

Intrallect intraLibrary 3.0: Advanced Searching

Revision: 5
Created: 4th August 2005
Last Revised: 12th February 2009
Contact: support@intrallect.com
Company: [Intrallect Ltd](#)
Product: [intraLibrary](#), [Digital Repository](#)
Copyright: © Intrallect Ltd 2003-2009. All rights reserved.

This document is made available to support Intrallect's customers and users of Intrallect's software. The text of these documents and the design of the intraLibrary software are both the intellectual property of Intrallect Ltd. Intrallect do not provide this document for any other purpose, and offer no warranty nor accept any liability for its use in any other context. Parts of this document are based on the documentation for Apache Lucene.

Table Of Contents

1. Introduction.....	3
2. Query Syntax.....	3
2.1. Basics	3
2.2. Case Insensitivity	4
2.3. Stopwords	4
2.4. Stemming.....	4
2.5. Plurals	4
2.6. Wildcard Searches.....	4
2.7. Fuzzy Searches	5
2.8. Proximity Searches.....	5
2.9. Range Searches	5
2.10. Boosting a Term.....	6
3. Search Operators	6
3.1. Boolean Operators.....	7
3.2. Grouping Operators	8
3.3. Special Characters.....	8
4. Advanced Search	9
4.1. Fields	9
4.2. Collections	9
4.3. Combining Searches.....	10
4.4. ALL or ANY	10
4.5. Saving Searches.....	10
4.6. Public Saved Searches.....	11
4.7. Reusing Saved Searches	11

1. Introduction

This guide describes two aspects of the search interface of intraLibrary - the syntax of search queries and the construction of complex search criteria through combinations of search terms. In addition, it describes how search parameters can be stored and reused.

Searches are conducted in intraLibrary either through simple search, which is always available using the search box at the top of the screen, or advanced search, which is always available from the *advanced search* tab near the top of the screen. In this manual the [query syntax](#) and [operators](#) sections apply to all simple searches and free text fields in advanced searches. The final section only applies to advanced searches.

Note: At present, searches in intraLibrary are performed on the metadata for each resource, not on the text contained in the resource itself. Full-text searching is planned for future versions. In addition, from v3.0 onwards, administrators have the option to switch on deep searching of content packages. When this is switched on, searches are performed on any metadata contained anywhere within an IMS or SCORM Content Package, including Manifest, Sub-Manifest, Organisation, Item, Resource or Asset level metadata. Otherwise, only metadata at the package manifest level will be searched.

2. Query Syntax

2.1. Basics

A query is specified by using query terms. There are two types of terms: Single Terms and Phrases:

- A Single Term is a single word such as *test* or *hello*.
- A Phrase is a group of words surrounded by double quotes such as "*hello dolly*". (If the double quotes are omitted the search will be treated as if it is a Boolean OR search. Matches will be found which contain the word *hello* or the word *dolly*).
- Multiple terms can be combined together with Boolean operators to form a more complex query (see [Operators](#) below.).

IntraLibrary provides a powerful and flexible search syntax. It automatically detects plurals and words with similar stems and ignores case. In addition there are many features that allow the search to be modified to narrow or broaden the terms used. These include: wildcards, fuzzy searching, proximity searching, range searching and boosting of terms, all

of which are described in this document.

2.2. Case Insensitivity

All searches are case insensitive. You can use upper or lower case and words in the other case will be found.

2.3. Stopwords

IntraLibrary uses a list of stopwords. Searches using words in this stop list are ignored. These are usually words such as *and*, *the* and *or*. Your intraLibrary administrator can modify the list of stopwords on your intraLibrary installation.

2.4. Stemming

Searches in intraLibrary are filtered using the Porter Stemming method. This means stems of words will be used in the search so that, for example, searches for *running* will match *run* and searches for *run* will match *running*.

2.5. Plurals

Plural searches are conducted automatically, so that a search for *bucket* will match *buckets*.

2.6. Wildcard Searches

There is support in intraLibrary for single and multiple character wildcard searches.

- To perform a single character wildcard search use the ? symbol.
- To perform a multiple character wildcard search use the * symbol.

The single character wildcard search looks for terms that match with only a single character replaced. For example, to search for *text* or *test* you can use the search:

te?t

Multiple character wildcard searches looks for 0 or more characters. For example, to search for *test*, *tests* or *tester*, you can use the search:

test*

You can also use the wildcard searches in the middle of a term.

te*t

Note: You cannot use a * or ? symbol as the first character of a search.

2.7. Fuzzy Searches

IntraLibrary supports fuzzy searches based on the Levenshtein Distance, or Edit Distance algorithm. To do a fuzzy search use the tilde, ~, symbol at the end of a single word term. For example to search for a term similar in spelling to "roam" use the fuzzy search:

roam~

This search will find terms like foam and roams.

2.8. Proximity Searches

IntraLibrary supports finding words which are within a specific distance of each other. To do a proximity search use the tilde, ~, symbol at the end of a phrase. The default is ~0. For example to search for the words *grilling* and *fish* within 10 words of each other in a resource's metadata, use the search:

"grilling fish"~10

2.9. Range Searches

Range queries allow you to match resources whose field(s) values are between the lower and upper bound specified by the range query. Range queries can be inclusive or exclusive of the upper and lower bounds. Sorting is done lexicographically.

[Ernani TO Falstaff]

When used to search the Title field this will find all resources whose titles are between Ernani and Falstaff, including Ernani and Falstaff.

Inclusive range queries are denoted by square brackets. Exclusive range queries are denoted by curly brackets.

{Aida TO Otello}

This will find all resources whose titles are between Aida and Otello, but NOT including Aida and Otello.

2.10. Boosting a Term

The relevance level of matching resources is based on the terms found in intraLibrary. To boost the relevance of a specific term use the caret, ^, symbol with a boost factor (a number) at the end of the term you are searching. The higher the boost factor, the more relevant the term will be.

Boosting allows you to control the relevance of a resource by boosting a specific term. For example, if you are searching for

Verdi Rigoletto

and you want the term *Verdi* to be more relevant boost it using the ^ symbol along with the boost factor next to the term. You would type:

Verdi^4 Rigoletto

This will make resources with the term Verdi appear more relevant. You can also boost Phrase Terms as in the example:

"opera by verdi"^4 "opera by puccini"

By default, the boost factor is 1. Although the boost factor must be positive, it can be less than 1 (e.g. 0.2)

3. Search Operators

In free text searches, including all simple searches, terms may be combined using Boolean operators. They may also be grouped together to create powerful combinations.

3.1. Boolean operators

Boolean operators allow terms to be combined through logic operators. IntraLibrary supports *AND*, *+*, *OR*, *NOT* and *-* as Boolean operators.

Note: Boolean operators must be in CAPITAL LETTERS.

OR

The OR operator is the default conjunction operator. This means that if there is no Boolean operator between two terms, the OR operator is used. The OR operator links two terms and finds a matching resource if either of the terms exist in a resource's metadata. This is equivalent to a union using sets. The symbol `||` can be used in place of the word OR. To search for resources that contain either "italian operetta" or just "opera" use the query:

```
"italian operetta" opera
```

or

```
"italian operetta" OR opera
```

AND

The AND operator matches resources where both terms exist anywhere in the text of a single resource. This is equivalent to an intersection using sets. The symbol `&&` can be used in place of the word AND. To search for resources that contain "italian opera" and "spanish opera" use the query:

```
"italian opera" AND "spanish opera"
```

+ (plus)

The **+** or **required** operator requires that the term after the + symbol exist somewhere in a resource. To search for resources that must contain "opera" and may contain "verdi" use the query:

```
+opera verdi
```

NOT

The NOT operator excludes resources that contain the term after NOT. This is equivalent to a difference using sets. The symbol `!` can be used in place of the word NOT. To search for resources that contain "italian opera" but not "spanish opera" use

the query:

"italian opera" NOT "spanish opera" **Note:** The NOT operator cannot be used with just one term. For example, the following search will return no results:

NOT "scottish opera"

- (minus)

The - or **prohibit** operator excludes resources that contain the term after the - symbol. To search for resources that contain "italian opera" but not "spanish opera" use the query:

"italian opera" - "spanish opera" **Note:** The "-" symbol is only treated as an operator when it has a blank space on at least one side. Otherwise it is treated as a hyphen. That means if you search for *x-ray* you can be hopeful of exposing some resources.

3.2. Grouping

IntraLibrary allows you to use parentheses to group clauses to form sub queries. This can be very useful if you want to control the Boolean logic for a query.

To search for either "italian" or "spanish" and "opera" use the query:

(italian OR spanish) AND opera

This eliminates any confusion and makes sure that *opera* must exist and either term *italian* or *spanish* may exist.

3.3. Special Characters

When you want to search for terms which include characters which have a special meaning in the query syntax you must **escape** these characters - you must let them escape from their normal meaning. IntraLibrary supports escaping of special characters that are part of the query syntax. The current list of special characters is

+ - && || ! () { } [] ^ " ~ * ? : \

To escape these character use the \ before the character. For example to search for (1+1):2 use the query:

\(1+1\) :2

Note: The exception to this is searching the IEEE LOM identifier fields (general.Identifier.Entry and metaMetadata.Identifier.Entry). These fields can be searched without needing to escape any special characters, allowing users to cut and paste entire identifiers into the simple or advanced search box.

4. Advanced Search

The Advanced Search facility in intraLibrary allows you to search specific fields and to construct complex searches from combinations of these specific searches. In addition, you can save these searches to use again at a later time.

4.1. Fields

You can select specific fields to search using the "Select search field" drop-down menu in the Advanced Search interface.

The options that appear on this menu have been configured by your intraLibrary administrator. Once you have selected a field you will be presented with an option to search that field. Some of these fields contain free text while others contain a limited vocabulary of terms. If the field you choose has a limited vocabulary then the available options will be presented in a second drop-down menu. If the field contains free text then you can include any of the search terms and operators that would normally be used in the simple search interface.

To conduct the search click on the *search* button.

4.2. Collections

There may be multiple collections of resources in the library. The system will search all collections you have access to. If you prefer, you can search a specific collection or collections.

To search by collection, select "Collection" from the drop-down "Select search field" list. A drop-down list will appear below, listing the collections which you have permission to

search. Choose the collection you would like to search. If you wish to search more than one collection, click the "add" link and select "Collection", then a specific collection from the available list again. If you only want resources that are in both collections, select "ALL"; if you want resources from either collection, select "ANY".

4.3. Combining Searches

When using Advanced Search to search specific fields you can combine the search terms. For example you might want to search the *Title* field for Mozart, and the *Technical Format* field for a suitable audio format.

To add another field to your search, click on the *add* link. You can add as many additional fields as required. If you wish to remove any field simply click on the *remove* link.

4.4. ALL or ANY

Choosing to match "ALL" criteria will give you a narrow search; it is the same as a Boolean "AND" search.

Choosing to match "ANY" criteria will give you a broad search; it is the same as a Boolean "OR" search.

Example:

- Choosing "ANY" criteria for a search for "title"= fish and "title"= chips will give you all resources in the library with either word in the title, including resources on computer chips and wood chips that don't even mention fish. Choosing "ALL" criteria for the same search will only return resources with the words "fish" and "chips" in their titles.

When searching using multiple constraints the default is that ALL criteria must be satisfied.

Remember to click on the *search* button to conduct the search.

4.5. Saving Searches

When you have created a set of search constraints that you may wish to use again some time in the future you can save the search by clicking the "Save" button. A pop-up window will appear allowing you to give the saved search a name. The saved search will then appear in the drop-down list when you select "Saved searches" from the "Select search field" drop-down under advanced search. It will also appear in your work area, if you are a contributor user, so that you can use it to search resources there.

4.6. Public Saved Searches

Administrators can save a search filter as a public saved search. This makes the saved search available to all users in the Preferences section of their Profile, where it can be applied to all searches they carry out. Public saved searches are also made available to contributor users in their work area to search the list of resources they are working on.

4.7. Reusing Saved Searches

You can use search filters by selecting "Saved searches" from the "Select search field" drop-down menu under Advanced Search. If you no longer require a saved search, click on the "Edit Saved Searches" button, then use the X button to remove it.

A saved search can be applied by default to all your searches if you wish. To set a default saved search go to the Preferences section of your Profile.

A saved search can also be used in your work area (contributor users only) to search the list of resources you are working on.