well in a highly computerised society like the US, with many public records online. ([Pretrieve](#), for instance, accesses all sorts of US public databases, letting you browse through things like criminal records.) Even there, you may have to pay to find what you need.

## PLACES

**TIME:** Knowing the time in other countries can be important if you are trying to contact people, or check when events happen. [Time Zone Converter](#) starts automatically with the time where you are, and you select the place where you need the current time. An alternative approach from [World Time Server](#) lets you set a time in one place, and then check the time somewhere else. These sites also indicate where they have taken into account seasonal changes due to daylight saving hours.

At [TimeandDate.com](#), you call up a continent and there is a display of the current time at all the cities covered. A 'personal world clock' lets you check at a glance the times for up to 25 cities. It also offers calendars for a country, showing public holidays. And if you want to count the days, hours or even minutes until some significant date, personal or public, like Christmas or a New Year, it will provide the figures.

For phone calls, you could use an Australian site, [White Pages](#) – write in the places you are dialling from and to, and it tells you the time at either end, plus the dialling code.

**WEATHER:** If you simply want to know the weather somewhere, write 'weather' and the city name into [Yahoo! Search](#). Or [CNN](#) is pretty efficient. For greater detail, try [The Weather](#)

[Channel](#). For the typical climate in places at different times of the year, see [Weatherbase](#); you can set it for Fahrenheit or centigrade.

**TRAVEL:** Roving journalists will appreciate [Steve Kropla](#)'s help site, with details of electrical and phone connections around the world, including images of the plugs you need and facts on computer modems. It also covers mobile phones and local TV standards. But the site had not been updated since 2004, so may become less reliable. An alternative for plugs and modem adaptors is the [Voltage Valet](#).

The [Cybercafe Search Engine](#) currently lists nearly 6,000 internet cafes and public access points, in over 160 countries. The [The List](#) has details of internet service providers (ISPs) around the world.

For those with international cash cards, find automatic bank machines (ATMs) at [Visa](#) or [MasterCard](#).

When seeking flights, hotels etc, there are more travel sites than you can shake a stick at: a local directory or search engine is probably the best starting point to find one suited to your needs. For information on particular countries, the [Tourism Offices Worldwide Directory](#) can be a useful launch point.

**MAPS:** To find a local map, writing the place name and 'map' into a search engine may produce better results than using the leading mapsites, which lack detailed coverage for many countries.

For city maps, with at least some street detail, the [Perry-Castañeda](#) collection at the University of Texas has quite a selection, along with links to other map sites. For more general maps, go to its overall [listing](#). Or another American source, [Embassy World](#), has links to a fair mixture for each country.

The potential level of information becoming available is evident from the latest US coverage by the major search engines.

[Google Earth](#), with downloadable software, offers the ability to swoop down from space, looking at ever more detailed satellite images, and then searching for streets or post codes on the results. Within days of Hurricane Katrina hitting New Orleans in August 2005, it was offering satellite pictures of the destruction.

It's not just in the US, either: Already there is coverage of such world cities as New Delhi, Beijing and Moscow.

[Google Maps](#) had already been offering, for some US urban areas, satellite views linked to its maps locator. (Not quite everything is visible: looking at Hawaii, the Pearl Harbour Naval Base was distinctly fuzzy; security considerations presumably affect coverage.)

Both systems enable you to drag and drop the maps and images – much easier than online map navigation in the past.

MSN has been trialling its [TerraServer](#) aerial photograph collection for some US cities, with even greater detail: you can see the cars in people's driveways. Its [Virtual Earth](#) system is not yet so smooth as Google, but still impressive.

On a trial maps query for an address in Santa Cruz, California:

[Google Maps](#) showed the best street map, with a big red pinpointer, and then a colour satellite image – and a hybrid version allowed the overlay of the street names on the photographic image;

[Yahoo! Maps](#) offered a reasonable map (but seemed somehow inadequate without an aerial picture to go with it);

[MSN Maps](#) had a more basic map (the TerraServer service was not covering it, yet.)

All provided links to local businesses and services – the main point of the systems.

**HEALTH & SAFETY:** Travelling abroad, you can use the [Traveler's Health](#) service of the US Centers for Disease Control and Prevention, with much on diseases, vaccines and precautions. It is designed for prosperous westerners, so the advice may not always be practical. Another possibility is [Mdtravelhealth](#), with useful links to other travel information sources.

The travel safety warnings from the US government are also aimed at US citizens, who may be a good deal more trepid than the average citizen from the developing world – let alone a reporter. [Consular](#) information sheets about each country cover issues such as security, crime, and internal travel. The [British](#) and [Australian](#) governments are among other states offering such a service.

A recently-established organisation devoted to keeping you in one piece, the [International News Safety Institute](#), has advice on covering conflicts and safety equipment. As well as bulletins from hotspots and information on attacks against journalists, it has links to groups for helping journalists in trouble, like the [International Press Institute](#) and the [Committee to Protect Journalists](#).

## 12 CHECKING WHAT YOU FIND

The web is not some orderly library, tended with care by diligent scholars; sometimes it's more akin to a garbage heap. As a reporter, you need to be expert at sifting out the rubbish, and at detecting the dubious claims and biased presentations that abound.

### EVALUATING WEBSITES:

In 2004, the city council of Aliso Viejo in California nearly passed a law banning foam cups because they contained a 'dangerous substance', Dihydrogen Monoxide (otherwise known as water).

A spoof site, entitled [Dihydrogen Monoxide Research Division](#), warned of this 'potentially lethal' constituent in the cups. The official-looking site (*Figure 7*) gave plenty of clues to its fictional character (the daft picture of a 'DHMO-tainted cow' should have given the game away). But the municipality initially swallowed the bait.

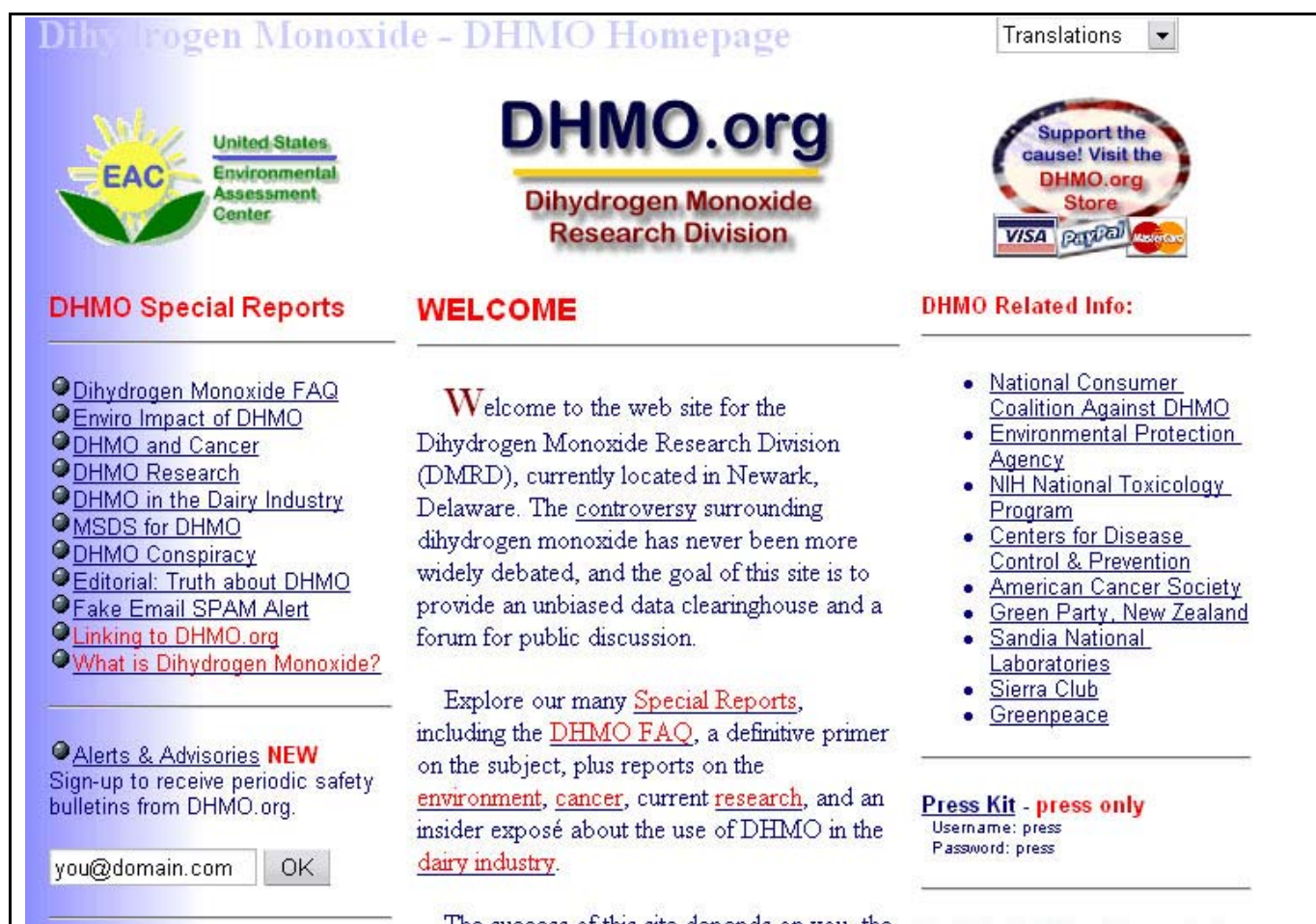


Fig 7: An official-looking warning that is entirely fictitious



It was a prank, to show users' gullibility, but other fakes on the net have worse intentions: conmen, crooks and terrorists all seek to confuse us, for their own ends.

Believing what you read on the net can kill you – like the medical site recommending a drug dosage 100 times more than normal. As a journalist, if you casually pick up material from the net (on racial issues, religion, or a current conflict), and paste it into a story, you may do damage on a much greater scale.

Once you've found a website, you need to work out whether it is worth using. The basis of website evaluation is simple: use your common sense, and normal journalistic caution, about any source you don't know.

**WEBSITE INTEGRITY:** A scale for determining the integrity of websites, on the [lines](#) drawn up by Steve Miller, assistant to the technology editor of *The New York Times*:

- 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.
- 2) Universities: Academic institutions offer a level of authority – this may vary, but it is something to depend on. Most studies by recognised experts are still reviewed by their peers, so the information is likely to be good quality.
- 3) Special interest groups: Non-governmental organisations and pressure groups may push a particular line, but if they are recognised bodies, you, and your readers, have some idea of what is being provided – it might be Transparency International, the Caracas Chamber of Commerce, or the Red Crescent. Companies and commercial sites could be regarded similarly, though the reliability of the site for an internationally-known brand would be different from an unheard-of dotcom.
- 4) Everything else: Unidentifiable organisations, personal sites, hobbies, obsessions etc. If it looks as though the person responsible is an expert, try to check them out independently.

But what you see is not necessarily what you get. Plenty of commercial sites try to pretend they are a genuine article – big celebrities always seem to have many "official" fan sites.

Lots of sites seek an authority they do not deserve; others go much further to mislead you. During the wars in the former Yugoslavia, there were sites which claimed to support one side, but were propaganda put up by their opponents. (Operators were clever enough to watch for when a prominent site went down for a period, and then immediately launched a site in its place.) Elections are prone to web activities from groups masquerading under false colours.

There were photos of the 2004 Asian tsunami which turned out to be from completely different events (and appeals for financial help on behalf of equally phoney people).

Digital technology means that it is increasingly easy to falsify the contents of photographs – it's no longer just people like Josef Stalin who can put themselves into a picture, or airbrush others out of it. A good introduction to photo fakery comes from [OEmagazine](#).

Computers can add to the problems. Automated news selectors occasionally come a terrible cropper, as when [Google News](#) in November 2004 led its coverage with the arrest of President Bush in Canada on 'war crimes' charges (*Figure 8*). Unfortunately, their software had picked up a spoof news item from a political satire site. (Perhaps we should blame human error for including the site in the first place.)



*Fig 8: Google News in November 2004 – tricked by a satire site*

So how do you assess what you find?

## TEN POINT WEBSITE EVALUATION CHECKLIST

(Developed from the system of the [Poynter Institute](#))

- **AUTHORITY:** Is this a recognised expert, or a body with a known reputation?
- **AFFILIATION:** Is the site connected with a reputable person or organisation?
- **ACCURACY:** If you can spot mistakes while reading the site, then start worrying.
- **APPEARANCE:** Is the site well put together? A reliable site may look old-fashioned, but a sloppy or amateurish presentation could indicate an individual, or some fly-by-night operation.
- **COMPREHENSIBILITY:** Does it make sense? (But remember: people writing in a second language may be less clear than native speakers.)
- **CURRENCY:** Is it up-to-date?
- **LINKS:** Look at respected sites on the subject: do they link to the one you are examining?
- **PAGE RANK:** If you use a Google Toolbar, it will offer a 1-10 rank on how a page is regarded on the web. Several other search services offer such rankings.
- **OBJECTIVITY:** Are there signs of bias?
- **CREDIBILITY:** A simple test: do you believe it? Does common sense tell you it is true?

Don't feel you should use these criteria every time you check out a site; just be aware of the pitfalls.



## **SIMPLE CHECKS YOU CAN MAKE:**

See who links to the site. Search in [Google](#), [Yahoo!](#) or [MSN](#) by writing the address preceded by the word 'link' – as with: link: [pajhwok.com/](#), to find who links to this Afghan news agency. (With [Yahoo!](#), you must write the full address: link: <http://www.pajhwok.com>).

Remember [A9](#) provides site info alongside its search results, including the number of links to that site and how long it has been in existence; an allied search service, [Alexa](#), usually provides even more about a site.

If you cannot work out how up-to-date a site is, write in its address bar: 'javascript:alert (document.lastModified)'. A box should then appear with the date and time it was last changed (sometimes, many months ago).

If in doubt, [Google](#) it. People on the net are compulsive communicators – in groups, blogs, sites, they are forever passing on information, offering critical judgments, analysing what they have found. You may learn enough to make up your mind.

Is there some way to contact those behind the site? With a reputable organisation, you would hope for a proper address and phone number. If there is not even an email address, treat it with caution.

Who runs a particular site? You can make a check at [MarkMonitor](#). For the most detailed information, consult the official registries: go to [ICANN](#), which has links to the Internic WhoIs registry and regional registries around the world (APNIC, LAPNIC, RIPE and ARIN). These WhoIs databases used to yield not only an address, but a telephone number for the administrator of a site; increasingly the most you get is a contact for a web hosting company, a proxy organisation, or perhaps a Cayman Islands bank.

**FAKES:** There are sites which specialise in uncovering fakes. One, correlating information, is a US government bureau with a sense of humour: [Hoaxbusters](#). It has a wonderful collection of scams, scares and phoney appeals, along with advice on recognising hoaxes and email frauds, plus what to do when you receive a virus warning which might not be genuine. Search its archive for anything that sounds dodgy. Links at [Purportal](#) explore other collections of hoaxes.

Unlike a library, the web doesn't have a big sign saying FICTION over one corner. A vast edifice of invention may have no warning for the unwary.

Some spoofs are such over-the-top fantasy that you presume no-one would be taken in. They have a life of their own: you not only find reports on the assassination of Microsoft founder Bill Gates in 1999 (oh, you missed it?), but arguments over the 'commission of inquiry' into this non-existent event, at a site entitled '[Citizens for Truth](#)'.

Hoaxes can be a useful corrective to media gullibility: in 2003, Computerworld magazine reported a radical Islamic group claiming responsibility for the Slammer virus, in a 'cyber jihad'. Unfortunately, the email interview with Abu Mujahid in Pakistan turned out to be with a freelance journalist in the US called Brian McWilliams, who was running the false [site](#). (It wasn't just media and security officials who were caught out; McWilliams also received emails from people wanting to join the terrorists.)

**ADDITIONAL ADVICE:** If you want more on evaluating websites, the library of the University of California at Berkeley offers detailed [instruction](#), along with links to further resources.

For more on checking domain names, try sources in [The Journalist's Toolbox](#).

A basic reminder: it is vital that information you get from a website is attributed to the source (just as in any other journalism). After all, it helps to protect you if something is wrong.

## 13 MORE HAZARDS ON THE ELECTRONIC SUPERHIGHWAY

### ETHICS:

A big 2005 survey, by the [PEW Internet and American Life Project](#), suggested two-thirds of net users were unaware of the difference between paid and unpaid search results. People do not appreciate that many search engines are full of concealed advertising: links get to the top of their search lists because companies pay to be there.

This guide has tried to avoid such offenders. The main search engines listed here now show separate links on their results pages signposted as 'sponsored' or 'featured'; even if it does not sound like it, presume these are all adverts.

Growing efforts by the search engine marketing industry to manipulate results, to ensure their product or service tops search lists, mean that what you see may still not be an unbiased choice.

Remember also that online operations do not necessarily have the standards you might expect from an established newspaper or broadcasting company. One head of UK editorial content for the [AOL](#) portal said it was his job to get readers to go to the areas promoted by the company's paying partners.

### SPELLING:

Do not rely on computer spell-checkers. They are good for highlighting errors, but they won't catch them all, and can introduce new mistakes. Dow Jones Newswires in the Americas banned reporters and editors from using spell-checkers, after too many incidents where the proper name of a person or company was changed into something unrelated, because of a spell-check suggestion.

If you've doubts, read: [Ann Owed two the Spelling Checker](#) (try pronouncing it). It's a spelling horror-show that sailed through a [Microsoft](#) checker without a hitch. ("There are know faults with in my cite.")

And, of course, there are names: [Google](#)'s search claims 50,000 results for: George Schultz "Secretary of State". A search for this 1980s American leader as Shultz brings 58,000 results – his correct spelling, only just ahead. Careless people endlessly recycle mistakes. Anything you are even slightly unsure about should be checked at the most authoritative official source.

### **WATCH THOSE GOOGLE STATISTICS:**

The sort of figures used above should be treated with caution. The net can encourage lazy journalism, and one particular habit among writers is using the number of results delivered by [Google](#) to demonstrate a subject's relative importance. The number of hits is adduced to show how little-known something is, or how popular a person might be.

Don't use such figures to create a headline. Studies have shown these numbers can be very rough, fluctuating considerably from day to day. You don't know what the totals include – there may be many results unrelated to your search subject.

### **THINGS CHANGE ALL THE TIME:**

This book will contain errors. Even if it was perfect when it was complete, the net is developing so fast that, by the time you read this, there are sure to be some URLs which do not work.

Websites vanish overnight. Companies change or go out of business at a speed unusual in the traditional media.

If something does not appear as expected, scout around. When a site is defunct, there may be a forwarding link.

If a long URL cannot be found, the material may have been shifted within a site, and given a new page name. Try cutting back the address: delete words from the end, back to a forward slash mark; if that does not work, keep cutting back, till the domain address if necessary, and then search within the site.

### **AND FINALLY...**

Research on how students used the net showed most would stop searching at the first answer they found. The report on this was entitled: ["Of course it's true; I saw it on the Internet."](#)

Try never to say the same.

## 14 NET CODES

The net has always been awash with initials, but things are getting worse. It's worth understanding the system of website addresses, but don't worry about learning the details.

### DOMAIN NAMES

Material appears on websites under domain names: starting usually with 'www' (representing the World Wide Web), then with the name of the operation concerned, followed by domain names or codes which may define its type and place, like '.com' or '.gov', or country locators, like '.cl' (Chile) or '.jp' (Japan).

The system was invented by the Americans, so, though '.us' is now being seen, most of their sites have no country code.

For such sites, the final set of letters – the top-level domain – supposedly defines the sort of organisation behind a website, using the following abbreviations:

.com	commercial companies
.edu	universities
.gov	US government sources
.mil	US military sites
.org	organisations
.net	operations linked to the internet

But the last two are very vaguely defined, and in practice you have much freedom in the domain name you wish to pay for.

Domain codes may tell you little about the site operator: '.com', by far the most popular, is used for anything and everything.

The 21<sup>st</sup> century has brought a great expansion of top-level domains, though you may not have noticed yet. A slew of new names were introduced (some with much tighter controls over who could use them):

.biz	businesses
.coop	cooperatives
.info	general use
.name	individuals
.museum	museums
.aero	aviation industry
.pro	professionals (currently being introduced for people like lawyers and engineers)

The recession in the dotcom world cut back on new registrations generally, and thus few examples of the new designators were spotted on the electronic highways. Some 3m '.info' names have been registered and about a million for '.biz', but they have remained almost invisible among the mass of '.coms' and the like.

Now another bunch are on their way. Among those already approved are:

.jobs for human resources managers  
.travel for airlines and hotels etc.

Argument continues (September 2005) over:

.xxx pornography

The system is overseen by [ICANN](#), the Internet Corporation for Assigned Names and Numbers – a co-operative body representing different parts of the internet world, though ultimately still subject to US control. ICANN's choices can seem idiosyncratic: Catalan culture and language gains a new top-level domain, '.cat' – not the most obvious priority.

There is '.int' for international bodies set up by government treaties, like the [World Health Organisation](#) and [Interpol](#). The [European Union](#) uses '.int' at present, but an '.eu' designation should be operational by 2006.

## COUNTRY CODES

States can choose the second-level domains they prefer, which means codes are not standardised. For instance, the UK uses '.ltd.uk' and '.plc.uk' for different types of companies, as well as '.co.uk'; India, meanwhile, is introducing '.firm.in'. India has '.ind.in' for individuals, while Britain has '.me.uk'.

Some countries use '.ac' for academic or university sites, unlike the American '.edu'. And there are other causes for confusion: '.ca' is Canada; '.ca.us' is California.

Country domain names are not set in stone: '.bu' disappeared long ago, replaced by '.mm' after Burma became Myanmar; likewise '.zr', for Zaire, has switched to '.cd', for the Democratic Republic of the Congo. Hong Kong, '.hk', and Macau, '.mo', survive as independent entities on the net, though they are now part of China, '.cn' (and also appear as '.hk.cn' and '.mo.cn').

A country code can have a dual use: the Pacific island state of Tuvalu may not have many surfers (of the electronic variety), but it has done well financially because its designation '.tv' has been snapped up by television companies.

Thus the country designator does not tell you whether a site is owned by a local citizen, or run by a business half a world away. Few countries demand you are actually a resident before letting you register a local site name. And it is hard to search all sites from a country, as local people using generic domains like '.com' will be excluded.

Don't think of memorising all these permutations: a complete list of country codes is at the [Internet Assigned Names Authority](#). But if you're wondering where a document with an address ending in '.aq' actually comes from, it's usually easiest just to open the file – and discover it stands for Antarctica.

## **URLs**

An individual file or page you can access via the web has a URL. (Universal, or Unique, or Uniform Resource Locator – quite a few names on the net vary because there is no central authority in charge of determining such matters. Custom and practice mean that, in time, one name usually becomes the standard.)

You don't need to concern yourself that HTTP means HyperText Transfer Protocol; just note that the initials are the start of a full web address. Nor do you have to bother, much of the time, about typing the full URL for a web page; these days 'http://www.' is often omitted from the start.

New designations, like 'www2.' or 'www3.', or 'www-sul.', are not significant – they are commonly used because the amount of traffic to that site requires extra computer servers.

If having difficulties with a particular URL, make sure that, in any words or characters after the first forward slash, capitals and lower-case letters are written exactly in the way shown.

## **STRANGE INITIALS YOU MAY ENCOUNTER**

Parts of the net use a sub-language of abbreviations and acronyms, and these can be prevalent in chatrooms and instant messaging. Things like CUL8R, see you later, and AFAIK, as far as I know, save time and space, and have taken on a new life in mobile phone texting.

Even if you never use them, be ready to check the meaning if they occur in a message that concerns you – acronyms can stand for a variety of things: a message ending LOL could be sending you 'lots of love', but LOL is more commonly 'laughing out loud'.

There are also smileys/emoticons, little images created out of punctuation and letters – the originals were :- ) and :- ( [Turn your head to the left to see happy/unhappy faces.]

For more on all this, consult the [NetLingo](http://netlingo.com) dictionary, netlingo.com.

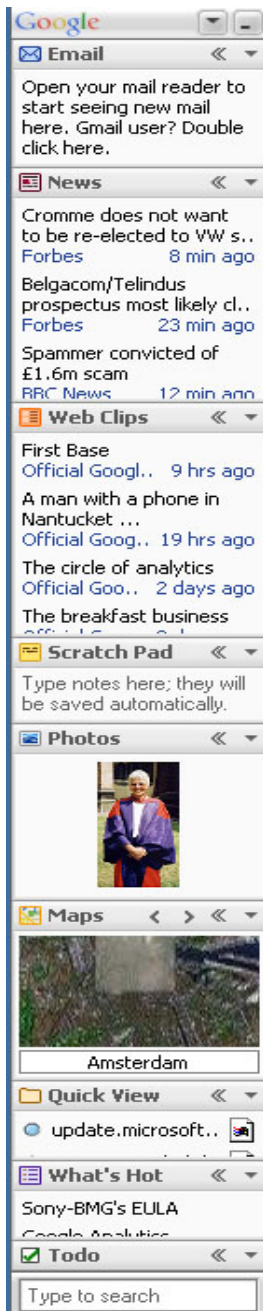
## **NET USAGE:**

About nine hundred million people, by mid-2005, were using the net – one in seven of the world population. Around a third each were from Asia and Europe, and a quarter from North America.

The proportion of users in each region was rather different: over two-thirds of North Americans used the net, and over a third of Europeans, but less than a tenth of Asians. It was about one in ten of Latin Americans and Caribbeans, about one in 13 in the Middle East and around one in 66 in Africa – but usage in these latter regions has been growing faster. (Source: [Internet World Stats](http://www.internetworldstats.com).)

## 15 DESKTOP SEARCH

Until recently, it could be easier to find something on the net half way round the globe than in a file on your own computer. Now desktop search is becoming the norm. The subject is listed separately here because you need to download special software.



It became urgent as people created and downloaded more material on their PCs. Even with efficient filing, it's hard to keep track of the categories chosen, let alone where particular files have been deposited. The main companies all offer desktop search now, usually integrating web search with a simultaneous check on your files. It often provides a valuable reminder of something you filed or viewed previously.

Google's [Desktop Search](#) has just sought to trump the opposition, with a new system displaying facilities on a customisable 4cm-wide sidebar at the edge of the screen (*Figure 9, left*).

It's a bid to take over part of your screen territory. In return, it displays news headlines (personalised on the basis of what you have looked at previously), RSS feeds, a rotating picture gallery, incoming Gmails, weather, real-time stock prices, and a list of your recently visited sites.

Click on an item, and a pane with more info slides out. You can have the latest news and alerts popping up in the sidebar – a bit like a ticker.

The Google Sidebar takes a fair chunk of space. You can hide it as necessary, but you may prefer the little Deskbar, with just a search box; it can go anywhere on your screen.

Cached pages from past web searches are included in results, so you can choose between images of the different versions you've seen. There is an advanced email search, selecting just, say, incoming messages.

It now accesses a wide mix of files, browsers and email services. Watch, though: it can index all sorts of personal material; it's not ideal for shared computers. The download takes about four minutes on a standard modem, but it will occupy a whole gigabyte of space on your PC, and it only works with recent versions of Windows.



Yahoo's [Desktop Search](#), now out of the beta trial, covers over 200 file types, including databases, spreadsheets, graphics and presentation formats. In its results, it shows large previews of files; it plays audio and music files without launching a separate player. It allows you to choose which parts of your PC's contents you want to search, with individual indexing options for different types of material, and maintaining your privacy by omitting your search history. Advanced search is possible too.

It also has a Live Words system: highlight a word or phrase and click on a button to search on that topic.

Microsoft's Windows Desktop Search, which requires installation of the MSN [Search Toolbar](#), is now out of its trial. It inserts a query box in the bottom taskbar, and can show you preview images of the files without opening an application. It also offers a smart folders feature: save a search and it will update the folder with relevant new emails and documents.

A stand-alone system, [Copernic](#), has rated highly with search experts. Download the tool from this Canadian firm (it takes about six minutes over a standard modem) and it sits on the taskbar.

An easy-to-use system, its clearly presented results have each search word highlighted in a different colour. Hover your pointer over a result, and it gives you information on the contents, so you do not have to open the application concerned.

[Blinkx](#) supplies a search tool integrating desktop, web and television search. Its smart folders automatically update content based on the ideas they contain as well as keywords.

Check whether a tool will search your particular emails. Not all, for instance, will search Outlook/Outlook Express, or inside email attachments

You might want more than one type of desktop search, to reach into all the types of file you have, or cover different facilities.

But they take up space on your system, and can use up processor capacity as they busily index what is on your machine.

If you need more facts, try CNET's [comparison](#) of desktop search facilities, with a chart of their features. (But it did not, at last look, include the new Google version.)

If you keep this book on your PC, desktop search will find things in its contents.

## 16 BROWSERS & TOOLBARS

### BROWSERS

The browser is the application you launch when you connect to the net. Most people use Microsoft's [Internet Explorer](#), but lately a newcomer has offered real competition: [Firefox](#). From the Open Source world of software development, Firefox is fast and lean – only 4.7 Mb in size, so a download should not take long.

It offers tabbed browsing: rather than open a new browser screen each time you look at a new page, it opens each page within the same screen, with tabs at the top naming the open pages. You can flip back and forth fast between pages. This keeps your desktop from becoming cluttered with browser windows, though if you have a small screen, you may regret losing another slice of space off the top.

Firefox has a customisable search bar: you can set this for [Google](#), and on the drop-down menu switch at a click to [Yahoo!](#), [MSN](#), [Ask Jeeves](#) and other search systems you want. RSS feeds will allow you to read headlines from a bookmark. You can re-arrange the buttons on the browser to fit your needs, and once you're used to the system, you can explore optional extensions: small 'plug-in' programmes, which add facilities like controlling your screen by moving your mouse.

Firefox will operate on Windows 98 and up, and with fairly low-powered PCs, though one drawback is that it does not run Active X controls – this is a security plus, but means certain sites will not load properly. For such cases, keep Internet Explorer to hand.

[Microsoft](#) is updating the latter. Version 7.0, currently being tested by developers, will offer badly-needed improvements to security, plus tabbed browsing and a reader for RSS feeds.

Other browsers are the onetime top contender [Netscape](#), which warns you when it comes across a site possibly containing spyware, and also the notably speedy [Opera](#) (now available free without adverts).

### TOOLBARS

You can set your preferred search engine as the home page which appears when you open your browser: from the main Start button, click Settings>Control Panel>InternetConnections, and write in the URL of the engine you want. You might choose the engine's advanced search page, to give yourself the most options.

Other possibilities, though, are now available, enabling you to install on your computer direct connections to search engines, to save a little effort each time you search. The toolbar is an extra strip across the top of your browser screen. Initially designed to work with a recent version of Internet Explorer, like 5.5 or later, many now work with Firefox. They may only operate on Windows 98 onwards.

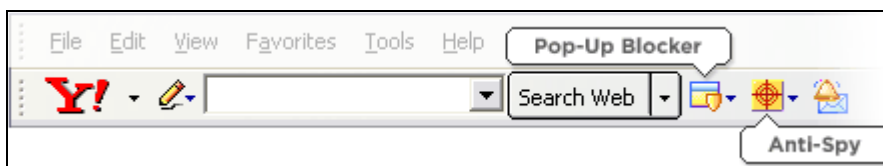
You can choose between proprietary toolbars from the main companies, or ones combining different search systems. Most are reasonably small files to download and easy enough to install. [Google](#) and [Yahoo!](#) offer toolbars that are among the best featured – though assume that a useful new option on one will soon be offered by competitors. Google's [Toolbar](#) offers a search box with options to search the whole web, your local country sites or the site you are on. It's customisable, so you can add what is important to you: icons for searching news, images, groups, or an I'm Feeling Lucky shamrock symbol.

The toolbar can save your search history; other options include page info and rank details, and a blocker for pop-up ads.

There's a highlighter to mark search terms on the pages you find, while a translation tool allows you to hover the mouse pointer over a word in English and get a translation of it in a language of your choice (at present Chinese, Japanese and half a dozen European languages). There's even a button which can add the web page you are looking at onto your blog.

A nice extra is that, in the browser address bar, you can write the main name for a site, rather than remembering the detailed URL, and go straight there.

Yahoo's [contender](#) (*Figure 10, below*) (which operates on Linux as well as Windows, and with Firefox) has most of the options [Google](#) offers. It supports the My Web facility to save pages you search, and you can access your bookmarks when you sign onto any PC with the toolbar installed.



Its Smart Tools facility means that when you highlight a word or phrase on a page, an icon appears, giving you the choice of copying it, translating it (non-English text into English), getting dictionary definitions or searching on those terms.

You can have the [Yahoo! Toolbar](#) with an anti-spyware system built in, but that boosts the size from 0.4mb to 3mb, and slows downloading.

MSN's developing [system](#) has yet to match [Yahoo!](#) or [Google](#), though it offers tabbed browsing, and provides clustering of results. If you want the toolbar, though, you need to take the whole suite, including desktop search – a bigger package to download.

Another possibility is the [A9 version](#). As well as much that comes with the other toolbars, including storing your search history, it will search reference services – the film search takes you to the very comprehensive [Internet Movie Database](#).

With Dogpile's [toolbar](#), simply right click on a word to make a search on it.

## DESKBARS

If space on your computer screen is at a premium, you won't want another bar across the top. But you might like a deskbar, a simpler addition, which installs a small query box on the basement taskbar. This lacks the features instantly available on a toolbar, but still enables you to type in a search without opening a browser.

You can get deskbars from [Google](#) and the like, but a strong independent contender is [Dave's Quick Search Deskbar](#), a small but speedy package provided by American Dave Bau. It initially searches [Google](#), but open the menu and choose from hundreds of search engines, reference sources and tools.

## 17 GROUPS & LISTS

If this book had been written a couple of years ago, it would have given greater prominence to the subject of groups and mailing lists, as two of the prime systems for communicating with others over the net. Blogs have stolen much of their thunder, and they may now be less important to journalists, even if traffic overall is still growing. But there are enormous communities out there, exchanging views and information – and you can listen in.

**HOW TO USE THEM:** Spend a little time getting to know the groups concerned, so you understand the type of people involved and can measure the quality of the material.

Apart from keeping up with trends, such networks can be a great way to find individuals involved in activities or controversies you are covering. (For the latest news event, blogs may now be the first place to turn for people's reactions).

Those affected by broader and longer-term matters can be hard to locate. If you are producing a programme, or preparing an article, on a type of pollution, a disease, or a problematic product, you will want people to film, or interview. Do a search for groups devoted to the subject in question.

Finding suitable 'ordinary' people can sometimes be the hardest thing on a story. But finding the right expert may also pose problems. Professionals and academics share information and discuss their concerns online. When you spot a likely prospect on the net, check them out: not just their qualifications and experience, but look at what else they have posted, so you can judge how articulate they are, how up-to-date – generally, how useful they might be.

If you want to contact someone, don't necessarily do so online. Is there a way to email them, or to find their phone number?

Before posting a question, check if someone has asked the same before – it's not only quicker to get the answer, but regular users get annoyed at newcomers wasting everyone's time.

You can ask things anonymously, but if you want to quote people, you owe it to them to say you are a journalist and explain the purpose of your query. And you might get a better response as a result.

Answers can take a while, so this may be unsuitable if you are on a tight deadline. And remember you cannot be sure a respondent will tell the truth – further checks could be necessary.

**HOW YOU SHOULD BEHAVE:** Be careful what you say online. Even if you are using a pseudonym, it is probably important that this 'personality' does not get a reputation for stupidity or nastiness. And if it is your own name, remember material on the net can have a half life longer than atomic radiation.

Of course, laws still apply to the net: if you break the law or libel someone online, you could be traced and dragged into the courts. Don't give your normal email address if there is any chance of it being published – you could be inundated with spam. Get a free alternative, like one from [Hotmail](#).

If you are posting something, think about the wording of the subject line in your message. Indeed, as a general rule (with email as well), regard the subject line like a headline – it has to make absolutely clear what the message is about and preferably make it inviting enough to stand out from the mass.

**THE DIFFERENT NETWORKS:** It's easy to get hung up on the terms for the various communities out there. But all this activity is essentially a way for people to share information and virtual materials. You need to understand the different systems, because the access methods vary. They usually come complete with help facilities and FAQs (Frequently Asked Questions) – which may answer both technical queries and the most common questions about the subject concerned. (There's even central storage at the [Internet FAQ Archives](#)).

The main systems are:

- mailing lists (otherwise known as list-servers, discussion lists, or email forums)
- forums (web or net forums, or message boards, which differ from groups because they are usually run on a specific site rather than shared across the world, and are accessible from the web without special software)
- newsgroups (or just groups)

**MAILING LISTS:** These can be more serious than many information-sharing systems; they are also much less afflicted by spam. You have to sign up to take part, and there is often someone in charge, a moderator who aims to keep contributions relevant and exclude junk. Joining up can mean a delay, but you may be able to read what was circulated previously.

If you want to keep in constant touch with a subject, like a reporting speciality, a mailing list can be useful: find the right one and you have an expert selecting the most important things, and then sending you the collection – normally once a day, or perhaps weekly. After joining, you may be able to get a list of members, which can be helpful in finding contacts.

Some lists are private or available to limited subscribers; others only send out announcements, without two-way communication. RSS feeds are replacing some email newsletters – which can be treated as spam by over-zealous security systems.

Be cautious what you sign up to, especially if the list offers to send posts as they appear: busy lists generate an awful lot of posts, and you could be swamped with emails. First, study some samples, to see if the material is really what you want. Ask yourself: have I time to read this every day?

If you are a member of a really active list, you might want to unsubscribe temporarily if you are going on holiday – to avoid returning to a great electronic pile of messages.

[Tile.net](#) has a comprehensive index, which is searchable or browsable alphabetically.

**FORUMS:** These tend also to require registration. You go to a site to see what's new, though there may be a system of email notification about relevant posts. As with lists, there is normally an administrator to keep things in order. Discussions usually take place in 'threads', which follow the train of an argument, linking the contributions.

Some forums allow anonymous posting, but if not, it's common for people to avoid using their actual names, and to choose user names. Readers care how reliable different contributors are, so systems have grown up whereby people's posts are appraised by other users. Regulars gradually build a profile indicating how respected they are.

To look for forums, try [Yahoo! Message Boards](#).

**NEWSGROUPS:** These are discussion groups – not usually concerned with news – and operate on the Usenet system linking computers. (Usenet stands for Users Network; it predates the web.) You generally require a type of software reader to connect with them, but easy access is available via [Yahoo!](#) and [Google](#) Groups, or Outlook Express.

[Yahoo! Groups](#) reckons to cover millions of groups. There are, for instance, 1,100 alone on 'Music that sucks' (for various artists, types of music, countries and languages), but there are also valuable information services. With 90,000 groups categorised by country, see what is being published about, and from, your country.

[Google Groups](#) now has more than a billion messages in its archive, in addition to current activity. Its directory picks out 50,000+ groups – ranging from some 70 forums about different aspects of Adobe software to several on Zippo lighters. Most include spam and other junk among serious discussions.

If you are searching on [Google Groups](#) and are interested in a particular contributor, click on 'show options' by his or her name, and then on 'Find messages by this author' to see everything they have posted on Usenet.

MSN is developing a valuable tool, [Netscan](#), which provides detailed analysis of the content of different newsgroups, updated monthly – you can see at a glance the number of users and posts each month. It has lists of the main discussion threads and profiles showing the numbers of postings by an individual since 2000.



Recognise that the main interests of groups on the net are sex, games, music, films and software – skip over these to reach subjects that matter for your stories.

Newsgroups are usually unmoderated, but have the plus that they are arranged on a worldwide system, even if there are local variations.

There are perhaps nine main categories or hierarchies, all with endless sub-sections:

alt.	alternative, the least organised, sprawling across anything (especially the 'main interests', see above),
comp.	computer-related
humanities.	arts, literature
misc.	miscellaneous – can include education, health and business (though there's also a biz.designation)
news.	meaning news about Usenet, not current events
rec.	recreation and entertainment
sci.	science
soc.	social discussions
talk.	debate on topics such as politics and religion

Thus a group entitled 'rec.autos.makers.vw.aircooled' is for people interested in types of Volkswagen vehicles. Check you have the group for the right meaning of a word: 'nuts' groups can include ones on nuts you grow, nuts who are enthusiasts, and people who are nuts. (There always seem to be plenty of the latter.)

Groups can be useful in providing help – much activity comes from people seeking answers to knotty questions, including computer problems.

If net access is difficult or expensive, use Outlook Express to work offline. It will let you download newsgroup postings and read them later on your computer. An alternative is the [Free Agent](#) software.

## CHAT

If you want to explore other areas where net users communicate, there are chat rooms – big among younger people – and instant messaging, both of which are live activities, like conversations. These tend to be concerned with personal exchanges, but to see what is on offer, try [Yahoo! Chat](#), on many subjects and in various languages. Be aware that, when you enter a chat room, your presence is immediately signalled to others using it.

IRC, [Internet Relay Chat](#), is a separate system, and can be more serious, with much emphasis on computer matters.

Chat rooms and messenger services are often exclusive operations run by net conglomerates like [Yahoo!](#) and [MSN](#), developing online communities. Newer operations like [MySpace](#) compete for the youth market.

## 18 BLOGS

Weblogs – blogs – have already attained almost iconic status in some Western countries (partly perhaps because quite a few are written by the trendy commentators who also write about them). Their explosive growth has attracted widespread attention. So what are they, and what use are they to you?

Blogs are a sort of online diary, written by individuals or groups, on websites with easy ways to add material. Some allow outside contributions, on a selective basis; others permit anyone to toss anything into the pot. Photos are much used, while audio and video blogs, and also moblogs, created by mobile phone users, add to the mix.

Blogging – writing blogs – is often treated as some special, coherent type of behaviour, but it is essentially a new communication method, and millions are taking advantage of these simple personal websites to get their messages out to friends, family, or the world.

It's not just English-speaking countries: China alone boasts some 14m blogs.

**WHAT'S OUT THERE?** What are these messages about? Pretty much anything you can imagine. In some countries, blogs have made an impact on politics and journalism, but the majority are individuals doing their own thing – only more publicly than before.

Thus [VeryPoisonousLady](#) is a young Singaporean clubbing and having fun; a more sober approach is offered by [The 3<sup>rd</sup> World View](#), from Bangladesh. An interesting compendium of blogs from different countries is [Global Voices](#).

Blogs are seen by some as important for empowering individuals, giving them the ability (potentially, at least) to let their voices be heard anywhere, without any control by editors. But the networking aspect seems especially valuable – enabling conversations, arguments and co-operation across the web; blogs also score in sharing ideas, at passing the word about matters and materials regarded as interesting, important or useful.

Commercial blogs are spreading on the web. Plenty of companies (especially those in the technology and media worlds) run blogs as adjuncts to their main operations: corporate blogs offer a less formal way to publicise activities and discuss matters with customers. Individual staffers are allowed or encouraged to describe and comment on their work.

Bloggers are already beginning to form networks in the hope that readers attracted to one blog will also look at its neighbours. Commercial operations run groups of blogs, like the New York-based [Gawker Media](#), with such titles as [Gizmodo](#) (gadgets) and [Wonkette](#) (an off-the-wall take on Washington politics). There's money to be made in the blogosphere, and the most popular offerings attract cash from advertisers and sponsors.

**HOW TO FIND THEM:** With plenty of free hosting sites, anyone can start a blog, which of course has its downside. So you need tools to sort out what to bother with.

In September 2005, Intelliseek's [BlogPulse](#), was offering to locate information on more than 16m blogs – a total increasing by sometimes tens of thousands per day. Individual posts were running at 450,000 a day.

Most blog sites report new postings to the search engines within minutes, so word can get round fast.

Major search engines are not suited to finding blogs, but Google is already trialling a dedicated [Blog Search](#); expect others to follow. For now, specialists dominate the field: Ask Jeeves' [Bloglines](#), searching 750m articles; [Technorati](#), searching 18m blogs; and [Feedster](#), covering 15m syndicated RSS feeds. The last two, with three others, can be looked at together by a metasearch on [Clusty](#).

Appearing in the lists or directories on the main blog search sites can make an enormous difference to a blog's visibility. BlogPulse's most popular blog was a perennial favourite, [BoingBoing](#), 'A directory of wonderful things'.

Once you see a blog you like, you should soon find more. Bloggers tend to link to like-minded sites – you won't see many right-wing bloggers linking to left-wing sites, or vice-versa.

A great merit of blogs is that you can easily jump between similar sites. But the popularity of blogs is partly judged on the numbers of links to them, so bloggers assiduously add links, in the hope that the other bloggers will return the favour. As a result, they list links to lots of blogs and other sites that they may not care about.

Collective blogs tend to be better at maintaining standards. They are not so subject to the perils of everyday life, when illness or a family crisis suddenly interrupt an individual's writing and editing. With personal blogs, there is a trade-off between quality and the frequency of posting. It's hard work keeping a blog going every day, let alone all day (especially as readers can expect something new seven days a week). Ambitious undertakings like video blogs soon find their creators worrying about where to come up with fresh programming.

**BLOGS AS JOURNALISM:** Comment is relatively easy; facts are harder work. The former *Financial Times* reporter, Tom Foremski, talks on his site, [SiliconValleyWatcher](#), about the graft involved in going out with a notebook to do the reporting necessary for a worthwhile blog. What was needed to get people to visit a site was "original content freshly served," he noted, and that was not found easily.

He sought first to produce a compelling product, and only then (very 21<sup>st</sup> century this) looked at how to make money out of it.

Some of the more successful blogs are like newspaper columns – indeed some of them are written by such columnists, whose experience enables them to provide a mix of knowledge, informed comment and lively writing.

But much of what appears, even in the field of current affairs, is from 'ordinary people'. Mindy McAdams, of the University of Florida, points out: "The man-on-the-street interview is now authored by the man on the street, and self-published, including his photographs."

[Technorati](#) registered 1.1m posts on the day of the London underground bombings in 2005, while the Asian tsunami caused traffic at [Blogger](#) to jump by three-quarters in the week. The latter outpouring was a major aid to broadcasters and newspapers – difficulties in obtaining coverage from remote disaster areas meant that they eagerly seized on videos, photos and first-person accounts appearing on blogs and elsewhere.

There are bloggers who want to be journalists, an aspiration derided by some establishment sources, who argue that just because you write something and publish it that does not make you a journalist.

Claims for the journalistic achievements of blogs have been exaggerated – the number of stories broken by them seems pretty small. Nevertheless, they follow an honourable tradition of independent gadflies, from pamphleteers of past centuries onwards, and their influence in some circles is growing. They can be less hide-bound than established media, and can quickly gather momentum for inquiries, campaigns and the like.

## **LAUNCHING A BLOG**

Tempted to start blogging? It is easy to begin: [Blogger](#) (now part of Google) should enable you to avoid the technical difficulties, and create your blog in a few minutes. [w.bloggar](#) is one of several downloadable services that will allow you to work on a blog offline.

It can be a lot of fun to say what you like, at whatever length you fancy, though there's no guarantee people will come and read your words.

It will be hard to match the success of the Baghdad blogger, Salam Pax, who, when Iraq was invaded, won fame for writings which displayed bravery, humour and perception. But he (currently blogging with [Shut Up You Fat Whiner!](#)) benefited from a situation where there was a hunger for information and little competition.

A blog can be a fine advert for yourself, but that requires quality – which takes time. Some media employers might allow you to do this as part of your normal work, but most will want you to devote your energies to your main job.

If you go ahead under your own name, bear in mind that whatever you write will likely stay

out there, somewhere – so you need more restraint than the average blogger. Avoid saying anything which could damage your future reputation.

(There is a general point about anything written on the net: whatever appears about you, as well as by you, could be around forever. Even if something is grievously wrong, it can be impossible to correct every instance of the offending material. That means if you have dealings with someone new, they could well Google your name and discover all sorts of things – including material you've long forgotten. Don't make matters worse by writing anything now that you will regret later.)

Caution is also required in anything you write about your work, or your employer. Bloggers have already been fired by bosses who don't appreciate freewheeling commentary on their organisations. Some writers suffer worse: there are bloggers in jail in several countries because their comments upset the authorities – as a blogger, you have even less protection than a normal journalist.

Reporters Without Borders has just published a downloadable [Handbook for Bloggers and Cyber-Dissidents](#), in such languages as Arabic, Chinese and Farsi, with detailed advice on how to avoid getting arrested.

For more on blogging generally, try the [Weblogs Compendium](#), from Canadian Peter Scott. [Cyberjournalist](#) has a list of the journalism blogs out there – 340 at last count.

## 19 MULTIMEDIA - PICTURES, AUDIO AND VIDEO

The net is finally becoming truly multimedia. High-speed net connections, PCs with far greater capacity, and, as hardware costs fall, web operations offering consumers much more – all combine to make obsolescent the idea of the net as a text medium.

This may add to the frustration of those with poor net connections and minimal computer facilities, but the benefits should spread, and, for broadcasters, more raw materials could become available.

In this chapter, we'll look first at how to find images, then audio, and then video, along with details of some of the main feeds and webcasts for such materials, and finally at some online facilities useful to broadcasters. Music, games and inter-active systems are left aside, as less important journalistically.

Search systems for multimedia have a long history, but they have faced great technical challenges in identifying content in ways that allow us to find what we want. Research is yielding results which are changing what is on offer.

### IMAGES

[Yahoo!](#) and [Google](#) dominate much of the picture search world. Find their search pages, click on Images.

[Yahoo!](#) Images currently searches 1.6b+ pictures, [Google](#) 1.3b+. Searches on either produce 20 thumbnail images, on which you click to learn more, and normally find a bigger image. When you search, it may say 50,000 images found, but they only let you look at around 1,000.

What they find varies: [Yahoo!](#) offered 3,700 results for Zimbabwean leader Robert Mugabe, [Google](#) 3,500. But for former Malaysian leader Mahathir Mohamad, [Yahoo!](#) had 1,900, [Google](#) under 1,000. Both included duplicate images from different sources. With advanced image search, they allow you to choose colour or black and white, and between small and larger images; [Google](#) can seek out some specific file types.

An alternative is the Swedish company, [Picsearch](#), which powers image search for [Ask Jeeves](#) and [MSN](#). It offers 16 images per page (*Figure 11*), but will let you search all it finds – 610,000 images if you search for 'man' (only 450,000 for 'woman'). Its database is unspecified, but may not compare to the big boys: a query for Mugabe brought 280 results, and Mahathir about 500. (The image totals seem to vary enormously over a period, so the numbers may be pretty imprecise.)



*Fig 11: Picsearch's collection of photos for the late southern Sudan leader John Garang*

As well as choosing black and white or colour, you can search for just images or animations, though the latter only selects some graphics, not cartoons. It has a very specific selector for the size of image.

If none of these have what you need, try [Search-22](#), which has links to a raft of picture search engines. There are also meta-search image engines, but a recent survey by [TASI](#), the British-based Technical Advisory Service for Images, concluded that they had little extra to offer.

Much research is continuing into how software can directly identify the content of pictures, rather than just index captions and catalogue material.

The French-based LTU Technologies categorises different types of image, like cartoons or graphics, excludes pornography, searches on both a visual and text basis, and seeks to find similar images. Try its [Image-Seeker](#), which is reasonable at finding suitable pictures. But when asked to find similar images, based on the original keywords and the visual characteristics of the image found, it can produce some seriously off-beam suggestions.

Still, watch this space: image recognition technologies are developing swiftly.



## IMAGES TO USE

What you find roaming the web may not be usable, simply because it is copyright material you must pay for. Many pictures are just tiny thumbnails or simply cannot be downloaded – which prevents temptation. But there are quite a few image databases which offer you material for free publication.

The Hungarian-based [Stock.XCHNG](#) has over 100,000 stock pictures on display – from fruits to furniture, bicycles to breakfasts. When using a photo, it asks you to let the artist know by email. Other collections may require a printed credit for themselves or the photographer.

Some databases offer images for use by graphic designers and illustrators – like the US [Morgue File](#). It offers pix such as mushrooms, marijuana, pavements and peeling paint; a credit to the photographer is requested, though not required. [Image\\*after](#) is a Dutch site providing free raw material for multimedia designers.

Recently the field has been extended by the growth of public collections, simple technology allowing people to store and display their photos on the web. A successful example is [Flickr](#), flickr.com, now owned by [Yahoo!](#) At last look, it had 17m+ photos. People tag their photos with what they think are relevant categories – done casually, this builds into a natural catalogue system (a 'folksonomy' – the latest bit of jargon).

We are not necessarily talking great art here, or important subjects – expect endless groups shots of friends and families. But there are worthwhile images, both generic subjects and current events, and the picture takers can indicate what restrictions they place on the re-use of their photos.

More current images are also becoming available as it becomes easier to upload photos, and cheaper to host them online.

The London underground bombings produced a mass of online images, partly due to the fact that cameras are now common features of UK mobile phones; Hurricane Katrina in the southern US produced much more.

## AUDIO

Speech has received less attention from the major players, and is often still an adjunct of video search.

A leading search system is the AOL-owned [Singingfish](#). The emphasis is on video and the sources can be mostly American; to concentrate on radio material, use the boxes on the left to click off formats you don't want. [AltaVista](#) has a useful audio search (click on the MP3/Audio button).

What you'll find is invariably a tiny percentage of the text results. A significant speech by a political leader may be available, but day-to-day comments which appeared in news bulletins may not be preserved, or not in an easily accessible form.

Most of what you find is likely, if you want to reuse it, to require permission, and often payment, from the broadcasting organisations.

Finding the words you want in audio can be hard, unless there is the text of a speech or script of a programme. But speech recognition systems are progressing. Not for the first time in the history of technology, security requirements are helping to drive research: much funding is coming from intelligence sources that want to scrutinise what is being said in billions of phone calls and other conversations across the globe.

A consumer use already planned is storing phone conversations, from mobiles or over the net, which can then be searched. [Google](#) is among those working on how to analyse the content of audio files.

Hewlett-Packard's [SpeechBot](#) audio search, can demonstrate the progress made. As with early machine translation, the results are still pretty rough, but you should at least be able to find a passage you want and then listen to the sound version to get the precise words. SpeechBot has indexed over 15,000 hours of content (tiny, of course, in search engine terms) from various US radio programmes.

Only a brief mention here for music, which has multitudinous options for finding, downloading, swapping and listening. The new [Yahoo! Music](#) site shows some of what is on offer. The company's new [Audio Search](#) will also seek other audio files and podcasts (see below about the latter).

## **RADIO**

Want to listen to a radio station in Angola or Anguilla, Nepal or New Caledonia? Go to Brian Buckley's London-based [Live-Radio](#) service to find thousands of broadcasters worldwide. Buckley works hard to ensure the links stay up-to-date. An American alternative, [radio-locator](#), has links to over 10,000 station web pages and 2,500 audio streams, from North Norway to the South Pole.

Radio is changing technically, and thus altering the way we listen to it. Increasingly, it is possible for people to hear programmes when they wish, rather than just at broadcast times.

For an example, try the [BBC](#), and click on Listen to World Service. Scroll down to World News Summary and you can get the latest news bulletin, the one broadcast on the previous hour or half hour. News and audio is also offered in more than 40 other [languages](#). As well as listening live, or to a recent bulletin, the system allows you to select past programmes (usually from the previous seven days) and not just from World Service: you can hear programmes from all the domestic networks, including music, sport, comedy and drama.

So when you listen to a particular station, keep an eye open for what extra services may be on offer.

A service of possible interest to broadcasters is the non-profit [Public Radio Exchange](#), a web-based marketplace for public radio programmes. While largely American in content, it could provide an opportunity to offer work to stations abroad.

## **PODCASTING**

A new way to receive radio programmes via the net is podcasting. The name is not too accurate, since it implies that audio is being sent out to an iPod personal audio player. The name has come about partly because of the great success of Apple Computers' iPod, but the audio files can be accessed by personal computers and many types of portable players.

You sign up to services which allow the automatic download of audio files to your computer, and then, if you wish, they can be transferred to a portable player. You don't necessarily need special software to hear them: many podcasters and directories provide 'streaming' audio, so you listen directly to the site.

iPods are pretty expensive, but their popularity in richer countries is driving an increase in the amount of audio material available online. As with blogs, much of it is produced by amateurs, but professional media organisations are also involved, partly as another way of distributing existing programmes.

There are lots of podcasts about computers and technology, music, and current affairs. Wider applications include listening to books online, while universities and schools have started producing podcast course materials, like foreign language instruction.

[Podcasting News](#) is one of the sites that have sprung up offering lists of leading podcasts, a podcast directory, lots of resources and links. Its FAQs provide a simple introduction to the subject. Podcasting is not an expensive process – a microphone and a sound recording programme (readily available free) are the basic requirements.

## **VIDEO**

This is another fast-changing field where speedier net links enable consumers to receive greater content. (Without a broadband connection, online video can be a palsied affair.) Video search exists at various levels of technology, and much that has been on offer comes from American TV networks.

[Yahoo! Video Search](#) speaks of searching millions of videos; it recently upped its international coverage. It has indexed programmes and news bulletins based on descriptive material about the files, and on closed captioning – text, usually for the hearing impaired or those learning English, which describes all the audio content. It provides a thumbnail image from each video, plus details of the source, length etc. Then you click to play it. An advanced search page allows you to select a particular format, and whether you want a smaller or greater image size or length of clip.

[Google Video](#) has a different approach. Its service, still in beta trial, does not trawl the web, nor initially was it offering actual video. Its coverage includes most of the US TV networks. It displays a still image from an indexed programme and then a section of text (once again from closed captioning), but you are not able to watch the video.

It now, though, offers videos made by individuals, whose films are uploaded to its site. To play them, you need to download a Google viewer. At present, the content probably concerns puppies more than politics.

[Singingfish](#) search has arrangements with Reuters TV, BBC and CBS, but also seeks out video (and audio) material across the web, like conventional search engines. It too categorises what it finds by an indirect method, on the basis of the titling and closed captioning. If you find something suitable, you may have a choice of speed for receiving the video, and also of media player to use (depending on what you have on your computer – Windows, RealPlayer and QuickTime are among the most common).

With [Singingfish](#), you can't actually tell what sort of video it is until you play it. It could be:

- *streaming video*, which starts playing soon after the transfer to you begins – you don't have to wait for the complete file to arrive before viewing it;
- a *download*, which means waiting till you have it all, but you get a copy left on your own computer. (Note: streaming video does not work with low bandwidth connections.)

Newer technologies index the actual content of videos, working initially on the words spoken. Several companies have developed such systems; [Blinkx](#) is available free. It uses speech recognition software to create a text transcription of the audio track, and then searches that document. It has been looking at 75,000 hours of TV and radio each month, and covers amateur products as well as content from commercial providers; podcasts and video blogs are included. For a search, you can specify which TV or radio channel you want to include.

[Blinkx](#) now has an RSS alert system, SmartFeed, which will message you when a search term crops up on a TV or audio channel. There are the regular [Blinkx](#) smart folders and even a conceptual search, which seeks material similar to the concept of the text that you place in the search box. A slider can give priority to how relevant the material is, or how recently it appeared.

[Yahoo!](#) plans to move into 'deep indexing' in the future, but [Blinkx](#) has stolen a march on the competition with its sophisticated offering.

A general point on video search: be prepared for dead links. Because video files are large, companies may not keep material available for long.

## **TV OVER THE WEB**

In the west, it is now commonly suggested that television in the future will be delivered via the net. TV, with the spread of broadband internet links, can be accommodated in a way previously impossible. Another set of initials, IPTV (Internet Protocol TV), is being inflicted on us – it covers both live TV and video-on-demand, plus other interactive services.

In practice, it is still early days and the big TV corporations are feeling their way over online services. Video on demand has been available from various sources, but it is only in the last year or so that serious efforts have begun to deliver TV over the net to a mass market.

As well as the entertainment offerings from US domestic networks, there is expanded online news provision from [CNN](#): it has introduced a very straightforward system, where you click on Browse/Search for Watch Free Video, and then you can flip back through the individual news reports for two weeks. There's not much world coverage, but [CNN International](#) offers a week of news reports. The downside is you will likely have to watch lots of adverts.

Across the Atlantic, the [BBC](#) is gearing up for a big push in 2006, to provide downloading across its TV output. Video news on demand is already available with main bulletins, individual items and (on major stories) previous news items and background material, all accessible via an interactive system. Another BBC scheme is the Creative Archive, aiming to provide downloads from the corporation's wide collection of past programmes.

The [BBC](#) hopes to keep down distribution costs by using peer-to-peer technology, whereby users share what they receive with others. There will be built-in controls so people cannot retain programmes beyond a week, nor burn them onto CDs.

Elsewhere, plenty of stations offer breaking news footage and video clips on demand, and there are more expansion plans afoot. Among the web companies, [Yahoo](#) has major ambitions as a TV provider.

Directories of world TV stations have varying lists: see [Comfm.com](#), a Paris site which also has links to world radio stations and even live webcams – feeds from cameras set up in (mostly) public places, for those wanting a voyeuristic look at (often boring) scenes around the globe.

Another directory is [TVRadioWorld](#), from the US, with a lengthy collection; [World Wide Internet TeleVision](#) has fewer, but offers some programming details.

The bad news is that all this activity is not likely to provide much help for broadcasters in developing countries – the material on offer is usually for personal use only. [JNet](#) lists online video archive companies and TV network archives, but these are invariably fee-charging.

One resource is [The NewsMarket](#) which offers free video (plus still images and digital audio) from big companies and NGOs. Newsroom managers and programme makers can access material from the UN, World Bank and big charities as well as commercial companies.

Looking further ahead, the growth of small stations on the net looks certain. In the UK, a TV station hosting company, [Narrowstep](#), reckons on start-up fees for a station of under \$10,000, and already hosts such netcasters as the official world badminton TV, specialist sports channels for golf and extreme sports, a travel company – offering different holiday videos on demand – and Glasgow TV, from the Scots city council. Education and training providers are other organisations already exploring the possibilities in different countries.

## BROADCAST SERVICES

**PRONUNCIATION:** With a speaker on your computer, you can hear the words you want.

Click on a [Google](#) word definition, and you go to [Answers.com](#), where each word is accompanied by an audio version, with a bit of an American accent. MSN's [Encarta](#) Dictionary provides pronunciations too, while [Yahoo's dictionary](#) has sound versions for more than a third of the 200,000 entries.

[Fonetiks](#) offers pronunciation practice in seven different regional versions of English, and nine other languages.

**SOUNDS:** If you require sound effects – be they bombs, bullets, bicycles, bongoes, or burping – go to [FindSounds](#), run from the US by the Comparisonics Corporation. It's a tremendous compendium of links to all sorts of natural and artificial noises. Before reproducing the sound of a didgeridoo, or a donkey, check the copyright situation for the individual audio files, as some may not allow free use.

**LISTEN TO THE TEXT:** [Speakwire](#) is a rather different sound service, from Scotland. It allows you to have selected RSS news feeds read out to you, using synthesized speech technology; it's run in conjunction with the company's speaking search machine, Speegle. (You may find it a little disconcerting to hear news from Australia or China read out with a Scottish accent.)

To turn documents into spoken form, download [Speech Expert](#), which supports over ten languages and 20 voices.

**GUIDELINES FOR BROADCASTERS:** In 2004, the [Commonwealth Broadcasting Association](#) produced a comprehensive set of editorial guidelines, drawing on the practice of some ten broadcasters around the world. This can be [downloaded](#) freely. The [BBC](#) has put online its new [editorial guidelines](#) – with practical advice as well as coverage of moral and legal issues. The site includes advice notes on specific issues, and full guidelines for online journalism.

### **COPYRIGHT NOTE:**

A great deal of multimedia material is available only if you pay – sometimes considerable sums – to reproduce it. But there is a growing movement for less commercial producers to publish content under a Creative Commons license, which defines the rights reserved by the writer, artist, film producer etc – somewhere between full copyright and public domain (freely available).

If you see something useful with such a license, [Creative Commons](#) can tell you more. Check out the individual sections there on video, images, text – they also list collections of freely available films, photos and more.



## 20 INTERNET PHONES AND EMAIL COMMUNICATION

### PHONES

Would you fancy a free telephone system, to make free calls to anywhere in the world? If you have broadband access to the net, you can set this up in minutes.

Few people in developing countries have broadband so it may seem unfair to take space here writing about such a service, but VoIP (Voice over Internet Protocol), as it is known, is challenging the established phone networks across the globe. The benefits it offers are such that your local internet points and cafes may adopt broadband if it is available.

To use it, the minimum necessary is a computer with a loudspeaker and a microphone, though a (cheap) headset with earpiece and microphone provides better sound. Software is available free and then you can contact anyone on the net with the same software switched on at their computer. The net carries your conversation as it would an email. Without a microphone, you can still use the system for instant message chat.

You can pay for add-ons to such free VoIP: a handset to make the calls; a system allowing you to call anyone at an ordinary phone number; or one which gives you a phone number where people can call you. (In the latter case, if the caller is in the same city as that number, the call costs no more than one to a local number – even if you are sitting on another continent.)

Net phone systems have been around for years but have taken off as broadband has spread. Japan already has many users. VoIP on dial-up access seems to be feasible only with problems like delays in the conversation.

**AN INITIAL SUCCESS:** A current phenomenon of VoIP in the West is [Skype](#), founded by two Scandinavians (who previously created the highly successful KaZaA software). It already has 60m registered users, and was recently bought by the [eBay](#) auction group for a cool \$2.6b+.

Once you have adopted a user name, you signify your presence online and are ready to make and accept calls. Anyone you list as a contact is shown if they are online at that point – last time I checked, there were 3m people online. (If you don't know someone's user name, but think they might be on [Skype](#), you can search via their email address.)

Calls via the net can offer facilities unavailable on landlines or even mobile services. Without the physical constraints of cabling, sound quality can be increased; privacy can be improved – [Skype](#) offers automatic encryption of phone calls and instant messages; conference calls between several people on different continents are fairly straightforward.



On the other hand, net connections have ups and downs, and low or variable bandwidth at one end of a call, along with other problems, can cause poor quality sound, delays, and online noise and echo. So don't abandon the regular phone just yet.

Another concern might be that [Skype](#) uses peer-to-peer technology, which means your computer can be co-opted into the process of passing on other calls around the system .

[Skype](#) incorporates a file transfer system, so when you have something too big to send with an email, or too large even to write to a CD, you can send it via [Skype](#) (including Windows systems to Mac or Linux ones, and vice-versa). It should handle up to four gigabytes (the hard disc capacity of an old computer).

As a journalist, note that people using VoIP with an associated phone number choose where they want that number. So if you ring someone with a phone number in Hong Kong or Manhattan, they might not be based there at all – the person or company might choose to have a number there to facilitate cheap calls from local people, or for other reasons, like disguising an identity.

**MORE ON THE WAY:** Other free services are being offered in different countries; success may depend on building networks large enough that users can hope to find people they want within the same system. The big net concerns are currently announcing services – instant messenger for voice as well as text is becoming common. These companies already have big networks of regular users, so they may have a head start; [Google Talk](#), though, is bidding for a wider audience with a system which will operate with other companies' programmes.

Paid services, like [Vonage](#), vonage.com, offer greater facilities, though monthly fees and call charges can mount up. Many of the big telecommunications services are announcing their own broadband phone services – they want to hang on to their paying customers. The net is commonly seen as a future way – perhaps the way – for delivering phone services.

The competition is forcing down the cost of international phone calls. Niklas Zennström, co-founder of [Skype](#), said: “We hope we can drive prices down to zero. That's the sweet spot for voice calls, zero.”

Other developments are under way. Asked about video calls over the net, he replied: “Not yet, but .....”

## EMAILS

Most of us jump into writing emails, even quite formal messages. The importance of thinking first is demonstrated in a [guide](#) to writing emails from Singaporean Wenkai Tay. And if you are hoping to talk to someone, do include a phone number – people often forget. You may want to add at the end of your emails a regular signature block, including your phone number, office address etc.

If you get lots of email, don't make yourself a martyr to your inbox. It's easy to feel you must see everything as soon as it arrives, and reply instantly. If you do, you won't get much work done; if you're feeling the pressure, schedule email sessions a couple of times a day. (Concentrate on what is important, not on what seems urgent.)

Keep the flood down by being ruthless about cancelling registrations to email lists and newsletters – do this once you realise you are just deleting stuff, because you've not time to read it all. And don't contribute to the flood by sending out messages to everyone in your address book, or sending memos to half the people in your office, when most won't be interested.

Don't send attachments if you can avoid it – they can take ages to download over a poor connection, and some offices block them, for fear of viruses. Better to offer to send something if the recipient asks for it. And watch out when sending pictures – it is easy to despatch something massive; instead, compress the files, converting, say, a TIF file to a JPEG. Zip files can help – see under software, below.

A small point: don't use signs like \$, £ or € in emails; you don't know how they will emerge at the other end, and you don't want to make mistakes with money. Say USD, GBP or EUR if you want to be precise, though writing the name of the currency, like Euros, probably makes it clearest to an ordinary person.

## 21 THE FUTURE

It's easy to get jaded about the way the net is developing – the speed, the capacity, it all blurs into a mass of barely intelligible figures. But note that one company recently announced a DVD disc which can store more than four million pages of text (amongst other things). That's the equivalent of more than 10,000 books – a sizeable library, on one disc.

The chief executive of [Google](#), Eric Schmidt, reckons that within a decade you will be able to store the entire world's information on a small handheld device the size of an iPod. Optimistic? Perhaps, but it is not hard to conclude: you ain't seen nothing yet.

This book will not try to provide a portrait of that future – there are plenty of prognosticators elsewhere. We won't look, either, at particular new developments which, typically, are described as something which could knock [Google](#) off its pinnacle, or challenge some other mighty concern. We have touched on a few earlier in the book, though most will never survive the ferocious competition.

In the current net world, a glut of innovative new products can overwhelm the market before demand builds up. However good a new device, it may be swamped by so much else coming out – after all, how much time will you have to try all the new things you've heard about in this book?

Worthy developments may simply fail to catch fire with potential users. Much research effort has gone into developing intelligent agents, software devices that can be programmed so they will go out on the web and find information for you – you do not have to wait online while searches continue. It sounds good (a real boon in countries with costly net connections) – yet public interest has been weak. It has not proved easy to programme the agents – people prefer something simple and intuitive, and, typically, they want information now, not tomorrow.

**THE SEMANTIC WEB:** There is a wider movement, on the other hand, which is bearing fruit. The semantic web represents the grandest aspect of the drive to enable computers to understand the data out there, rather than just present it to us. The semantic web aims to allow search engines to draw intelligent inferences about what a site contains, and whether it is significant to the current inquiry.

Its greatest proponent is probably the web's inventor, Sir Tim Berners-Lee, director of the [World Wide Web Consortium](#) (W3C). He explained in 2004 that the first phase, finding languages which enable computers to exchange such information, had been completed. Research in the life sciences is an area with early progress.

But there are allied developments already becoming available, in the field of contextual and conceptual search. An engine such as [Blinkx](#), reads what is open on your desktop and finds results related to that. The big engines are developing 'personalisation', seeking to build records of your search interests and thus select more suitable results.

Networks are developing which can pass on the collective search choices made by their members – one pioneer is [del.icio.us](#), a 'social bookmarks manager'. A research manager like [Furl](#), or Yahoo's trial [My Web](#), are other examples of what is on offer.

A comment from [Ask.Jeeves](#) was that the future of search rested on relevance: it was not a question of technology – it was simply how much information people would share in order to get a better search engine. But how much effort will the average searcher devote to supplying additional data?

And how willing are we to let some company know our interests, preferences, and keep records of where we go on the net? You need to trust a global concern not to start peeking at what it finds, for some immediate commercial gain, nor to let others rummage through your records.

Luckily, you have choices – already there are competing systems. It also makes sense to be part of a network of like-minded people – is there any advantage in sharing info selections with techno-geeks in Europe, games enthusiasts in East Asia, or music-obsessed teens in the USA? If you are interested, look at the material collected by a network – is it attracting people from your part of the world, or those related to your interests and professional needs?

Taking part may have other merits. As one writer put it, such networks enable people to tap into 'group intelligence', while Peter Hirschberg, of [Technorati](#), the blog tracking site, concluded: "The giant brain is us." (The cyber world is not known for understatement.)

**MOBILE WEB:** The [W3C](#) is also keenly involved in efforts to make the web fully available to mobile devices. Its [Mobile Web Initiative](#) is not so concerned with invention as with co-ordinating the activity in technology, telecoms and phone manufacture with the information and software providers.

**VOICE ACTIVATION:** One development that could soon come into its own is voice activation, allowing you to control a machine by speaking to it. Giving orders verbally is a normal way of behaving – once you adjust to the idea of talking to an inanimate object.

Such services already exist, like [Yahoo! By Phone](#). This is a paid-for service which allows you to use spoken commands to access email, voicemail and other info over the phone. [Google's](#) test version of voice-activated search technology is no longer on display, though [Google Labs](#) still has the [research](#) showing what can be done. Research elsewhere in the world suggests that more progress in this field can be expected.

**INFRASTRUCTURE:** Particularly in the US, there are interlinked developments, such as the Next Generation Internet and Internet2, seeking to upgrade the fundamentals of the net – an aging system, strained by the many unforeseen demands placed upon it.

Amongst the software and hardware developments, the service area could be the most important to users.

**UNDER THE HOOD:** If you ever doubt that we are still in the infancy of computers, think how users are expected to sort out all sorts of technical problems by themselves. It's rather like the days when motor cars were new inventions: the driver also had to be a mechanic, plunging under the bonnet to keep the temperamental machine going.

Computers are similar, with frequent glitches which require us to fiddle with the electronic innards to get them operational again. Even if available, professional mechanics can cost money, so we waste time and perspiration, wrestling with help programmes and online advice, trying to keep the machines working as they should.

With an old car, you probably still get covered in grease, but a new one now runs for ages with minimal attention. How long before a computer does the same, or works like a TV – you turn it on and that's it? Users require hardware and software they can rely on, and a supplier who provides the security and warrants a clean supply of information and material from the net (even though the latter might mean ceasing to visit dubious websites).

[Microsoft](#) is clearly aware of the problem – its founder, Bill Gates, has spoken much of the need to upgrade the PC experience. They're testing a service, [One Care Live](#), to remotely manage your computer and deal with problems that occur. This costs money, but it's a sign of what may come in the long run.

Other indications of progress are the widespread automatic updating of software, security files and the like, and the ability with current Windows versions for someone elsewhere to access and control your PC remotely, so they can diagnose and sort out a problem.

Self-repairing computers may be further down the road.

An extra problem is the 'jargon barrier': we know what we want, but we do not know the right words to use – how to phrase our demand in a way which will activate the relevant programme. Help is out there but we cannot locate it – though if we do get there, the standard of advice available is improving.

Until such problems are properly addressed, life on the net is not going to be easy – even if you do have a decent PC and a good net link.

## 22 SLOW NET ACCESS, AND HOW TO GET ROUND IT

How is life on the electronic superhighway? Does it feel as if your information is being collected by an ancient cart, with squealing wheels – one which not only takes forever to arrive, but eventually dumps a load of not very useful material in a confusing heap?

Your experience is common in many parts of the world. This book has no magic wand to turn your information delivery vehicle into a Formula One racing model, but perhaps we can oil the wheels a bit, and provide some carrots for the horse.

This chapter is concerned with the mechanical barriers you face. Your net connection is like a series of pipes: from the international net, through the link from the local internet service provider to you, the pipe narrows. The capacity of the modem on your computer can reduce things further, so only a trickle of information actually emerges.

A modem at 56 kilobytes per second may allow an adequate flow; 28kbps is getting problematic; 14kbps means real delays. Broadband is at least ten times as fast as the standard modem, at 500+, and higher speeds, of one or two megabytes per second, are becoming common in developed countries. (South Koreans, the greatest net enthusiasts, are starting to use 100 megabytes.)

Another factor is the efficiency of your PC – one with an old 386 or 486 processor handles everything pretty slowly. And, of course, irregular electricity supplies and other hazards interrupt the work. So you have to optimise your use of the net – particularly because it may also be costly.

**ELECTRONIC ROBBERY:** One blight on net users in developing countries is that they are often charged far more than citizens of developed nations – higher prices for what is usually a worse service. Universities, schools, small businesses, let alone ordinary citizens, are priced out of access. State-controlled telecoms monopolies are particularly prone to such price gouging.

World Bank/ITU data from 2003 (reported by [The Globalist](#)) showed the cost of 20 hours a month on the net was \$12 in Japan, and \$14 in France and Germany – none of the 40 richest countries charged as much as \$40. The average in Africa was \$65.

It's not a matter of development: in Azerbaijan, the cost was \$108, next door in Armenia, it was \$45; in Tanzania, it was \$117, while in Togo it was \$7. In the vast majority of African countries, such a net connection costs more than the average income. Journalists need to keep pressing governments to stop blocking progress for short-term gain.

So, you may want to spend as little time as you can on the net, and get as much as possible done while you are there.

**EMAILS:** With such high costs, allied to slow speeds, it is no surprise that the lower cost email-only services are continuing to attract subscribers in Africa.

With internet in your office, an email client, like Microsoft Outlook Express, means you can write your messages offline, and store them until you get online. This has another advantage over writing emails online – you don't lose your work when the link is broken by a power cut or system crash.

Using a computer elsewhere, take a floppy disc or some other storage device with your prewritten material, which can then be cut and pasted into web-based services like [Hotmail](#).

Before signing up to a service (and sending a new address to everyone), find out how many pages you have to click through to check email. Some are not designed for places where every page takes an age to load.

If you need to download messages with large files attached, can you do it at night, when connections may be better? Updating software should likewise be carried out at the quietest times – weekends, or holidays.

Sending large emails? (Radio journalists now regularly send sound files by email.) Try using a zip file, which compresses the material into a smaller file; it can be done with free software (see Chapter 24, Practical Help).

**SPAM:** In mid-2005, the [OECD](#), the Organisation for Economic Co-operation and Development, reported that developing countries were losing out in the battle with spam.

It is crucial to avoid spam as far as possible, especially as internet service providers in such countries have a poor record on blocking it. Do not put your email address on a website, nor when commenting in a group or message board – it's an open invitation to the spammers, who have software constantly trawling the web for addresses. If you or your organisation do need to publish an address online, use a different one from your regular address.

**INTELLIGENT BROWSING:** You want your browser to bring back the maximum info in the minimum time – that probably means cutting out all images.

On Internet Explorer, click on Tools at the top, then on Internet options, Advanced, and, under Multimedia, clear the boxes for Show pictures, Play animations, Play videos, and Play sounds. If subsequently you do still want to look at a particular image, right-click on its little icon.

On a Firefox browser, clicks on Tools, Options, Web Features and then click off Load Images.



If you're just temporarily on a PC, you can still take action about a slow-loading site: as soon as the text is visible, click Stop, to avoid the images loading. Generally, if a page seems to have halted during loading, hit the red X or similar Stop button, and then Refresh/Reload at the top. It can kick a stalled browser back to life.

If your connection is excruciatingly slow, watch out for anything that seems to arrive a little faster – a clean search screen like [Google's](#) will be quicker to load than an advanced search page, let alone a portal front page. Glance down the search results and try to choose ones which are smaller files.

It's not just the size of the file, though; it may be a slow website. You can use [A9](#) and then on a result hover your mouse over Site info, to see whether it is a fast-loading site.

Watch out for thumbnails. Some search systems now work on the assumption you have broadband, and provide little images of each page found – the extra time these take to load erects another barrier between you and your goal.

If you make the same search regularly, add the first page of results to your bookmarks; then next time it's one click to rerun the search.

It used to be hard to access some sites around the time Americans were logging on in the morning. But world capacity now usually seems sufficient to cope with such peaks. Individual sites can be busy for inexplicable reasons; if so, try again later.

Don't download files or software without checking the size first. Well-run sites normally give an estimate of how long something will take to download over either a dial-up or a broadband connection.

(News organisations with access problems need a rigid policy on personal net use, otherwise the downloading of pictures, music and videos can greatly slow connections – as well as filling up office computers.)

**WORK OFFLINE:** If it costs a lot to stay online, you should download pages to read offline. Browsers offer various opportunities to vacuum up web pages as you visit them – including pages to which those ones link. Look at Work Offline on your browser's help pages, to see how to set up the system.

**PATENT REMEDIES:** There are many companies out there selling software to speed up your net connection. They're a bit like cures for baldness: you are unlikely to see too much benefit for your money.

[Google](#) is testing a [Web Accelerator](#) which makes some difference. Unfortunately, it is not only optimised for broadband, but it brings other problems in its wake. The system stores copies of frequently used pages, to make them easily accessible, and downloads just updates to pages you saw previously. It also collects links to a page you have opened on the web, downloading some of these in advance, in case you want to look at them – which you can then do very speedily.

The downside is that you relinquish control over what is imported to your PC – you might even find you have pornographic material on it (unless you have a good filter in place) or politically dangerous (if you live in a country where the rulers do not favour the freedom of information).

As well as net accelerators, there are download managers that can make some difference, but there are few easy fixes to the problems of speeding up net connections.

**OTHER SERVICES:** With poor net connections, video really is not feasible. Probably the lowest speed modem over which you can receive radio is 28 kps, but some stations require double that.

A sign in mid-2005 that matters might improve was the introduction by the big Indian portal, [Rediff](#), of an instant messenger service offering 'good quality' voice calls over a connection of just 14 kps (even if you need 33kps to download the software). It will work on Windows 98 (the operating system of four-fifths of Indian computer users).

It is significant that such a provider is putting serious effort into a service for those with poor connections. There is a big market ignored by the western giants; hopefully more such developments will follow.

As you will have gathered from this section, the net offers all sorts of goodies which may be impossible for you to reach – you're left with nose pressed to the shop window, unable to do more than admire the contents. But consider how things have changed over the past decade; benefits do trickle down – eventually.

## 23 SECURITY

Wrap up well on the digital highways – there are plenty of nasty ailments you can catch.

The threats to your computer system are many and varied – and not just to your files or software. Hackers are now increasingly targeting instant messaging systems and even mobile phones. But there are also threats to you as a journalist, if your reporting activities via the net anger or threaten others. We shall look at some of the precautions you may need to protect your anonymity, your sources and the information you have gathered.

Using an unprotected computer on the net is like sleeping out in the tropics with no mosquito net. It's not a surprise when you get bitten, and those bites can lead to long-term harm.

When, in a survey by the British security firm [Sophos](#), computers were connected to the net without things like anti-virus protection or a firewall, half were infected within ten minutes, and 90% within 40 minutes (and that was without visiting websites or opening emails).

There is some good news: poor access to the net means you are less in danger than someone connected with 'always-on' broadband, though dial-up access does not exclude attacks. As your connections to the net improve, so you may share more of the problems that afflict others. You can hope, though, to be saved by the big battalions. Once a threat is seriously hurting global business, an awful lot of resources are devoted to countering it, and their efforts can benefit all net users.

On the other hand, the planned [Windows Vista](#) operating system should start making computers more secure against invasion. If it lives up to the claims, then the bad guys could turn their attention to older, less well-protected machines.

### THREATS, AND PROTECTIVE MEASURES

What is it you can catch?

#### VIRUSES

The traditional threat has been the virus, a type of software that can invade your system, and damage or even destroy its contents. (It won't affect the hardware.) Viruses come in many varieties – there are well over 100,000 recognised.

Viruses are normally programmes or pieces of computer code attached to executable files (especially with the suffix .exe), which can create a nuisance, or even havoc, if you open them. Unopened, they can still be passed on, if a file or email attachment containing a virus is sent elsewhere.

A worm is a type of virus which replicates itself, and typically sends copies to all the people in your address book, while a Trojan Horse hides itself within some more acceptable software and then goes to work nefariously. (Technically the latter is not a virus, since it does not reproduce itself.)

Viruses have usually been distributed by email, often with attractive subject lines encouraging users to open something from an unknown source; they may also arrive with an address which tricks you into believing it is a reliable source.

Users have finally been getting wise to the hazards of email attachments, and become less willing to open them; so the virus producers are increasingly using websites to spread their infections. Malicious code is embedded in web pages, and now even in images – which had been regarded as completely safe. It is also possible to pick things up via a preview pane.

Five hundred or more viruses are discovered monthly. An initial virus will likely mutate quickly: as virus scanners discover the threat, and develop counter-measures, so the virus-writers launch new variants (on a daily basis sometimes). Thus if you get a worm like Netsky (written by a German teenager, whose work caused about half of the virus incidents reported in 2004), it might be one of dozens of versions. Still, big virus attacks, or at least very successful ones, may now be in decline.

### ***HOW TO BEAT THEM***

First, don't open email attachments from anyone you don't know. If you are not sure about something, query the sender about it. If you do have to open something you are unsure about, do it on a floppy disc or similar media, so it is reasonably isolated from the rest of your system.

Second, don't just click on links in the body of an email from someone you don't recognise. Do you know where you are going? Or what you might pick up when you get there? Type an address manually into your browser, to avoid redirection scams.

Third, get proper protection. Software from one of the big consumer suppliers, like [Norton AntiVirus](#) or [McAfee](#), is a common basic purchase. They are not light on your computer's resources – they take up memory and slow the speed, but if you buy a subscription, you will get automatic updates, so you are covered against new threats as soon as the companies get on top of them. As well as receiving signatures for the latest viruses, the scanning engine may be updated as well.

If cash is short, one option is a system developed in the Czech Republic, the [AVG Anti-Virus, free edition](#). It is designated for non-commercial home use only. When you consider an anti-virus service, make sure it will scan the email you receive. Web-based email systems, like [Yahoo!](#) or [Hotmail](#), and some proprietary systems, are not always covered.

Fourth, if you do pick up something nasty, all the big security firms have public listings of the different virus types, with advice on what to do next. They usually show the different subject lines used in the emails, and the texts found in the messages. If you are unsure what you've picked up, try Googling relevant phrases – you usually find that what is happening to you has hit others.

## **SPAM**

The threat from junk email has been mentioned previously, but it is worth realising that most email is now spam. A US company, [Email Systems](#), put the percentage of email on one day in February 2005 as 97.6% spam. Their website shows current levels.

It comes from all over the world, though a lot has been originating in the US, frequently forwarded via computers in Asia.

Pornographic spam has decreased, and efforts now seem devoted to more general sales, such as medical goods, financial products and dating services.

Some of the big internet service providers are beginning to implement controls on mass outgoing emails, and there are many other attempts to attack the problem. Laws against spammers have been introduced in some countries, but it can seem as if legislators are playing catch-up on a fast-moving target. The introduction of reliable ways of authenticating emails looks a likely long-term solution.

### ***HOW TO BEAT IT***

Don't display your email address on the net if you can possibly avoid it, especially not on a public website.

If you do have to do so, try to put it in a human-readable fashion, like: info (at) martinhuckerby (dot) com. Or [Privacy.net](#), [www.privacy.net](http://www.privacy.net), can turn email addresses into ones the spam robots cannot read.

Look for an email filter. If your email account provides such a filter, activate it – they are getting a great deal more efficient. Many have facilities where you indicate, as you delete unwanted email, that it is spam – this educates the system to stop future arrivals from the same source. It is also important to check any 'This is not spam' button when genuine mail appears in a spam folder by mistake. (You may never know what you have missed, as a result of emails deemed to be spam and disposed of without even reaching your spam folder.)

Don't send it back to where it came from, as most spam now comes from phoney addresses – though if it does reach the senders, they'll probably be pleased to hear it's a live address, and worth using again.

## SPYWARE

Spyware, also known as malware or adware, is overtaking viruses as the biggest threat. By and large, this is least serious in developing countries, since it thrives on broadband access to your computer.

It is software that gathers information on a user, and is installed without proper notice or consent. [Symantec](#) announced that in the first half of 2005, the number of viruses had more than doubled, but spyware programmes and the like had quadrupled. [Microsoft](#) believed spyware was at least partly responsible for half of the application crashes reported to it.

In a 2005 US survey, over two-thirds of computer users believed they had suffered from software problems due to spyware; nine out of ten had changed their behaviour as a result, avoiding potentially dangerous sites, and not opening email attachments.

Spyware ranges from passing on information about your net use, or causing pop-up ads to appear, to hijacking your browser – taking you to some site every time you go onto the web. Nastier versions will do things like log your key strokes, which means that if you were using a credit card or accessing a bank account, it could discover details and passwords.

Argument continues over what is spyware. Some of the adware, which passes back information on how you use your computer, is claimed to be legitimate since it is mentioned somewhere in the conditions you accept when downloading other software. (But who reads the small print of the EULA – the End User Licensing Agreement?)

The levels of technical sophistication are considerable. In certain cases, you can download individual components that are benign, but which later combine into something much less pleasant.

An email purporting to carry news of the suicide of singer Michael Jackson directed people to a site for more information. The site stated it was busy, and while they waited, it downloaded its illicit freight onto their hard discs.

In other cases, it is more like extortion: sites dump spyware on visitors, and then sell them software to remove it. One recent appearance is outright 'ransomware': a computer network is infected, and the perpetrator then demands money to clean the system, or unlock documents which have been randomly encrypted.

Multiple attacks are becoming common, sometimes including invaders that are meant to be caught, accompanied by others designed to lie dormant for later activation. Don't worry too much about whether something is spyware, spam or a virus – sometimes it's all of them; the lines are blurring between the different varieties.

As an example of how bad it is getting, one security report found that if you went to the site googkle.com (a common mis-keying for [Google](#)), and clicked on a pop-up page, you could unwittingly acquire two back doors, two Trojan droppers, a proxy Trojan, a spying Trojan and a Trojan downloader. (Don't ask what they all do – you can be sure none of them bring happiness.)

### ***HOW TO BEAT IT***

1) Install a firewall, a barrier between your computer and the net, and keep it on all the time you are connected to the net. This is designed to prevent unwanted intrusions from outside, and control what programmes can access the net from your computer. It will normally ask you for permission before allowing something new to access the net, thus catching spyware trying to send out information.

[Zone Alarm](#) is a decent free system; an alternative is the [Sygate Personal Firewall](#). You can install more than one – which will probably catch more intruders, though sometimes one firewall conflicts with another and affects the way your computer operates. Hardware firewalls are available, and are better at resisting the most virulent attacks, which can disable software versions.

2) Install an anti-spyware system, preferably more than one: it is hard to find a service that will catch everything. They find unwelcome arrivals and eject them from your computer. Set each one to run a regular scan of your computer – do it on a daily basis if you spend much time on the net,

Two free programmes with good reputations are [Ad-Aware](#), from the Swedish-based Lavasoft company (click on Download), and [Spybot - Search & Destroy](#), from Patrick Kolla in Ireland, [www.safer-networking.org/en/home/index.html](http://www.safer-networking.org/en/home/index.html).

Microsoft has [Windows AntiSpyware](#) under trial – though my experience doesn't suggest it is too great at catching anything. Big internet service providers may offer protection systems as well.

Normally, the anti-spyware programmes will simply dispose of the invader, but spyware is being designed to counter such efforts. Some of it is very hard to evict from your system – people developing such materials advertise 'rootkits', supposedly invisible to the user and undetectable to security programmes.

In the US, users have been reported throwing away computers rather than face yet more hours trying to clean them up. My experience suggests that even in quite severe cases, it is possible, by following the advice from security sites and collecting information from [Google Groups](#) and other fellow sufferers, to rid yourself of a spyware problem – just so long as you have time to spend on the task.



## ZOMBIES

One by-product of computer invasions is that software can seize control of an insufficiently protected PC and use it to send out spam, viruses or spyware. Such computers can also be used in 'denial of service' attacks, when an organisation is damaged, or threatened with damage, by having its site so bombarded with messages that it is overwhelmed and has to shut down. The US messaging security company, [CipherTrust](#), said in May 2005 that it was finding daily more than 170,000 new computer zombies worldwide.

This can happen without you realising: if your machine has suddenly become slow and unresponsive, it might be that a zombie programme has been installed.

### ***HOW TO BEAT IT***

Use that anti-spyware programme – and try checking with another if nothing is found by the first.

## ROGUE DIALERS

These take control of your PC and dial-up to the net, not on your regular connection, but to a number which charges extra high rates. Rogue diallers are only a problem if you have such premium lines in your country, and neither a legal system nor phone companies that care much about stopping these activities.

### ***HOW TO BEAT THEM***

[Google](#) the subject if you meet trouble – there is plenty of good advice out there, and it is usually possible to configure your software to avoid further problems.

## PHISHING

This is a process by which spyware and other tactics are used to find out information on people, usually with the aim of fraudulently using credit card and bank details. Full-scale identity theft occurs: criminals acquire enough personal information to pass themselves off as another individual, and start making money. (Communications are also vulnerable to criminals, who can scan emails and access instant messaging.)

In the first half of 2005, [Symantec](#) reported 5.7m phishing messages a day – almost double the previous six months.

### ***HOW TO BEAT IT***

Do not give out personal financial details unless you are completely sure the website you are dealing with is genuine. There are awfully professional looking spoof sites out there.

Choose your passwords carefully. There is spyware designed for hijacking passwords, along with programmes for cracking passwords. If you need to preserve anything of value, don't use an obvious name – indeed, try not to use a name at all, as systems have been designed to find passwords by running through long lists of names; they also run complete dictionaries through password checkers.

For anything important, choose a password which is a mix of letters, figures and punctuation – and change it every 90 days.

Watch out for over-generous special offers. The age-old advice with regard to fraud is still true in the electronic age: if it looks too good to be true, that's because it's false. (It's a maxim which applies to apparently great news stories, too).

## **BACK DOORS**

Among the programmes you can pick up are ones that open a 'back door' in your computer, allowing easy access to your system. Security experts are worried that a lot of what is happening is not being exploited immediately – they believe hackers are planning ahead, intending to take advantage of these weak points in future.

The real problem computer users face is that the threat no longer comes so much from disaffected or ambitious teens, out to cause trouble or show how clever they are. As the net has become important to business, so it is attracting serious criminals, who regard it as a place to make serious money. Online bank accounts are an especial target.

Developing countries are likely to remain on the fringe of the action, since they do not provide so many wealthy targets (though Nigeria is struggling with the other side of the coin, in that its citizens have gained a worldwide reputation as perpetrators of financial scams on the net).

### ***HOW TO BEAT THEM***

Keep your software up-to-date. [Microsoft](#) has been much hit by hackers exploiting gaps in its defences, but it offers automatic updates to Windows, and security patches for other applications. If not using automatic updating, go regularly to the Windows [update](#) site.

The hackers are now taking an interest in other targets, such as the Linux and Macintosh operating systems. So whatever system you use, check how to keep it up to date.

Security companies now offer integrated services, a single product which aims to prevent all the different threats, though it is hard to find one equally efficient in all aspects. It will likely be a while before a worthwhile combined system is available for free.

## PEER-TO-PEER SOFTWARE

Films, music and software have been speeding around the web through the use of file-sharing software (on broadband, again). Programmes like [BitTorrent](#) or [eDonkey](#) aid the process by using your computer, as it downloads material from elsewhere, to share that material with others. It is an egalitarian idea, which has become immensely popular (not least because much of the material is pirated). But it also surrenders some control over your computer.

Your bandwidth is occupied by a supply being funnelled out onto the web, and can be a lot slower as a result, so you need to close your email client (Outlook or whatever) once you've made your contribution to redistributing what you have downloaded. The system is designed that the more you upload, the faster it will allow you to download. Learn more at [BitTorrent](#). It does not work well on a dial-up connection.

Some types of this software deposit spyware on your computer, and it can also take over more than you bargain for: officials at one university discovered that, as a result of someone using peer-to-peer software, their system was simultaneously uploading 30 copies of the Hollywood film, *Kill Bill 2*, and using 80% of the institution's bandwidth in the process.

### ***HOW TO DEAL WITH IT***

Organisations need tight control over their computers; an individual needs to be aware of the implications of downloading some inviting software, film or music.

One general precaution: back-up your files on a regular basis, onto an external storage medium. There are all sorts of programmes and hardware for this, but if you can write them to a CD, that would be a good start – so long as you store these somewhere other than next to the computer. A theft, a fire – replacing all that work and all the software would be a real pain.

A final thought: the internet is well-protected, but could still be seriously disrupted. When in 2004 leading computer experts were asked their expectations for the coming decade, more than two-thirds predicted at least one devastating attack on the net infrastructure. So it's possible the network might go out of action for longer than just a jam caused by virus attacks.

After reading all the above, it might feel as if the sensible course is to hide under the desk. But security companies do have a vested interest in talking up threats. Take precautions, and you should be able to avoid most of the hassles. The criminals are after bigger fish than journalists, and usually concentrate on the richer countries.

And if one German teen was responsible for both the Netsky and Sasser worms, it suggests that the number of other youths trying to cause generalised havoc may not be too great after all.

**MORE?** The US Government's [Cyber Alert System](#) has pretty comprehensive information for ordinary users. As well as the security companies mentioned above, [ThreatNet](#) has good detail on the many types of spyware. Anyone getting serious about the subject can turn to [Astalavista](#) and [Virus Bulletin](#).

## **SPYING – AND BEING SPIED ON**

As a reporter, you may want to discover secret or private information, while simultaneously protecting what you are doing from the scrutiny of others.

Assume anything you send over the web can be monitored and recorded (and that goes for mobile phone calls, landline calls, text messages and other communications) – possibly by criminals; probably by one or more governments.

Quite a lot can be discovered even by public methods. Email systems can be configured to notify users when a message is received, and Outlook can obtain information on whether a message was deleted without being read. More sophisticated tracking is being marketed which can tell senders when an email was read, and how long was spent doing so.

If you worry about covert surveillance of your messages, you may be able to communicate with a contact more safely via public internet points, between one anonymous webmail address and another, but internet cafes and the like can still be monitored, especially by governments.

**CONCEALMENT:** Scrambling your data offers much greater security, but has its complexities. If you are using Windows XP, this provides an encryption system. PGP (Pretty Good Privacy), described as “the closest you are likely to get to military-grade encryption”, comes in [free versions](#) for different operating systems.

The new [Ciphire Mail](#), [www.ciphire.com](http://www.ciphire.com), developed in Germany, offers a reasonably straightforward version of this: people wanting to exchange private messages use the other's 'public key' from a Ciphire directory to encrypt messages, which can then be decrypted by the recipient's private key. The initial [Ciphire Mail](#) version will remain free not only for individuals, but also to the press.

A problem is that reporters may need to do things not only without being detected, but without it being apparent that they are doing something underhand. Encrypted messages are a signal you have something to hide. Remailers are one way around this: [EmailPrivacy.info](#) has information on how to send email which hides the sender's identity, though something approaching real anonymity takes a bit of work; the site includes a list of free remailers.

If you want to see how easily your presence can be registered on the web, go to [Privacy.Net](#) and click on Analyzer; it will trace the electronic route back to your machine, as well as detecting things like what the time is where you are. Go to the associated beta site, [network-tools](#), to run a trace on any computer address.

One way to disguise your involvement is using third-party systems, which provide anonymous surfing of the web. Normally when you visit a site, you reveal your computer address (your Internet Protocol, IP); anonymising sites seek to get round this, but can arouse suspicions.

The American digital rights organisation, the [Electronic Frontier Foundation](#), aims to go further with the [Tor](#) system, which bounces data packets round a network of servers until it is hard for anyone to know where they've come from. Still under development, [Tor](#) covers instant messaging and other applications as well as browsing. (Do read the manual carefully before trying to install it.)

**YOUR PC:** Covering your tracks on your own computer is extraordinarily hard. There are just so many places in your PC that store information about where you have been on the net, material you have received, downloaded, and worked on. You can delete items, but they may also be listed in internet histories, desktop indexing, system restore and backups.

Since deletion only means that eventually the data may be over-written, you cannot be sure it will all be erased. There are programmes that promise total cleaning of a hard disc, but I'm not sure I would bet my freedom on them.

If there is any likelihood of your computer being examined by experts, seeking something you wish them not to find – then the safest thing to do is not use it in the first place. Find another computer, one to which you cannot be connected, or, if you are doing really sensitive reporting, ask yourself whether you can avoid electronic means altogether.

You can encrypt the whole contents of a computer, but if you are in a situation where you need to appear an innocent, that will just arouse serious suspicions – quite enough to be thrown in jail by plenty of governments today.

**THERE ARE PLUSES:** Journalists have benefited, though, as well as suffered, from the fact that a computer file can carry with it all sorts of metadata – like information on the original author, who edited it, and when; a revision log may show who else has looked at the document, and much more can be discovered. Which is what happened in the UK, to the embarrassment of 10 Downing Street, when someone started looking at the metadata on the government's controversial dossier on Iraq's security. (See the [article](#) on metadata.)

A different vulnerability was demonstrated when the US military published a report into how they had mistakenly killed an Italian intelligence chief in Iraq. They blanked out the name of the soldier responsible, but an Italian, who cut and pasted the file into another application, found the electronic blacklines had gone, revealing the soldier's name.

Trace! software, available from [Workshare Inc](#), will provide information on what has happened to Microsoft Office documents.

The UK-based Institute for War and Peace Reporting has a detailed [analysis](#) (click on Web Security under Subject-specific) of possible threats to a journalist's privacy online, and how to counter them. It sensibly stresses the importance of studying the motivation of those who might probe your work – would it be worth their time and effort to discover all you have done via the net?

[Journalism Net](#) has good material on the subject of spying and being spied on – some of it is a little dated now, but there's plenty to worry about.

## 24 PRACTICAL HELP & SOFTWARE

As a net user, you are sure to need help in dealing with individual problems, and you also require the best possible tools.

This section looks at the basics of finding help, and then examines the free software you can obtain – we have included here some of the most important, so you don't have to worry about downloading it.

### WHEN THINGS GO WRONG

Remember the favourite mantra of computer support staff around the world: “Switch it off, and switch it on again.”

If you don't have an expert in your office whom you can call on, the essential advice when you hit trouble is twofold: use the help systems that come with individual applications and operating systems, and turn to the collective knowledge and experience available on the net.

Help systems vary in user-friendliness, but usually cover the common problems, whether it is computer equipment or software. If you are not finding what you need, consider whether you can rephrase how you described the difficulty. What something is called in one system can be completely different in another.

Operating systems like Windows or Linux, software like Microsoft Office or OpenOffice, have all sorts of facilities you will have forgotten about (or never knew existed). So spend a few minutes checking things out in the help directory and perhaps glossary, as well as doing a search.

Sometimes, though, you may not even know where to start. Don't throw up your hands in despair. If you have a few words to describe the problem, or an error message, or the name or title of something specific, put them in [Google](#) search – even a set of initials can turn up a link which will begin to explain what you are faced with. If you get results indicating something might be a virus or spyware, you can then work out more precisely where to look for help.

Two other good places to check:

[Google Groups](#), where the Computers section shows how others have been tackling the problems;

[CNET](#), where you can search articles, reviews and news, as well as its forums.

If you cannot see what you need, you may want to ask a question in one of the groups or address a query directly to someone who seems to know. Do, though, make a proper search first – the answer may well be in a FAQs list.



If you enlist someone's personal aid, do thank them – a lot of people on the net give freely of their time to help others. If you've been helped in the past, you may want to assist others in return – and there's nothing like that feeling of superiority when you realise you actually do know more (very slightly more) than someone else and can help a novice sort out a problem.

## SOFTWARE

In many countries, pirate versions of [Microsoft](#) are the standard software. With prices the way they are, it's hard for an impoverished news organisation to go legitimate – but rather than be stuck with illegal, and possibly faulty, versions of [Microsoft](#), there is an alternative: [Open Source](#). This type of free software used to be the preserve of expert users, but much effort has gone into making it user-friendly.

There is an extra plus in migrating to [Open Source](#): [Microsoft](#) products have been castigated as 'bloatware', for their sheer size. Windows XP has over 40m lines of computer code; its replacement Vista threatens to have twice as many. There are reasons for this – such as that it has to support a couple of hundred thousand different devices. But if you've a computer with only limited capacity, you want the smallest packages of software. GNU/Linux, the Open Source operating system, has only about 5.7m lines of code.

While you may not want to get into changing the operating system on your computer, you may want to consider some of the major software available through [Open Source](#).

If you have Windows already, but balk at the price of Microsoft Word, its word processor (usually bundled expensively in [Microsoft Office](#)), then look at the comparable [OpenOffice.org office suite](#). You won't find it too alien. It looks and works similarly to MS Office – Writer is the word processor, Calc the spreadsheet, Impress the multimedia presentation tool and Draw the drawing tool. OpenOffice will open Word, Excel and PowerPoint files, and you can save in the Microsoft formats – it makes sense to save word processing documents in the .doc format, so they can be read by people all over the place.

The suite is not as slick as MS Office, and some aspects of the help system are heavy going, but overall it works effectively. It's better, for instance, at ensuring that when you copy something in one application, it can be pasted into just about anything else you have open. There is information and support in more than 30 languages – Hindi, Tamil, and Turkish among them. A version 2.0 is already under test.

## FREE SOFTWARE DOWNLOADS

[Open Office suite](#)

[Firefox](#) browser

[Thunderbird](#), for serious email users, a fast email client that complements Firefox

[FreePops](#), an email aggregator, collecting mail from various accounts

[Audacity](#), an audio recorder and editor which will handle a variety of file formats

[Dave's Quick Search Deskbar](#), a searchbar with many varieties of search on offer

[VLC Media Player](#), to play most every kind of file

[W.Bloggar](#), for working on blogs offline

[7-Zip](#), for compressing files and opening ZIP files

[Adobe Acrobat Reader](#), for reading and searching PDF files

[AVG Free edition](#), anti-virus protection

[Spybot – Search & Destroy](#), spyware protection tool

[Ad-Aware Personal Edition](#), spyware protection (non-commercial use only)

[ZoneAlarm](#), firewall (non-business use)

[Outpost Firewall](#), a Russian product, with a basic version for free

## SOURCES FOR DOWNLOADS

[SourceForge](#) is a central point for information and downloading of [Open Source](#) software, but if you want something more suited to the ordinary user, two reliable sites are CNET's [Download.com](#) and [Tucows](#) – both provide some explanation and assistance. For a listing of 4000+ packages of free software, there's the Free Software Foundation/UNESCO [directory](#).

## FURTHER ADVICE

One problem with the net, and the world of computers generally, is finding time to stay up-to-date with all the changes. Keep an eye on the local press for a columnist who surveys new developments and offers advice on problems and purchases. [Jack Schofield](#) in *The Guardian*, and [Walt Mossberg](#) in *The Wall Street Journal*, are two good advisors, but a knowledgeable writer who understands the local scene, and what is feasible and available in your country, would be ideal.

- For understanding processes, a useful starter's guide is [How Stuff Works](#)
- For practical advice, a good source is [About](#)
- For detailed information and definitions, try the [Wikipedia](#), or [Webopedia](#)
- For comparative information on products and services, analysing many different reviews, see: [ConsumerSearch](#)
- For delving deeper into the world of search, [Pandia](#), run by Norwegians Per and Susanne Koch, offers straightforward introductions to many aspects of web searching; [Search Engine Watch](#) concentrates more heavily on search engine marketing, but its team – Danny Sullivan, Chris Sherman and Gary Price – are three of the best-informed writers in the field.

## 25 WHAT NOW?

If you found useful things in this book, you need to work out how to take advantage of the suggestions – there are probably too many things to act on straightaway, or remember easily.

Try making a 'to do' list, in rough order of priority, for matters like browser settings, installing software etc.

Then draw up a second list, of things to try on a daily basis, like searching techniques you want to practice – so they become a everyday habit.

The best way to find anything again is to keep a copy of the text on your desktop. Then just hit Ctrl & F, or use a desktop search.

Then, in a few months, come back to skim through the book again. There are sure to be suggestions that did not concern you on first reading. With more experience, you will probably be keen to tackle more complicated procedures.

A final thought: The net is a tool you may be using for the rest of your life – put in a little work now, and you could save an awful lot of time and effort in the decades to come.

## TOP TEN TIPS for better searching

- Use more words in your search – try six or eight
- Use an exact phrase (when it is likely to be in the pages you seek)
- Use quotation marks around phrases and full names
- Choose unusual search words
- Choose nouns and objects, in preference to verbs, or other types of word
- Put the most important term first among the search words
- Check your spelling (it's the biggest problem among searchers)
- Turn off images in your web browser, so pages load faster
- Do not believe all you read on the net – check anything doubtful
- You cannot find everything on the net – be ready to look elsewhere

Anyone with suggestions about how to improve the book, or complaints, corrections, or other ideas, can contribute to an online forum about the book, at [www.unesco.org/webworld/internetforjournalism](http://www.unesco.org/webworld/internetforjournalism), or otherwise write to [h.padhy@unesco.org](mailto:h.padhy@unesco.org).

Punctuation: after some URLs in the book, you may find a comma or full stop, placed there for grammatical reasons – always delete this if typing in the address.

Spelling: British English is the basis for spelling in this book, but American English is used where required in names, titles and addresses.

## ADDITIONAL RESOURCES

### A. DO-IT-YOURSELF NET TRAINING

Seven chances to improve your prowess on the net  
Training exercises for individuals

Gain some practice in using the net to find information. The first exercises are designed to try out finding and assembling facts and figures; later ones will test your skill in obtaining information and assessing its quality. In the course of the activity, you should get the chance to experiment with some of the different tools available for the net.

If you cannot think how to find something, try searching the contents of the book for the technique you need, or the type of information you want to find.

**1) TRY GOOGLE:** Start out by trying [Google](#) search, to look for information which should be readily available.

Choose the public figure in your country whom you most admire, and start asking Google for information about him or her. (*Note: if there is nothing about them available in English, choose an alternative, or a figure from some neighbouring country.*)

See if you can find out their:

age

birthday

birthplace

family

education

career

achievements

Have you enough to write a brief biographical note about them?

Can you find their address? Phone number?

You may think the easy way is to go to a single reliable source which will have all or most of the information needed. But is all the information correct? Look at alternative sources, to see if everyone agrees on the facts.

While collecting all the details, did you come across anything interesting about the person? Any ideas for an article about them? (These may be exercises testing out your net skills, but never miss opportunities to pick up story ideas.)

**2) TRY AN ALTERNATIVE:** Choose a different search engine – one you have not used before.

What is the biggest company in your city? Does it have a website?

*(Note: If the company has no site, find another big concern, in the city, or elsewhere in the country.)*

See if you can tell from the company site,

- its annual earnings or turnover
- its annual profits
- its total value (in shares or some other measure)
- its total staff

Can you compare this information with the previous year's results? Are things getting better, or worse?

If the site does not give any of the above information, can it be discovered elsewhere?

Even if the site has provided all the necessary information, check the details with another source – is the company giving an accurate picture to the outside world?

While looking at the site, check when it was last updated.

Did this search engine prove better/worse than Google – if there was a difference, remember where this engine might be valuable in future, or when it should be avoided.

Any story possibilities? Not just from the info you found, but from what you didn't find – the absence of details about profits, ownership, future plans can be significant.

**3) COMPARE AND CONTRAST:** Try different search tools in pursuit of the same information – to compare their usefulness. Use [Google Search](#) & the [Yahoo Directory](#) & the [Ixquick](#) metasearch.

Spend five minutes on each of the three looking for how many people in your country have AIDs.

*(Note: If your country only keeps figures for those who are HIV-positive, or combines both sets of figures, pursue those totals. If your country has absolutely no published figures, pursue the total for a neighbouring country.)*

At the end of five minutes, in each case, have you got a result you believe is as accurate as possible?

Even if the source is the same, was it easier to find with one of the three approaches?

Which was quickest? (It's possible it was the last one simply because, by then, you knew what you wanted!)

Do you, as a result of looking at different sources, now have any doubts about the reliability of the public figures?

Can you find a source, in your country or elsewhere, which might tell you how best to assess the accuracy of what you found? (It might, for instance, be a world organisation, a medical research group, or an NGO.)

Think about the merits of the different tools you used – whether any had obvious advantages for you.

Look at how much the findings of the metasearch engine varied from those of Google. Were the specific pages largely different? If they were, did what was produced by Google nevertheless cover most of this material in other ways?

Ask yourself: if I ran into trouble trying to find the information, did I think about asking the query differently? Or did I just continue clicking between the initial results? Did I waste time in the process?

**4)JOIN A GROUP:** Subscribe to a newsgroup, a mailing list or a message board.

Select one which could be useful for your work..

Choose three possible alternatives that might be useful, and make an assessment of them, based on who runs the group, the sources of information (the quality of the contributors or of the material distributed), the timeliness of the discussions or material, the reliability of the service.

Select one, and join up. After two weeks, make a further assessment: Did it live up (or down) to your expectations? What have you received or learned as a result? How reliable was the information? Do you feel you've spent the time usefully, or wasted it? Could you have obtained the information elsewhere – quicker or with less effort?

And, did you get any story ideas as a result?

**5)DETAILED COMPARISON:** Next time you have to do a search for work, make time to do exactly the same search with two of the main search engines, and compare the results alongside each other – not just to see how they differ, but to get a better feel for which might best suit your needs. You can use [Dogpile](#)'s comparison tool, which covers [Google](#), [Yahoo!](#) and [Ask Jeeves](#).

**6)PROBE A SITE:** Choose a foreign site that interests you. It can be one which would be useful for work, but it might just be a subject of personal interest. Then you must discover what happens behind its (glossy?) facade.

*(Chapter 12 has relevant advice, but there are other ways of discovering much of this information.)*



Look first at the domain name, to see what you can guess about the website, and then answer the following:

- Where is the site based? (There may be dual answers – the location it is hosted on the net and the location where the people running it reside.)
- Who owns it, or runs it?
- Is there a phone number, address or email address to contact?
- Who or what is the officially registered contact for the site?
- When was the site started?
- When was it last updated?
- How fast is it, in terms of speed of loading?
- How popular is it? How authoritative? Find three different assessments of how many people use it or how it is regarded.

(Don't worry if you cannot discover every answer – sometimes, the information is not all available.)

## **7) FINDING FACTS**

How many people have been displaced by the Three Gorges Dam in China?

This is a test of analysis as well as search techniques. You need to find the latest information, from the most reliable sources – which could be local, national or international.

Even if there is an agreed official total for the final number of people who will be displaced, how many have been displaced so far? Whose figures are the most accurate?

The basic information sought seems simple, but try to refine your search – use different search queries, or even different engines, to see if you find extra material.

You will doubtless find it hard to come up with a single figure which you are satisfied is definitely correct. But assess the various statements and the sources they come from – study the veracity of the organisations involved (and their past records for accuracy).

What, at the least, is the very best estimate you could offer?

## **B. ONLINE TRAINING COURSES**

### **THE BASICS OF COMPUTER USE**

The BBC has a very simple guide for those starting out on the net, originally produced for family viewing. The [Web Wise](#) online course uses flash animation, so, for the individual sections, you will probably need to click on 'text only' (unless you have a good net connection). But the illustrations certainly help you get acquainted with what everything looks like.

### **JOURNALISM TECHNIQUES**

[News University](#), the online training programme from the Poynter Institute, offers over 20 free short courses, which take an hour or two per module. They cover such subjects as writing story leads, revision and rewriting, interviewing, learning to be a better listener, and maths for journalists.

### **INTERNET & INFORMATION RESOURCES**

[REPORTSTAGE-eLearning](#) is a new web-based training site, funded by the European Union, and bringing together a wide selection of online courses for journalists, from a variety of countries. These cover everything from aspects of web searching to digital photography and creating a website; there are links to sites with further resources.

### **ONLINE JOURNALISM**

It is not strictly a course, but the Cyberjournalist site's [Online Journalism](#) has a good introduction to the different forms of writing for the net, with lots of examples; it has further [Online Writing Tips](#).

## FOR CLASSROOM TUTORS

### C. LESSON PLANS

Before you begin, you may want to put round this short questionnaire, to find out what people have done on the net before.

Frequent use of the net is usually the best indicator of someone's abilities.

(I suggest not asking people to put their names, to reduce the temptation to exaggerate about their experience.)

#### NET USE QUESTIONNAIRE

*Tick all answers that apply. (Don't worry about putting your name on it.)*

Do you have an email address? \_\_\_\_\_

Do you use the net regularly? \_\_\_\_\_ Occasionally? \_\_\_\_\_ Not at all?  
\_\_\_\_\_

Do you use the net for work? \_\_\_\_\_ For personal matters? \_\_\_\_\_

If you're a regular user, do you use it daily? \_\_\_\_\_ Several times a day?  
\_\_\_\_\_

Do you have an 'always-on'/broadband net connection at work? \_\_\_\_\_

If you have the net at your office, is there a net-equipped computer for your sole use?  
\_\_\_\_\_

Is there one search engine you normally use? If so, which one? (write in)  
\_\_\_\_\_

Do you belong to any internet groups or mailing lists? \_\_\_\_\_

Have you ever bought anything over the net? \_\_\_\_\_

*(Vary the questions as necessary for your local situation. Delete the reference to broadband if it doesn't apply. If everyone is from the same organisation, it probably won't be necessary to ask about the computer facilities they have.)*

## INTRODUCTION:

The lessons are graded in terms of difficulty, growing harder as people continue. Groups tend to have mixed levels of experience, so sessions are often structured as a sequence of activities. There should be no pressure on the slower ones to complete all the questions. The process is designed to accustom people to using the net – it's better that someone happily finds out one piece of information than struggles to find lots. Whenever a person is having real trouble, switch the task to something simpler. For this reason, it usually makes sense to avoid sessions where people call out the answers when they find them – a few are quick with the results; the rest are left feeling inadequate. Expect to spend plenty of time on the mechanics of the search process. You are likely to devote more energy to helping people with controls, or getting them out of problems as they navigate the net, than you are on analysing their results. But it's all part of helping them to feel at home in a strange environment.

Most of the exercises are not quizzes, with right answers, but more open-ended, where the task is to find the best answer, perhaps the most acceptable estimate. Where there are questions demanding specific answers, the latter are interleaved in italics. The times given here for lessons are very approximate – a novice group with slow computers could easily take twice as long as these timings. Each lesson should include a substantial period (perhaps 20%) for discussing the answers – not just the actual results, but how people found them. They can learn a lot from the problems met, not just from the successes. All the lessons should be adapted as much as necessary to fit local and specific needs.

THINGS TO WATCH FOR: Inexperienced users may believe that, if they cannot find the information (usually with the one search tool with which they are familiar), then it isn't there. They need to be encouraged to expand their options. With more experienced users, over-confidence is common. Reliance on the first speedily-found answer needs to be replaced by a willingness to study how to use the full potential of search tools.

As I have tried to do in the book, it's necessary to keep a balance between turning people on to the immense benefits of the net, while ensuring that they don't reach the stage of believing it has all the answers.

### LESSON 1 **STARTING OUT** (60minutes?)

A few fairly simple queries, just to make sure people know how to carry out some basic searches – use [Google](#).

1 Which country has an internet code .bt?  
(*Bhutan – IANA will have the info, but an ordinary web search on the initials should provide the answer.*)

2.If you have a pint of beer in Britain, are you going to drink more or less than if you have a pint of beer in the US?

*(More – the Imperial pint is bigger than the US one; the Google calculator should help, though a straight web search should offer the info.)*

3.Translate 'Je t'aime' from French into your language.

*(It means 'I love you' – if that might cause embarrassment rather than laughter, substitute 'Je ne comprends pas' – 'I don't understand'. If it proves impossible to find a direct language translator between the two languages, encourage people to translate the phrase into English and then into the second language. See Translators in Chapter 11.)*

4.On President Bush's Texas ranch, where is the swimming pool?

*(There is a Google aerial photo of the Crawford, Tx, ranch, available on [www.eyeball-series.com](http://www.eyeball-series.com), but searches for maps as well as aerial photos may provide the answer. First, though people have to find the place name – a simple web search.)*

5.Participants should choose the city they would most like to visit, and find out the weather there. Is that hotter or colder than usual? *(See Weather in Chapter 11.)*

6.Look for the official Paris tourist office, using three different search engines: [Google](#), [Ask Jeeves](#) and [Look Smart](#)

*You may have difficulty in working out which is the government site for Paris tourism (likely the main and most reliable source for information on the city). It is the [Paris Convention and Visitors Bureau](#), but you may not find it very prominently. See which of the engines gets closest to the official site.*

*Then look at other results: there will be commercial sites, using different tricks to get to the top of the results. Look at how many times a page may repeat 'tourist office'. Also consider how a site like Fat Tire Bike Tours (or some other, equally unlikely, top selection) gets so high up – it might be, on at least one search engine, that they have paid to do so.*

## LESSON 2:

### **TRY AN ALTERNATIVE**

(75 minutes?)

People should try a different search engine – one they have not used regularly before (perhaps [Yahoo!](#), [MSN](#) or [Ask Jeeves](#)).

Assign each person to research a company, in your city, which has a website – if the group is less experienced, you may wish to have everyone look at the same concern.

*(If there is a shortage of companies with a significant website, choose large companies from elsewhere in the country. If absolutely necessary; choose companies in a neighbouring country.)*

Ask them, if they can tell from the website, the company's

- annual earnings or turnover
- annual profits
- total value (in shares or some other measure)
- total staff

Can they compare this information with the previous year's results? Are things getting better, or worse?

If the site does not give any of the above information, can it be discovered elsewhere?

Even if the site has provided all the necessary information, they should try to check the details with another source – is the company giving an accurate picture to the outside world? While looking at the site, they should check when it was last updated.

Did this search engine prove better/worse than Google? If there was a difference, remember where this engine might be valuable in future, or when it should be avoided.

### LESSON 3

#### **COMPARE & CONTRAST**

(40 minutes?)

Try different search tools in pursuit of the same information – to compare their usefulness. Use [Google Search](#) & the [Yahoo! Directory](#) & the [Ixquick](#) metasearch (if necessary, remind the group briefly of the difference between these).

Ask the group to spend five minutes, with each tool, looking for how many people in the country have AIDs.

*(Note: If organisations in the country only keep figures for those who are HIV-positive, or combines both sets of figures, pursue those totals. If the country claims to have absolutely no figures, pursue the total for a neighbouring country.)*

If you prefer not to look for AIDs figures, choose another statistic which might be debatable, or still uncertain. How many people, for instance, were injured in the last calamity to strike the country? (Injured, not killed – the latter figure is more likely to be kept accurately.) The aim is not just to develop skills in seeking info, but to increase people's ability to analyse what they find.

At the end of five minutes, in each case, have they a result they believe is as accurate as possible? (You don't need to be rigid about five minutes – they might need twice as long – but sometimes people need reminding to switch over.)

Even if the source was the same, was it easier to find with one of the three approaches? Did they notice if any of the tools had particular merits?

Which was quickest? (Sometimes, though, it was the last one simply because, by then, people knew what they wanted!)

Look at the sort of questions people asked – did they go about the search query in the most efficient way? If people ran into trouble, did they redo the query, or simply continue clicking between the initial results?

Ask people to look at how much the findings of the metasearch engine varied from those of Google. Were the specific pages largely different? If they were, did what was produced by Google nevertheless cover most of this material in other ways?

As a result of looking at different sources, are there any doubts about the reliability of the public figures?

If people feel there are doubts, ask them to do a further brief exercise: Can they find a source, in the country or elsewhere, to say how best to assess the accuracy of what they found? (It might, for instance, be a world organisation, a medical research group, or an NGO.) [*This latter exercise may not work for a subject different from AIDs.*]

#### LESSON 4:

##### **PROBE A SITE**

(60 minutes?)

Let people choose a foreign site that interests them (*you can select a topic, but you might wish to let people just pick a hobby or something personal – it does not have to be work-related*) and then get them to find the basic information about it.

What happens behind the facade?

*(Chapter 12 has relevant advice, but there are other ways of discovering much of this information. Ask people to write up their answers for later review. Or otherwise make this a group exercise, with everyone using the same site – and a discussion at the end comparing their results.)*

Ask people to look first at the domain name, to see what they can guess about the website, and then answer the following:

- Where is the site based? (There may be dual answers – the location it is hosted on the net and the location where the people running it reside.)
- Who owns it, or runs it?
- Is there a phone number, address, and email address to contact?
- Who or what is the officially registered contact for the site?
- When was the site started?
- When was it last updated?
- How fast is it, in terms of speed of loading?
- How popular is it? Find three different assessments of how many people use it or how it is regarded (such as the number of visitors, its page rank from some established source etc)

(Not every question may be answerable; if so, people should explain what they tried in pursuit of the information.)



## LESSON 5:

### **OBTAINING REFERENCE INFORMATION**

(80 minutes?)

This is a test of people's ability to check out ready reference questions (*see Chapter 11*). Some queries are deliberately tricky, and may best be answered by a variety of methods.

*Ask people to jot down their answers, how long each took to find, and a brief note of how they located the information.*

*(These should be just rough notes – for discussion, not to be handed in)*

1) How many togrogs are there to the US dollar?

*Togrogs are the national currency of Mongolia – but the currency can be spelled in all sorts of ways (which include Tugrik/Tughrik/Tugrul). The value obviously changes – ask people what rate they found, where they got it from, and which day it was for. Praise goes to the one with the latest result (though someone who gets a particularly good rate might be the better business person, if they can change lots of cash at that rate).*

*The National Bank, [mongolbank.mn/rates.asp](http://mongolbank.mn/rates.asp), could be the most reliable source – but might not be the most up-to-date.*

2) What is the telephone number in your city for .....?

*For this question, write in the name of someone whose number is available online (if there is not a person, choose an organisation), but select someone whose number could be found more quickly without using the net. Preferably it should be by looking in a local phone book (or some other well-known reference source), if there is one in the room; otherwise, it could be by asking someone in the office/training centre. The aim is to remind them to use common sense about the net – they may be doing a net exercise, but if someone next door is sure to have the answer, why jump through unnecessary hoops?*

3) If you are in Hong Kong and need today a copy of that valuable journalism handbook, *The Universal Journalist* by David Randall, where would you find one?

*This is an opportunity to get people searching suitable libraries and bookshops. (They can't go to Amazon to buy the book online – it is needed straightaway.) Last time I checked, there were copies in two branches of the public library – the one in the Central Library was on the shelves. But there are also university libraries and bookshops which can yield results.*

4) What are the official recommendations for travel to Indonesia?

*Whose advice? It is up to the participants to decide: they can take the travel advisories of the US, or UK, or Australian governments, mentioned in Chapter 11, or they might take the advice of the country's tourist ministry – not surprisingly, these have varying views on visiting the country. And what about health recommendations? Did people also look at the travel health sites?*

5)What is at nineteen degrees north and ninety-nine degrees west?

*Mexico City, though the sharper ones may come up with specific settlements to the south of the city itself. You can just search on the figures above, apart from looking at more sophisticated sites. Check how they found the information: if there's an atlas in the room (or an encyclopaedia with a map section), it could take just seconds to find the answer there.*

6)What is the distance between Kabul and Kandahar?

*This may yield widely differing results – there are plenty of web references, to road distances (some about the lengths of roads being built, rather than the precise distance), to traditional distances (like the forced march made by a British army during their Second Afghan War) and to figures from official sources and old guidebooks. Statistics in Afghanistan tend to be somewhat flexible – 491 kilometers may be right, but listen to anyone with a cogent argument that the figure is something else.*

*There are various sites claiming to show the distances between vast numbers of cities, but they don't seem to cover Kandahar. Mathematically-inclined journalists can try finding the latitude and longitude for the two places and thus calculate the distance.*

*(If anyone finishes early, refer them back to Question 4, and ask them to find tourism recommendations for Basra, in southern Iraq. If they find a Guardian article from June 2005, it will include the ineffable advice from the city's head of tourism, urging visitors to come, and saying that if they dress like locals and hire armed escorts, then there is "a 70 to 80% chance" they will be OK.)*

#### LESSON 6:

#### **FIND YOUR WAY IN L A**

(90 Minutes?) Only use if broadband internet connections are available.

This exploration of local search uses systems which are not yet relevant to many countries, so you may want to skip this lesson. But it helps journalists to think laterally – to complete these exercises, they have to keep switching between different search tools. And it tests their ability to discover what is in places they cannot reach physically.

Increasing numbers of net facilities help find locations. The latest are still only being trialled in the USA, but see if it is possible to traverse one part of a street in the city of Los Angeles, via the net.

*(People should make a brief note of the information found, and how they got there – not to be handed in.)*

1)Where would you find dire wolves in Los Angeles? (Clue: a ranch)

*Dire wolves no longer exist – but their remains can be found at the tar pits of Rancho La Brea. A general web search, using Google and the like, should reveal this.*

2)What is the name and address of the building?

*It's the [George C. Page Museum](#), 5801 Wilshire Boulevard. The web search should also yield this info.*

3)What is the public institution, of world importance, just to the west along that street?

*There are various ways to discover this – with a search on Wilshire Boulevard on numbers between, say, 5000 and 6500. A tourist map of the city might prove useful. But if people remember the [A9 Maps](#) yellow pages search, which brings up photos of buildings, then it is possible to scroll along the street looking at the images until reaching the next major building – which is the [Los Angeles County Museum of Art](#) (LACMA), at 5905.*

4)What were they advertising there recently?

*Using A9, it is possible to read the banner outside the main entrance, which used to read “Experience the world through art”. Or use the gallery's [website](#) to see what is their latest exhibition (King Tutankhamen, at last look).*

5)What similar institution is on the other side of the street, at the intersection to the west?

*A9 would show, by scrolling through the photos along the opposite side of the street, the [Petersen Automotive Museum](#) at 6060 Wilshire. But a tourist map might be the quickest way to spot this (Not to be confused with yet another two museums in the same three block stretch – the California Craft & Folk Art Museum or the Kaye Museum of Miniatures.)*

6)What notorious event happened there in March 1997?

*The obvious thing to do here is to go to the Petersen museum website and search, but this will do little good. The event in question was the murder of the rap singer, Notorious B.I.G, outside the building – it's not something that the museum would wish to advertise, but a web search back on one of the main search engines should bring it up*

*You'll probably have to help many people finding the tools for this lesson, but others will complete their tasks faster.*

*Ask the early finishers:*

*How would you get to these museums from Los Angeles International Airport?*

*(Tell them not to worry about the details – just to note where they got the information.)*

*And if they proudly announce they've done that, too, check if they have got driving instructions – the normal type of directions you'd receive in the USA. If so, ask if they can afford to hire a car, and then tell them to find out how to get there by public transport (the L.A. [Metro](#) should get them there by bus for about \$2.50).*

NOTE: People may use better ways than those mentioned here to discover the information – especially with the different map tools being trialled, an inventive searcher could find more efficient methods to obtain the required data.

## LESSON 7

### **FIND THE UNFINDABLE**

(120 minutes?)

Where is Shangri-La?

This could be a philosophical question, but people should concentrate on the geographical aspects. There is no correct answer, since it is a mythical place, but that does not stop localities claiming ownership.

Use the net to collect information on the claims.

It's not quite as easy as it sounds – it is necessary to avoid all the results about the Shangri-la hotel chain, rock albums and more. So, in this lesson, participants (using whichever search tool they prefer) should practice refining their search. If receiving loads of irrelevant material in the initial results, they should not just start searching one page that might be useful. Instead, they should rewrite their search query more precisely – perhaps including extra words, phrases in quotes, or possibly excluding some terms. They should do this several times, if necessary, until they have a good selection of potentially useful results. And they should keep modifying the queries as they chase further information.

If, after much searching, they (or you) are still not satisfied the field has been properly covered, they should go back to the beginning: either try a completely different search query, using an alternative approach to the subject, or consider whether to use another search engine, or a different search tool. (Indeed, even if people are satisfied, each should have at least one last try with a different engine, to see if something new does turn up.)

Other answers they should seek:

How many localities have made claims to the title?

Are there any criteria for judging which claim has the greatest support? Among the public at large? Among any other group which might be relevant – travel writers? literary critics?

They should think of the work as research for a potential feature or travel article – it should produce sufficient information to write an outline for a prospective editor. (*It would not be acceptable for participants just to find one source purporting to say where Shangri-La is, and quote all that person's assertions – that would be more like plagiarism than journalism.*)

At the end, find out whether they felt it was useful to refine the search, or whether they would have preferred just to pursue some of the first results – discuss whether their views are coloured by the fact that the first way involves less initial effort.

As well as the techniques used, you may wish to discuss the search conclusions, and perhaps conclude the session with a vote on which place is the best candidate for Shangri-La.

IF FURTHER TRAINING OPTIONS ARE REQUIRED, CONSIDER TWO OF THE SEGMENTS FROM THE DO-IT-YOURSELF PROGRAMME:

Join A Group (Number Four)  
Finding Facts (Number Seven)

Both can be adapted for classroom use.

## **D. Eight ways to save words and write better English**

- 1) Never use a long word where a short one will do.
- 2) If it is possible to cut a word, always cut it.
- 3) Make every word count — treat words like a miser. Think of each word as a coin you could save.
- 4) Don't use long complicated words to impress people; use short, simple words to inform them.
- 5) Avoid clichés "like the plague".
- 6) Try to use active rather than passive verbs: 'The train hit a car' rather than 'A car was hit by the train'.
- 7) Try not to reverse sentences by putting a subsidiary clause first.
- 8) Break any of these rules rather than write something which is unclear, inelegant or downright foolish.

*Credit for much of the above goes to journalist and author George Orwell.*

## E. FREELANCING VIA THE NET

The net offers new opportunities for journalists to write or broadcast for the international media.

Instead of wrestling with expensive foreign phone calls, or telexes, or hoping a mailed report will eventually get through, you can send something cheaply and swiftly to news organisations everywhere.

Overcoming such communication hurdles, though, is not the only advantage: you can now also access foreign media to see and understand the sort of material they publish or broadcast .

There are journalists across the world who believe their stories only have to be delivered to a newsdesk in some foreign capital, and, lo, the material will be published.

Unfortunately, the outside world is rarely that interested in what happens where you live. There are exceptions – but most people in the rich countries are interested in themselves, their immediate neighbours and just a few places around the world they've heard about.

They can be attracted to something from your homeland, but it needs to be styled and presented in a fashion tailored precisely to their needs and interests.

This is where the net scores: you can look at stories, watch or listen to broadcasts, from most major news organisations. A prospective writer or broadcaster can study what these organisations use and how it is presented – the subject matter, the length, the style, the language.

If you've checked on a possible outlet and think you can offer something that would interest them, what next?

- Don't waste time drafting a whole story or article, let alone producing a video news package. Ask them first if they are interested in the subject.
- Don't bother to approach them unless you have good English (or French, Spanish, Russian, if you are seeking an outlet in such countries). Newsdesks have no time to plough through pieces in poor English, and won't want to spend costly time editing your contribution.

- Find an email address for the appropriate editor, and tell them, briefly, what you have to offer. Don't exaggerate, or shade the truth – editors need reliable correspondents, and will drop you like a hot potato if they think you are inventing things. Provide brief details about yourself and your experience.
- Most reputable organisations will not ask you to send something unless they are genuinely interested – if only because they don't want to waste time reading or listening to yet more material.

## WHAT SORT OF STORY SHOULD YOU PRODUCE?

### **Keep it short**

If a length has been suggested, don't write more. (You'll probably ignore this injunction – most reporters seem to – but that just angers editors. A would-be correspondent should not start by annoying a potential employer.)

Brevity is even more important if you are offering online journalism.

You may have wonderful things to say, but if you write screeds or send a great long tape, an editor is likely to put it aside for inspection some other time – and it may never even be looked at.

If necessary, add a note explaining you have more facts/quotes/colour, or actuality – if this is required.

### **Keep it clear**

Ensure that someone knowing nothing about your country can understand what you are presenting. Explain where places are and what customs or activities consist of – not lots of detail, but enough that people can follow the story.

### **Keep it unbiased**

An editor overseas won't want something from an unknown correspondent pushing a particular line. If you feel an issue or an event requires exposure in the outside world, let the facts speak for you – keep the adjectives and the adverbs in their box.

### **Keep it simple**

It's hard enough for a staff writer to produce something clever or humorous; it's extraordinarily difficult to do this from thousands of miles away.

### **Keep in touch**

Provide good contact details. If reaching you by email or phone is hard, give alternatives where a message can be left.



## START SMALL

Journalists in developing countries ask questions like: how can I write for the New York Times? The answer? With inordinate difficulty.

The most prestigious news organisations are not places for relative beginners. They already have networks of staffers and correspondents, and they receive far more material than they can use.

So why not start at the other end of the scale? Try publications which may not have a dozen pieces competing for any space.

If your country is a sizeable producer of carpets, or copper, or cocoa, check out the international trade magazines in those fields. Do they have any coverage of your local industry? If not, offer them a brief piece on some aspect of current activity – after checking the style and likely length for news items.

They're not looking for great prose, and might overlook problems of weak English, so long as the facts and figures are good.

They won't pay a great deal, but even fifty dollars for a short item can go a long way for many a third world journalist, and, most important, you will have something in your resumé – editors overseas will see you are capable of meeting the demands of a foreign publication.

Look at speciality publications or listen to programmes via the net on areas that interest you. A good sports story, or an interview with a local musician, may find a much readier audience than a general news piece.

Western news outlets can be depressingly uninterested in important news developments in your country, but may be tempted by the offbeat or entertaining.

Be alert to stories concerning local activities by foreign citizens.

Travel articles sound easy to write – but you need much more understanding of the mindset of the average reader in Birmingham (Alabama or England) than you're likely to gain sitting in Hanoi or Havana. However, an item on some foreign celebrity who is holidaying in your country might well interest a gossip column.

Business stories, sports teams, charitable activities can all provide opportunities.

Check precisely where someone is from. A provincial daily in a western country may not pay much, but will be more interested in what its residents are getting up to in faraway places than will a big national paper. Likewise, the radio or TV station in a particular city could be a better bet than a national network.

Look east as well as west, to Japan or Seoul, to Singapore or Hong Kong. Their media have decent budgets and could be more interested in your country.

Media from your neighbouring states may be less well-heeled, but could have a real enthusiasm for material about their citizens in your country, or articles about events linked to issues of importance to them.

## **IMAGES AND PRESENTATION**

A bonus of the net is the ease for sending images by email. A story with a photograph will be read more than just text. Try carrying a camera, so you can at least offer a head shot of someone featured in your report.

Are pictures available from a professional photographer? Don't include these, except perhaps a sample in the body of an email (big picture files may not pass computer security systems). But tell editors what is on offer. Presentation can be all-important, so the chance of a good picture display may multiply your story's chances, especially for a feature or magazine article.

Appreciate that the more prominent a piece, the more likely it is a news organisation will want a staffer involved – perhaps commissioning a photographer themselves, or seeking to use your material as the basis for a staffer's article or broadcast package; if so, negotiate both prominent credit when the material is used, and decent payment.

## **STUDY MORE**

Learn a bit in general about the media of a particular country. An optimistic TV reporter might think the big US networks are a market to aim at. The reality: a survey a few years ago, of the content of the networks' evening news bulletins, found each broadcast on average one foreign story every other evening.

Since 9/11 and the invasion of Iraq, things have changed a bit, but I wouldn't waste time trying to sell them foreign news.

Understand that an awful lot of the media in developed countries devote their energies to entertaining their readers, rather than informing them. And not just in the west – lifestyle journalism rules in Bombay and Shanghai as in Los Angeles or Madrid.

If you want to understand how newspapers operate, don't just look at the news sections, but at all the feature and specialised sections, and magazines. It may be harder to watch all the TV broadcasts, but do realise that the current affairs documentary is almost extinct in the US – and is an endangered species in some other places, too. News and current affairs agendas in many countries are widely regarded as 'dumbed down'.

This may be dismaying to a serious journalist, but that does not necessarily mean a lack of opportunities for you. There is such a proliferation of media outlets – of magazines, of radio stations, of websites – that there's space to fill. Scour the net for possible recipients of your work. If you are inventive, it could be you who fills those gaps.

All the above is advice from just one journalist. It won't apply in every country or every case.

But if you want a career, or a boost to your earnings, with work for foreign outlets, begin by spending as much time studying those markets as ever you do pursuing the initial stories.

## The Author

Martin Huckerby spent much of his career as a journalist in London's Fleet Street, as a reporter and news editor with *The Times*, and later as Foreign News Editor of the Sunday paper, *The Observer*, during the era of Tiananmen, the collapse of Soviet communism and the First Gulf War.

Since then, he has worked around the world, initially in post-communist Eastern Europe, as Editor of the English-language newspaper in the Czech Republic, *The Prague Post*.



Subsequently he became an editorial consultant, training journalists and advising news organisations in such places as Russia, China and India, and elsewhere in Asia, Africa, Central America and Europe.

Most recently, he spent half a year in Afghanistan, as Managing Editor for the newly-established national news agency, Pajhwok, and then began a project in Beijing to produce a guide for Chinese journalists on reporting human rights issues.

*This book grew out of a short manual, Internet for Hacks, originally written in 1999 for journalists in China. The net changed so fast that a revised version was needed a year later; a third followed in 2002.*

### AUTHOR'S ACKNOWLEDGMENTS

For initial help in understanding the net:

Mike Holderness (UK); Nora Paul (US)

For sharp and perceptive advice on my first manual:

Pieter Wessels (Australia)

For enthusiasm in bringing about this book:

Hara Padhy (UNESCO)

And for thoughtfulness, consideration and encouragement, in many ways:

John Ryan (Thomson Foundation)