

**E101**  
**eSearch Pro\***  
**AI-Mode**

**DRAFT**

Issue: 18 March 2025

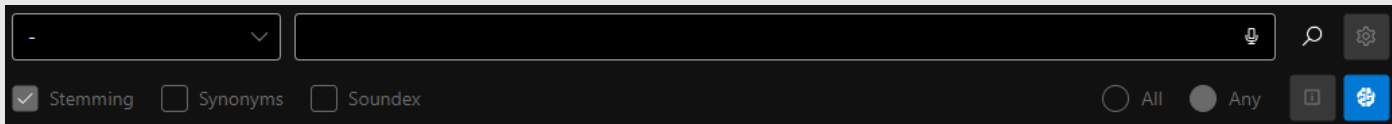
\*Name subject to change before general release

## 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
Other Online AI LLM API's .....	8
System Prompts.....	9
User Prompts.....	9
Perplexity Sonar.....	9
Sonar Reasoning.....	11
Sonar Deep Research.....	11
Large Language Models (LLMs).....	12
The future.....	12

# Introduction

eSearch Pro has an online AI-Mode. When first installed, if you click on the AI-Mode button (in the red box 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 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.

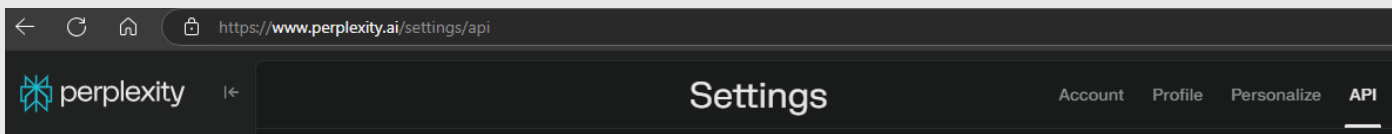
### Who is an owner?

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

### How to allocate a key to a project



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.

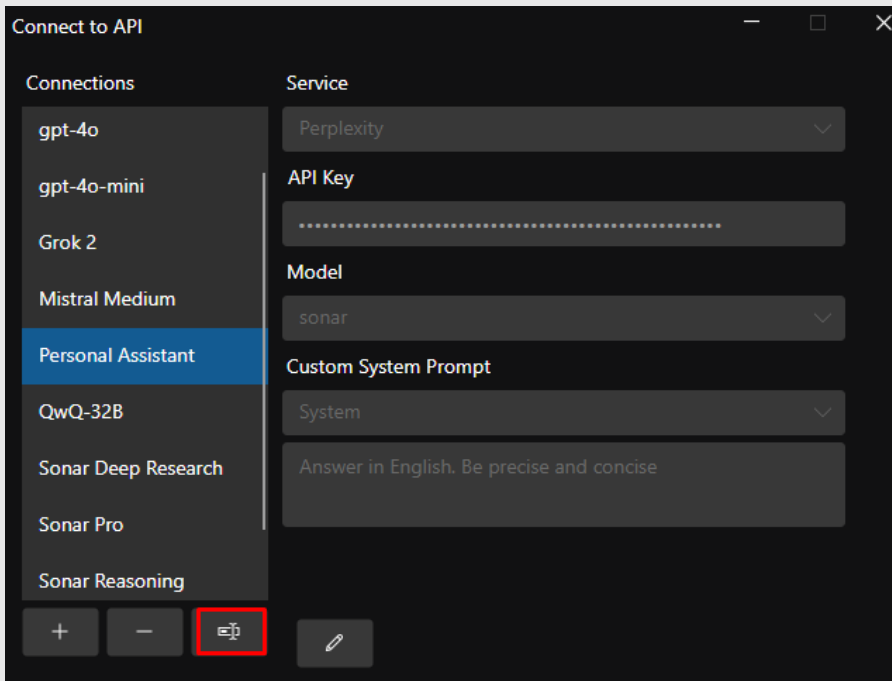
1. Click on + to create a new connection.
2. Select Perplexity from the Services drop-down control.
3. Select a model that will be used on the project.
4. Enter the API key you have allocated for the project.
5. Click the Test button, wait for the 'success' message to activate the connection.
6. Close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

You can use the same API key for all models or generate a new API key for

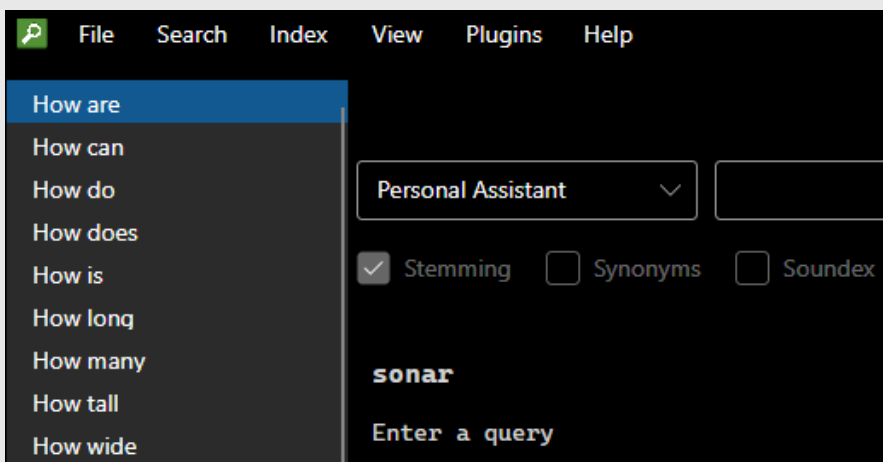
each model. Activate the models for each user in turn and [rename](#) as required.

## 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.

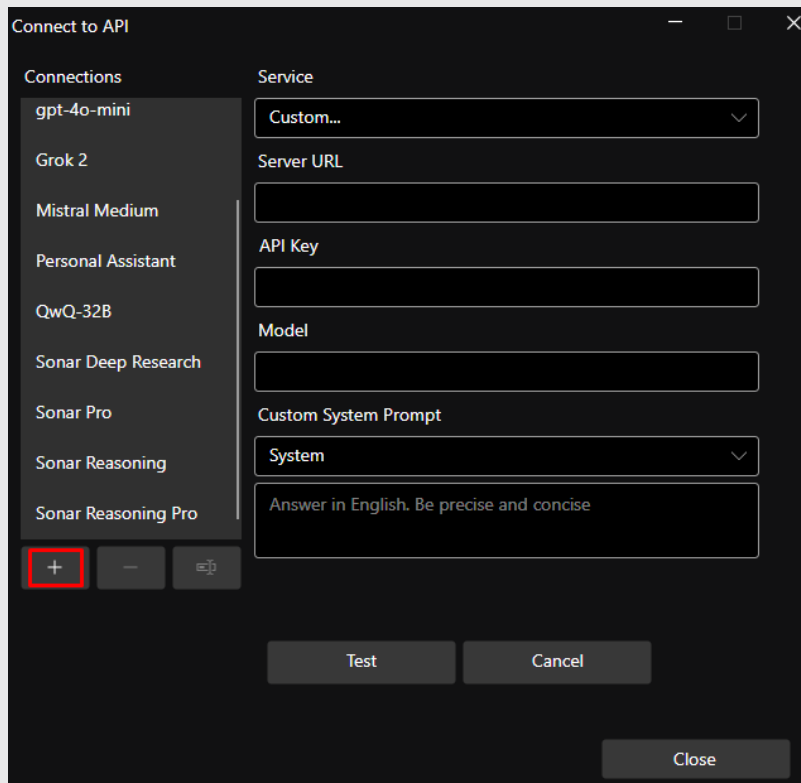


The model's name is displayed in the viewer:



## 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).



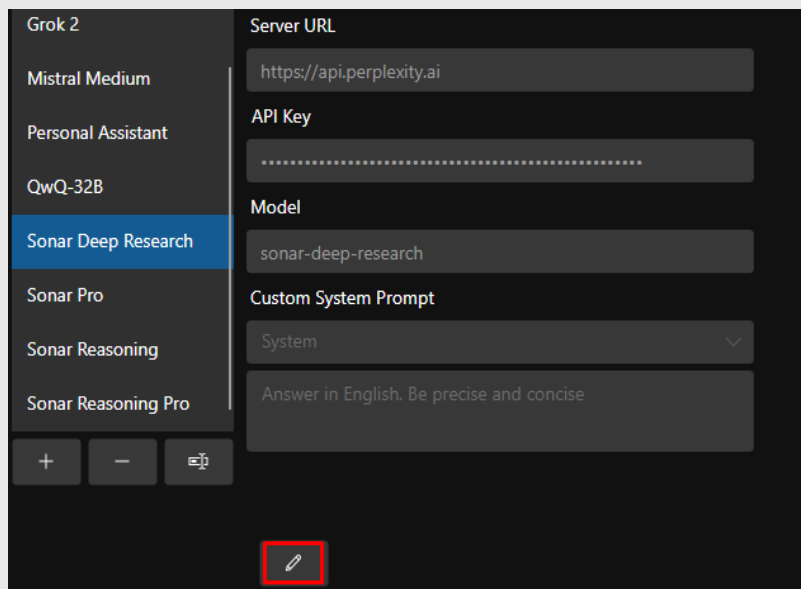
1. Enter <https://api.perplexity.ai> in the Server URL textbox.

2. Enter the model's name exactly as listed here: [Supported Models](#)

3. Enter the API key.

4. Click on Test to activate the connection.

5. [Rename](#) the connection as required, it does not have to match the model's name.



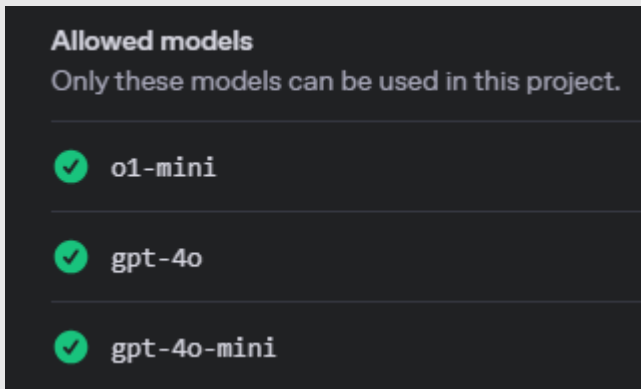
6. Edit the system prompt later if required.

## OpenAI

The OpenAI service connects to the URL: <https://api.openai.com/v1>

Setup an account at <https://platform.openai.com/>

You can allocate models and API keys to specific projects:



eSearch Pro has been tested with these models:

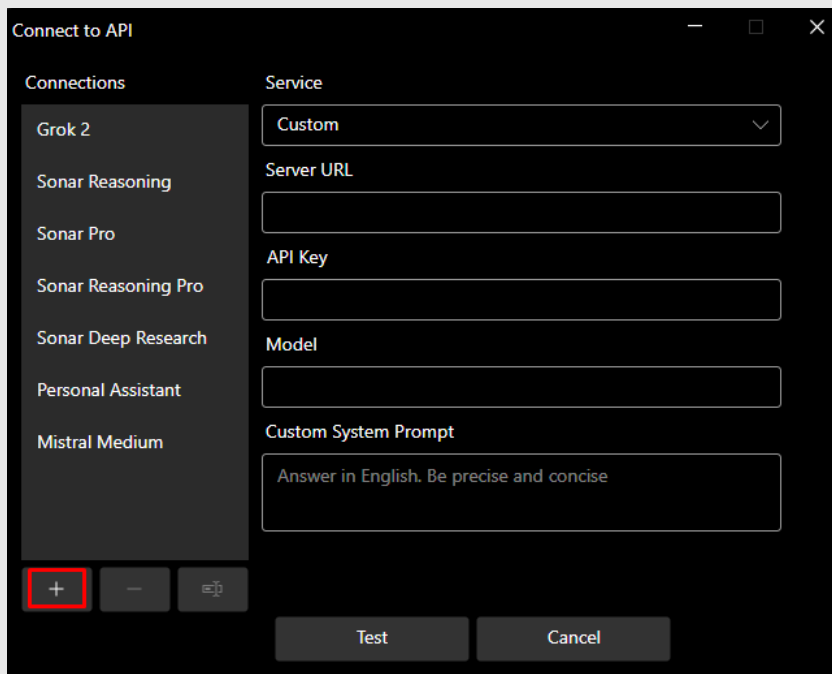
Model	Notes
chatgpt-4o-latest	Oct 23 text model
gpt-4o-mini	Oct 23 cutoff
gpt-4-turbo	Dec 23 cutoff

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

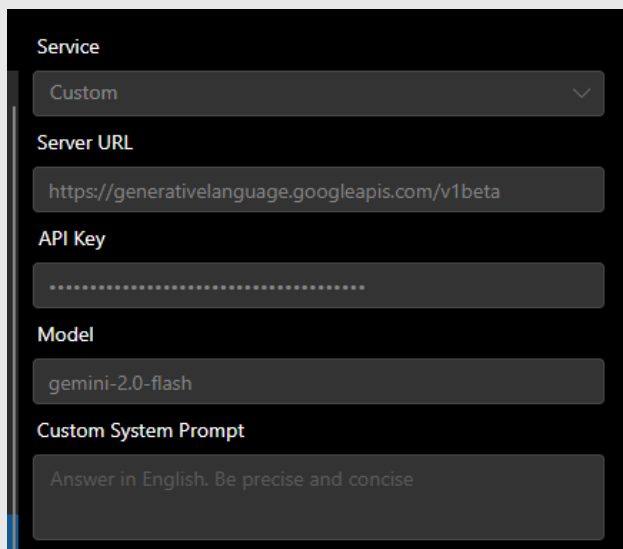
You can set up as many connections as you wish. Some services offer free personal accounts to evaluate models. For production or commercial use most services operate on a pay-as-you-go basis.

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



Below models have been tested with eSearch Pro as of 4 March 2025.

Service	URL	Model
Anthropic	<a href="https://api.anthropic.com/v1">https://api.anthropic.com/v1</a>	claude3.5
X.ai	<a href="https://api.x.ai/v1">https://api.x.ai/v1</a>	grok-2-latest
Mistral	<a href="https://api.mistral.ai/v1">https://api.mistral.ai/v1</a>	mistral-medium
Novita	<a href="https://api.novita.ai/v3/openai">https://api.novita.ai/v3/openai</a>	qwen\qwq-32b



Google gemini-2.0-flash example, note long server URL.

If you need assistance with setting up any model listed, please contact [support@electoart.co.uk](mailto:support@electoart.co.uk)

## Other Online AI LLM API's

Venice.ai	<a href="https://api.venice.ai/api/v1">https://api.venice.ai/api/v1</a>	Needs a Pro account.
Cohere	<a href="https://api.cohere.com/v1/chat">https://api.cohere.com/v1/chat</a>	
AWS Bedrock	<a href="https://docs.aws.amazon.com/general/latest/gr/bedrock.html">https://docs.aws.amazon.com/general/latest/gr/bedrock.html</a>	



## 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

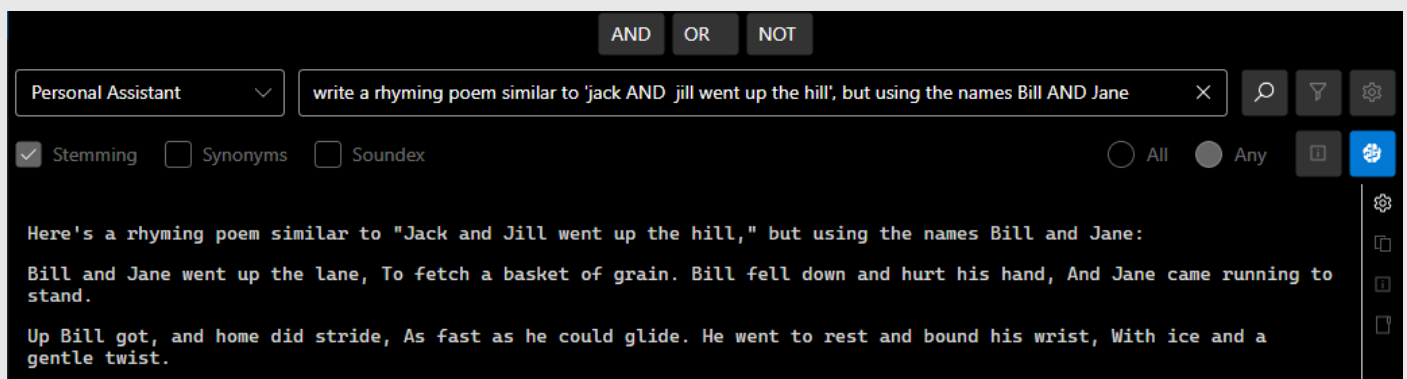
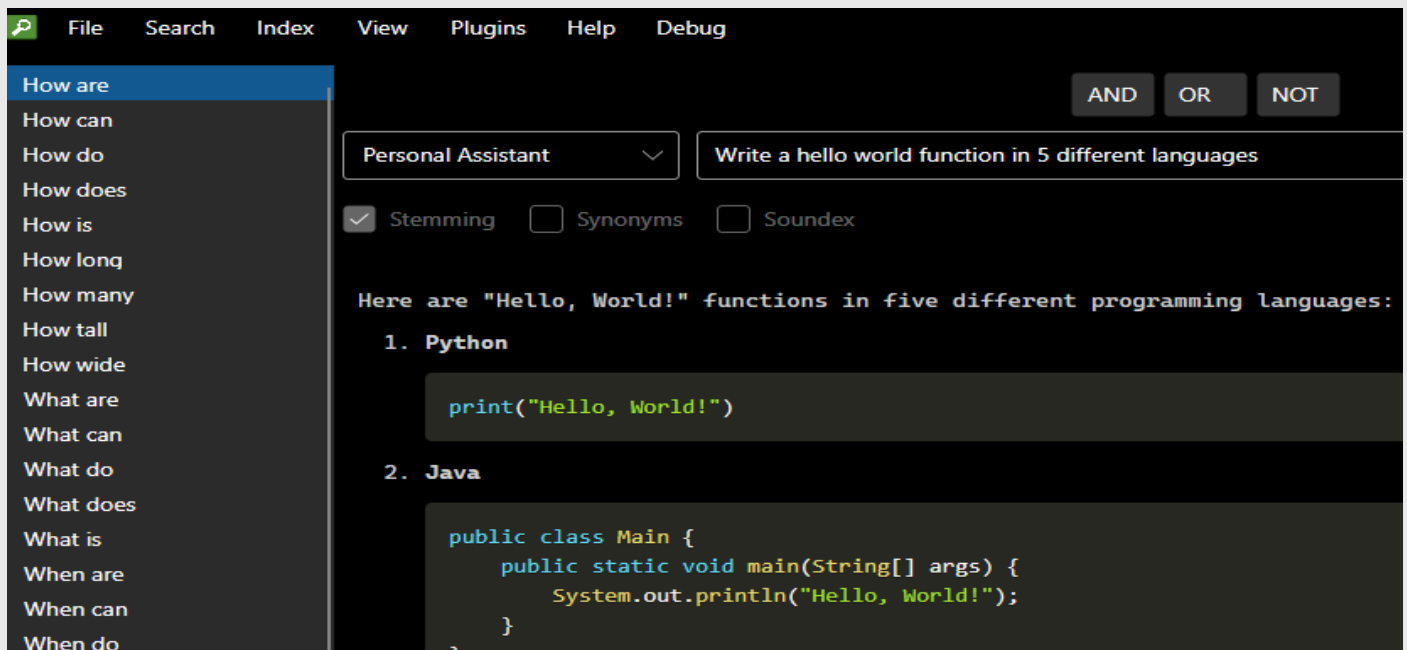
When you start to type a query beginning with the first letter or two from 'how', 'what', 'who', or 'when', the word list will scroll to some common query phrases. You can then double-click on a query stem to complete the query. It isn't necessary to use the listed query stem phrases.

You can also use voice input from eSearch Pro version 1.0.9097 onwards.

## Perplexity Sonar

### Examples

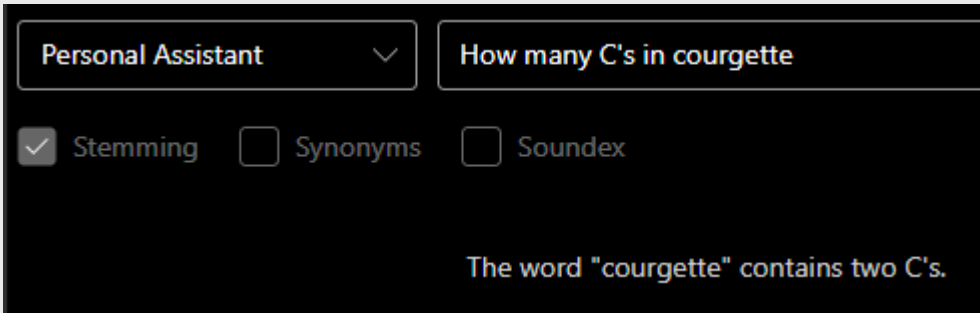
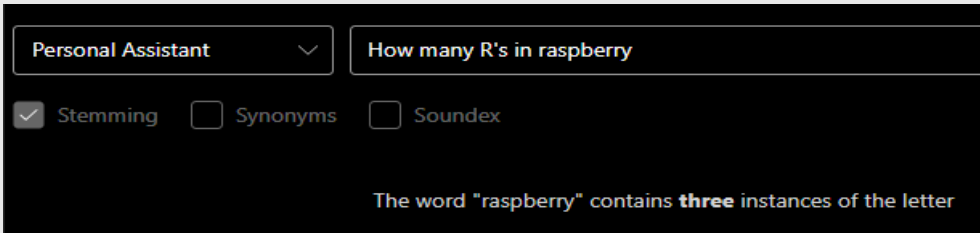
1. You can ask it to write code or a poem:



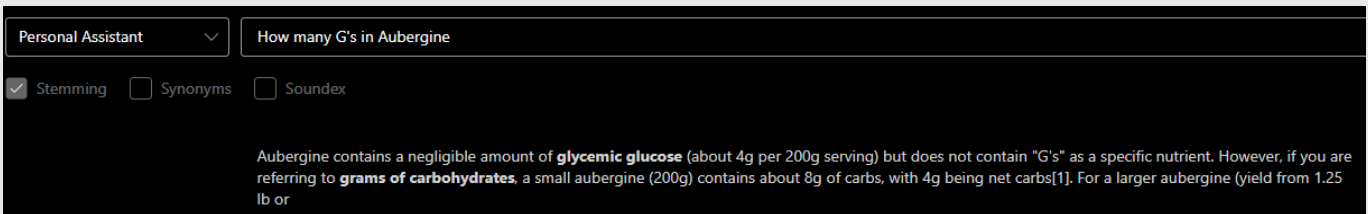
The results may be useful but are clearly not perfect. Reasoning models are more useful in business or educational settings, since the user can

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



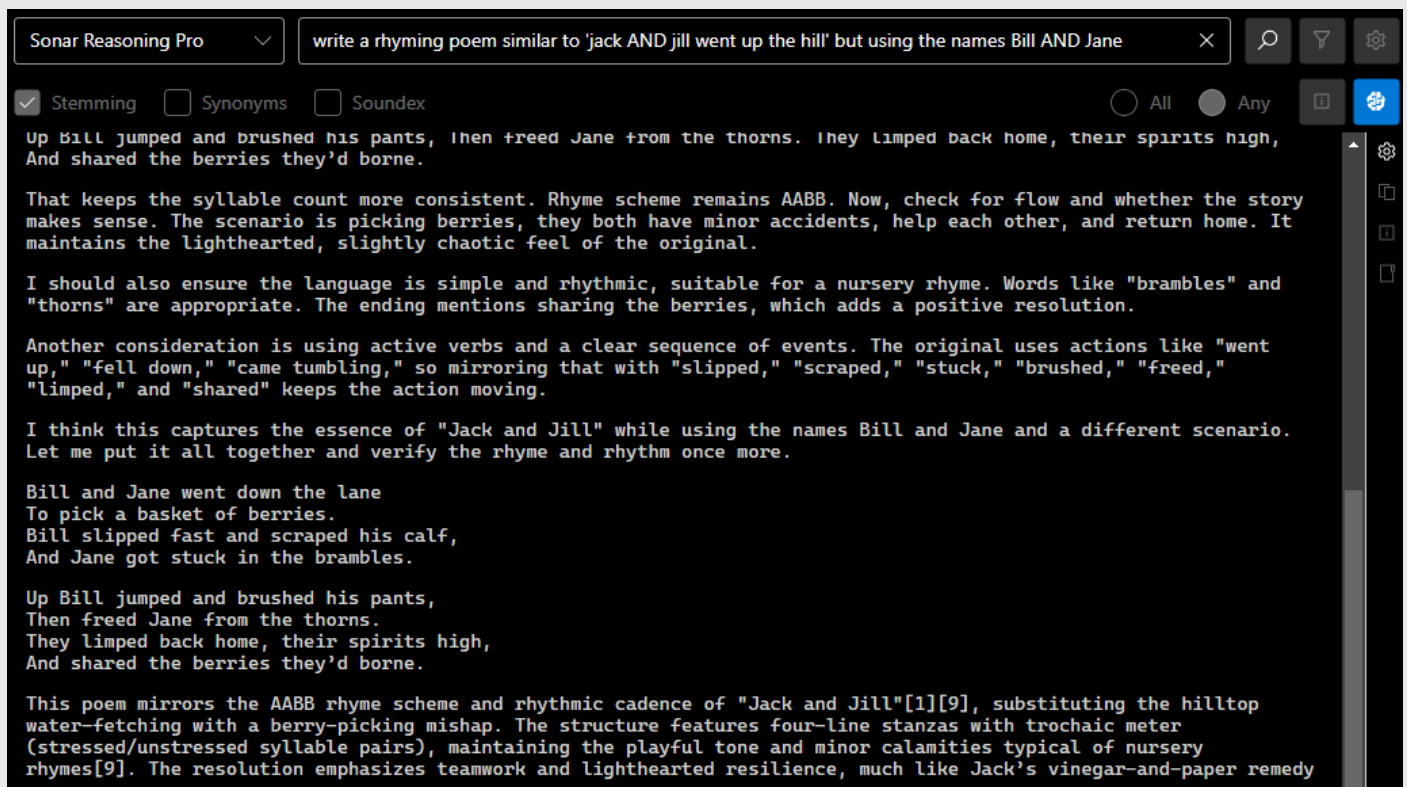
But non-reasoning models can get facts wrong. Note: Zucchini (US) vs. Courgette (UK)



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

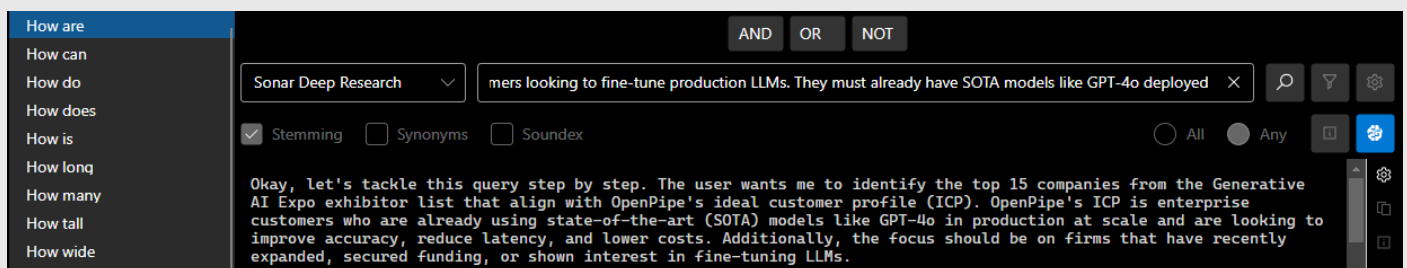
## Sonar Reasoning

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



## Sonar 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.



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](mailto:support@electronart.co.uk), we may have it the pipeline or will put it on our roadmap!

-oo0oo-