E101 eSearch Pro AI-Mode

P	File	Search	Index	View Plugins Help Debug	
Hov	w are			AND	OR NOT
Ηοι	v can				
Hov	v do			Personal Assistant V Write a hello world function in 5 different lan	guages
Hov	v does			, (
Ηοι	v is			🖌 Stemming 📄 Synonyms 📄 Soundex	
Ηοι	v long				
Ηοι	w many			Here are "Hello, World!" functions in five different program	nming languages:
Ηοι	w tall			1 Python	
Ηοι	w wide				
Wh	at are			<pre>print("Hello, World!")</pre>	
Wh	at can				
Wh	at do			2. Java	
Wh	at does				
Wh	at is			public class Main {	
Wh	en are			<pre>public static void main(String[] args) {</pre>	
Wh	en can			<pre>System.out.println("Hello, World!");</pre>	
Wh	en do			}	

Issue: 29 March 2025 *Name subject to change before general release

Contents

Introduction	3
Perplexity	3
Who is an owner?	3
What are the steps to open an account?	3
How to allocate a key to a project	3
Rename	4
Add New Perplexity Models	5
OpenAI	6
Custom	7
Perplexity	7
Anthropic	7
Google	7
Intern	7
Mistral	7
Novita	7
X	7
Other Online AI LLM API's	8
System Prompts	9
User Prompts	9
Perplexity Sonar	9
Sonar Reasoning	
Sonar Deep Research	
Large Language Models (LLMs)	
The future	

Introduction

eSearch Pro has an online AI-Mode. When first installed, if you click on the AI-Mode button (Blue button in the image below), there will be nothing listed in the drop-down list to the left of the search bar. To use AI-Mode, you need to connect to the Perplexity API service or a service that offers comparable functionality.



Perplexity

eSearch Pro features several of Perplexity's Sonar language models already set up for you. The Sonar models provide real-time access to the internet and up-to-date information, making them valuable tools for businesses.

Before you can use the Perplexity AI models, the 'owner' needs to open a Perplexity account to obtain one or more API keys. It is recommended to allocate API keys to projects or users to control usage and costs.

Who is an owner?

This refers to the person or entity (e.g. IT department) that controls and manages the account.

What are the steps to open an account?

Visit <u>https://www.perplexity.ai</u>, click on the sign-up button. Choose a sign-up method and enter your details.

How to allocate a key to a project

\leftarrow	С	ର (Ô	https://	/www.perplexity.ai/settings/api				
棥	per	olexit	y		Settings	Account	Profile	Personalize	API

Navigate to Settings (Bottom left, cog icon), click on the API tab, copy the existing API key or generate a new one. Before using the API key, you need to setup a payment method and add credits to your account (\$5 min.).

In eSearch Pro, for every user that will be using AI on a project, navigate to the Search > AI Search menu to open the 'Connect to API' window.

- 1. Click on + to create a new connection.
- 2. Select Perplexity from the Services drop-down control.
- 3. Select a model that will be used on the project.
- 4. Enter the API key you have allocated for the project.
- 5. Click the Test button, wait for the 'success' message to activate the connection.
- 6. Close the form

You can use the same API key for all models or generate a new API key for each model. Activate the models for each user in turn and <u>rename</u> as required.

Rename

Rename each connection either with the model's name, project name, or purpose (e.g., Personal Assistant, Researcher). These names will also be listed in the dropdown control alongside the search bar.



The model's name is displayed in the viewer:



Add New Perplexity Models

To add a model that is not currently listed in the Model drop-down, click on the + button (shown highlighted in red below).

Connect to API			-	-		
Connections	Service					
gpt-4o-mini	Custom				\sim	
Grok 2	Server URL					
Mistral Medium						
Personal Assistant	API Key					
QwQ-32B	Model					
Sonar Deep Research						
Sonar Pro	Custom System Prompt					
Sonar Reasoning	System					
Sonar Reasoning Pro	Answer in English. Be pro	ecise and concise				
+ – EĎ	L					
	Test	Cancel				
				Close		

1.Enter

https://api.perplexity.ai in
the Server URL textbox.

2.Enter the model's name exactly as listed here: <u>Supported Models</u>

3.Enter the API key.

4.Click on Test to activate the connection.

5.<u>Rename</u> the connection as required, it does not have to match the model's name.

Grok 2	Server URL
Mistral Medium	https://api.perplexity.ai
Personal Assistant	API Key
0.000	
QwQ-32B	Model
Sonar Deep Research	sonar-deep-research
Sonar Pro	Custom System Prompt
Sonar Reasoning	
Sonar Reasoning Pro	
+ - ײ	

6.Edit the system prompt later if required.

OpenAI

The OpenAI service connects to the URL: <u>https://api.openai.com/v1</u> Setup an account at <u>https://platform.openai.com/</u>

You can allocate models and API keys to specific projects:



eSearch Pro has been tested with these models:

Model	Notes				
chatgpt-4o-latest	0ct	23	text model		
gpt-4o-mini	0ct	23	cutoff		
gpt-4-turbo	Dec	23	cutoff		

The above GPT (Generative Pre-trained Transformer) models, do not use live web search, the knowledge cutoff date is shown in the Notes column.

Custom

You can set up as many connections as you wish. Some services offer free personal accounts to evaluate models. For production or commercial use most services operate on a pay-as-you-go basis.

Perplexity: Note, it is NOT necessary to subscribe to a Pro account, all models are pay-as-you-go, with tiered limits dependant on cumulative payment.



Support

If you need assistance with setting up any model listed, please contact support@electoart.co.uk

Other Online AI LLM API's

Venice.ai <u>https://api.venice.ai/api/v1</u> Needs a Pro account.

Cohere <u>https://api.cohere.com/v1/chat</u>

AWS Bedrock https://docs.aws.amazon.com/general/latest/gr/bedrock.html

Note that Meta open-source Llama models are available via these partners: https://www.llama.com/docs/getting-the-models/405b-partners/

System Prompts

You can use the system prompt to provide instructions related to style, tone, and language of the response.

For the Perplexity service see: https://docs.perplexity.ai/guides/prompt-guide

User Prompts

When you start to type a query beginning with the first letter or two from 'how', 'what', 'who', or 'when', the word list will scroll to some common query phrases. You can then double-click on a query stem to complete the query. It isn't necessary to use the listed query stem phrases. You can also use voice input for eSearch Pro version 1.0.9097 onwards.

Perplexity Sonar

Examples

1. You can ask it to write code or a poem:

File Search Inde	x View Plugins Help Debug	
How are	AND OR NOT	Г
How can		
How do	Personal Assistant V Write a hello world function in 5 different languages	
How does		
How is	Stemming Synonyms Soundex	
How long		
How many	Here are "Hello, World!" functions in five different programming lang	uages:
How tall	1. Python	
How wide		
What are	<pre>print("Hello, World!")</pre>	
What can		
What do	2. Java	
What does		
What is	public class Main {	
When are	<pre>public static void main(String[] args) {</pre>	
When can	System.out.println("Hello, World!");	
When do		
	AND OR NOT	
Personal Assistant V	write a rhyming poem similar to 'jack AND jill went up the hill', but using the names Bill AND Jane $ imes$ $oldsymbol{arsigma}$	\$ \$
Stemming Synonyms	Soundex All Any	• 🕹
		\$
Here's a rhyming poem sim	ilar to "Jack and Jill went up the hill," but using the names Bill and Jane:	G
Bill and Jane went up the stand.	a lane, To fetch a basket of grain. Bill fell down and hurt his hand, And Jane came running t	to
Up Bill got, and home did gentle twist.	stride, As fast as he could glide. He went to rest and bound his wrist, With ice and a	

The results may be useful but are clearly not perfect. Reasoning models are more useful in business or educational settings, since the user can see the reasoning behind the feedback and judge if it is appropriate or learn from the process. 2. Sonar non-reasoning models can also answer questions

Personal Assistant V How many R's in raspberry	
Stemming Synonyms Soundex	
The word "raspberry" contains three instances of the letter	\checkmark
Personal Assistant V How many C's in courgette	
Stemming Synonyms Soundex	
The word "courgette" contains two C's.	×

But be aware, non-reasoning models can get facts wrong! Note: Zucchini (US) vs. Courgette (UK)



The example above shows it interpreted "G's" as glycaemic glucose (GI).

These types of error can sometimes be avoided by providing more explanation of your intent. For instance: "How many letter g's are in the word aubergine?" would be more specific. Note: The acronym for glycaemic glucose is GI, which stands for Glycaemic Index, a measure of how quickly carbohydrates in food raise blood glucose levels. The American English spelling of Glycaemic (British English) is Glycemic.

Sonar Reasoning

Seeing the reasoning helps users learn from the process. Here is an example of the same poem task:

Sonar Reasoning Pro 🗸	write a rhyming poem similar to 'jack AND jill went up the hill' but using the names Bill A	ND Jane	×	Q	Y	
Stemming Synonyms	Soundex			Any	ī	۲
Up Bill jumped and brush And shared the berries t	ed his pants, Then treed Jane trom the thorns. They Limped back home hey'd borne.	, their spir	its h	iigh,		^ ‡
That keeps the syllable makes sense. The scenari maintains the lightheart	count more consistent. Rhyme scheme remains AABB. Now, check for flo o is picking berries, they both have minor accidents, help each othe red, slightly chaotic feel of the original.	w and whethe r, and retur	r the n hom	story e. It	y	Ē
I should also ensure the "thorns" are appropriate	e language is simple and rhythmic, suitable for a nursery rhyme. Word a. The ending mentions sharing the berries, which adds a positive res	s like "bram olution.	bles"	and		
Another consideration is up," "fell down," "came "limped," and "shared" k	s using active verbs and a clear sequence of events. The original use tumbling," so mirroring that with "slipped," "scraped," "stuck," "br meeps the action moving.	s actions li ushed," "fre	ke "w ed,"	ent		
I think this captures th Let me put it all togeth	e essence of "Jack and Jill" while using the names Bill and Jane and er and verify the rhyme and rhythm once more.	a different	scen	ario.		
Bill and Jane went down To pick a basket of berr Bill slipped fast and sc And Jane got stuck in th	the lane ties. craped his calf, re brambles.					
Up Bill jumped and brush Then freed Jane from the They limped back home, t And shared the berries t	ed his pants, thorns. heir spirits high, hey'd borne.					
This poem mirrors the AA water-fetching with a be (stressed/unstressed syl rhymes[9]. The resolutio	BB rhyme scheme and rhythmic cadence of "Jack and Jill"[1][9], subst rry-picking mishap. The structure features four-line stanzas with tr lable pairs), maintaining the playful tone and minor calamities typi n emphasizes teamwork and lighthearted resilience, much like Jack's	ituting the ochaic meter cal of nurse vinegar-and-	hillt ry paper	op remed	dy	

Sonar Deep Research

User prompts can include specific website URLs. The length of the query is not restricted by the search bar; the query text will scroll left as you continue typing.

How are		AND OR NOT				
How can						
How do	Sonar Deep Research V mers l	looking to fine-tune production LLMs. They must already have SOTA models like	GPT-4o deployed	x p	\mathbb{Y}	
How does						
How is	Stemming Synonyms S	Soundex		Any	i	3
How long					1	<u>^</u> @:
How many	Okay, let's tackle this query s AI Expo exhibitor list that al:	step by step. The user wants me to identify the top 15 compar ign with OpenPipe's ideal customer profile (ICP). OpenPipe's	iles from the Ge ICP is enterpri	nerative se		
How tall	customers who are already using	g state-of-the-art (SOTA) models like GPT-40 in production at	scale and are	looking	to	
How wide	improve accuracy, reduce Laten expanded, secured funding, or s	cy, and lower costs. Additionally, the focus should be on fin shown interest in fine-tuning LLMs.	ms that have re	cently		I

eSearch Pro connects using streaming, so you don't have to wait for the complete answer. It starts providing feedback as soon as it begins analysing the query.

Large Language Models (LLMs).

AI is developing at a rapid pace. Early non-reasoning models provided fast answers but were prone to errors. They are being superseded with reasoning models and deep research models, which have a slower response but are more suited to business use. When combined with real-time web search, these models can overcome the knowledge cutoff date and provide links to sources, allowing users to verify information or conduct further research.

The future

Our aim is to incorporate the latest advances in LLMs to assist with information retrieval. Currently, Perplexity and certain other models enhance LLM knowledge with real-time web search.

In non-AI Mode, we index local files with a plug-in architecture that allows connection to external data sources like databases and third-party storage such as Dropbox, OneDrive, or Google Drive.

Retrieval Augmented Generation (RAG) combines LLM knowledge with data from your private local or externally hosted sources. This, combined with reasoning or deep research models, will provide a powerful tool for retrieving and analysing a wealth of information that was not possible with older technology.

With these tools, you can make smarter data driven decisions and quickly enhance your knowledge in unfamiliar topics.

AI agents take this a step further; they are designed to autonomously perform tasks, make decisions, and interact with their environment.

AI agents integrate various AI forms and remain goal-oriented, taking targeted actions to achieve specific objectives. Think of it as having a smart virtual assistant or office system that can, for example, locate alternative suppliers, get quotes, and perhaps even produce an order ready for human approval, or book travel and accommodation for a business trip given your availability.

We will be pleased to quote for custom plugins to connect to your secure data sources, including automating tasks using agentic AI.

If you have specific information retrieval requirements, now or in the future, please contact our support at: support@electronart.co.uk, we may have it in the pipeline, or we'll put it on our roadmap!

-00000-