Resource Sharing Alliance NFP
715 Sabrina Drive
East Peoria, IL 61611
866-940-4083

31 August 2022

Subject: Advanced Searching in Online Articles (CloudSourceOA) Search Tab

Introduction

This document provides users with background information on how CloudSource performs searches and is intended to mitigate confusion over why certain results are returned and also provide power users with the tools they need to get the specific information they want.

The Basics

CloudSource makes three independent searches as part of each request, and then combines them into a results page with the results ordered based on a weighting algorithm.

The first is a phrase search that looks for the entire search term together. The second is an AND search that looks for all the words in the search in any order and not necessarily next to each other. The third is an OR search where the more words the search has the fewer words are required to match. (See the list of search fields below.)

Any article with a field that matches on any of the three searches will be returned. The more fields that match and the more searches they match the higher up the page the record will be.

Boolean logic is used to form the database requests from the search terms provided by the user. This means that the order of the terms and Boolean operators used directly determine how the search engine interprets the request, and in turn, affects the results returned.

Additionally, author names have been indexed in a way so that searches return results for names that are phonetically similar. Priority is still given to names with exact matches, but similar names are also included.

Boolean Operators

Boolean operators are special characters and phrases that are used to specify relationships in advanced searches. CloudSource uses the following Boolean operators:

Boolean operators and Special Characters

Operator/Character	Explanation
AND, &&, +, MUST	These operators are used when the results to be returned need to contain terms from both sides of the operator. For example, **pandemic AND flu** returns results that contain both pandemic and flu, but they do not have to be in a specific order. Both *Bird flu and pandemic flu and Pandemics and protectionism: evidence from the "Spanish" flu* appear in the results.
NOT, -	These operators are used when the results to be returned cannot contain the term following the operator. For example, **occultation NOT religion** returns results about stellar observations while excluding references to occult practices of religions.
OR, ||	These operators are used when the results to be returned need to contain at least one of the terms from either side of the operator. For example, **hydrogen || helium** returns results that contain one term or the other, with higher priority given to those results that contain both, like *Hydrogen contamination in liquid helium.*
" "	Double quotes are used to specify exact phrases. These operators are used when the results to be returned need to contain the term exactly as it is written within the double quotes. For example, **"seedling reaction"** only returns results that contain an exact match within the article or the article's metadata. This includes *seedling reactions* (the plural form), but not *reactive seedling*.
()	Parentheses are used to group words together in a looser format than double quotes. These operators are used when the results to be returned need to contain at

	least one of the terms within the parentheses. For example, **(medical massage) AND sports** returns results that contain at least one of the terms within the parentheses.
:, =	A colon or equal sign is used when specifying search fields. The term to the left of the operator indicates which search field to match against. (See the list of search fields below.) The term to the right of the operator is the search being made. For example, **author.affiliation:Harvard** returns results for articles written by scholars affiliated with Harvard. Using specific schools or fields of study can be used to narrow the results returned.

Creating Boolean Phrases

Boolean operators and special characters can be used in combination to create phrases for precise searches. For example, you may want to search for articles about war and peace, but do not want Leo Tolstoy's "War and Peace" to clutter the results. In this case, use a phrase with the following search grammar:

war AND peace NOT (leo tolstoy)

Note that the author's name is in parenthesis. This is because Boolean operators only use the first word, unless you use double quotes or parenthesis to combine multiple words together. If the above example did not have parenthesis, the search would use the Boolean "NOT leo", which is not fully correct for the context of this search.

To better see how this works, consider these two searches, where search results are identical:

dog training AND high energy

dog (training AND high) energy

Boolean operators, parenthesis, specifying a search field, and exact searches with double quotes cause CloudSource OA to break the search into chunks and then combine the results from the chunk searches together. The above search has three

parts: dog, training AND high, energy. Both of the above searches are therefore equivalent to this phrase:

dog AND training AND high AND energy

It is important to use double quotes or parenthesis to make sure that multiple words are kept together in a search. For the context of this example, the search would work best by using a phrase with this search grammar:

(dog training) AND (high energy)

Search Fields

The following search fields can be specified when forming a search query:

Search Fields

Name	Description
abstrakt	The article's abstract, it includes a summary of key findings.
allAuthors	This field is a shorthand for searching across all author fields, including: • author.affiliation • author.name • author.orcId
allSubjects	This field is a shorthand for searching across multiple other fields, including: • abstrakt • chemicalName • fieldOfStudy • keyword • meshTerm.meshHeadingName • meshTerm.qualifierName • subject • chemicalId • clinicalTrialId • meshTerm.meshHeadingId • meshTerm.qualifierId
author.affiliation	Usually a university or research institution, this field indicates the author's backing or credentials.

author.name	The author's name.
author.orcId	The Open Researcher and Contributor ID (ORCID) is a 16-digit alphanumeric code that is used to identify the same author or contributor across publications.
chemicalId	Scientific articles may reference chemicals by an ID, as assigned by the National Library of Medicine. This field represents that ID.
chemicalName	Scientific articles may reference chemicals by name. This field represents that name.
clinicalTrialId	Clinical trials registered with the National Library of Medicine are assigned an ID that can be linked to PubMed journal articles. This field represents that ID.
doi	The digital object identifier (doi) is a standardized way of uniquely identifying digital works such as articles, reports, publications, and other media content. This field represents that ID.
fieldOfStudy	Some articles include a field of study, such as Engineering or Psychology. This field represents that metadata.
format	The article's format. This may include HTML, PDF, etc. [Need definitive list]
general	This includes all possible searchable fields and is the default behavior when no search field is specified.
isbn	International Standard Book Number (ISBN) has a numeric id of either 10 or 13 numbers. Hyphens should be removed before entering ISBNs as search terms.
keyword	Many articles include keywords relative to the topic. This field searches only against an article's keywords.
license	The Creative Commons (CC) open license used. Examples include:

	• CC0 • CC BY • CC BY-SA • CC BY-NC • CC BY-NC-SA • CC BY-ND • CC BY-NC-ND • other-oa
meshTerm.meshHeadingId	Medical Subject Heading (MeSH) is used by the National Library of Medicine to index articles, such as those found in PubMed. This field represents the ID of a mesh heading associated with an article, for example "D007239".
meshTerm.meshHeadingName	This field represents the mesh heading name associated with an article, for example "Infections".
meshTerm.meshId	This field represents a shorthand for querying both the meshTerm.meshHeadingId and the meshTerm.meshQualifierId at the same time. An example search would be meshTerm.meshId:D007239:Q000469
meshTerm.meshName	This field represeents a shorthand for querying both the meshTerm.meshHeadingName and the meshTerm.meshQualifierName at the same time. An example search would be meshTerm.meshName:Infections:parasitology
meshTerm.meshQualifierId	This field represents the ID of the mesh qualifier associated with a mesh heading, for example "Q000469".
meshTerm.meshQualifierName	This field represents the name of a mesh qualifier associated with a mesh heading, for example "parasitology".
publication.issn	Much like how books are assigned an ISBN at publishing, journals, magazines, and other scholarly publications have an eight digit International Standard Serial Number (ISSN).
publication.name	The publication name.

publicationCountry	The country in which the journal publishes.
publishDate	The date the article was published.
publisher	The name of the publisher.
reference.doi	The DOI of articles that reference any given article.
reference.lensId	The Lens.org ID of articles that reference any given article.
source	There are a number of sources that have been combined to provide the best metadata for an article. This field can be used to narrow results by which sources were used.
title	The title of the article.

Note: The field names are case sensitive.