

Cyberlex

by Jason Schoolmeester



As the Internet's growth continues, so too does the breadth of legal material available via the Internet. Legal material is being provided by commercial publishers (i.e. www.butterworths.com.au) and public interest publishers (i.e. www.austlii.edu.au). The provision of any substantial amount of legal information via the Internet creates large amounts of data that is not suitable to search using the general Internet search tools such as AltaVista or Yahoo. For this reason, large Internet sites provide a site search engine. In this issue I will be looking at SINO, the engine used by Austlii.

Diving into Austlii

SINO or ('size is no object'), written by Andrew Mowbray, is the power behind Austlii. Text retrieval systems are not new to lawyers, and most of us would have used or at least heard of Lexis, SCALE or Info-One. SINO has been designed so that the user (that is you) can use the techniques learned from other text retrieval systems. In other words, SINO will accept a search query styled on SCALE or most of the others (see full documentation for more information: http://www.austlii.edu.au/help/sino_full.html).

Tip #2: You can bookmark the Austlii search interface <http://www.austlii.edu.au/cgi-bin/sinofrm.pl> and avoid the need to go through the main page every time you want to perform a search.

When you look at the Austlii search page you will see five (5) elements for user input: Search for; Database; Searchscope; Searchtype and Quantity.

'Search for' is where you type in your words or phrases to be searched for. 'Database' determines the breadth of your search. The most common entry here is 'All AustLII Databases'. This is the widest search possible. You can target your search to specific groups of databases such as 'All Legislation Databases' or 'All Case Law Databases' or to specific databases such as 'Northern Territory: Supreme Court'. The wider your search the more likely you are to have a large quantity of hits to

digest.

Searchscope defaults to 'Complete Documents', however, if you know part of the title, case name or legislation name, you can restrict your search to just the names of documents. In this way you can limit your results to only those documents that contain the search word or phrase in their name. Search type enables you to choose the way in which your search is conducted and the results displayed. There are two types of searches, Boolean and Freeform. I will focus on Boolean in this article, but Freeform offers, according to Andrew Mowbray, "a 'quick and dirty' way of finding things". The last element is the quantity of hits. It does not mean I only want '100' hits, but rather how many to display to me at a time. The default is 100, and if there are more hits, the results screen allows you to request the next set (i.e. if the default is 100, the next 100). The other reason for limiting the quantity is that the more hits you request displayed, the longer the download time for the results page. As most users do not have direct links to the Telstra backbone, any more than the default 100 just increases the time spent waiting for results.

What do I type into the 'Search for' element?

George Boole (1815-1864) developed Boolean algebra which has formed the basis of text retrieval systems. SINO is no exception. The basic unit of a SINO search is the word, a continuous string of alphanumeric characters. Words are not case sensitive and are all searchable except for a small list of common words such as 'and', 'the', 'is', 'but' and so forth. So, if you have a distinctive word (i.e. an uncommon name), simple type this in and hit search. However, not all searches are this

easy, and will involve a combination of words and phrases. There are three types of operators that you can use in your search request: boolean; proximity; and truncation. I have already introduced you to the boolean operators of 'and' and 'or' (see last issue). SINO will recognise 'AND', 'and', '&', '+', 'OR', 'or', ',' and '|'. Another boolean operator is 'NOT' ('-', '^') which will find documents that contain one search term and NOT another. For example 'negligence NOT medical'. To search for a phrase (a group of words) you simply encapsulate the phrase in parenthesis. For example "reasonable man", "family trust", "on the balance of probabilities"

Proximity operators are used to find documents where your search terms are near one another. Thus you are trying to find word(s) in relation to their position to one another. The most basic operator you can use is 'near'. The 'near' operator returns a hit if the search terms are within 50 words of each other. SINO supports other styles of proximity operators such as 'pre/n' where n equals the number of words and the first term must precede the second. SINO also supports the Lexis operator 'w/n' (where n is the number of words) allowing you to vary the 50 words provided by 'near'.

Truncation operators enable the user to search against the stem word. For example if you are researching a matter involving negligence, you might want to search negligent, negligently, negligence. Using the truncation operator '*' you can search for 'neglig*'. '*' matches any string following the truncated term, while '?' matches only any single character. You might use '?' when searching for women or woman (e.g. 'wom?m'). An advantage of SINO is that it automatically searches for English plurals! (i.e. car and cars). This means you do not have to search for 'solicitor*' and 'lawyer*'

The only other important concept to remember is the order of operation. SINO evaluates from highest binding strength to lowest as follows: word; phrase; proximity and segment operators; 'or'; 'and' and 'not'. Where the operators or terms have the same precedence, then they are evaluated from left to right. If you want to alter the order of evaluation, you can use brackets (any type, square, squiggly or

continued page 21



Society Noticeboard

MOVEMENT AT THE STATION

Australian Institute of Judicial Administration

Has moved to:
Level 1
723 Swanston Street
Carlton VIC 3053
Phone: 03 9347 6600
Fax: 03 9347 2980
Website: www.aija.org.au

Professor David Weisbrot

Elected President of the Australian Law Reform Commission. Three year term

Saul Harben

formerly of De Silva Hebron is moving to:
Freehill Hollingdale and Page
AMP Building
440 St George Terrace
Perth WA 6000
Phone: 08 9211 7777

Mr Damian Bugg QC

Appointed Commonwealth Director of Public Prosecutions. Appointed begins August 2, 1999

Bridget O'Brien

is returning to her position in Adelaide
Executive Assistant
Law Society of South Australia

Julie Davis

Administration Manager
Law Society Northern Territory
is welcomed back to her office.

CLA Legal Writing Competition

The deadline for the receipt of papers has been extended to: **July 18, 1999.**

Please mark your entry to the attention of:

Commonwealth Lawyers Association
International Directorate
The Law Society
113 - 114 Chancery Lane
London WC1A 2PL
United Kingdom
Fax: 0015 44 831 0057

Entries must be typed, no more than 2,000 words and unpublished, original work. Entries require a proof of eligibility (copy of practising certificate and proof of age is required).

EVERYONE'S AN EXPERT *from page 13*

obtained from persons who find themselves in police custody regardless of whether the material has any relevance whatsoever to the charge that put them there.

As lawyers we have failed in our obligation to ensure that the Government has been properly taken to task over the issue. More importantly, pressure must now be applied in order to ensure that the process of gathering, sorting and using DNA samples is open to public scrutiny. Unfortunately that approach has all the effect of shutting the gate after the horse has bolted.

An independent forensic laboratory is a practical step to ensuring open access to sensitive information about ourselves. The idea of freedom is worth only as much as the work we put into maintaining it.

Leaving individual liberty in the hands of the Government is the equivalent of leaving a beautiful landscape by Matisse hanging on Mike Reed's lounge room wall.

Cyberlex continued from page 19

round). Search terms enclosed in brackets will be evaluated first.

The only other operators that you may come across are segment operators such as @. AustLII provides the searchscope element in its search form, which has essentially two options: the whole document or the title. In other databases segments have a more useful role.

So now you should be able to string together very complicated search requests such as: neglig* and (solicitor or lawyer or "legal practitioner") near beneficiary
Of course you could narrow the above search to 'Hawkins v Clayton' and set the searchscope to case name. The first has 87 hits while the second has 1, but then I already knew the answer.

What you have learned here can be applied to any text retrieval system. The only difference will be the expression of their operators (e.g. 'AND' or '+'). I encourage everyone to

experiment. I know that old systems such as InfoOne charged by the search, so experimentation was costly, however AustLII is free. I will end with one word of caution, and this applies to any database, always ensure you know what the database contains (and thereby what it does not contain) and the level of currency! Otherwise you may be searching in vain.

When the boss is not looking...

Yes, Australia is in the final of the World Cup Cricket 99. For all the news, facts and figures visit www.bol.net.in/cricket/. Have you wondered how the point system works? What about the words to the official song? It's all there!!

Jason Schoolmeester is a legal practitioner at Parishes.

He can be reached at jasons@legalmnet.net.au (for those without email you can call on (08) 8948 1155)