

CABI Training Materials

CAB Direct

Advanced Searching of CAB Abstracts

Contents

Why use Advanced Search features.....	3
Field Searching.....	3
Title Fields	6
Author Fields.....	6
I. Personal Authors.....	6
II. Corporate Authors.....	7
Index Terms or “Descriptors”	8
CABICODES.....	11
The CAB Thesaurus	12
Subject Codes Field.....	17
Search Limits	18
Additional Search Fields	19

Why use Advanced Search features

In the Simple Searching session, we looked at single and multi-word searching of CAB Direct using the Free-Text index. However, in a typical CAB ABSTRACTS 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 includes.

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

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 the breeding of maize but, by searching for **Maize** and **Breeding** in the Free-Text index, you may get papers about the breeding of cattle fed on maize. In order to improve the quality and relevance of your search 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 and is used as an advanced search feature.

Field Searching

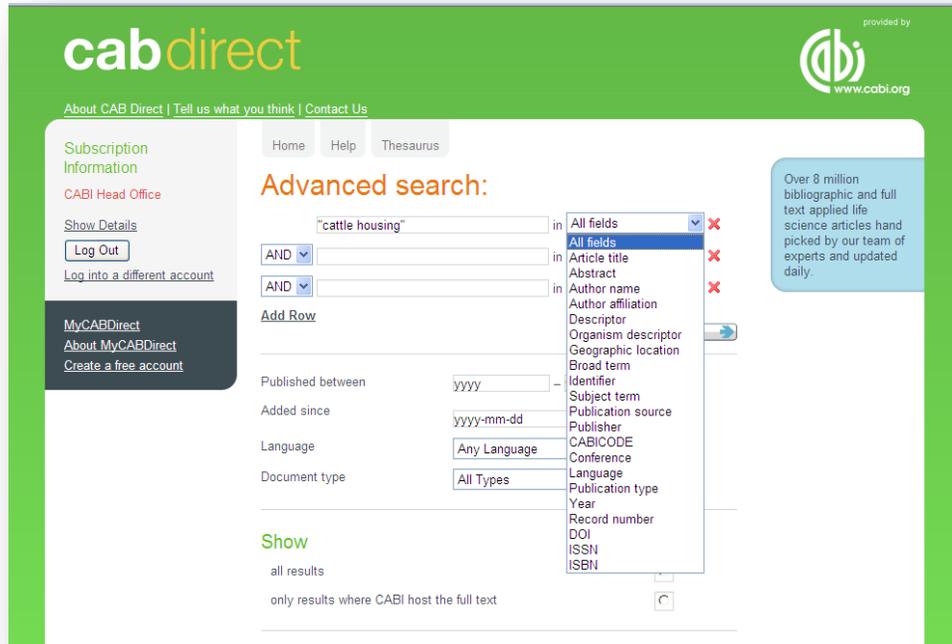
All the fields that appear in the Free-Text index, shown above, are individually searchable. This is very useful for refining your search.

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:"cattle housing"

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 field tags that can be used to select the tag for the field that you want to use.



The drop-down list includes the tags for all 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 the housing of cattle and sheep in the UK. The search terms here are **housing**, **cattle**, **sheep** and **UK**. In the Quick search screen, we might have performed four 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:

Home Help Thesaurus

Advanced search:

in

in

in

[Add Row](#)

Published between –

Added since

Language

Document type

The terms **sheep or cattle** have been entered in the top search box. The two terms have been combined using the Boolean Operator OR. Next to **cattle or sheep** the “Organism Descriptor” index has been selected from the drop-down menu. More about these individual indexes later. In the second search box, the term **housing** has been entered, alongside which has been chosen the field tag “Descriptor”. Finally, in the third search box, the term **UK** has been added, and the field tag, “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 CAB ABSTRACTS records have an English Item Title (Article Title). 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, for searching purposes.

The image shows a search interface with three rows of search criteria. Each row consists of a text input field, a connector dropdown, and a field selection dropdown. The first row has the text 'sheep or cattle' in the input field, 'AND' in the connector dropdown, and 'Article title' in the field selection dropdown. The second and third rows have empty input fields, 'AND' in the connector dropdown, and 'All fields' in the field selection dropdown. Each row ends with a red 'X' icon.

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:

author: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 hyphens or dashes, as shown below:

author:smith-j-a

or

author:"smith a j"

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:world-health-organization

ca:who

Note, again, the use of double quotes or the use of hyphens (dashes) 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:

Fields	Tags
Organism Descriptor (od:)	OD:
Geographic Location (gl:)	GL:
Descriptor	DE:
Broad Term (Up-posted Term)	UP:
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 CAB ABSTRACTS 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 Beetles, you would need to search for the scientific index term, **Coleoptera**, rather than Beetles. 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 a paper is a general paper about Beetles, for example, it will be indexed with the Organism Descriptor term **Coleoptera**. But, if the paper is about a specific beetle species, it will be indexed with the species name, and not the word **Coleoptera**. In the past, this policy has made searching for broad concepts like "beetles" very difficult because, in order to find every record, the user needed to search not only for Coleoptera, but had to include all the specific names of individual beetles. This is clearly a difficult if not impossible task.

The problem was solved, several years ago, when CABI began using the CAB Thesaurus to add additional index terms, automatically, to a new field call the **Broad Term** 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 Coleoptera, what this means is that

every time a beetle species name appears in the **Organism Descriptor** field, the broader term **Coleoptera** is automatically added to the Broad Term field. What this means is that a user can search for the term Coleoptera, in the Broad Term field, and they will find, automatically, all the records containing the narrow beetle species names. The field tag for the Broad term field is **up**:

up:coleoptera

Other search examples:

od:cattle

gl:(france or germany or spain)

od:rice and up:"south east asia"

The last indexing field, not yet mentioned, is the Identifier field. This field is used for non-controlled index terms; terms that do not appear in the CAB Thesaurus. This field is important for papers that discuss 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 tag is used for which field. It can also reduce the amount of typing, if you use brackets, as in the following example:

subject:(rice AND irrigation 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 also “indexed” with a classification system known as CABICODES. The CABICODES are a hierarchical list of classification codes that divide the subject coverage of the CAB ABSTRACTS database 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 Forestry, for example, has its own set of codes, as shown below.

- KK000** Forestry, Forest Products and Agroforestry (General)
- KK100** Forests and Forest Trees (Biology and Ecology)
- KK110** Silviculture and Forest Management
- KK120** Forest Mensuration and Management (Discontinued March 2000)
- KK130** Forest Fires
- KK140** Protection Forestry (Discontinued March 2000)
- KK150** Other Land Use (Discontinued March 2000)
- KK160** Ornamental and Amenity Trees
- KK500** Forest Products and Industries (General)
- KK510** Wood Properties, Damage and Preservation
- KK515** Logging and Wood Processing
- KK520** Wood Utilization and Engineered Wood Products
- KK530** Chemical and Biological Processing of Wood
- KK540** Non-wood Forest Products
- KK600** Agroforestry and Multipurpose Trees; Community, Farm and Social Forestry

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:kk160AND de:"urban development"

cc:kk* AND de:management

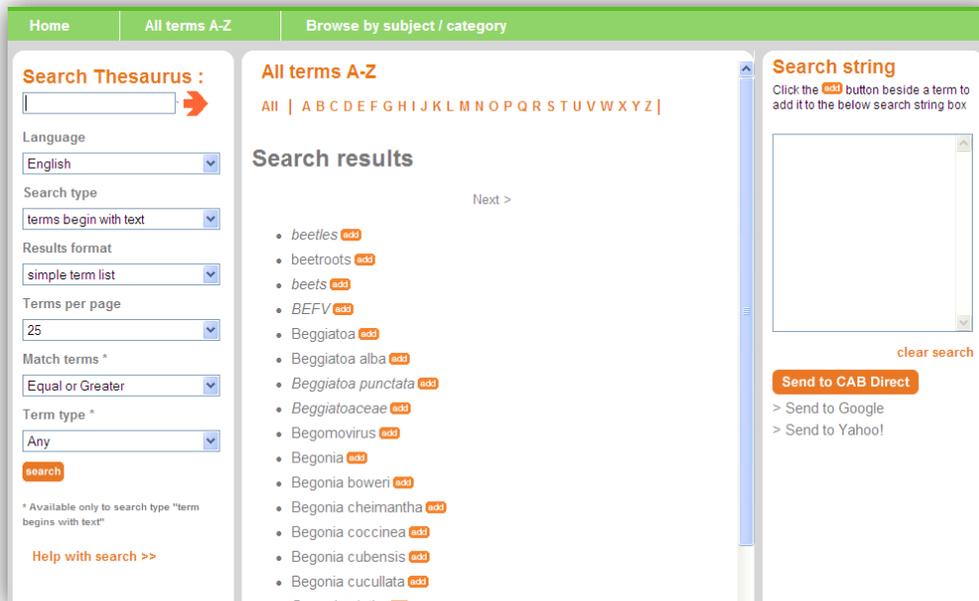
Note the use of truncation in the second example.

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 on the next page:

The screenshot shows the CAB Thesaurus 2011 website. The header is green with the CAB logo and the text 'CAB Thesaurus 2011'. Below the header is a navigation bar with 'Home', 'All terms A-Z', 'Browse by subject / category', and 'Cookie Information'. The main content area is divided into three columns. The left column contains a search interface with a search box, a language dropdown set to 'English', a search type dropdown set to 'terms begin with text', a results format dropdown set to 'simple term list', a terms per page dropdown set to '15', a match terms dropdown set to 'Equal or Greater', and a term type dropdown set to 'Any'. The middle column contains an 'About CAB Thesaurus' section with a list of strengths including controlled vocabulary, regular updates, broad coverage, and specific terminology. The right column contains a 'Search string' section with a text input box and buttons for 'Send to CAB Direct', 'Send to Google', and 'Send to Yahoo!'. The footer contains an 'About CABI' section with a brief description of the organization's mission.

In the **Search Thesaurus** : search box, at the top of the left 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, but the default options will provide you with the broadest result. In the screen, below, we have searched for the term “Beetles”.



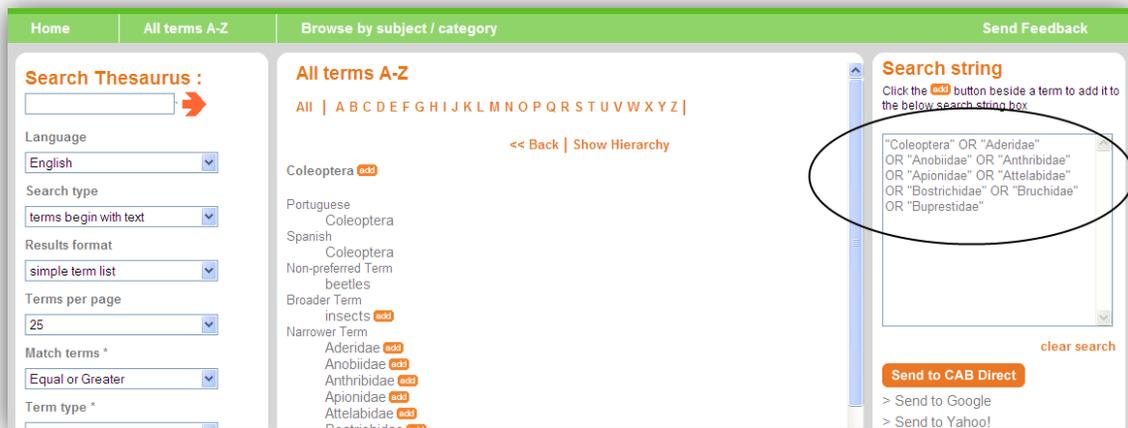
In the central frame, we see an alphabetical list of terms, from the Thesaurus, starting with our term Beetles. To see the Thesaurus entry for any of these terms, simply click on the term of interest. In this example, we are interested in the term Beetles. Clicking on a term will display the Thesaurus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

The screenshot shows the 'All terms A-Z' page. On the left is a 'Search Thesaurus' panel with various filters: Language (English), Search type (terms begin with text), Results format (simple term list), Terms per page (25), Match terms (Equal or Greater), and Term type (Any). A red circle highlights the 'Preferred Term' 'Coleoptera' in the list. The right panel contains a 'Search string' box and buttons for 'Send to CAB Direct', 'Send to Google', and 'Send to Yahoo!'.

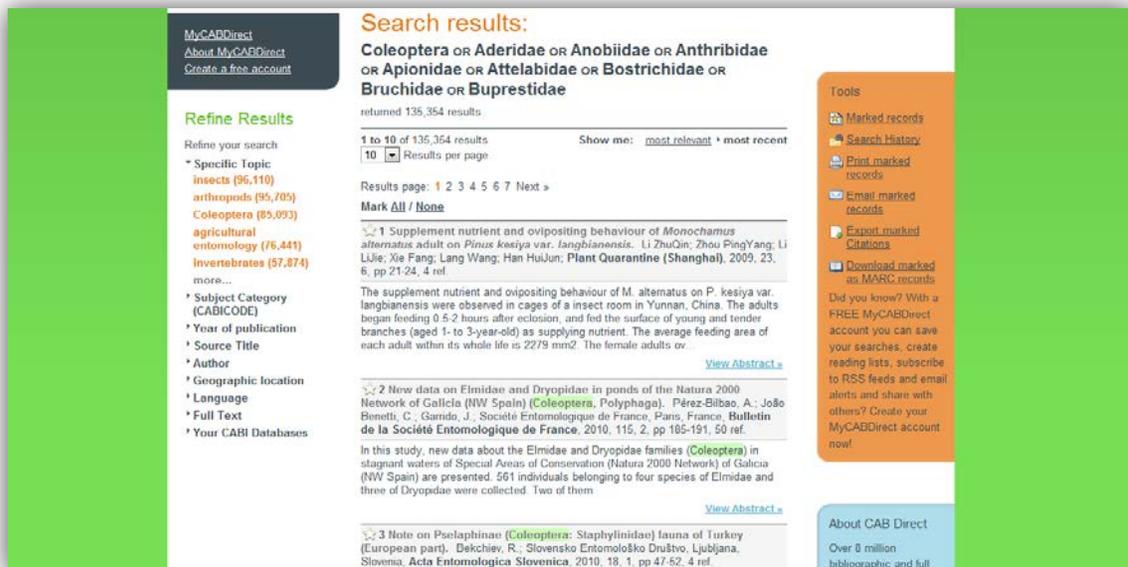
Here, we see that the Thesaurus is telling us that the preferred indexing term for the concept “Beetles”, is actually Coleoptera, not Beetles. We can now click on Coleoptera to see its Thesaurus entry, shown below:

This screenshot shows the detailed Thesaurus entry for 'Coleoptera'. The left panel remains the same. The central panel displays a hierarchical list of terms: Coleoptera (preferred), Portuguese (Coleoptera), Spanish (Coleoptera), Non-preferred Term (beetles), Broader Term (insects), and Narrower Term (Aderidae, Anobiidae, Anthribidae, Apionidae, Atteblabidae, Bostrichidae, Bruchidae, Buprestidae, Byturidae, Carabidae, Cerambycidae, Chrysomelidae, Cleridae, Coccinellidae, Cryptophagidae, Cucujidae, Curculionidae, Dermestidae, Dryophthoridae). The right panel is identical to the previous screenshot.

We now see the term Coleoptera at the top of its hierarchy. This page shows us the Broader term, above Coleoptera (Insects), and all the Narrower Terms, one level below Coleoptera. These are the more specific terms, used to index these particular concepts. Next to each term, there is the **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 in the screenshot below.

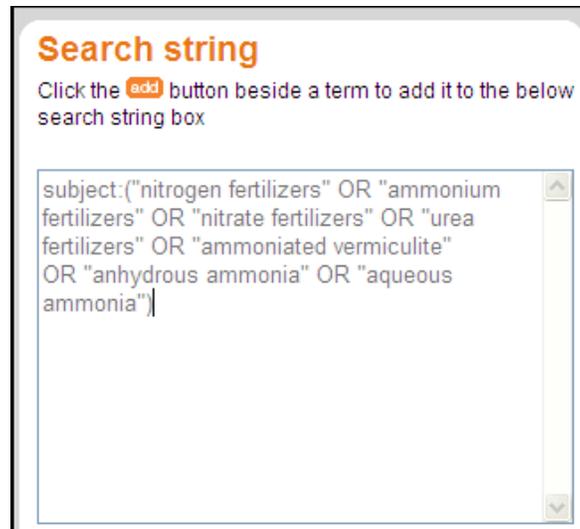


Once you have selected all the terms that you want to search, simply click the **Send to CAB Direct** button to perform the search. CAB Direct will then run your Thesaurus search and display the results, as shown below.



The search is actually performed in the Free Text Index, and is not restricted to the CABI Indexing fields, to 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 **ORs** to **ANDs** 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, CAB Abstracts 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 Horticulture, Soils and Fertilizers, 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 allows for seamless, full text linking from a database record through to the CABI full text PDF files. If, for example, a database user also subscribes to the CAB Reviews full text database, they could search for **"transgenic plants"** and **sc:(fr or fa)** and this would retrieve records about Transgenic Plants that had links through electronic, full text Reviews on the CAB Reviews database. The following screenshot shows a CAB Abstracts record with a CAB full text link button to a CAB Review.

The screenshot displays the CAB Direct search results page. The search query is "transgenic plants" and sc:(fr or fa). The search results show 82 results, with the first result being "Increasing the vitamin E content of food by in-plant production" by Farré, G., Zorrilla, U., Zhu, C. F., Christou, P., Capell, T., CABI, Wallingford, UK, CAB Reviews, 2012, 7, 030, pp 1-10. A red circle highlights the "View Full Text" button for this record. The page also includes a sidebar with "Refine Results" and "Tools" sections.

For institutions subscribing to CAB Abstracts, who also have a subscription to CAB Reviews full text database, this link will take the user straight through to the full text PDF file.

A full list of the CABI Subject Codes can be found at the following Web site:

<http://www.cabi.org/Uploads/File/User%20Guides/cabisubjectcodes.pdf>

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 between 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 records for original articles published in a specific language.
- Document type** Will limit results to records for original articles of 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.

The screenshot displays the 'Advanced search:' interface on the CAB Direct website. At the top, there are navigation links: 'About CAB Direct | Tell us what you think | Contact Us'. Below this are 'Home', 'Help', and 'Thesaurus' tabs. On the left sidebar, there is a 'Subscription Information' section with links for 'CABI Head Office', 'Show Details', 'Log Out', and 'Log into a different account'. Below that is the 'MyCABDirect' section with links for 'About MyCABDirect' and 'Create a free account'. The main search area features three input fields for search terms, each with a dropdown menu set to 'All fields' and a red 'X' icon. Below these is an 'Add Row' button and a 'Submit' button. Further down, there are filters for 'Published between' (with 'yyyy' - 'yyyy' placeholders), 'Added since' (with 'yyyy-mm-dd' placeholder and a calendar icon), 'Language' (set to 'Any Language'), and 'Document type' (set to 'All Types'). At the bottom, there is a 'Show' section with two radio button options: 'all results' (selected) and 'only results where CAB host the full text'. A final 'Submit' button is located at the bottom right. A blue callout box on the right side of the page states: 'Over 8 million bibliographic and full text applied life science articles hand picked by our team of experts and updated daily.'

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
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

Description	Field Tag/Name
Identifiers	ID
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