

Cyber Journalism

(Excerpts from the book [*Information and Communication Technology: Recasting Development.*](#))

Data mining:

In a large newspaper, an archive of its own reports dating back to a few decades would serve much of the needs of its reporters for background information. As far as current and relatively new developments are concerned, material would be available on the Internet. Internet will, in fact be, much more than that. It can be a mine house of information for reporters working on in depth and investigative reports. It demolishes the boundaries of beats and specialisations to a great extent and expands the range of reporter.

However, a lot of 'mining' would be needed before journalists could actually make use of data on the Internet. It is, in every sense, like mining. Worthwhile information will have to be churned out of massive quantities of data on the Web. The grains will have to be separated from the chaff. Some of the gems would be found only if we look at the Invisible Web.

This calls for a deep understanding of the Web and training. However, many journalists, as hinted earlier, are not used to systematic use of Internet. Training and guidance are lacking. Often, they learn in an exploratory manner. This sometimes leads to grave mistakes. On the other hand, better accuracy in reports can be achieved using the Internet, if the journalist has the necessary skills.

Techniques of evaluation:

Many techniques have been suggested for evaluating content on the Internet. However, none is foolproof. The domain name could be the first pointer to the authenticity of the source. Reserved top level domains such as *.gov*, *.gov.in*, *mil* indicate authority and hence reliability of the material. On the other hand, top level domains such as *.com* and *.org* would not even indicate the status of an organisation as anyone can register these domains. Organisations having commercial character have registered *.orgs* which were

meant for non-profit organisations and voluntary services. More than that *.coms* need not always be commercial in character. For example, several Government Web sites have a *.com* label. In Kerala and elsewhere, this happened because some bureaucrats thought that *.com* is the most popular domain in India and hence the most appropriate for Government sites!

In some respects, India has stricter norms for allocation of top level domain names. Documents ought to be produced for registration of domains such as *.ac.in*. However, these domains can belong to academic institutions that maintain standards as well as those that cheat the students. Though country codes indicate probable location and ownership of the site, owner could be from outside the country indicated by the code. The site could also be hosted outside the country. The only assured aspect is the legal jurisdiction regarding the registration.

Who owns the domain is important in evaluating the accuracy and objectivity of the information on a Web site. Tools are available on the Web to probe the registration and other details about Web sites. Apart from basic *whois* services, some Web site offers probing tools that can be used for a mini detective work on domains and their owners, coupled with search tools. (See [Who Own the Domains](#)).

Determining the currency of data could be a problem for those making a quick reference on the Internet, if the information is time sensitive. If the site itself has no information about updates, one may even have to do some research to find out whether the material is current. Though browsers will give the date of modification of the file on the Web server, this may not necessarily relate to the last update. It could be the date when the file was last uploaded to the server, after, say, a crash.

Many things on a site such as the design, *about us* and biographical information, the links, style of writing and the currency of information (date of last update) will give indications about the dependability of the site, though a well designed site may not necessarily be authentic or vice versa. Quality of information can also be assessed by checking whether other quality sites including human compiled subject directories provide link to the site and whether the site gets a good ranking on certain search engine listings (other than sponsored listings). The search engine ranking is particularly relevant in case of [Google](#). It employs an algorithm that ranks a site on the basis of how many

other important sites links to the ranked site. A listing of linked sites can be obtained from search engines such as Google by doing a link search (This involves prefixing *link:* to the keywords you enter in the search engine dialog box). Use of common sense is perhaps the most important thing in the evaluation of a Web site. Sites lacking valid contact addresses and information about the content provider should be considered as suspect. In fact, there are many sites on the Internet that are just there to fool people.

(Excerpted from Chapter on Cyber Journalism by [Roy Mathew](#))

For further reading:

University of California, Berkeley (2003). Evaluating Web Pages: *Techniques to Apply and Questions to Ask*, UC Library— Teaching Library Internet Workshops, <<http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/Evaluate.html>>.

[BACK Home](#)