CAN YOU TRUST ONLINE INFORMATION?

Sabine Louët

Independent journalist and editor at Alphagalileo

Adreça: Alphagalileo, c/o BAAS, 23 Savile Row, London, WIX2NB, UK.
Adreça electrònica: *sabine.louet@alphagalileo.org*

RESUM

L'arribada d'Internet ha ofert una gran abundància d'informació científica especialitzada en xarxa. Per això, és essencial aprendre quina informació hi ha disponible i si es tracta d'una informació de qualitat. Els problemes més comuns van des de la inexactitud i la desactualització fins a la irrellevància de la informació disponible. Els reptes que han de superar els professionals que depenen de la informació són diversos: han de saber esbrinar la identitat d'un lloc d'Internet; distingir la informació de primera mà d'entre la que ha estat processada; identificar els apuntadors i ser conscients de la importància de les dades, així com de la rellevància cultural del material publicat a Internet.
Paraules clau: Internet / informació / qualitat / recursos.

SUMMARY

The advent of the Internet offers a wealth of specialised scientific information online. Hence, learning what information is availabel and estimating whether it is quality information has become essential. Common problems range from inaccuracy, out of date and irrelevance of available information. Identify the site identity, distinguish between raw information and processed information, cross reference sites, be aware of the time sensitivity and cultural relevance of material published on the Internet are among the challenges that information dependent professionals need to meet.
Keywords:Internet / information / quality / resources.

BEING DIGITAL IS ONE THING, BEING CRITICAL IS BETTER

Retrieving information on the Internet is getting increasingly easier. However, it is equally important to have an idea as to the reliability of the source of information. There are several steps to take to ensure that information found on the Internet can be re-used for professional purposes by taking a critical stance.

Domain

The first clue to the reliability of a web site is its address. Hence, the location of a web site in a specific country can be identified from the ending of the URL. For instance, sites located in Spain end in *.es*, and in the United Kingdom in *.uk* etc. In the same way, it is possible to distinguish between a commercial site *.com* and a non-profit organisation *.org* or government *.gov* web site, whereas universities can sometimes be recognised by specific endings such as *.edu* or *.ac*. A list of domain names is available at: http://*www.alldomains.com/*. Arriving at a web site via the address given by a search engine does not always give away the site's identity. Nevertheless, it is possible to narrow down the address to its root location by only keeping the letters between the http:// and the first slash (/). It normally leads to the home page of the site. Learning about the host site tells you a lot about reliability of the site by giving clues about its identity. Hence, a government site would target a different audience than an educational site and thus have different uses. In addition, cultural differences intervene. For example, a site in Mexico would have a different angle than a site in Spain in the same way that Mexican and Spanish newspapers differ in content.

Primary vs secondary information

The second step is to identify whether the information contained on the site is either first- or second-hand. Scientific papers or government policy are an example of primary sources of information. Second-hand information has been processed, analysed, put into perspective by a third party and its meaning can therefore be examined under a different light.

Access to primary sources of information is one of the key advantages of the Internet. For instance, AlphaGalileo (http://*www.alphagalileo.org*) is a site specialising in gathering specific types of information such as press releases issued by scientific research centres which can then be turned into secondary information, such as newspaper articles by journalists using the site. Although the site is primarily aimed at journalists and posts press releases concerning the latest progress in scientific research across Europe, it can be used by people interested in keeping up with the latest scientific progress, since all the press releases are available for free on-line. Only releases protected by an embargo date are available to journalists. Such sources of information, which are regarded as a primary source by journalists, give them a basis for cross-checking and putting the information into perspective.

Cross-Checking

Putting information into perspective is equivalent to validating it with a second source. For instance, finding information on topics like GM food is relatively easy, but given its controversial nature, all information on the subject needs to be read with a critical eye and cross-checked with different sources. Hence, for instance, an environmental campaigning group such as

Greenpeace (http://*www.greenpeace.org*) or the British organic food campaigning group, Soil Association (http://*www.soilasso-ciation.org*), offers information on GM food at their site. This can be contrasted with the information contained on CropGene (http://*www.cropgen.com/*), a site funded by the crop biotechnology industry which claims to operate independently, or even the UK government GM information web site (http://*www.gm-info.gov.uk/*). In both cases, vested interests have to be taken into account when assessing the content of the sites. A third type of sites, such as the Non-Governmental Organisation OECD Food Safety (http://*www.oecd.org/subject/biotech/*) web site, attempt to provide independent information. Nevertheless, any source of information has a different bias.

Time dependence

Another pitfall to avoid stems from the constantly changing nature of the Internet, where a piece of information can quickly become out of date. Hence the need to check the date of the last update on a particular page, if and when it is available. Scientific information is particularly time-sensitive, as research results provide continuous updates on knowledge. Moreover, information can unexpectedly disappear from a page without any warning. There have been cases, in start-up companies where staff, once made redundant, were completely removed from the web site. The Internet cannot always be used as a reliable archive.

Authentication

Finally, the Internet favours the propagation of rumours and false information which can even be relayed by reputable sources of information. A recent example is the availability of top model's ovule which the American photographer and film-maker Ron Harris offered for sale over the Internet pretending to offer beauty features to future offspring. Later exposed as a spoof, this information was relayed by newspapers and ended up advertising the actual site owner's activities. This example raises the question of authentication. There is nothing easier than pretending to be somebody else on the Internet as there is no real way to check. It is no longer reliable to depend on first-hand perception, as anybody can set up a website containing unchecked information. In the face of such problems, preference for well-established 'brands' of information providers would prevail over individual web sites, wich can still be of great quality.

WHAT RELIABLE SCIENTIFIC INFORMATION IS AVAILABLE?

Overall, university web sites can be considered as reliable, although individual pages do not always bear the same credibility. Most universities now feature a web site containing scientific information describing the research work carried out within their laboratories. They can be located through resource pages such as Braintrack (http://*www.braintrack.com*), which provides a list of universities from all over the world. In addition, most university pages contain an integrated search engine enabling users to access appropriate department pages through a keyword search. In the meantime, looking at the page's URL will help to ensure that a page belongs to a university since universities have an address ending in *.edu* when based in North America, whereas universities in the United Kingdom would typically have a URL ending in *.ac.uk* On the other hand, Spanish universities are listed at the following address:

http://www.rediris.es/recursos/centros/univ.es. html.

Back to brands

Reliability of scientific information is often associated with primary sources such as peer reviewed journals which are becoming increasingly more present on the Internet, while access to such journals is password-protected and aimed at subscribers only. Nevertheless, most of these sites are a mere translation of the paper version onto electronic format which only enables faster access to research results while offering the flexibility of electronic documents, such as keyword search and cross referencing between relevant articles. Some of the regular scientific publishers such as the American Chemical Society (http:// *www.acs.org/*) offer an optimised on-line version of their regular publication. Hence, their policy, called «Articles ASAP» (As Soon As Publishable), consists of publishing articles on-line 2 to 11 weeks prior to their publication in print as soon as they have been approved for publication.

New resources

Internet users tend to trust existing 'brands' such as well-established scientific journals, like Nature or Science, to judge the quality of a web site, so paper-based publications are no longer unique sources of reliable information. These established 'brands' now face competition from alternative web sites making on-line publishing of research freely available on the Internet. An example of such burgeoning on-line resource is the American National Institute of Health (NIH) site, which hosts a free electronic repository called PubMed Central for peer-reviewed research papers in the life sciences. These resources will be operational

as of May 2000. In addition, a sister site called BioMed Central (http://*www.biomed-central.com*) also intends to provide help and tools for researchers wishing to set up their own autonomous publications. The credibility of such sites will depend on the reliability of their promoter, since they have not yet had the opportunity to establish themselves as a 'brand'.

Gateway to scientific information

Such examples show how the Internet is bound to impact the content of scientific publications in a way that current scientific journals might not have expected. Web-based-only sources of information such as the World Wide Web Virtual Library (http://*vlib.org/*) or the SciCentral (http://*www.scicentral.com/*) site offer gateways to information classified by categories for easier access. They avoid the tedium of looking up information via a search engine and offer the quality control that the editors of such sites impose on their selection. Such gateway sites are likely to become the new reference for the Internet. Hence, in the same way that 'brands' guarantee quality, specialised Internet guide sites are acting as gateways to useful and reliable sites.

Nevertheless, every single web user can also build their own list of reference sites thanks to the bookmarks facility. Individuals can select good quality sites through the use of meta-search engines, such as Copernic (http://*www.copernic.com)*, which accurately pinpoint sites related to very specific keywords. Searching through about ten search engines at a time, it gives a ranked list of relevant sites without duplicating them. This type of tool tends to save time. Nevertheless, it is important to learn how to distinguish between first- and second-hand information. In the meantime, there are also sites specialised in

first-hand scientific information such as MedLine (*http://www.nlm.nih.gov/medlineplus/medline.html*), which is the US National Library of Medicine's database of references, containing more than 11 million articles published in 4,300 biomedical journals. Searching though such resources facilitates the location of relevant published material.

CONCLUSION

Using the Internet as a tool for searching for specialised information offers time flexibility and multiple sources of information to choose from. However, this information can only be used if the origin, the cultural framework, the nature of the source —be it primary or secondary— the author and the time of the last update, can be clearly identified. This amounts to a lot of additional checking compared to traditional sources. In order to avoid this inconvenience, it is time-saving to rely on 'branded' sources of information or on gateway sites subject to quality control.

INDEX OF CITED SITES

Domain and country names:
    (http://www.alldomains.com)
AlphaGalileo:
    (http://www.alphagalileo.org)
Greenpeace:
    (http://www.greenpeace.org)
Soil Association:
    (http://www.soilassociation.org)
CropGene:
    (http://www.cropgen.com)

UK government GM information web site:
    (http://www.gm-info.gov.uk)
OECD Food Safety:
    (http://www.oecd.org/subject/biotech)
Braintrack, worldwide universities index:
    (http://www.braintrack.com)
List of Spanish Universities:
    (http://www.rediris.es/recursos/centros/univ.es.html)
American Chemical Society:
    (http://www.acs.org)
PubMed Central:
    (http://www.pubmedcentral.nih.gov)
BioMed Central:
    (http://www.biomedcentral.com)
World Wide Web Virtual Library:
     (http://vlib.org)
SciCentral:
    (http://www.scicentral.com)
Copernic:
    (http://www.copernic.com)
MedLine:
    (http://www.nlm.nih.gov/medlineplus/medline.html)

Further References

Internet Detective:
    (http://sosig.ac.uk/desire/internet-detective.html)
Thinking critically about information on the Web:
    (http://omni.ac.uk/agec/vine.html)
Information Quality WWW Virtual Library:
    (http://www.ciolek.com/WWWVL-InfoQuality.html)

ABOUT THE AUTHOR

Sabine Louët is Editor of AlphaGalileo, the Internet-based press centre for science, medicine and technology reporting on the latest progress in European science. A physics graduate, she studied Science Communication at the Imperial College in London and writes articles for scientific magazines on a freelance basis.