# eSearch Pro User Guide

# Contents

Welcome!	3
Installation & System Requirements	4
Quick Start	5
Index	6
Indexing Options	7
Filetype filter	7
Search Queries	8
Scrollable Word List	9
Search-as-you-type	9
Intelligent indexing	9
Use List option	10
Stemming & Synonym options	11
Settings	12
Search Settings	12
Stemming	12
Synonym Files	13
Results Settings	14
Viewer Settings	15
PDF Options	16
Geolocation in images	17
Audio & Video files	18
Data & Source Code files	19
Index Settings	21
Stop words	21
Maximum word length	21
Maximum File Size	21
Case Sensitive	21
Keyboard Shortcuts	23
User Translation	24
View menu	24
Language File Editor	24
License	24
Registration	25
eSearch Pro	25

Evaluation	25
eSearch Lite	25
Support	26
Feedback	26
Known Issues	27
Issue 20 draft	28
Beta status The initial 1.0 (beta) has had limited testing and I file format support or limited performance with large files, lafiles in certain languages, etc	ırge indexes,

# Welcome!

We hope you enjoy using eSearch! eSearch is a full-text information retrieval application. Our aim is to provide a useful and modern user experience (UX) in YOUR language to streamline your access to digital information (ESI)!

#### • Scrollable word list

Quickly discover the vocabulary or spot typos in a document collection, gain quick insight to a case, project, or prior research. Click on a word for entry in the search bar to avoid typos!

- AND OR NOT buttons for quick & foolproof entry of Boolean operators. Typed operators are automatically uppercased!
- Legal professionals <u>Use List</u> of keyword search queries to share with others for Early Case Assessment (ECA), case reviews, etc.
- Sort search results by filetype, title, size, chronologically & automatically extracted meta data.
- Expand search with multilingual <u>stemming</u> and <u>synonyms</u> with predefined and user defined thesaurus files.
- Export timestamped search results & reports in XML, comma or tab separated formats, with selected metadata included.
- Viewer supports plain text with keyword highlighting, built-in <u>PDF</u> <u>viewer</u>, colour syntax highlighting and line numbers in <u>source code</u>, tables for <u>csv files</u>, images with <u>geolocation on map</u>, player for web format audio & videos files.
- User Interface (UI) in YOUR language. <u>User translatable.</u> French and German supplied. Free Language File Editor.
- Dark mode theme option for eye comfort. Choose Horizontal or Vertical Results | Viewer layout.

# Installation & System Requirements

eSearch is a self-contained 64-bit .NET 6.0 application, you can install on any Windows 10 or 11 machine without worrying about the .NET runtime version installed on that machine. It includes the .NET runtimes and all necessary libraries within the application itself to ensure it runs with the exact version it was tested with.

Download the esearch\_installer.zip file, this consists of a zipped Microsoft Software Installer (.msi) file. To extract the file right-click and choose Extract All... double-click on the .msi file and follow the instructions to install eSearch Pro.

Microsoft Software Installer (.msi) files are the preferred installer format for enterprise environments due to its standardization and best practices structure. All MSI packages support the same installation command lines to install, uninstall, update, or repair the application.

The default install path is under the Program Files folder. User settings are stored in the Users\Username\AppData\Roaming folder. (The Portable version stores user settings in the install folder).

Disk Space required: 2Gb minimum. Additional space is required for the indexes.

Unlike some consumer grade search products which simply index everything on the installed drive, eSearch not only allows you to select the folders and files to include in an index, you can also quickly select by all or part of a category of file types to include using a tree view control - see <u>Filetype Filter</u>.

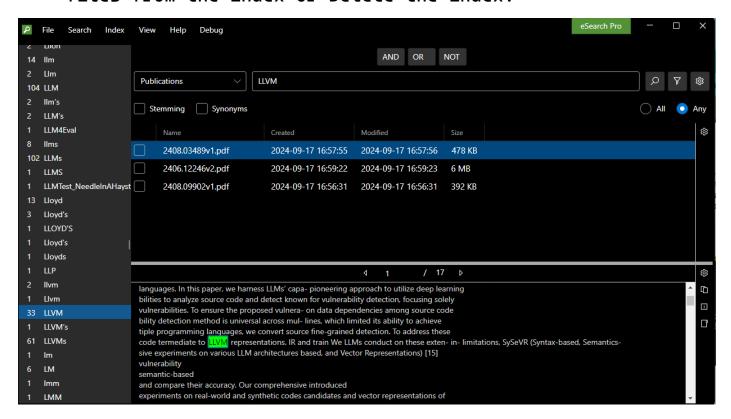
# Quick Start

- Go straight to the View menu, choose your language, theme and layout! Need to translate? See: <u>Translation</u>
- eSearch uses indexes to locate information fast!
  - Go to the Index menu. choose New...
  - Enter a meaningful name for the index
  - Add Files or Folders
  - For more information see: Index

You're now ready to search!

See: Search Queries

• If you later edit, add or remove documents, go to Index menu > Index Update... to Update, Rebuild, Rename, remove files from the index or Delete the index.



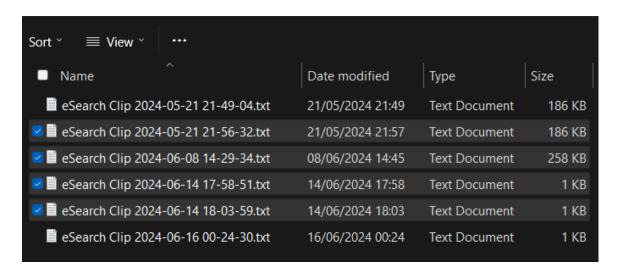
# Index

Enter a name for the index e.g. Closed-Cases, Recipes-Vegetarian, Projects-old.

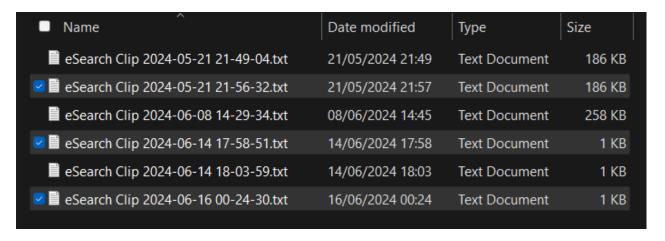
We recommend keeping each index small and specific. Keep archived or material that is not changed often in separate indexes from files that maybe edited frequently to make updating easier.

Choose Folder or File, you can add multiple files and multiple folders in the same index and apply <u>filetype filters</u> as needed. You should also check that <u>Settings...</u> are appropriate.

To select multiple files in a range, select the first file, hold down the Shift 1 key then select the last file.



To select multiple files, hold down the Ctrl key, select the files by clicking on them with the mouse or arrow keys and enter.



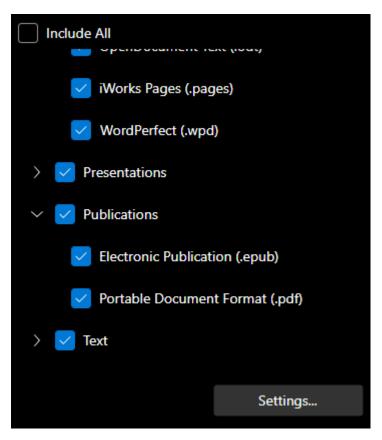
# Indexing Options

### Filetype filter

(Pro version. Lite version defaults to Include All)

When creating or updating an index you can select which file types to include. This is useful when indexing a Folder that may have nested folders containing filetypes that are not of interest and is quicker than having to select Files to index.

You can select all or parts of a category or individual filetypes from the tree view:



If Include All is selected it will only exclude a list of known executable extensions. It will attempt to parse every unrecognised filetype as well as all the filetypes in the tree view.

If you need to index a Folder with unknown filetypes, we recommend a trial index of a single File to check it will be handled, especially if the file(s) are large.

Click on <u>Settings...</u> for more options that may be needed.

# Search Queries

By \*default, eSearch Pro will list all the documents in the selected index without having to enter a search query. (\*You can change this from <a href="Search settings">Search settings</a>)

Enter a search query to limit the results list to just the files of interest. Search queries can just be a list of keywords like dog hound pooch (DON'T separate with commas or other characters).

Choose ALL if you want the words joined by AND, this will narrow your search the more words you add; remove words if you get no results. Word order does not matter.

Choosing ANY is a last resort for a broad search.

For more precise searching use Boolean AND OR NOT operators.

	Example Query	Will find files containing:			
Boolean	dog AND cat	both words dog and cat.			
	dog OR cat	dog or cat or both words.			
	dog NOT cat	dog and not cat.			
Proximity	"cat mouse"~10	mouse within 10 words of cat.			
Phrase	"dog kennel"	the exact phrase 'dog kennel'.			
Wildcards	dog*	<pre>dogma, doggy, dogfish, etc. (* replaces several</pre>			
(Do NOT use		characters)			
at the	moderni?e,	modernize or modernise			
start of a word.)		<pre>(? replaces a single character)</pre>			
Fuzzy	word~ 500~	similar words or numbers.			
Fields	Some files may contain fields you can search like Fieldname:				
	e.g. Content-Length:500~	File size of approximately 500 bytes.			
Regular Expressions	/[mh]otel/	either motel or hotel			

Note: Synonyms includes a Numeric Pattern file that has a regular expression for finding IPv4 addresses.

#### Scrollable Word List

#### Search-as-you-type

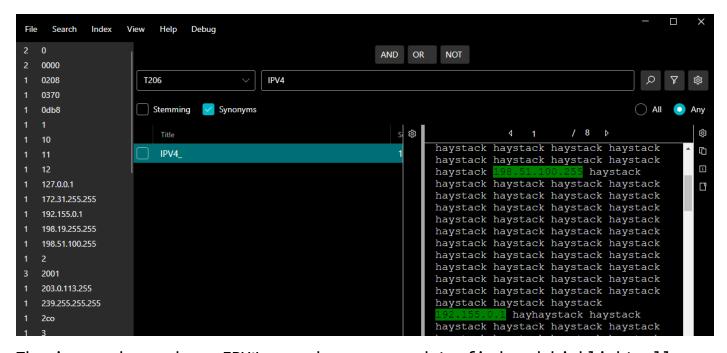
As you type a word in the search query bar, it will 'search as you type'. At the same time, the word list will scroll to the matching word. An alternative to typing, especially if you are unfamiliar with the content of the index, is to scroll through the word list and double-click on a word of interest. It will then be entered as a search guery.

#### Intelligent indexing

The word list keeps numbers and letters separated by punctuation together. For example, complete IP addresses and sums of money appear in the word list unbroken, as do possessive apostrophes and contractions.

Email addresses are split at the "@" symbol, so domains and email aliases appear separately in the word list. You can search for all users at contoso.com by entering that as the search query (do NOT use a wildcard). Alternatively, you can enter an alias like jon.doe to find all emails with that alias at any domain, such as jon.doe@gmail.com or jon.doe@contoso.com. To search for the exact email, use both the alias and domain in quotes with the "@" sign, e.g., "jon.doe@contoso.com".

eSearch Pro includes a useful regular expression in the pre-defined synonyms list for searching for IPv4 addresses. To use it, go to Search Settings and select the Sample Numeric Patterns, select Synonyms and just enter "IPV4" as the search query.

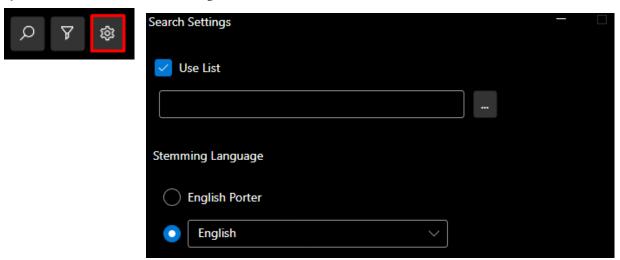


The image above shows IPV4 search query used to find and highlight all IPv4 addresses.

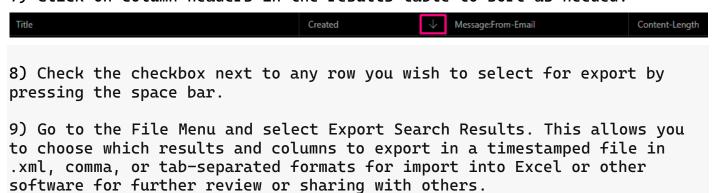
### Use List option

Intended for using a 'Keyword List' during eDiscovery (eDisclosure) by case reviewers. To create a list the first step is to sample the dataset.

- 1) Create an index of a sample of the data and use the fully scrollable word list to quickly identify the vocabulary, jargon & abbreviations used.
- 2) Create a file in NotePad or similar text editor and save it in Unicode format. The file should contain a list of keywords (words that you identified as important from the word list) or valid search queries, one per line. Leave no blank lines. Save it in your Documents folder or with your case files, we recommend the filename should have a timestamp and suitable file name e.g. Case479\_2024-04-01.
- 3) Open the Search Settings window and check the Use List checkbox.



- 4) Click on the inavigation button and open the file you created above. Click OK. The search query will appear in the search bar.
- You can expand the search using <u>stemming or synonyms</u> to discover additional responsive documents.
- 6) Choose suitable columns from Results Settings:
- 7) Click on column headers in the results table to sort as needed.



# Stemming & Synonym options



Check either of these to expand a search.

### Stemming

Stemming will expand a search term such as "apply" using different word endings such as applied, applies, applying.

You need to select the correct language from <a>Search Settings</a>.



There is a choice of English Porter Stemming or stemming in English or 30 other languages. A vocabulary file for evaluating the Porter Stemmer is available from here: https://tartarus.org/martin/PorterStemmer/

Note: if using the Case sensitive index option, Stemming should be deselected. If Stemming in eSearch Pro is selected it will expand a search by making it case-insensitive as well by using different word endings.

#### Synonyms

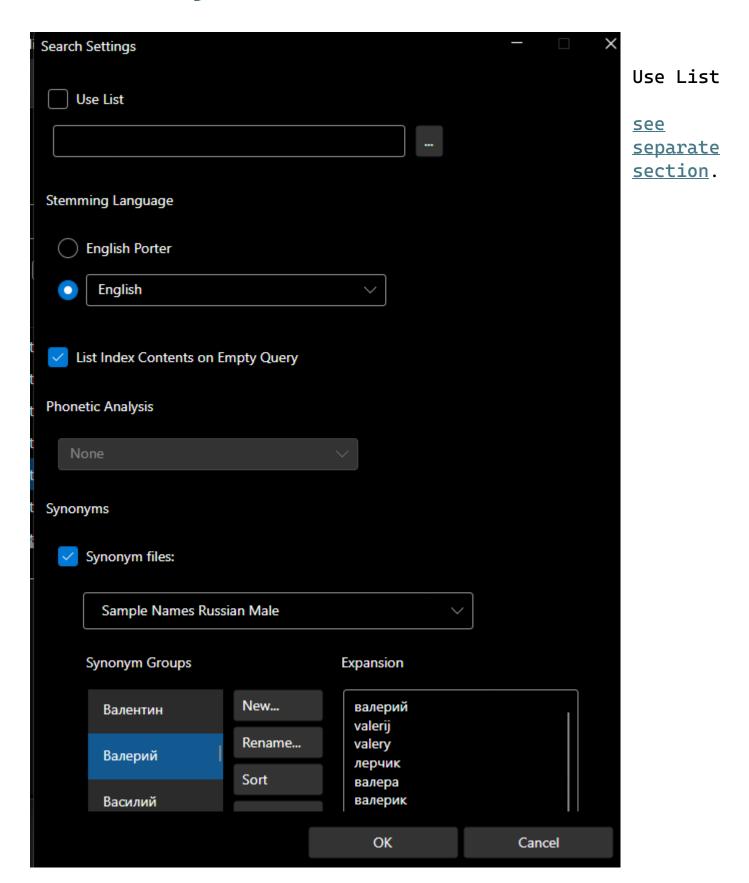
Synonyms will expand a search with user-defined or the supplied predefined synonym files in several languages.

You need to select one or more of either type from Search Settings.



# Settings

# Search Settings



# Stemming

Ensure that Stemming Language matches the language of your search queries.

### Synonym Files

Pre-defined synonym files are supplied in multiple languages. You can also select 1, 2, or 3 from the Synonym Files list to add your own user-defined synonym groups.

You can add not just synonyms, but also related or equivalent words or phrases in other languages. Regular expressions are also supported; the sample numeric patterns file contains a regular expression to find IPv4 addresses.

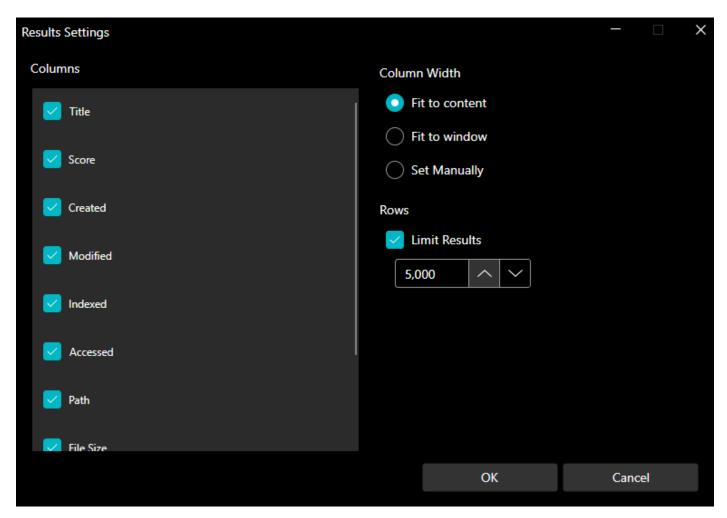
You can select multiple synonym files. For example, if you are searching through a mix of English and Spanish texts, you might want to select the English irregular verbs and nouns, as well as the Spanish irregular nouns and verbs. Additionally, you can add your own subject-specific synonym group.

Danish irregular nouns Sample months Danish irregular verbs Sample names cross lingual Dutch irregular nouns Sample names genealogy Dutch irregular verbs Sample names political English irregular nouns Sample names Russian female English irregular verbs Sample names Russian male French irregular verbs Sample numbers Italian irregular verbs Sample numeric patterns Norwegian irregular nouns Sample Prenoms Français Masculins Norwegian irregular verbs Sample trade Sample Unregelmäßige Deutsche Verben Sample colours Sample currencies Sample Verbe Irrégulier Français Sample days Sample Wochentage Deutsch Spanish irregular verbs Sample fashion Sample geographic Swedish irregular nouns Sample Jours de la semaine français Swedish irregular verbs Sample legal Sample medical

Note: the sample numeric patterns file contains regular expressions for IP addresses and credit-cards. The IPv6 example will not currently work with eSearch Pro indexes because the colon (:) character is treated as a space with no option to change it.

The Amex, Visa and Diners card regular expressions may work; however, they are not foolproof. They might detect patterns that are not credit card numbers or miss valid card numbers, as there is no checksum validation.

# Results Settings



The Columns list includes not only the standard columns shown above but also metadata automatically collected from the files in your current index. For example, if emails are included, it may show 'To', 'From', and 'Subject'. If images are included, it may show geolocation, DateTime, and camera details of a photo.

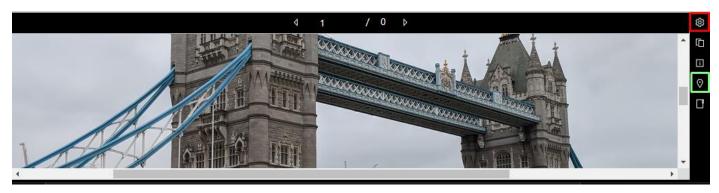
You can sort your results by any column by clicking on the column header in the results list. You can also click and drag column headers to reorder the columns.

Column width is remembered if you choose 'Set Manually'."

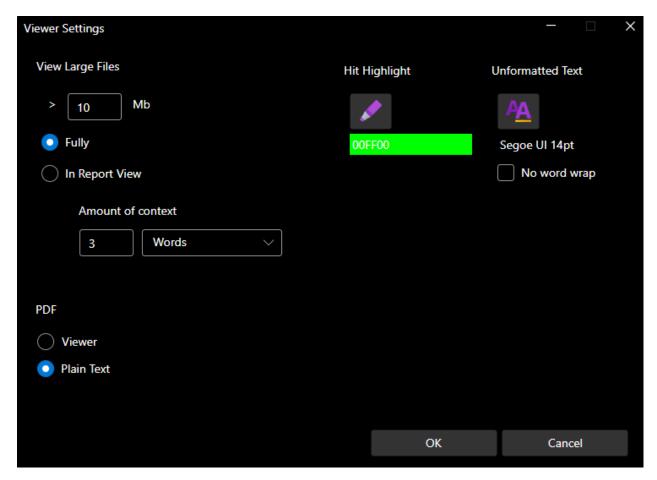
Rows. You can set the number of rows to display in the results table. 5000 is the maximum, but you may prefer to set to a lower number for a quicker response. (Pro version. Lite version is fixed at 10 rows.)

# Viewer Settings

Click on the settings button (top) in the taskbar next to the viewer window to open Viewer Settings.



The map pin button (outlined in green in the image above) is a Pro version feature, it is displayed when images with geolocation metadata are viewed.



View Large File: Large files may be very slow in showing in the viewer, choose report view to quickly show just keywords in context (KWIC).

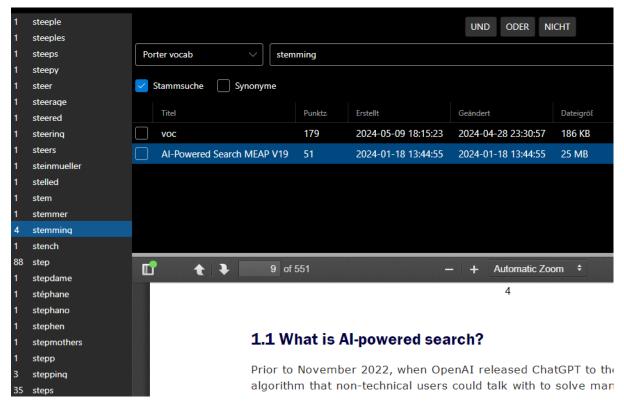
You can select the colour of the highlighted words or set the size and font of the text in the Plain Text view.



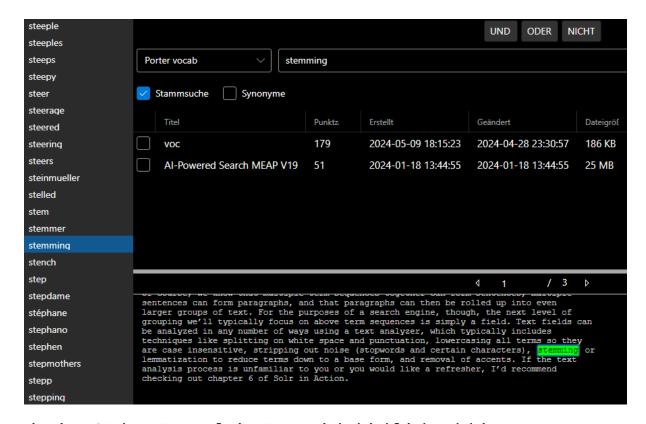
PDF: See <a href="PDF">PDF Options</a>.

### PDF Options

You can view as text with hits highlighted in your choice of colour or use the built-in PDF viewer as shown below.



Built in PDF Viewer.

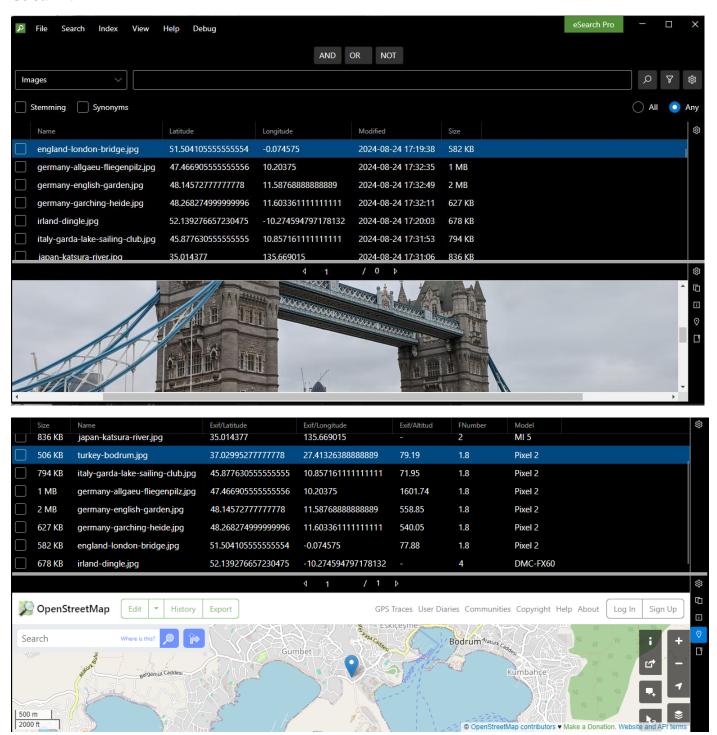


Viewing 25Mb PDF as Plain Text with highlighted hits.

Illustrations show sample German UI translation.

### Geolocation in images

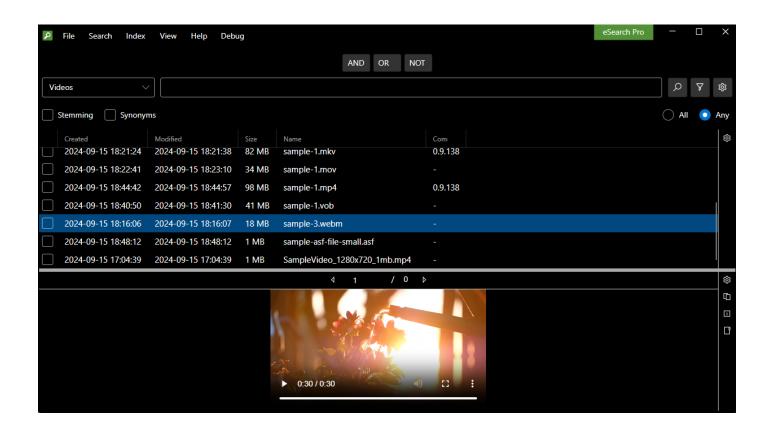
From <u>Results Settings</u> select the metadata you want to display. Images in Web formats are displayed in the Viewer. Automatically extracted metadata includes full geolocation, camera details, etc. You can sort on any column.



Click on the Map Pin button to view the location on a map. (Pro only) From the File menu, choose Export Search Results. Select metadata, then choose .csv for output of a timestamped file to Excel.

Images above use geolocation test photos from <a href="https://www.geoimgr.com/sample-photos.html">https://www.geoimgr.com/sample-photos.html</a>

#### Audio & Video files



Web format audio and video container files, such as (.wav, .mp3, .mp4, .ogg, .webm), contain different audio and video codecs, not all of which may play in the viewer. For files that don't play in the viewer, we recommend installing VLC Media Player and clicking on 'View in folder'.

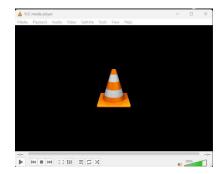


Ensure that VLC Media Player is set as the default for opening the filetypes in your indexes.

VLC is a free and open-source cross-platform multimedia player, it can be downloaded from <a href="https://www.videolan.org/vlc/">https://www.videolan.org/vlc/</a>

If your collection contains little meta data, use the default search setting to return all files in the index with an empty search query.

The results are limited to 5000. For larger collections split your files up so that different formats, subject areas, projects, dates, etc. are in different folders. Create a separate index for each folder.

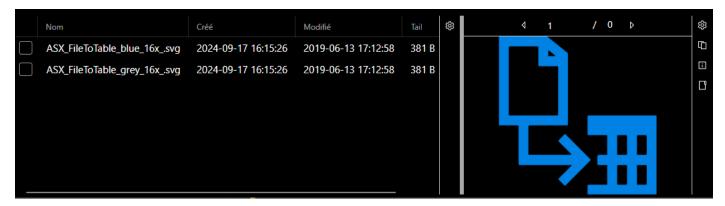


VLC Player can also be used for static image files.

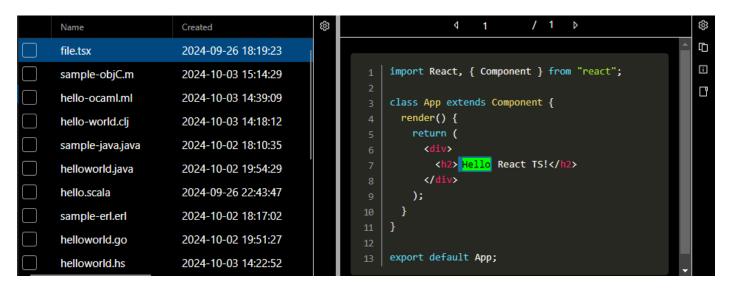
#### Data & Source Code files

2017-12-19 1	0:01:52 96 k	KB bahamas_lea	ks.nodes.interme	diary.csv					
2017-12-19 1	0:01:52 111	KB bahamas_lea	ks.nodes.address	.csv					
2024-09-06 1	7:05:26 140	KB eSearch.dep	s.json						
				<b>√</b> 1	/ 0	<b>&gt;</b>			
"labels(n)"	"valid_until"	"country_codes"	"countries"	"node_id"	"sourceID"	"address"	"name"	"jurisdiction_description"	"sen
" [""Address""]"	"The Bahamas Leaks data is current through early 2016."	"ВНЅ"	"Bahamas"	"2400001"	"Bahamas Leaks"	"ANNEX FREDERICK & SHIRLEY STS	P.O. BOX N-4805	NASSAU	ВАН
	"The								

.csv files are handled as text, with table layout preserved.

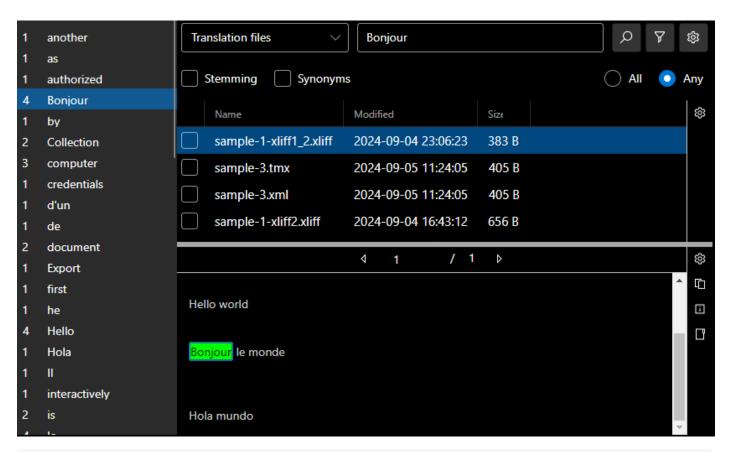


Small images may be better viewed in horizontal mode to increase size.



Hit highlighting in source code with syntax highlighting & line numbers.

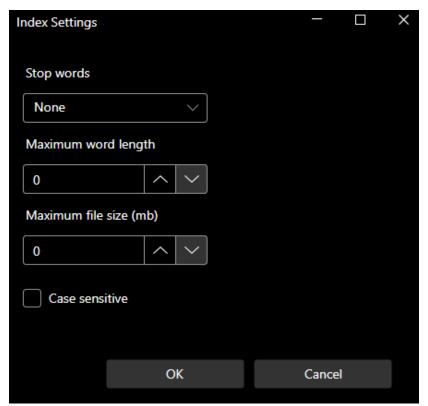
### Translation file support



Translation industry file formats XLIFF 1.2, XLIFF 2.0, .tbx, and .tmx are supported for small files only. The image shows an index with the Case Sensitive option enabled, ensuring that acronyms and abbreviations are found. Additionally, contractions like d'un are preserved in the index and appear in the word list.

eSearch Pro 1.0 currently lacks file segmentation functionality, which splits large data files, such as .csv and translation industry XML formats, by rows or nodes, allowing them to be indexed as separate files.

# Index Settings



Settings will not take effect until you Rebuild an index.

#### Stop words

Stop words (also known as noise words) are frequently used words that are not included in an index to save indexing time and storage space.

Choose from a list of stop words in 30 languages or you can choose 'None' (at top of the list). If you change the stop words used by an index, you will need to rebuild the index for the change to take effect. You can edit the stop words from the ..AppData\Roaming\eSearch\Stop folder.

# Maximum word length

The default setting is 0, which means no limit. (The actual limit is currently 125). You may prefer to lower this if your documents contain many long strings of letters or numbers.

#### Maximum File Size

The default setting is 0, which means no limit.

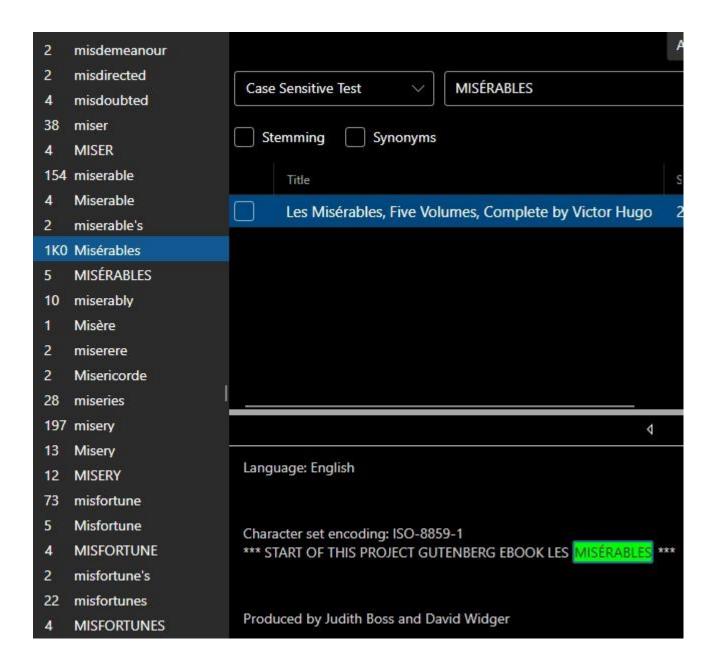
#### Case Sensitive

If Case sensitive is selected it will index language, LANGUAGE and Language as three separate words; it is useful for more precise searching, for example distinguishing wax polish from Polish (a person from Poland),

chapter titles, brand names (Apple NOT apple) or acronyms and abbreviations (PIN NOT pin). Indexes will be larger with this option.

Note: Stemming should be disabled for case-sensitive searches.

In eSearch, stemming will broaden the search by making it case-insensitive and expanding it to include words with different endings. (e.g. a search on 'title' will find TITLE, Title, title, titles).



With the 'Case sensitive' option selected the word list shows alphabetically sorted uppercase words and letters as well as accented words. eSearch also correctly retains possessive apostrophes ('s) and contractions like d'un in the word list.

# **Keyboard Shortcuts**

Menus - use conventional ALT + key shortcuts, press the ALT key to see the underlined character to use as the selection key. The English UI uses the first letter as the key by default, but other languages may differ to avoid duplicate keys. Users may also create their own English language file alternative to the built in English (GB) defaults (e.g. US, AU, CA or even GB if you prefer your alternative words or shortcuts).

Online Help - F1 key.

Results grid -

Up/Down keys for navigating rows.

Space bar to select/unselect a row (check/uncheck the checkbox)

Shift + Up/Down keys for choosing multiple rows + space bar to check/uncheck (select) rows in bulk.

**UI** Navigation

Tab or Shift + Tab keys.

Word List

Up/Down keys + Enter to enter the selected word as a search query

# **User Translation**

#### View menu

The User Interface (UI) Language is English by default, choose from the supplied languages by selecting From File...

The UI language changes immediately, no need to restart!

### Language File Editor

Our aim is to make eSearch available in many languages, we provide sample translation files in several languages. You can edit these, not only to add missing translations or correct errors, you can also change shortcut keys or change user interface wording to suit your own way of working, we provide a free Language File Editor to make this a quick and error free experience!

Download the free Language File Editor from <a href="https://searchcloudone.com/universal-desktop-search/">https://searchcloudone.com/universal-desktop-search/</a>

#### License

The eSearch license allows you to translate the user interface for your own personal or organisations use, you may not sell or license the translated files to others. If you would like to donate translations for use by others, please send them zipped to support@electronart.co.uk We will publish or distribute them free of charge on as AS-IS basis, with no warranty of accuracy or completeness.

# Registration

#### eSearch Pro

eSearch Pro is available in Desktop or Portable versions.

See <u>SearchCloudOne.com</u> for more information and pricing.

#### **Evaluation**

eSearch Pro may be downloaded from SearchCloudOne.com for a 30-day trial period.

After the trial period it will revert to free Lite mode with some limited features.

If you have purchased a Pro license, enter the serial number that we provide to remove any limits and to be entitled to support and updates\*.

Enter the serial at: Help menu > About > Serial...

#### eSearch Lite

eSearch Lite is free for personal and commercial use. For support and updates please register your email address with support@electronart.co.uk. The Lite version has limited functionality, for example limits to the number of results or features that are in the Pro version. These limits may change over time and will be published on our website or user guides.

\*Serials are not available while product is in beta test phase. After the general release date, beta testers, early evaluators and users of similar products will be offered discounted pricing. Lite version will remain free for private and commercial use.

# Support

Support is given by email to a single point of contact within an organisation. Always supply your serial number or when and where purchased if not obtained directly from SearchCloudOne.com or ElectronArt Design Ltd.

Use the contact form at SearchCloudOne.com or email support@electronart.co.uk

# Feedback

We welcome feedback, suggestions, translations at any time!
Use the contact form at SearchCloudOne.com or email support@electronart.co.uk

#### Known Issues

As of 20 Sep 2024: Pro and Lite version

### Viewer Settings

#### View Large Files:

- a) All files: if Report view is selected no keywords in context are displayed. Only metadata is displayed.
- b) Large files such as 160Mb Reuters headlines (TREC data) if "Fully" is selected eSearch will become non-responsive until file is read in (several seconds).

#### **PDF**

a) Viewer - no hit highlighting

#### Word List

g, j, y, p, q descenders are clipped.

Words longer than 23 characters aren't fully visible.

# Phonetic Analysis

Nonfunctional.

# Query Filters

Nonfunctional.

#### Viewer

Context menu - no text. Limited media support.

# File support

File parsers are in continuous development, if you experience issues, download a later build to check if the issue has been fixed. Contact <a href="mailto:support@electronart.co.uk">support@electronart.co.uk</a> if you have a particular need or that an issue has not been fixed.

# Accessibility

Generally keyboard accessible. All controls respond to Windows Narrator Screen Reader (turn on/off with keyboard shortcut Windows + Ctrl + Enter)

Keyboard shortcuts and screen reader prompts not finalised.

### Issue 20 draft

# eSearch Pro

#### Beta status

The initial 1.0 (beta) has had limited testing and may have limited file format support or limited performance with large files, large indexes, files in certain languages, etc.

Copyright © 2024 ElectronArt Design Ltd. All Rights Reserved.