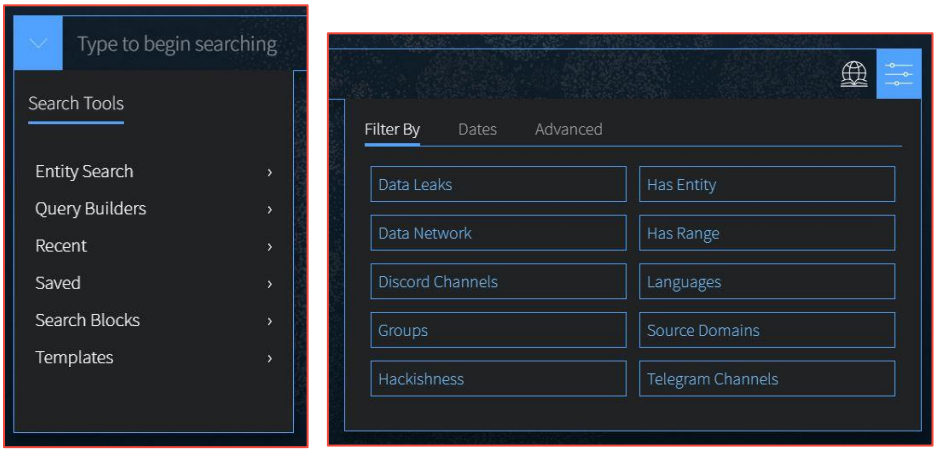
Searching in DarkOwl Vision

Using the Search Bar

1. The search bar works like most search engines; simply type words, phrases, numbers, or characters. The Search Techniques section goes into more detail and options for searching, but here a few quick start tips:
    - Use Boolean operators when searching more than one keyword. More Information: Searching with Booleans.
    - Use straight quotations ("**Jane Doe**") to send the query as one phrase, Jane Doe.
    - Use the **exact:** search operator (**exact:fullz**) to prevent word stemming, and search for exact matches of that term. See: Stemming and Searching for Exact Terms.
2. Use the left drop-down menu to open **Search Tools**, which include these options:
    - **Entity Search:** the best way to search for Emails, CVEs, Credit Cards, Cryptocurrency Addresses, IP Addresses, and Social Security Numbers.
    - **Query Builder:** a helper for search variations or advanced formatting for commonly search items.
    - **Search Blocks:** Pre-populated keyword as well as any custom search blocks that you create, are accessible from this menu for easy access.
    - **Templates:** pre-populated search templates to help you get started quickly.

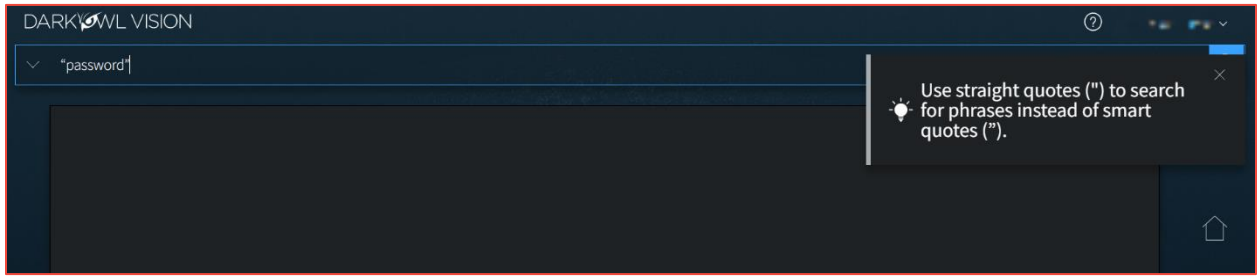


Search Tools drop-down menu (left); Search Bar Filter menu (right)

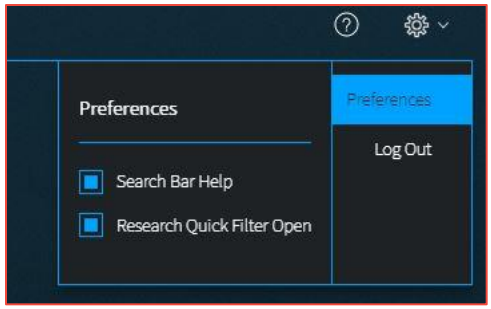
3. Once you've started searching, **Recent Searches** and **Saved Searches** will also appear in the Search Tools menu, for easy access.
4. Click on the **right Lexicon icon** to open the Lexicon and filter to (or exclude) darknet sites of interest. More information is provided in this section: Lexicon.
5. Click on the **right Filter icon** to open the Filter menu. More information about these is provided in this section: Filters, Dates, Advanced Search Options.
6. **Guidelines for using multiple search components (Filters, Lexicon entries, and free text) in the same search.** Using multiple *different* search components AND together; using multiple of the *same* search component (with different values) OR together:
   - Lexicon Market, Lexicon Forum in same search—OR
   - Search Bar Free Text and Lexicon Paste in the same search—AND
   - Network (Tor), Network (Discord) in the same search—OR
   - Entity Search, Search Bar Free Text in the same search—AND

Search Bar Help

Search tips will appear in the upper right corner when a search could be optimized or includes a character the field doesn't recognize.



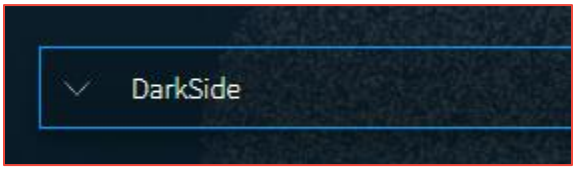
You can turn this feature off by clicking on your name in the upper right corner, selecting Preferences in the drop-down menu, and unchecking the toggle for Search Bar Help.



Search Techniques

Single Terms

To find any document with a specific keyword, simply put that keyword into the search bar:



Phrase Searching

To find two or more keywords in a specific order, place the keywords within double quotes:

- **"AES 256"**

Without the double quotes, the search would be sent as: *AES OR 256*. This OR search is an inclusive search and may return results that only include the term AES, that only include the term 256, that include both terms but not next to each other, and that include both terms next to each other.

Boolean Searching

Use Boolean operators **AND, OR,** and **NOT** to specify inclusions, alternate terms, or exclusions. (You can substitute AND, OR, and NOT with **&&,** ||**,** and **!** respectively.) Keywords and field searches can be separated by any of the above in order to fine-tune your results.

- **drugs OR crime** – find documents with either 'drugs' or 'crime'
- **drugs AND crime** – find documents with both 'drugs' and 'crime'
- **DarkOwl AND (hack OR ddos OR 0day)** – find documents with DarkOwl and any one of three hacking keywords

Note: Boolean operators must be in all caps. If they aren't in all caps, DarkOwl Vision will consider the word "and", "or", etc as keywords.

Using Subqueries/Boolean Order of Operations

You can group together phrases to form subqueries, using parentheses ( ) to indicate each clause. This is especially important when both ANDs and ORs are used, to designate the correct order of operations for your subqueries.

- **DarkOwl AND (drugs OR crime)** – find documents with DarkOwl plus one of either drugs or crime
- **("AES-256" OR "AES 256") AND ("RSA-4096" OR "RSA 4096")** – find documents with one of AES-256 or AES 256 plus one of RSA-4096 or RSA 4096

Stemming and Searching for Exact Terms

DarkOwl Vision supports a process called stemming, which tries to reduce a word to an approximation of its stem or root form. Usually, terms are stemmed to plural/singular versions or different tenses. This means that searching will return matches on related forms of a word, unless you specify otherwise:

- **Hack** may return **Hacked, Hacker, Hacking,** in addition to **Hack**

When you want to search for a specific term, including special characters and punctuation, use the **exact:** operator to prevent word stemming:

- **exact:hack** – will return only documents containing the word hack

Excluding Items from a Search

Keywords can be excluded in multiple ways:

- Using the NOT Boolean operator
- Prefacing the term with a hyphen (-)
- Prefacing the term with an exclamation mark (!)

For example, the three searches below are equivalent and will find documents that contain 'DarkOwl' but not 'drugs'. Note that when excluding a keyword via hyphen or exclamation mark, it must be placed directly before the keyword with no space in between.

(DarkOwl NOT drugs) is the same as **(DarkOwl -drugs)** is the same as **(DarkOwl !drugs)**

You can also exclude values in other fields in the same way:

- DarkOwl NOT domain:drugs.onion
- DarkOwl -domain:drugs.onion
- DarkOwl !domain:drugs.onion

Searching for Entities in the Search Bar

When searching for specific entities, such as an email address or credit card number, directly in the search bar, we recommend searching with the appropriate operator. The Search Tools (in the drop-down on the left-side of the search bar) are a shortcut and automatically convert your query to the correct syntax.

- email:first.last@company.com
- ccn:11111111111111111
- cryptocurrency:15ivMrk8VzaK9TEN85XYssVbU3Yd6tLzb9
- cve:cve-2022-12345
- ipAddress:127.0.0.1\/24

- ssn:123-45-6789

When searching for multiple entities, use the search operator and a Boolean OR, as follows:

- email:(first.last@company.com OR last.first@company.com)
- ccn:(1111111111111111 OR 2222222222222222)
- cve:(2022 OR cve-2021-12345)

Searching for both Keywords and Entities in the Search Bar

When searching for both keywords and specific entities (such as an email address or credit card number) directly in the search bar, use the following format:

- ("First Last" OR Nickname) AND email:first.last@company.com

Using Wildcards

Wildcards (***** or **?**) are currently allowed in limited usage, in the middle or end of terms only. An asterisk (*****) is used to find zero or more unknown characters; a question mark (**?**) is used to find any one unknown character. Examples:

- **dar*** – will find "dar", "dart", "darkowl", "daredevil", etc
- **d?rk** – will find "dark", "dork", "dirk", etc; will not find "drk" (however, d*rk would)

Note: DarkOwl Vision does not support leading wildcards. In other words, a search term cannot begin with either one of the wildcard characters.

Using Proximity Searches

You can find words near to each other by using quotations and selecting a maximum distance allowed: **"password hack"~2**. The maximum distance supported is 9.

Using Pattern Matching / Regular Expressions

Lucene-based regular expressions are allowed and should be wrapped by forward slashes (/). Not all functionality you may be familiar with may be supported. Additionally:

- **These queries may time out,** particularly when searching for a high volume of unknown characters. Regex searching is computationally heavy and will result in slower, less performant searches.

To use a regular expression in Vision, place the expression between two forward slash characters:

- **/r[0-9a-zA-Z]{24,34}/** – to find results matching the pattern of a Ripple cryptocurrency address (which starts with 'r', then has anywhere between 24 to 34 alphanumeric characters)

Note: Not all regex functionality you may be used to is supported by our system.

Using Special Characters

The following characters are reserved:

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

If any of the above characters are in a keyword or phrase being searched, you can escape the character with a backslash: \. For example, to search for mentions of a URL within a document, such as *https://darkowl.com/darkint-blog*, you must escape the colon, forward slashes, and hyphen, otherwise the search will return an error.

You can perform this search multiple ways:

- Escaping the special characters: https\:\/\/darkowl.com\/darkint-blog
- Putting the whole keyword in quotes: "https://darkowl.com/darkint-blog"

Without escaping the special characters, this search will be interpreted as:

1. Searching within a field called 'https' (which doesn't exist)
2. An empty regular expression (// signifies the start and end of a regex with no content)
3. The keyword 'darkowl.com'
4. The start of a regular expression starting with 'darkint-blog'
5. No end to the regular expression (will return an error)

Field Searching (Search Operators for Metadata)

Every search performed will look in one or more fields for the keyword(s) being searched. By default, the search bar will search both the 'title' and 'body' fields of documents. This means that results will be returned if the keywords you're looking for are found in either the body of the document or the title (or both). This means:

- A search of **the word 'drugs' in the search bar** is the same as
- **title:drugs OR body:drugs**.

Most searches will not require specifying a field name, since title and body are automatically searched. However, other metadata fields can be searched, including specifying only the title or body field, for example:

- title:alphabay
- hackishness:1
- domain:drugs.onion

The list of metadata fields is below. When searching within these fields, type the following search operators in the search bar, and then the query content:

- inUrl:

- contentType:
- headers.server:
- headers.last-modified:
- title: (to search within this field exclusively)
- body: (to search within this field exclusively)
- domain:
- leak:
- network:
- hackishness:

Multiple values within the same field can be searched in a number of ways. The following examples are equivalent:

- **domain:(drugs.onion OR crime.onion)** is the same as
- **domain:drugs.onion OR domain:crime.onion**

You can also look for phrases within specific fields using double quotes:

- **title:"Forum rules"**

Subqueries within fields are supported:

- **title:(darkowl AND (drugs OR crime))**

Note: when searching using fields, there should not be a space after the ':' character.

Field Searching (Search Operators for Chat Networks)

There are several search operators that can help you filter to content from Discord or Telegram servers, channels, and users.

- The **telegram:** operator filters to channels, IDs, or users from Telegram.
- The **discord:** operator filters to servers, channels, IDs, or users from Discord.
- The **user**: operator filters to usernames or user IDs from either Telegram or Discord.
- *To exclude telegram, discord, or user operators:* preface the search operator with a hyphen: ***-telegram:"EVILX.su Leaks Chat"***

Several examples are presented below:

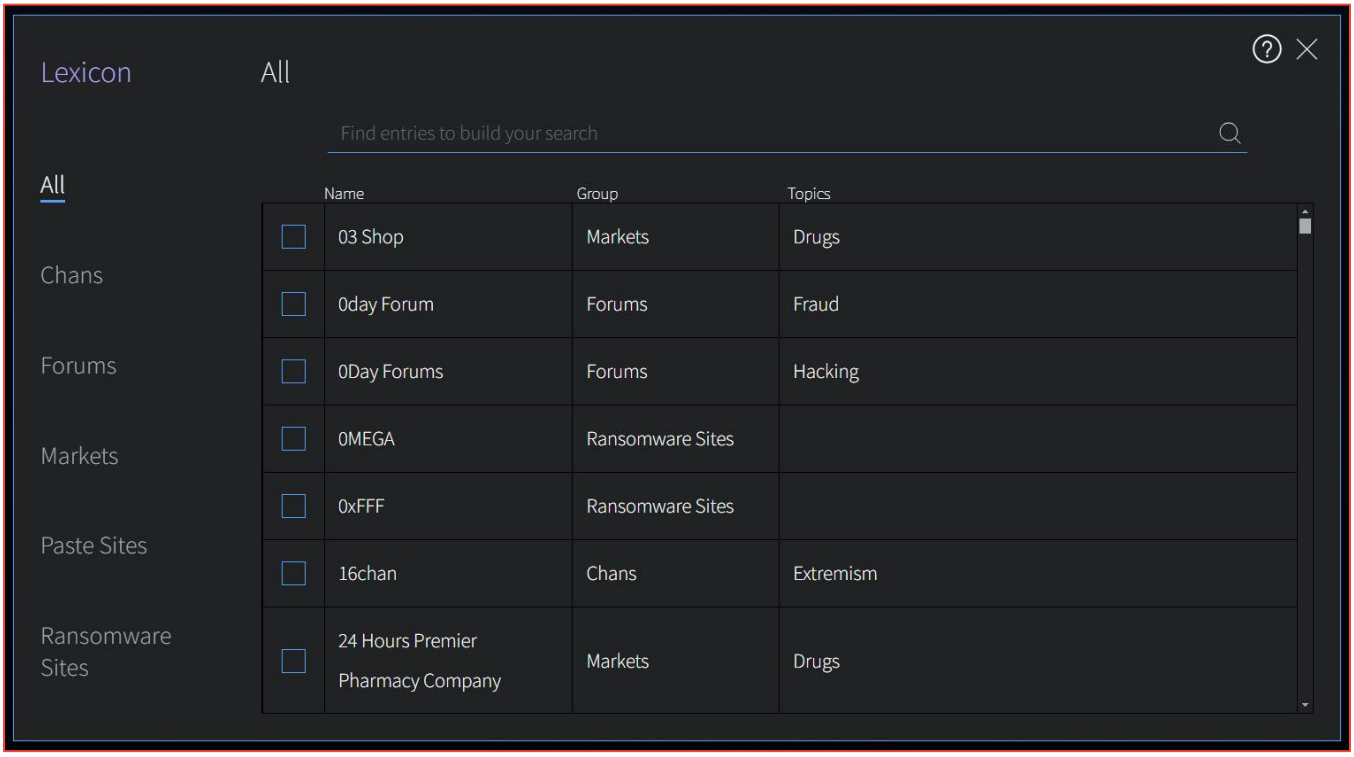
Search	Description
telegram:"Чат Военкоров Русской Весны"	Search for content from this Telegram channel (quotations are used as it has spaces in its name)

telegram:"***DEMON HUB*** 1 😈 ~ ***#THEDEMONNETWORK***"	Search for content from this Telegram channel (note quotations for spaces and special characters)
telegram:lapsus*	Search for content from any Telegram channel or username that starts with the characters *lapsus*
telegram:"-1001228309110"	Search for content from the Telegram channel with this channel ID. Include the - prefix with straight quotes.
discord:"HELLU!"	Search for content from this Discord server
discord:funny-sb	Search for content from this Discord channel
discord:tylerdurdan710	Search for content associated with username (Discord only)
user:tylerdurdan710	Search for content associated with this username on either Discord or Telegram

Lexicon

DarkOwl Vision's Lexicon is an easy-to-use tool intended to help you find interesting content from hacking forums, marketplaces, and other darknet sites. While not an exhaustive list of sites in our data, it's a good place to get started. Make a suggestion at any time for sites you'd like us to add at: https://www.darkowl.com/lexicon.

Click on the **right Lexicon icon** on the Search Bar to open the Lexicon. To find an entry in the Lexicon, start typing to filter through all entries, or select a Group from the left and scroll through the section. The search field will search all columns (Name, Group, Topics).



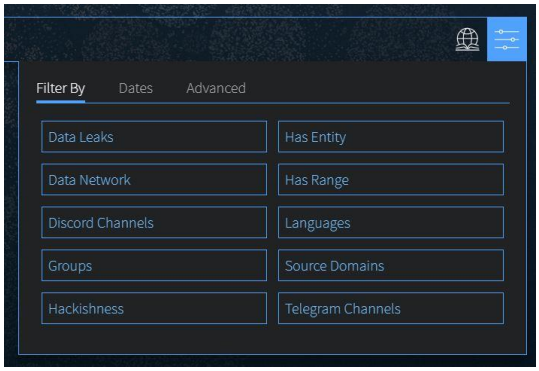
Clicking once next to your desired entry(ies) will immediately add the entry(ies) to the search bar. **Click once to include, click twice to exclude.**

Click the > icon, or anywhere outside of the modal, to go back to the Search Bar. In the screenshot below, we've selected Dread and Exploit.in forums, which will filter results to content from only those two forums. You can run the search with just the Lexicon entries, or add more keywords to your search.

Groups You'll Find in the Lexicon

Section	Filter	Description
Chans	Filter: Domain	A chan imageboard is considered a type of bulletin-board-like forum that revolves around the posting of images, often alongside text and discussion. Read more about Chans here.
Forums	Filter: Domain	Forums are online places where people discuss specific topic threads; some require authentication to access.
Markets	Filter: Domain	Markets include both small vendor shops to big-name marketplaces; some require authentication to access. Some marketplaces have previously been taken down by law enforcement, though we may still have historical content.
Paste Sites	Filter: Domain	Paste sites are online temporary content-hosting applications that allows for users to share text online anonymously.
Ransomware Sites	Filter: Domain	Ransomware sites are domains administered by ransomware groups. It is common for these groups to post their victims, and data stolen from their victims, on these sites.

Filters, Dates, and Advanced Search Options



Refine your search using the Filters icon on the right side of the search bar. This includes three tabs for:

Filters

Dates

Advanced options

Data Leaks

Filter or exclude content from any known data leaks or breaches. **Click the box next to the desired option.** To find individual leaks, see the Leak Explore section or the Lexicon.

- For free text searching, type *leak:leakname* in the search bar
- Prefix with **-** to exclude, i.e. *-leak:leakname* in the search bar

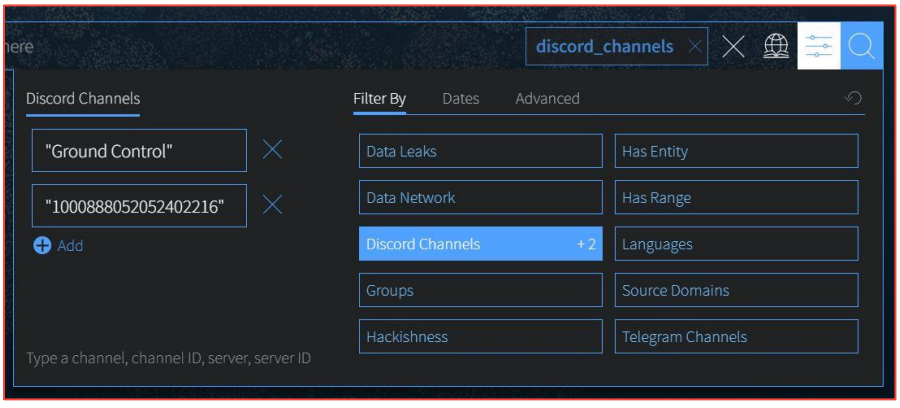
Data Network

Filter to content from a particular DarkOwl data collection network. Options include: Discord, FTP, I2P, IRC, Onion, OpenNIC, Telegram, Zeronet. **Click once to include, click twice to exclude.** More information: Data Networks.

- For free text searching, type *network:networkname* in the search bar
- Prefix with **-** to exclude, i.e. *-network:networkname* in the search bar

Discord Channels

Filter to content from one or more specific Discord channels or servers. Click the **Add** button, and then type either one channel name, channel ID, server name, or server ID in the box. You'll note that quotation marks are automatically applied when you start typing. You can add up to 10 channels or servers per query. You can remove previously-entered channels or servers by clicking the X next to the channel or server to remove.



- For free text searching:
    - *To filter to one channel or server:* Type ***discord:"Ground Control"*** in the search bar
    - *To filter to one channel ID or server ID:* Type ***discord:"757762132962967624"*** in the search bar
    - *To filter to multiple channels/IDs:* Type ***discord:("757762132962967624" OR "Ground Control")***
    - *To exclude discord channels or servers (free text searching only):* preface the search operator with a hyphen: ***-discord:"Ground Control"***

Groups

Groups are combined filters that narrow your search to specific categories; **click to include**:

- **Authenticated Sites:** Filter to content from sites requiring credentials or other challenges.
- **Blogs:** Filter to content from sites identified as blogs.
- **Chans:** Filter to content from a curated set of chan/imageboard forums selected by our analysts.
- **Darknet:** Filter to content from the Tor, I2P, and Zeronet darknets.
- **Forums:** Filter to content from sites identified as forums.
- **Forum Posts:** Filter to content specifically from forum posts (the **Forums** group will also include content from forum sites that may not be posts such as member profile pages, etc).
- **Markets:** Filter to content from sites identified as darknet marketplaces or vendor shops.
- **Ransomware:** Filter to content from known ransomware sites.
- **Paste Sites:** Filter to content from a curated set of paste sites.

Hackishness

Hackishness assigns a rating to every piece of content collected, indicating the likelihood to which the information could be used for criminal activity. The lower bound of hackishness is .01 and the upper bound is 1.0; the UI shows these as percentages on search results. You can quickly filter to results with hackishness by **using the slider** on the Hackishness filter to select a desired hackishness threshold.

You can also filter to hackish results using hackishness: in the search bar, which supports searching as range. This means, you can narrow down to values between two parameters, inclusive or exclusive, for example:

- **hackishness:[.01 TO 1]**
- **hackishness:{.01 TO 1}**

Note the '[' and '{' characters used above. In Lucene range queries, '[' and ']' are inclusive so the first query above would return values from .01 to 1, including both .01 and 1. The second example would return values from .01 to 1 not including.01 or 1. '[' and '{' can be combined:

- **hackishness:{.5 TO 1]** (this will find values greater than .5 and up to and including 1)

Has Entity (Credit Cards, Cryptocurrencies, CVEs, Email, IPs, Social Security Numbers)

Filter to content that have at least one selected Entity. **Click next to the Entity name to select.**

Has Range (Credit Cards, Cryptocurrencies, CVEs, Email, IPs, Social Security Numbers)

Filter to content that have a certain number of selected Entities. This filter is helpful in finding "dumps," as many threat actors will post multiple instances of PII on a singular site or document. **Type values next to a selected Entity**: Enter a lower bound (minimum 1), upper bound (maximum 999999), or use both fields to form a range (50 to 1000).

Language

Filter to content in a particular language. Languages are detected by DarkOwl Vision at the time of ingestion, using natural language processing. **Click once to include.**

- For free text searching, type *language:languagevalue* in the search bar

Source Domains

Filter to content from one or more domains, or exclude a particular domain by typing a hyphen in front of the domain. Type only the domain portion in the filter box (example: **arch3rsecgjqcmjb.onion**; no need for the www or http:// prefix). Be sure to remove any trailing slashes or paths from the domain.

- For free text searching, type *domain:domain.onion* in the search bar
- Prefix with **-** to exclude, i.e. *-domain:domain.onion* in the search bar

Telegram Channels

Filter to content from one or more specific Telegram channels. Click the **Add** button, and then type either one channel name or one channel ID in the box. You'll note that quotation marks are automatically applied when you start typing. You can add up to 10 channels per query. You can remove previously-entered channels by clicking the X next to the channel to remove.

Note: when typing a channel ID, use the hyphen prefix, i.e. -1001556588508.

- For free text searching:
  - *To filter to one channel:* Type ***telegram:"EVILX.su Leaks Chat"*** in the search bar
  - *To filter to one channel ID:* Type ***telegram:"-1001556588508"*** in the search bar

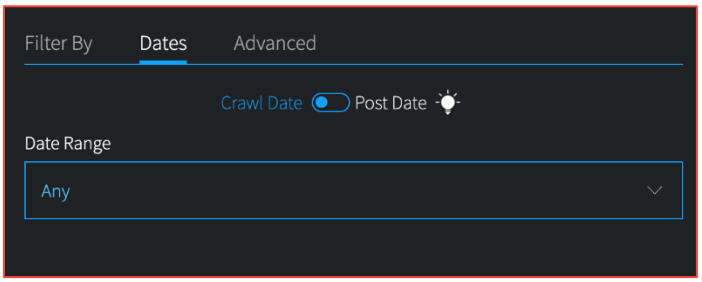
- *To filter to multiple channels/IDs*: Type ***telegram:("-1001556588508" OR "EVILX.su Leaks Chat")***
- *To exclude telegram channels (free text searching only):* preface the search operator with a hyphen: ***-telegram:"EVILX.su Leaks Chat"***

Dates

Use the drop-down to quickly select a time range for search results, or select Custom to choose a specific start/end date.

Use the toggle to choose between filtering by **Crawl Date** or **Post Date:**

- **Crawl Date:** Filters to any content that DarkOwl collected between the specified dates.
- **Post Date:** Filters to any content that was posted between the specified dates. Filtering by Post Date will filter to results where the post date is known. Other search results will be excluded.



Dates can be searched in the search bar as a range using the following date format: YYYY-MM-DDTHH-MM-SSZ:

- **crawlDate:[2021-07-01T00:00:00Z TO 2021-07-10T23:59:59Z]**
- **published:[2021-07-01T00:00:00Z TO 2021-07-10T23:59:59Z]**

As the Date filter supports range searching, you can narrow down to values between two parameters, inclusive or exclusive, as discussed in the hackishness section. In Lucene range queries, '[' and ']' are inclusive, and '{' and '}' are exclusive.

Advanced Options (Sort By, Show Similar, Empty Bodies)

Use Advanced Options to select a Sort option, or to show all results (including duplicates).

- **Sort options.** Use the drop-down to sort your results by **Relevance** *(default)*, **Hackishness, Crawl Date,** or **Post Date.**

  Note: Sorting by Post Date will filter to results where the post date is known. Other search results may be excluded.

- **De-duplicate your results.** You can choose to de-duplicate your result sets *(default)* or see all results, including similar results. The default is to de-duplicate result sets; you can toggle this on or off.

- **Empty bodies.** Our collection includes documents collected that do not contain any text characters; text content is stored in the Body field. Use the drop-down to select one:
    - **Any document** (documents can include text in the body, or no text in the body)
    - **Results must have body field** *(default)* (to only return documents that had text content on them)
    - **Results must not have body** (to see only documents without text)

You can personalize your experience and change the above defaults. When you select a new option in any of the above settings, a button will appear to **Save as Default**. Clicking to save will remember these settings for future searches.