SEARCHING IN SIGHTLINE V RELATIVITY

Sightline Help Documentation

Table of Contents

Briefing Summary

Sightline by Consilio and Relativity by kCura® utilize different underlying search technologies that have different syntaxes – or languages. Sightline employs a Lucene-based search engine called, ElasticSearch, but Relativity employs dtSearch. Thus, Sightline users who are familiar with Relativity need a "Rosetta stone" to interpret the search language they know into the search language that may be less familiar to them. This document serves such a purpose so that case teams can understand how searching in Sightline works in order to uncover the documents of interest in their project database.

Body Content Searching

The below matrix shows a side-by-side comparison of search syntax for body content (text) searches in Sightline v. Relativity 8.0 Keyword Search, based on this Relativity Searching guide, with an example query in each.

Category	Relativity Syntax	Relativity Example	Sightline Example	Sightline Syntax	Sightline Considerations
Exact phrase search	Quotation wrapped phrase	"cubs win"	"cubs win"	Quotation wrapped phrase	Same as Relativity Keyword Searching.
Phrases with Wildcards	Quotation wrapped phrase	"cub* win"	"Cub* win"	Quotation wrapped phrase	Same as Relativity Keyword Searching.
OR Boolean operator	OR	cubs OR sox	Either: cubs OR sox or cubs sox	Capitalized OR or space separator between terms	OR must be capitalized, otherwise, or is treated as a term instead of an operator. If you have two adjacent terms separated by a space, this is also treated by Sightline as an implied OR operator between the terms. This is different than Relativity that interprets two adjacent terms separated by a space as an implied AND operator between the terms.
AND Boolean operator	AND	cubs AND sox	cubs AND sox	Capitalized AND	AND must be capitalized, otherwise and is treated as a term instead of an operator.
NOT Boolean operator	NOT	cubs NOT sox	cubs NOT sox	Capitalized NOT	NOT must be capitalized, otherwise not is treated as a term instead of an operator.
Wildcard	*	*hicag*	*hicag*	*	Same as Relativity
Stemming	~	Cub~ for lexicon-	Cub*	*	In Relativity, tilde ~ stemming logic is lexicon-based and

		based stemming

Or

Cub* for wildcard-based stemming | | | therefore an English-only function.

In Sightline, there is no lexicon-based stemming logic, so only the wildcard (*) is used.

In this case, cub* will return docs with hits with root "cub" such as cub, cubs, cubbie, cubicle, etc. |
| **Single Character Wildcard** | ? | Chica?o | Chica?o | ? | Same as Relativity |
| **Nested Arguments** | ( ) | (cubs OR sox) AND (Chicago NOT Boston) | (cubs OR sox) AND (Chicago NOT Boston) | ( ) | Same as Relativity.

It is a best practice to use parentheses in order to specify the desired order of operations for any query that includes more than two terms. If you do not do so, Sightline will always give priority to NOT, then AND, and last, OR. |

Proximity	w/	cubs w/5 Chicago	"cubs Chicago"~5 "("Chicago Cubs") baseball"~2 "("Chicago Cubs"~2) baseball"~3	~	The two terms inside the quotations are those that are considered for the proximal connection. The number after the ~ character determines the maximum proximal distance allowed between the terms.
Directional Proximity	PRE/	Chicago PRE/5 Cubs	"Chicago cubs"~2! "cubs Chicago"~2!	!	To use directional proximity, please the terms in the order you want them within the quotation marks and then add the ! after the standard proximity logic.
Fuzzy for common misspellings	%	Chic%%go	Chicago~ Chicago~1	~	Sightline's ElasticSearch engine uses the tilde to determine the allowable amount of fuzziness. The default allowable character changes is 2. A character change can be changing a character (box vs fox), removing a character (black vs lack), inserting a character (sic vs sick) or transposition two adjacent characters (act vs cat). The user can switch to only 1 allowable character change by adding a 1 after the tilde as per the example to the left.

Fielded Searches

Metadata attributes are indexed as well in both Sightline and Relativity. By default, query searches in Sightline will go against BOTH document content as well as all fielded attributes. If a user wants to search only the body content, the user may use the query *DocumentContent.Content: term* in order to isolate the search to just the document body.

Searches may also be done against specific fields only. Fields are either text type fields, numeric type fields, date only fields or datetime fields, principally. The standard structure for field-specific searching in Sightline is [FIELD]: (query). For example, searching the EmailSubject field for the term "Rosetta" would be EmailSubject: (Rosetta).

Numeric fields, date fields and datetime field types have a series of special queries allowing for exact, greater than, less than, before, after and between searches. The below matrix shows a side-by-side comparison of these special field searches in Sightline v. Relativity with an example query in each.

Category	Relativity Syntax	Relativity Example	Sightline Example	Sightline Syntax	Sightline Considerations
Not Null Value	Must be a structured query against SQLServer database using Is Set attribute		_exists_: MasterDate MasterDate(*)	_exists_: [FIELD] [FIELD](*)	This search returns all documents that contain any populated attribute for the specified field
Missing Value	Must be a structured query against SQLServer database using Is Not Set attribute		NOT _exists_: MasterDate NOT MasterDate(*)	NOT _exists_: [FIELD] NOT [FIELD](*)	This search returns all documents that contain any missing/null attribute for the specified field
Is Exact Date	=	MasterDate >= '7/24/2008'	MasterDate: [2008-07-24 TO 2008-07-24]	FIELD: [YYYY-MM-DD TO YYYY-MM-DD]	Notice the proper separators between Year, Month and Day is – character, not / character. This is true in all date or datetime field searches.

Is Exact Date and Time of Day	=	MasterDate = 07/27/2008 1:23 PM	MasterDate: [2008-07-24 13:14:00 TO 2008-07-24 13:14:00]	FIELD: [YYYY-MM-DD HH:MM:SS TO YYYY-MM-DD HH:MM:SS]	Time of day is stored in 24-hour clock format in Sightline. Also, users must be aware of the time zone the data was processed in (UTC by default) and whether Daylight Savings Time is a consideration based on the day of the year and the originating source of the data. Any datetime field search can optionally include time of day
Is Before Date	<=	MasterDate <= 07/24/2008	MasterDate: [* TO 2008-07-24]	FIELD: [* TO YYYY-MM-DD]	
Is After Date	=>	MasterDate => 07/24/2008	MasterDate: [2008-07-24 TO *]	FIELD: [YYYY-MM-DD TO *]	
Is Between Dates	AND	([FIELD VALUE] >= '7/24/2008') AND ([FIELD VALUE] < '8/25/2008')	MasterDate: [2004-07-06 TO 2008-07-24]	FIELD: [YYYY-MM-DD TO YYYY-MM-DD]	
Is Exact Count	=	DupeCount = '8'	DupeCount: [8 TO 8]	FIELD: [# to #]	

Is Greater Than Count	=>	DupeCount => '8'	DupeCount: [8 TO *]	FIELD: [# to *]	
Is Less Than Count	<=	DupeCount <= '8'	DupeCount: [* TO *]	FIELD: [* to 8]	
Is Between Count		(DupeCount >= '8') AND (DupeCount <= '20')	DupeCount: [8 TO 20]	FIELD: [# to #]	

Default Stopwords and Reserved Characters

Stop words are high frequency words that do not materially add value to the content and are therefore not indexed by the search engine. Below is a side-by-side comparative of English language stop words between the two search engines.

Relativity Stopwords	Sightline Stopwords
- Any single letter word - Any single numeric digit 0 – 9 - Select punctuation: Period (.), Dash (–), Colon (:), Semicolon (;), At Symbol (@), Hyphens and Dashes (-) - The following English words: about, after, all, also, another, any, are, as, at, be, because, been, before, being, between, but, both, by, came, can, come, could, did, do, does, each, else, for, from, get, got, has, had, he, have, her, here, him, himself, his, how, if, in, into, is, it, its, just, like, make, many, me, might, more, most, much, must, my, never, no, now, of, on, only, other, our, out, said, same, see, should, since, so, some, still, such, take, than, that, the, their, them, then, there, these, they, this, those, through, to, too, under, up, use, very, want, was, way, we, well, were, what, when, where, which, while, who, will, with, would, you, your	- Select punctuation Period (.) Colon (:) Semicolon (;), Left Single Quote (`), Comma (,), Right Single Quote ('), Ampersand (&), Percent (%), Forward Slash (/), Carat (^) - Select currency symbols ($ and £) - There are no English stopwords in Sightline

In addition to the words and characters above not included in the default index, select characters are reserved for their special properties. These characters are also not indexed by default because they are interpreted by the search engine as performing discrete functions within the search engine itself.

Relativity Reserved Special Characters	Sightline Reserved Special Characters
" ( ) * ? % @ ~ # & : =	( ) " [ ] ~ ! \ * ? # (when used in Regular Expression searching)
	Sightline Automatically Escaped Characters
	+ - : \ > < = { } | #

Performing Work Product Searches in Sightline

Work Product (ex. tags, folders, assignments, productions) are also searchable in Sightline. Below is guidance on how to execute Work Product searches in Sightline.

Work Product searches are performed using the Advanced Search interface within Sightline and can be imported via Batch Search Upload as well.

Basic Tag, Folder, Assignment, Redaction Label, Saved Searches and Productions Work Product searches take the form of ATTRIBUTE: [name: ["A", "B"]] where the comma is treated as an implied OR operator. For example, tags: [name: ["DEVTagDuplicate", "DSG_Tag11"]] will return any document that is tagged with DEVTagDuplicate tag or DSG_Tag11 tag.

Certain Work Product Searches can contain additional defined attributes to further narrow the results set. For example, Assignments Work Product searches can optionally define which reviewers the documents are assigned to, what state the documents are in and a time range when the documents were assigned or when they were completed. For example, the query assignments: [name: ["ASGN1", "Comms"], distributedTo: [16, 21, 14], completed: "true", completeddate: ["from" : "2016/05/02", "to" : "2016/05/12"]] will return any document that is in Assignment ASGN1 or Comms that is presently in the individual reviewer batch of reviewer user 16, 21 or 14 and is in a Completed state and was marked Completed between May 2 2016 and May 12 2016.

Getting Help with Searches

Searching is not always straight forward, and approaches may vary that yield improved or degraded performance. If you are unsure how best to structure a search, please contact your Consilio Project Manager who may consult with the technical/engineering team to provide guidance.