



Information Overload

Issue 24, August 2004

Welcome:

Welcome to this month's issue of Information Overload. This month we will be looking at the issue of ethics in an online world. And in particular what do you do when you come across documents that you know should not be viewable in an open environment but are?

Your comments are always welcome, and if you would like to see us cover any other topics, we would love to hear from you. Just send an e-mail to training@iea.com.au.

We would like to thank you in advance for forwarding this edition to friends and colleagues.

We hope you enjoy reading, have a great week.

In this Issue we will be looking at:

- The Invisible Web and legal/ethical issues for librarians and library technicians relating to confidential documents.
- Google Hacking
- Hacker/Cracker – definitions
- But I found it by accident – now what?
- A Thought to Ponder.

The Invisible Web and legal/ethical issues for librarians and library technicians relating to confidential documents.

I will begin this month's newsletter with a subject raised by Elizabeth Swan on the alialNFOG List Serv a few days ago. She says: "An interesting issue has been raised by professional colleagues on an overseas e-list regarding the availability of "company confidential" or "internal use only" documents found on the open web.

Even in Google and presumably other search engines, searching on the phrases "internal use" or "company confidential" retrieves documents on some Australian websites that apparently are not for public viewing.

There are also documents on the web saying how important it is to carefully check if documents being made available via the web are confidential or not, but apparently there are webmasters in Australia (as in other countries) that are not doing that rigorously.

As information professionals what are our obligations regarding these documents? What are the ethical and legal issues? Do we have an obligation as information professionals to contact the website owners? Or the website designers? Or is it the responsibility of the website owner to ensure the security of its internal or confidential documents?"

Google Hacking

Elizabeth mentions the use of Google (and other search engines) to locate

© IEA 2004. All rights reserved. You are free to use material from the Information Overload newsletter in whole or in part, as long as you include acknowledgement of source.

documents that could be described as "sensitive". According to cybersecurity experts they say "an increasing number of provocative or putatively secret documents are online in out-of-the-way corners of computers all over the globe, leaving the government, individuals, and companies vulnerable to security breaches."

Online Search Engines Help Lift Cover of Privacy, Yuki Noguchi. Washington Post Feb 9, 2001: Page A01
<http://www.washingtonpost.com/ac2/wp-dyn/A24053-2004Feb8>

A quick search of Google gives hundreds of articles, tips, tricks and how to's on the subject of searching for this kind of material, and is known affectionately by the "Googledorks" as "Google Hacking" or "Google Hack". For instance:

The **site:** operator instructs Google to restrict a search to a specific web site or domain. The web site to search must be supplied after the colon.

The **filetype:** operator instructs Google to search only within the text of a particular type of file. The file type to search must be supplied after the colon eg., **archiving filetype:pdf** note there is no full stop (period) before the file extension, gives 193,000 hits (25/8/04).

The **link:** operator instructs Google to search within hyperlinks for a search term.

The **cache:** operator displays the version of a web page as it appeared when Google crawled the site. The URL must be supplied after the colon.

The **intitle:** operator instructs Google to search for a term within the title of a document.

The **inurl:** operator instructs Google to search only within the URL (web address) of a document. The search term must follow the colon.

Information taken from "Google Hacking Mini-Guide" Johnny Long, May 7, 2004
<http://www.informit.com>

More information can also be found on the advanced search options page on Google itself.

Hacker/Cracker - definitions:

It is perhaps wise at this stage to give a couple of definitions of the term "Hacker" and "Cracker". These definitions have been taken from the Free Online Dictionary of Computing.
<http://wombat.doc.ic.ac.uk/foldoc/foldoc.cgi?hacker>

A **Hacker** is "a person who enjoys exploring the details of programmable systems and how to stretch their capabilities, as opposed to most users, who prefer to learn only the minimum necessary. The term "hacker" also tends to connote membership in the global community defined by the net. It also implies that the person described is seen to subscribe to some version of the [hacker ethic](#).

The "Hacker Ethic" is the belief that information sharing is a powerful positive good, and that it is an ethical duty of hackers to share their expertise by writing free software and facilitating access to information and to computing resources wherever possible. And that:

The belief that system-cracking for fun and exploration is ethically OK as long as the cracker commits no theft, vandalism, or breach of confidentiality.

The dictionary goes onto say that "Sense 2 is more controversial: some people consider the act of cracking itself to be unethical, like breaking and entering. But the belief that "ethical" cracking excludes destruction at least moderates the behaviour of people who see themselves as "benign" crackers (see also [samurai](#)). On this view, it may be one of the highest forms of hackerly courtesy to (a) break

into a system, and then (b) explain to the sysop, preferably by e-mail from a [superuser](#) account, exactly how it was done and how the hole can be plugged - acting as an unpaid (and unsolicited) [tiger team](#).

A **cracker** on the other hand is an individual who attempts to gain unauthorised access to a computer system. These individuals are often malicious and have many means at their disposal for breaking into a system."

But I found it by accident – now what?

There are a lot of people who routinely search the internet and find interesting items quite by accident, as opposed to deliberately trying to find material that may be of a sensitive nature.

In response to Elizabeth's original message, a lady from New Zealand gave her perspective by saying "I tackle these sorts of issues by referring back to the professional body, in my case, LIANZA, which has quite clear guidelines about what is and isn't ethical in its document 'Principles Applying to Consulting Librarians'. It doesn't mean that I don't question the guidelines, but they are a most useful test of what is 'right'. In this case, I would suggest item 6 applies after a fashion

"Consultants should not disclose or use for other purposes any information or material of a confidential nature that has been made available to them in the course of their duties..."

From memory, ALIA has [had] a similar document. I also ask myself what would I do if it were a printed document/ file/ piece of information and question whether the approach should be any different because it has been accessed electronically.

It is surprising how often one comes across confidential information

accidentally in the pursuit of something else (or maybe not, serendipity rules!). So, if I found a confidential file in the street, would I return it to the organisation or use it? Simple, really, from an ethical point of view.

However, the industry we are in does not undertake 'competitive' intelligence without some "street smarts", I have no doubt that knowing the 'right' thing to do does not always equate with doing it!"

Another responded:

"This *is* an interesting issue. If we find such a document or site should we notify the web owner, should we indeed use the information in competitive intelligence? My own personal inclination would be that as I would only turn up something of this nature by accident, if it was an organisation I had an association with I would be likely to drop them a quick e-mail and tell them, but otherwise not. Is this socially responsible enough, or is it not? Has anyone on the list sought a legal opinion on this? Intranet seeping onto the Internet?

I do think that IT and IS security is an issue that libraries (and the rest of the world) are going to have to face seriously in the next little while, and this of course is just one of the components of that."

It has to be said that once a document has found its way onto the visible or "surface web" it is very hard to shoe horn it back into a secure site, and could be likened to trying to get toothpaste back into a tube, in other words it is virtually impossible to do. One of the main reasons for this is simple; once a robot or spider has found your site and merrily indexed all that it can find, then unless you have devised a program to delete every cache on every pc, server and ISP, chances are your information will always be available to those who want to find it, or stumble across it by accident.

Is it the fault of the site owner, "Just stick it on the web for goodness sake", the

host, or the web designer? There are some people who are simply not aware that what they are trying to link is "private and confidential" for instance some IT professionals, and web designers will not check whether the document or item is of a sensitive or confidential nature, assumptions will be made, they haven't been told, nor have they bothered to have a look. To the younger IT person, the problem may stem from the fact that they do not have the business "smarts" to know that what they consider to be an abstract problem "Can I put this document on the web or can't I?" may actually be a breach of privacy. Or is the fault that this kind of document can be found be a result of the "person responsible for the web", be it the "IT" department or a.n.other for not putting up a secure enough firewall?

Of course we also have to contend with "insiders", those people who deliberately set about leaking information into the worldwide domain of the surface web, and the well-meaning individuals who email confidential documents to colleagues with little or no thought as to the consequences of their actions. As we

are constantly advised – writing an email is akin to writing a postcard and putting it on the worldwide bulletin board. A secret is no longer a secret if more than one person knows about it.

So what is the answer? Well part of it lies with everyone who has input to the design and content of the company's web site. Do not make assumptions, and make sure that you brief your web designer(s) and IT professionals properly. Don't assume someone else will do it for you.

As to what do we do with the information that we find? Well it seems to me that ultimately it comes down to individual choice. Do we choose to use it and gain competitive advantage over others, or do we not use it, send the document owner a message and advise them to plug the holes?

I personally have to agree with our New Zealand colleague who said "I have no doubt that knowing the 'right' thing to do does not always equate with doing it!"

But in the mean time - Happy searching!

A Thought to Ponder:

"In life we are to show up, stand up, step forward and speak up; not to cringe in fear, choosing to sit down, to step back, allowing the voices of silence to speak so loudly"

Father Brian Cavanagh

Your comments and suggestions on the subject of this newsletter are most welcome. Or if you would like to see other issues covered in future editions, please email training@iea.com.au

Please feel free to pass on this newsletter to your colleagues' friends and associates. To subscribe they should send an e-mail to training@iea.com.au with "subscribe newsletter" in the subject line.

If you would prefer not to receive this newsletter, please send an email to training@iea.com.au with "unsubscribe newsletter" in the subject line. If you have any suggestions as to what should be included in future editions, then please send an email to training@iea.com.au.

© IEA 2004. All rights reserved. You are free to use material from the Information Overload newsletter in whole or in part, as long as you include acknowledgement of source.

Information Enterprises Australia Pty Ltd
Unit 4, Upper Level, 201 High Street, FREMANTLE WA 6160
Tel: 08 9335 2533 Fax: 08 9335 2544 e-mail: training@iea.com.au