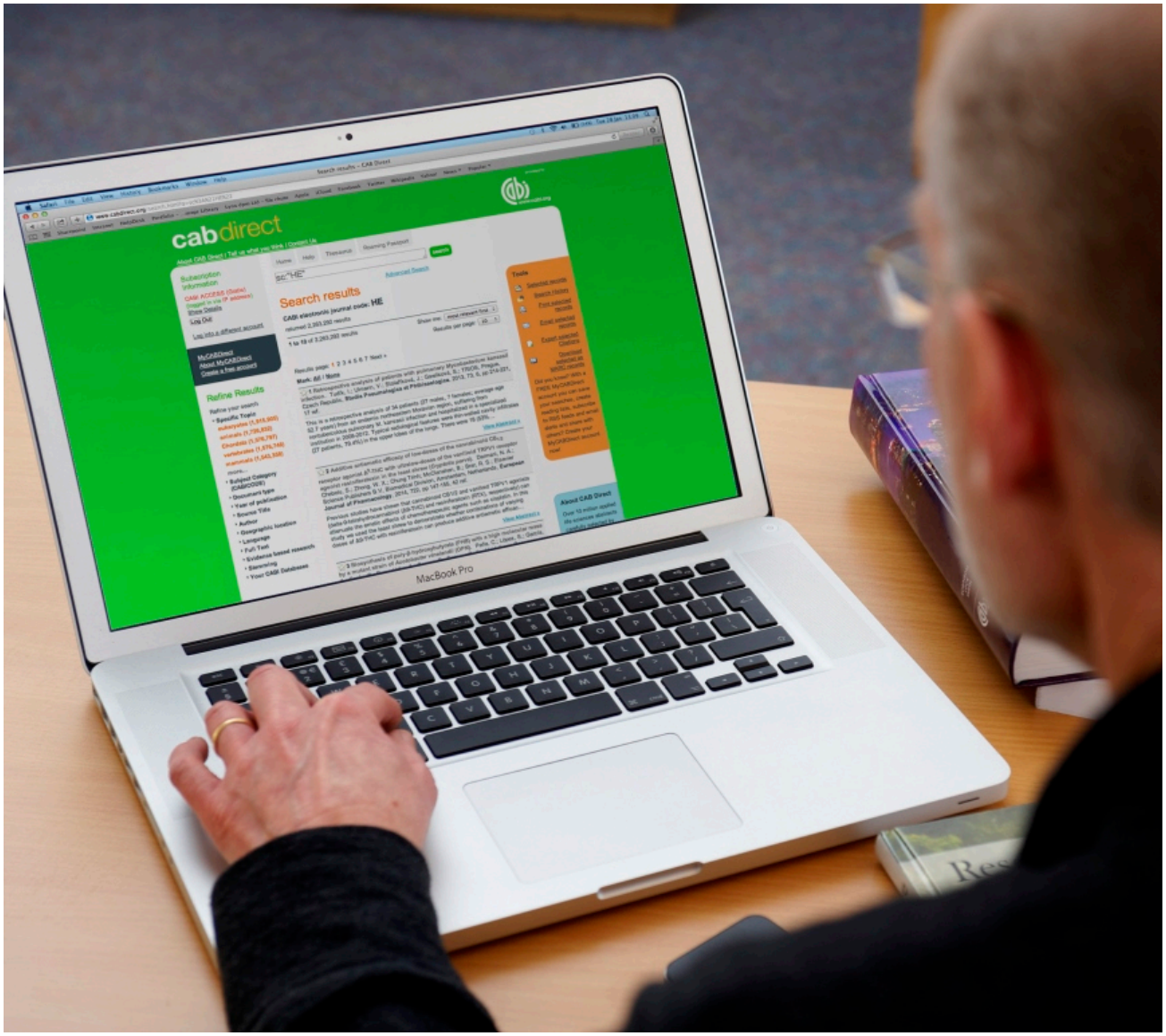
www.cabi.org



CABI Training Materials

CAB Direct

Advanced Searching of Global Health

www.cabi.org

KNOWLEDGE FOR LIFE

Contents

Why use Advanced Search Features

In the Simple Searching guide, we looked at single and multi-word searching of Global Health database on CAB Direct using the Free-Text index. However, in a typical Global Health database record, there may be twenty or more separate data fields. The Free-Text index has been compiled from the words that appear in many of these fields. The list of CABI indexing fields in the Free-Text index includes:

English Item Title
Original Item Title
Conference Title
Personal Authors
Editors
Corporate Authors
Source Title
Publisher
Abstract
Descriptor
Organism Descriptor
Geographic Descriptor
Identifier
Broad Term
CAS Registry Numbers
CABICODES
CABICODE Headings
ISSN
ISBN
DOI
CABI Record Number.

The Free-Text index is the default index, and its use will retrieve the maximum number of records. However, because it includes fields like the Title and Abstract, it is also likely to produce the highest number of irrelevant records, simply because the search terms that have been used appear in the record without any specific meaning. As an example, you may be searching for important papers about allergies to dogs but, by searching for **Dog** and **Allergy**, in the Free-Text index, you may get papers about allergy to cats, where the word dog is also mentioned; i.e. cats not dogs. In order to improve the quality of your search (its relevance) it is often better to restrict your search to a specific data field like the Title field or the Organism Descriptor field. This is known as Field Searching.

Field Searching

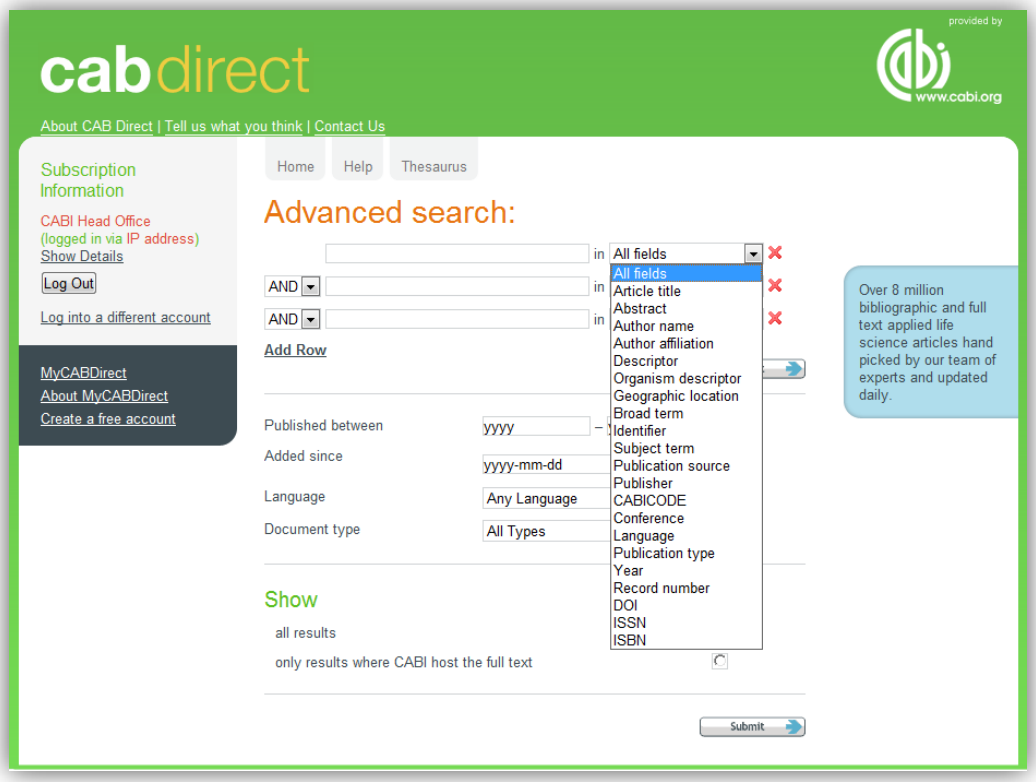
All the fields that appear in the Free-Text index, shown above, are individually searchable. This is very useful for refining your search.

Field searching, with CAB Direct, can be done in two ways. In both the Quick Search and the Advanced Search screens, field tags can be used in front of your search terms, to limit your search to a specific search index, or record field, like Title, Abstract, etc., as in the following example:

title: "heart disease"

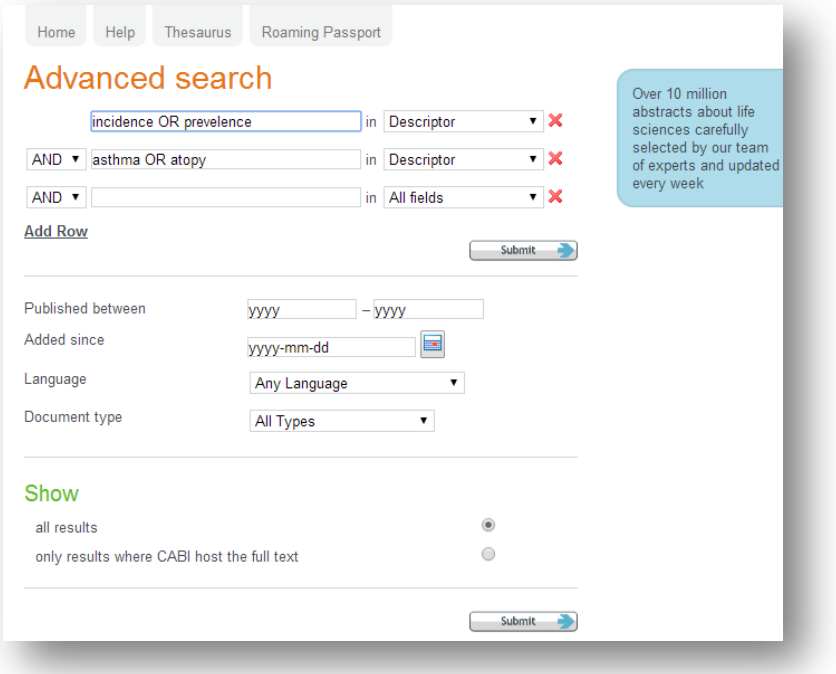
The field tag is entered in front of the search term, and separated from it by a colon (**:**). Note also that, in this case, we are searching for a phrase which has to be enclosed in double quotation marks.

In the Advanced Search screen, shown below, the three search boxes each have a drop down list of fields that can be searched.



The drop-down list includes the major searchable database fields. To choose a field tag, simply click on the tag you want, and it will be displayed in the field tag box.

Let's look at how we might use the Advanced Search screen to build a more complex search for records about asthma or atopy in the UK. The search terms here are **incidence**, **prevalence**, **asthma** and **atopy**. In the Quick search screen, we might have performed five separate searches, for each term in turn, and then combined these together using the Search History Screen. You would get the same result, but it would take longer than doing the whole search in one go in the Advanced Search screen. Let's see how our search would look:



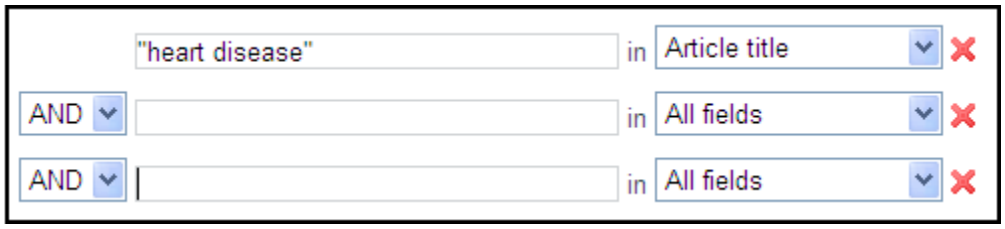
The terms **incidence OR prevalence** have been entered in the top search box. The two terms have been combined using the Boolean Operator OR. Next to **incidence OR prevalence** the "Descriptor" index has be selected from the drop-down menu. More about these individual indexes later. In the second search box, the terms **asthma OR atopy** have been entered, alongside which has been chosen the field "Descriptor". Finally, in the third search box, the term **UK** has been added, and the field "Geographic Location" chosen. The three separate search statements will be combined using the Boolean Operator **AND** but, for other searches, you may wish to combine these statements with **OR** or **NOT**.

This can be done by choosing the appropriate operator from the drop-down lists to the left of the search boxes.

In order to search efficiently, within the CAB Direct interface, it is important to understand the structure of the database and what the individual fields are used for. We will now look at the various, important data fields in turn.

Title Fields

All Global Health records have an English Item Title (Article Title or et). This is the English version of the title of the article that has been abstracted. Most of the original articles will be written in English, so the “Article Title” is usually the title of the original article. If the original article is written in a non-English language, the Article Title field will contain an English translation of the original title. Also, for non-English articles, that are written in a “Roman” script, an original language title will be provided in the “Original Item Title” field. For example, you may see a French article with an original title in French and an English translation of this title in the Article Title field. Although the English Title and the Original Item Titles are entered as two separate input fields, they are merged into one field, the Article Title field in advanced search.



Titles are particularly useful when searching for a paper when all or part of the title is known, and you are only looking for the additional bibliographic data.

Author Fields

There are two types of Author; individuals, who are often referred to as personal authors, and organizations like the World Health Organization, who would be referred to as Corporate Authors. All personal authors and editors are searched using the Author field (Author or au).

I. Personal Authors

The Author field actually includes data from 4 separate, personal name fields. When CABI creates a record for a paper written by a personal author or authors, the policy is to include the names of all the authors. When adding authors’ names to a record, they are added as Family Name, First Initial. Second Initial.

e.g. Smith, T. A.

These are entered into the Author Field. Many authors' names fit this format, but many do not. So, for names that do not fit this standard pattern, CABI will often include variations of an author's name in another field called Author Variants. Where a paper has an Editor, the Editor's name(s) will also be added to the record. When searching CAB Direct, all the personal authors and editors names have been put into the one Author Index, so that they can be searched in one place. So, you can use the Author search index to search for Personal Authors and Editors.

When searching in the Author field, you can search for just the Family name as:

au:smith

This will find all authors called Smith.

If you know their initials, you can include these, but they must be linked to the family name with quote marks as shown below:

au:"smith,j.a"

Note: if you include just the first initial, the system will automatically find all the names with this first initial, i.e. **smith,j.a**, **smith,j.t**, etc.

II. Corporate Authors

The names of organizations that publish papers are entered into the Corporate Author field, at the database input stage. Corporate Authors are searched in the Corporate Author field (ca), as shown below:

ca:"World Health Organization"

ca:WHO

Note, again, the use of double quotes between the words, to get an exact phrase match. You can also use round brackets which will offer a slightly broader search, as it searches not for the exact phrase but simply for the occurrence of the individual words somewhere in the Corporate Author field, and in any order.

ca:(World Health Organization)

Because it is not possible to apply strict rules for adding Corporate Authors to a record, it is often necessary to search for several variations, as shown above.

Index Terms or “Descriptors”

If you are looking only for important papers on a particular subject, where you want a high level of relevance, you should restrict your search to one or more of the CABI indexing or Descriptor fields. Every record on the database is indexed with terms that describe all the important concepts within a paper. The index terms may be added to one of 5 different indexing fields.

The indexing fields that CABI uses are:

Organism Descriptor (od)

Geographic Location (gl)

Descriptor (de)

Broad Term (Up-posted Term)

Identifier (id)

All the terms appearing in the Organism Descriptor, Geographic Location, Descriptor and Broad Term fields are controlled by the CAB Thesaurus, CABI’s indexing authority file. The advantage of having a controlled vocabulary is that users need only use one term to search for a concept rather than using lots of terms. The **Organism Descriptor** field is used for animal and plant names, the **Geographic Location** field is used for country and other geographic names and the **Descriptor** field is used for all the “other” terms that are neither animal, plant nor geographic. The entries in these three fields are added to the records manually, by the CABI Indexers.

Because Global Health is a scientific database, it is very important to remember that most animal and plant concepts will be indexed with their scientific names. All animals, except for commonly managed livestock like Cattle, Sheep, Goats, etc., are indexed with their scientific names. For example, if you want to search for papers about mosquitoes, you would need to search the descriptor fields for the scientific name, **Culicidae**, rather than mosquitoes. However, plants are

indexed with both their scientific and their common names, so the searching of plants is somewhat easier.

In general, index terms are added specifically to a concept within a paper. If, for example, a paper is a general paper about Europe, it will be indexed with the Geographic Location term **Europe**. But, if the paper is about a specific European country, it will be indexed with the specific country name, and not the word **Europe**. In the past, this policy has made searching for broad concepts, like Europe, very difficult because, in order to find every record, the user needed to search not only for Europe, but had to include all the specific country names as well. This is clearly a tedious and, sometimes, impossible task.

The problem was solved for some indexes several years ago, when CABI began using the CAB Thesaurus to add additional index terms, automatically, to a new field called the **Broad Term (up)** field. Because the CAB Thesaurus is hierarchically structured, all the terms are included in a hierarchy with all their broader terms above them and all their narrower terms below them. Since 1984, the electronic CAB Thesaurus has been included in the database production system, and has been used to automatically add broad terms, from the CAB Thesaurus, to the Broad Term field. This is only done for animal names, plant names and geographic terms, i.e. all the terms that appear in the **Organism Descriptor** field and the **Geographic Location** field. If we take our example of Europe, what this means is that every time a European country name appears in the **Geographic Location** field, the broader term **Europe** is automatically added to the Broad Term field. What this means is that a user can search for the term Europe, in the Broad Term field and they will find, automatically, all the individual country names. The field tag for the Broad term field is **up**:

up:Europe

Other search examples:

od:culicidae

gl:(France OR Germany OR Spain)

od:Rice and de:diet and up:"South East Asia"

The last indexing field, not yet mentioned, is the Identifier field. This field is used for terms that do not appear in the CAB Thesaurus. This field is important for papers that discus new concepts that, currently, do not have their own Thesaurus term. This would include new chemicals, new species, etc. The record has to be indexed with an appropriate term but, because it is not in the Thesaurus, this term can not be added to the Descriptor, Organism Descriptor or Geographic Location fields. It would be rejected. Instead, it is added to the Identifier field, where it can be searched using the **Identifier** field tag (id). Clearly, if you are not sure whether a term is an Identifier or a Thesaurus term, you need to search both fields.

For example:

od:speciesabc OR id:speciesabc

In a complex search, with lots of terms that may appear in different index fields, the CAB Direct interface offers an extra field tag, **subject**, which combines the Descriptor, Geographic Location, Organism Descriptor and Identifier fields, and which searches them all at once. This can make life a little easier, as you don't have to remember which field to search. It can also reduce the amount of typing, if you use brackets, as in the following example:

subject:(rice AND diet AND "south east asia")

Note: The Subject field is also available in the drop-down list of fields available on the Advanced Search screen.

CABICODES

In addition to adding index terms to a record, broad concepts are coded with a classification system known as CABICODES. The CABICODES are a hierarchical list of classification codes that divide the subject coverage of the CABI databases into 23 major sections. Each section then includes a series of codes that divides that subject into more specific subjects. The codes themselves are typically used to code for subjects that would be difficult to describe with keywords alone. The area of Human Health, for example, has its own set of codes, as shown below.

VV000	Human Health and Biology (General) (Revised June 2002)
VV050	Human Physiology and Biochemistry
VV055	Immunology and Allergology (New March 2000)
VV060	Human Reproduction and Development
VV065	Human Sexual and Reproductive Health (New March 2000)
VV080	Human Genetics and Molecular Medicine (New June 2002)
VV100	Human Nutrition (General)
VV110	Diet Studies
VV120	Physiology and Human Nutrition
VV130	Nutrition-related Disorders and Therapeutic Nutrition
VV140	Animal Models of Human Nutrition
VV200	Parasites, Vectors, Pathogens and Biogenic diseases of Humans (Discontinued March 2000)
VV210	Prion, Viral, Bacterial and Fungal Pathogens of Humans (New March 2000)
VV220	Protozoan, Helminth and Arthropod Parasites of Humans (New March 2000)
VV230	Public Health Pests, Vectors and Intermediate Hosts (New March 2000)
VV300	Public Health and Nuisance Pests (Discontinued March 2000)
VV400	Animal Models of Human Diseases (New March 2000)
VV450	Animals and in-vitro Models for Pharmaceuticals (New March 2000)
VV500	Human Health and the Environment.
VV550	Rural Health (New March 2000)
VV600	Non-communicable Human Diseases and Injuries
VV610	Human Injuries (Discontinued March 2000)
VV700	Human Treatment and Diagnosis (Non-drug) (Discontinued March 2000)
VV710	Non-drug Therapy and Prophylaxis of Humans (New March 2000)
VV720	Diagnosis of Human Disease (New March 2000)
VV730	Phamacology (New March 2000)
VV800	Human Toxicology, Poisoning and Pharmacology (Discontinued March 2000)
VV810	Human Toxicology and Poisoning (New March 2000)
VV820	Toxicology (New March 2000)
VV900	Occupational Health and Safety

All database records have at least one CABICODE but, according to the coverage, two or more codes are common. The codes are added in addition to the index Descriptors already described, not instead of them. The CABICODES

can be searched just like any other keyword, but using the tag **cc**, as in the following examples:

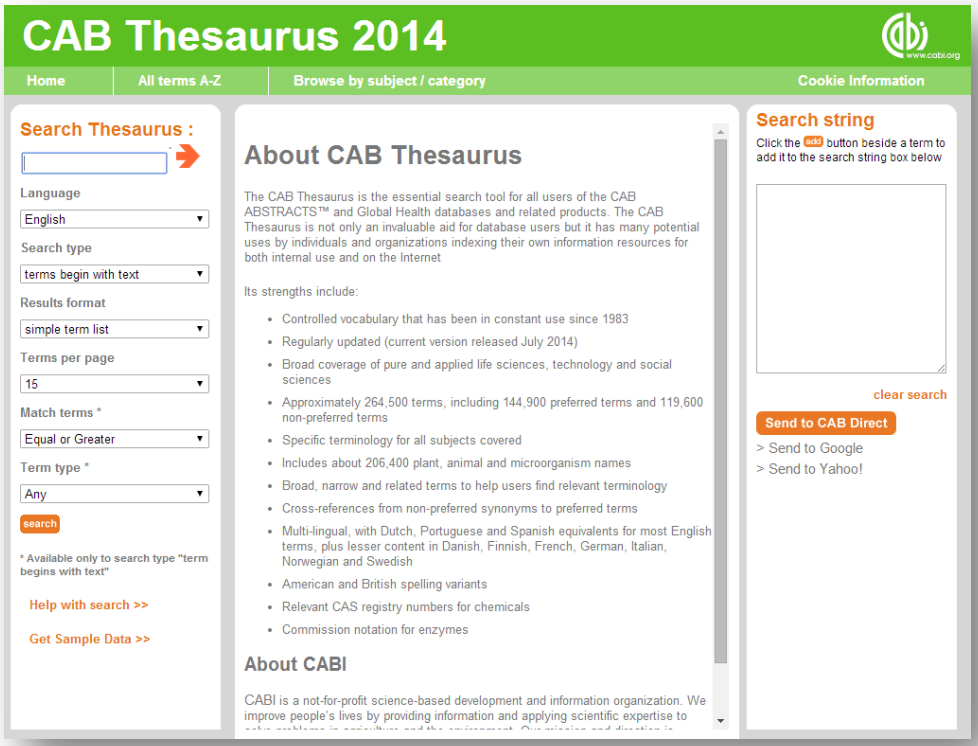
cc:VV900 AND up:Europe

cc:**VV2* AND de:Treatment**

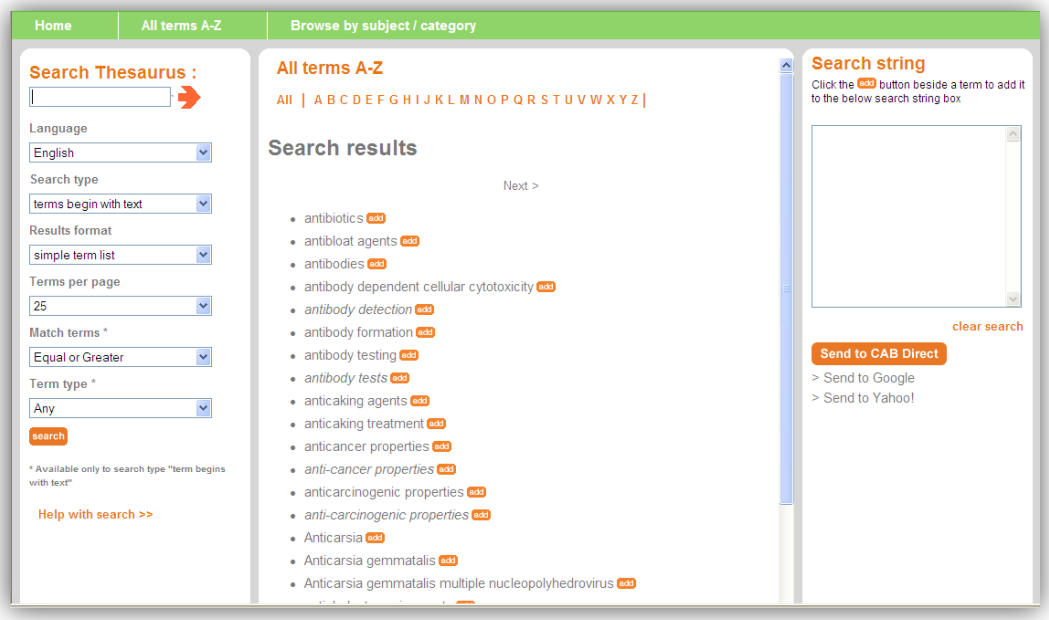
Note the use of truncation (*) in the second example.

The CAB Thesaurus:

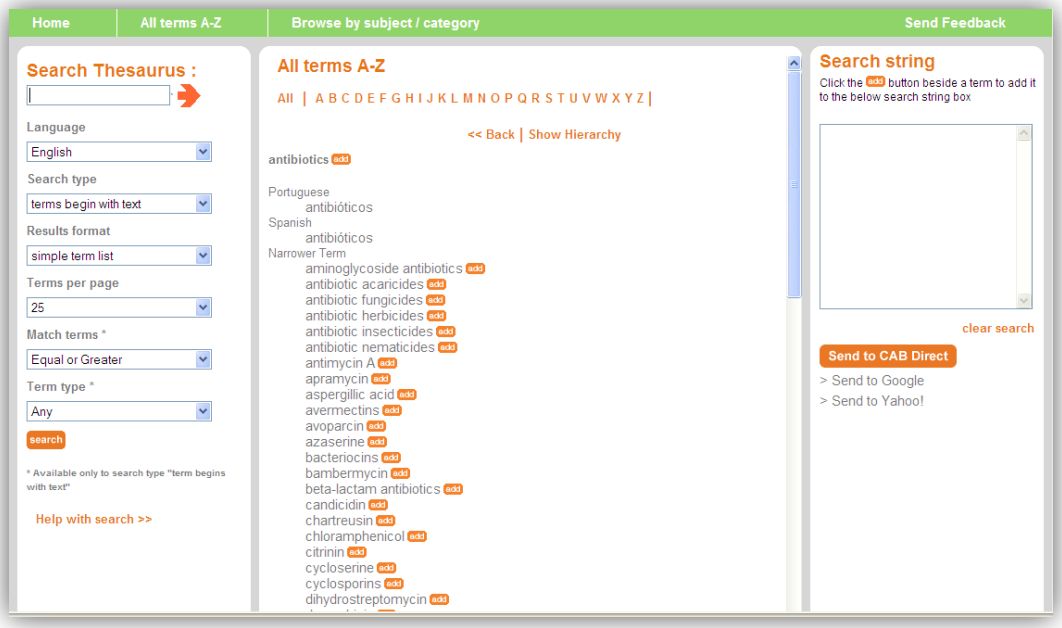
The CAB Thesaurus, CABI's controlled indexing vocabulary, now contains nearly 100,000 preferred indexing terms, and is used to index all the key concepts within the original article. These indexing terms are entered into the four indexing fields, Descriptors, Geographic Location, Organism Descriptors and Broad Terms, as described earlier. The CAB Thesaurus is provided as part of the CAB Direct platform, as an integrated search guide. You can use it to check for the correct terms to use in your search profile. You can also use it to automatically select terms and add them to your search. To browse the CAB Thesaurus, simply click on the **Thesaurus** tab, in the top menu. This will open the Thesaurus browse screen, shown below:



In the Search Thesaurus : search box, at the top of the left-hand column, type in the term that you want to look up, and click the search button. There are various parameters that you can set, down the left-hand side. In the screen, on the next page, we have searched for the term "Antibiotics".



In the frame above, we see an alphabetical list of terms, from the Thesaurus, starting with our term Antibiotics. To see the Thesaurus entry for any of these terms, simply click on the term of interest. In this example, we are interested in the term Antibiotics. Clicking on a term will display the Thesaurus entry for that term. The entry for Antibiotics is shown below:

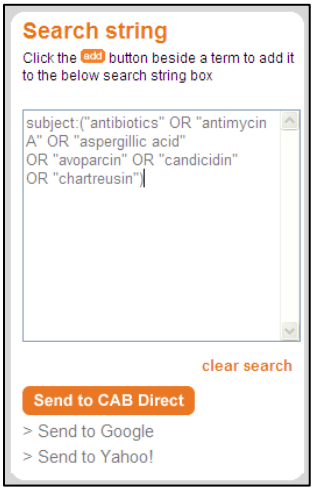


We now see the term Antibiotics at the top of its hierarchy. Below Antibiotics, we see all the Narrower Terms, one level below. These are the more specific terms, used to index these specific concepts. Next to each term, there is an add button, which can be used to add any selected term to the Search string box, on the right of the screen. Clicking on add will add it to the Search string box. If more than one term is clicked, the terms will be combined with the OR operator, as shown on the screenshot below.

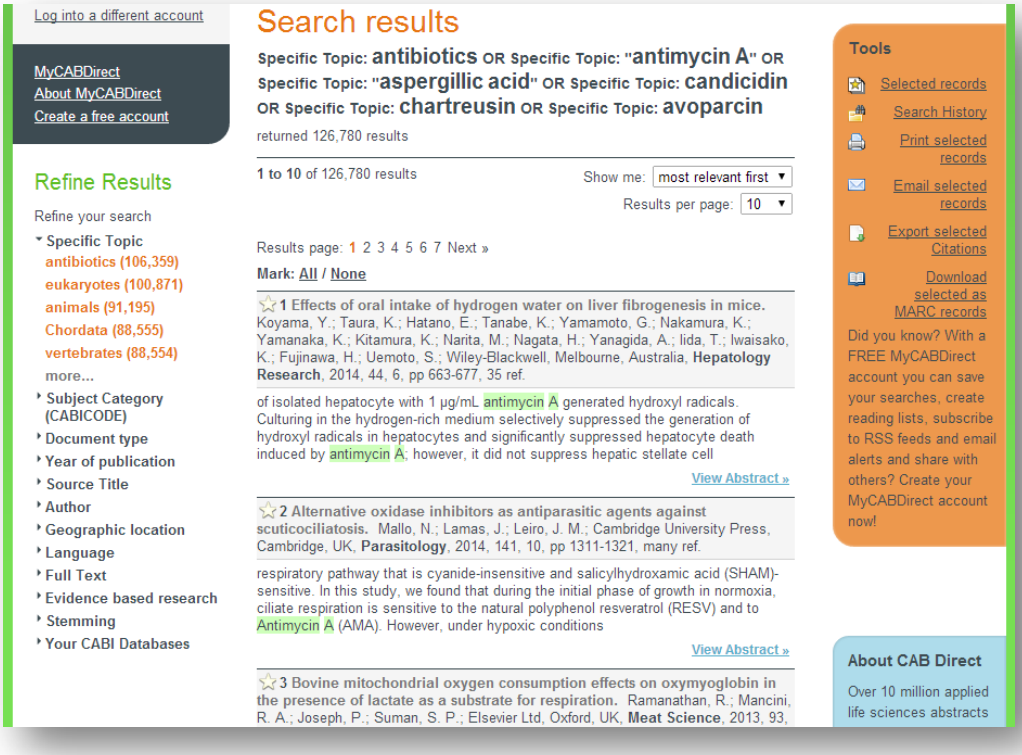
Once you have selected all the terms that you want to search, simply click the Send to CAB Direct button to perform the search. CAB Direct will then run your Thesaurus search and display the results, as shown below.

The search is actually performed in the Free Text Index, and is not restricted to the CABI Indexing fields, so this will provide the broadest result.

If you wish to modify the search statement, before you run the search, the **Search string** box, in the Thesaurus screen, is fully editable. You can add extra terms, change **OR**s to **AND**s, or add brackets, quotation marks and field tags. In the example below, we have added the field tag **subject** and a set of round brackets.



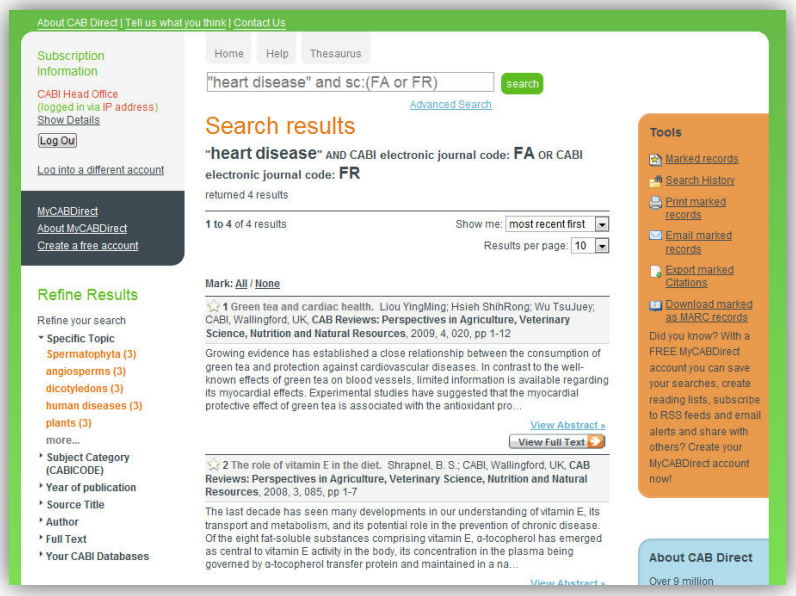
Here, we see that CAB Direct has searched for the selected terms in the **Subject** field, rather than the **Free Text** index.



This new CAB Thesaurus is very flexible and powerful. There is a separate, in-depth Tutorial available on using the CAB Thesaurus.

Subject Codes Field

In addition to the CABI indexing fields and the CABICODES, Global Health records are classified using a set of two character Subject Codes. Initially developed as a production tool for the printing of the 46 printed Abstracts Journals, these Subject Codes have been expanded to code records for broad subject areas like Parasitology, Human Nutrition, Plant Pathology, etc. Database records will have at least one code, but may have several, coding for different concepts within the original paper. The Subject Code (sc) field is also used to code database records which have links to the CABI Full Text databases and the CAB eBooks, which are available as separate databases. This coding helps when trying to limit to full text records or records in various database subsets. If, for example, a database user also subscribes to the CAB Reviews full text database, they could search for **"heart disease" and sc:(FR or FA)**, for example, and this would retrieve records about heart disease that have links through to electronic, full text Reviews on the CAB Reviews full text database. The following screenshot shows two Global Health records with CAB full text links through to two CAB Reviews.



For institutions subscribing to Global Health, that also have a subscription to CAB Reviews full text database, these links will take the user straight through to the full text PDF file. Click to see a full of the CABI Subject Codes.

Search Limits

On the Advanced Search screen, there are a number of limit options including:

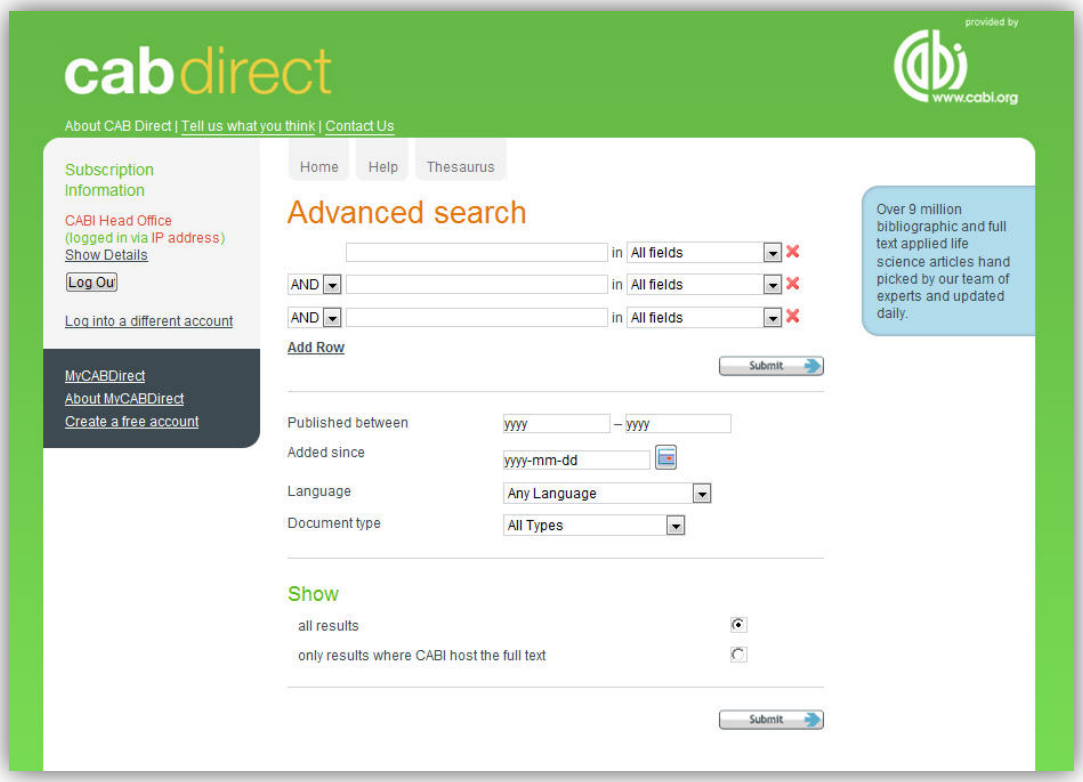
Published between: To limit results to records from original articles published within a range of years.

Added since: This limits the search to records added to the database since a specified date, and is useful for updating searches.

Language: Limits the search to records for original articles published in a specific language.

Document type Will limit results to records for a specific document type, such as Journal or Book.

At the bottom of the Advanced Search screen there is also an option to Show only database records that have links through to the free CAB Full Text articles.



Additional Search Fields

We have looked at some of the most important search fields, but there are many more that can be useful on occasions. A list of all the fields can be found in the short help file, accessed from the CAB Direct screen by clicking on the Help tab at the top of the screen. Here is a table of those additional fields.

Description	Field Tag/Name
Additional Authors	ad
Author Affiliation	aa
CABICODES	cc
CAS Registry Numbers	ry
Conference Dates	cd
Conference Title	ct
Corporate Author	ca
Country of Publication	cp
Descriptors	de
Digital Object Identifier	oi
Document Editors	ed
Document Title	do
Email	em
English Item Title	et
Non English Item Title	ft
Geographic Location	gl
Identifiers	id

Description	Field Tag/Name
ISBN	bn
ISSN	sn
Item Type	it
Language(s) of Summary	ls
Language(s) of Text	la
Location of Publisher	lp
Main Abstract	ab
Organism Descriptors	od
Pan Number	pa
Personal Author	au
Personal Author Variants	av
Publisher	pb
CABI Product Code	sc
Up-posted Descriptors	up
Web URL	ur
Year of Publication	yr