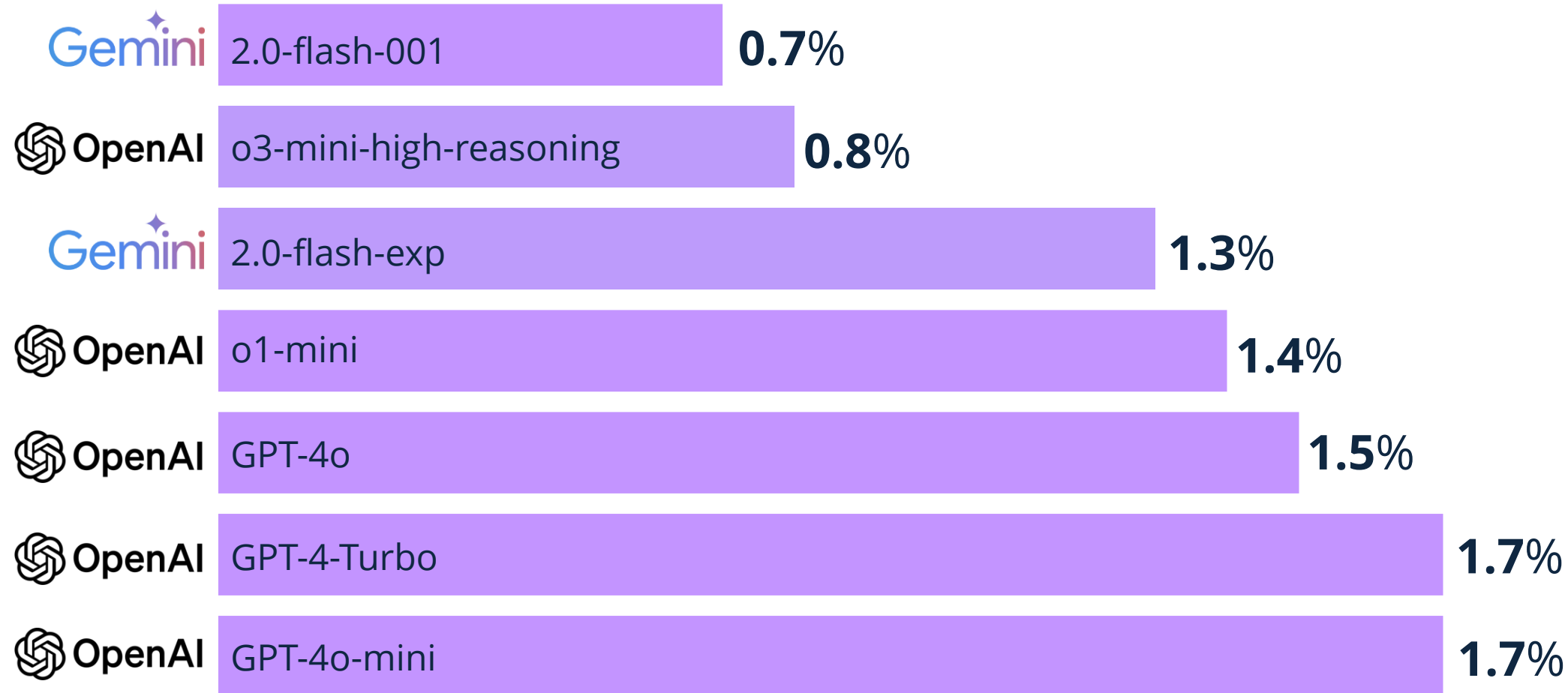


DATA ON STAGE JUNE 2025

# From Hallucination to precisions – Why Your AI Needs the Right Data

# Leading AI models by **HHEM hallucination rate** 2025

(percentage)



Note(s): Worldwide; May 2025  
Source(s): Hugging Face; [Link](#)



A man with a grey beard and hair, wearing a light-colored suit and a bow tie, stands in the center of a grand, ornate Victorian-style parlor. He has his arms outstretched in a gesture of explanation or presentation. He is surrounded by several people seated on red velvet sofas and chairs. The women are wearing elaborate Victorian dresses and hats. The room features high ceilings with ornate moldings, large mirrors, and chandeliers. The lighting is warm and ambient.

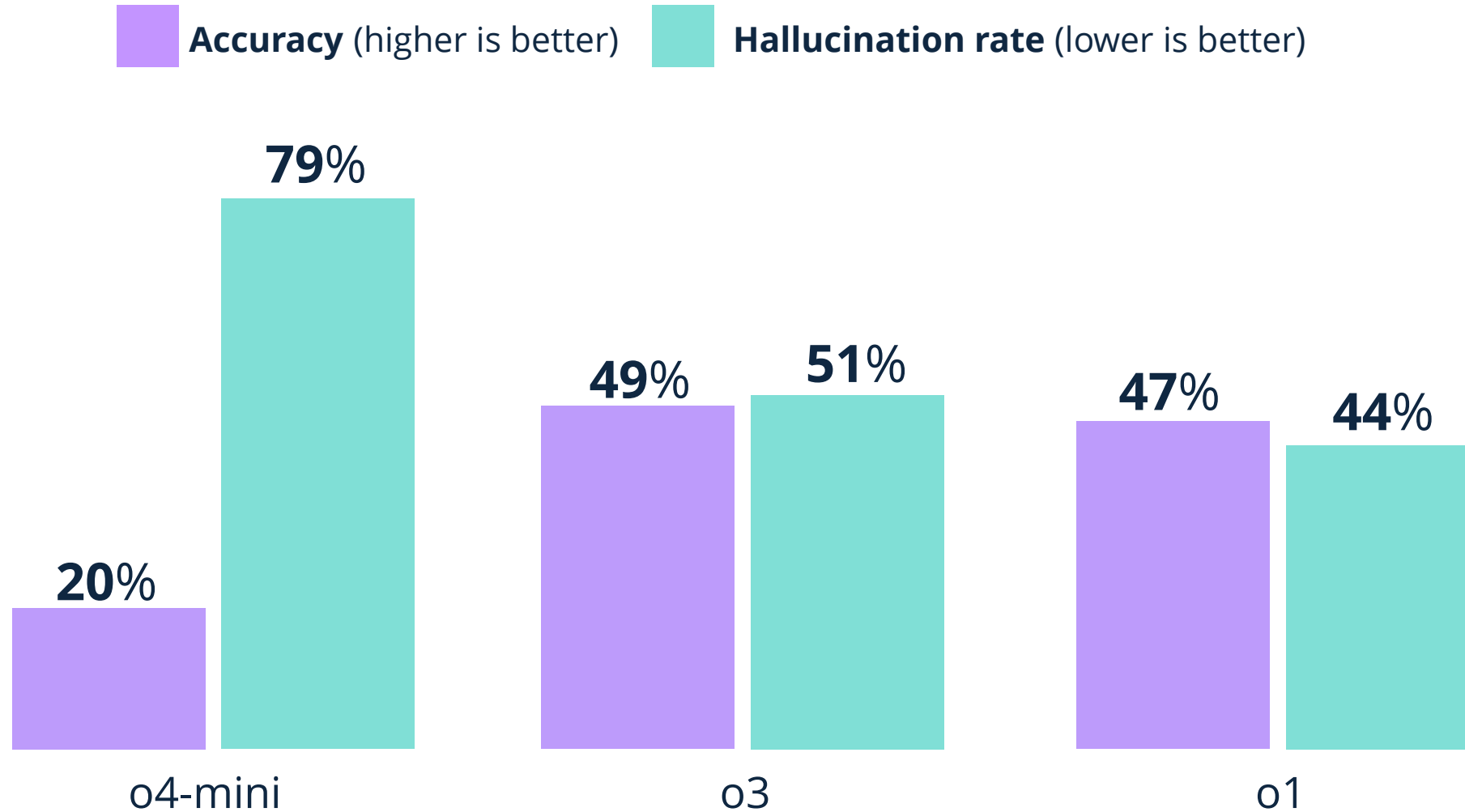
"There are three kinds of lies: **lies,**  
**damned lies, and statistics,..**"

*- Mark Twain*








# SimpleQA evaluation metrics of select models

(percentage)



# Multiple choice text benchmark for LLM's scoreboard











(Average score percentage)

Rank	Organization	Model	Score
-		Human baseline	83.7%
1st		Gemini 2.5 Pro (06-05)	62.4%
2nd	ANTHROPIC	Claude 4 Opus (thinking)	58.8%
3rd	 OpenAI	o3 (high)	53.1%
4th		Gemini 2.5 Pro	51.6%
5th	ANTHROPIC	Claude 3.7 Sonnet (thinking)	46.4%
6th	ANTHROPIC	Claude 4 Sonnet (thinking)	45.5%
7th	ANTHROPIC	Claude 3.7 Sonnet	44.9%
8th	 OpenAI	o1-preview	41.7%
9th	ANTHROPIC	Claude 3.5 Sonnet 10-22	41.4%
10th	 deepseek	DeepSeek R1 05/28	40.8%

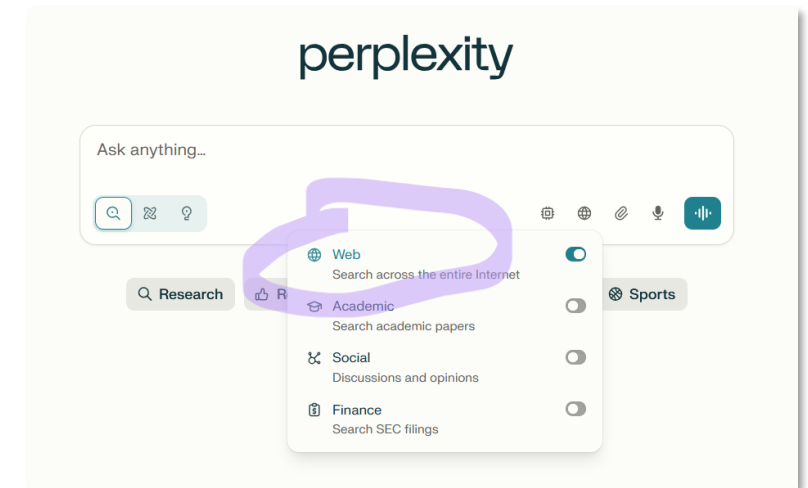
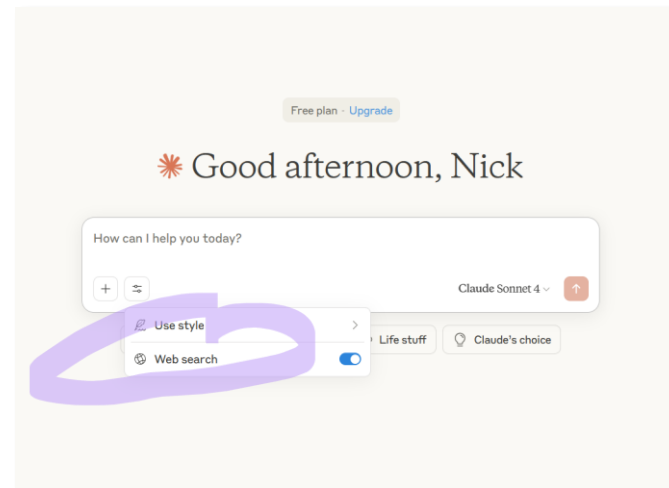
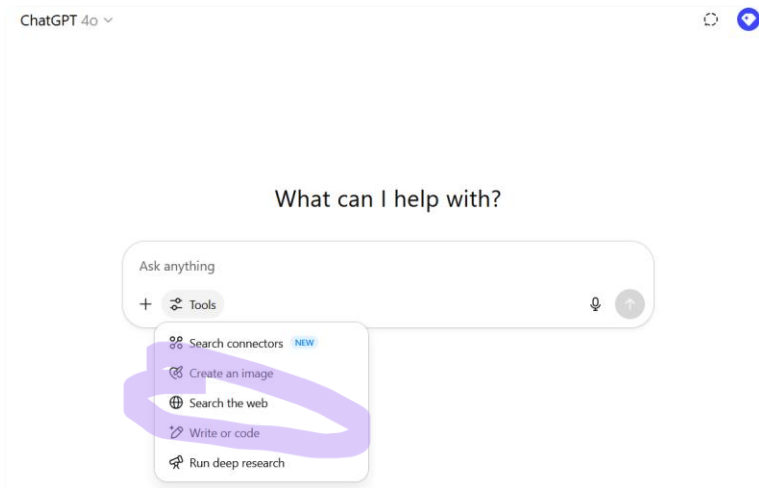
Note(s): October 2024  
Source(s): SimpleBench

# Multiple choice text benchmark for LLM's scoreboard

(Average score percentage)

Rank	Organization	Model	Score	
28th	 OpenAI	o3-mini (high)	<div></div>	22.8%
29th		Grok 2	<div></div>	22.7%
30th	 MISTRAL AI	Mistral Large v2	<div></div>	22.5%
31st	 Meta	Llama 3.3 70b instruct	<div></div>	19.9%
32nd	 deepseek	DeepSeek V3	<div></div>	18.9%
33rd		Gemini 2.0 Flash Exp	<div></div>	18.9%
34th	 OpenAI	o1-mini	<div></div>	18.1%
35th	 OpenAI	GPT-4o 08-06	<div></div>	17.8%
36th	 cohere	Command R+	<div></div>	17.4%
37th	 OpenAI	GPT-4o mini	<div></div>	10.7%

# All major LLMs **now include web search** capabilities



Information obtained from **LLMs always needs to be verified** as:

## HALLUCINATIONS

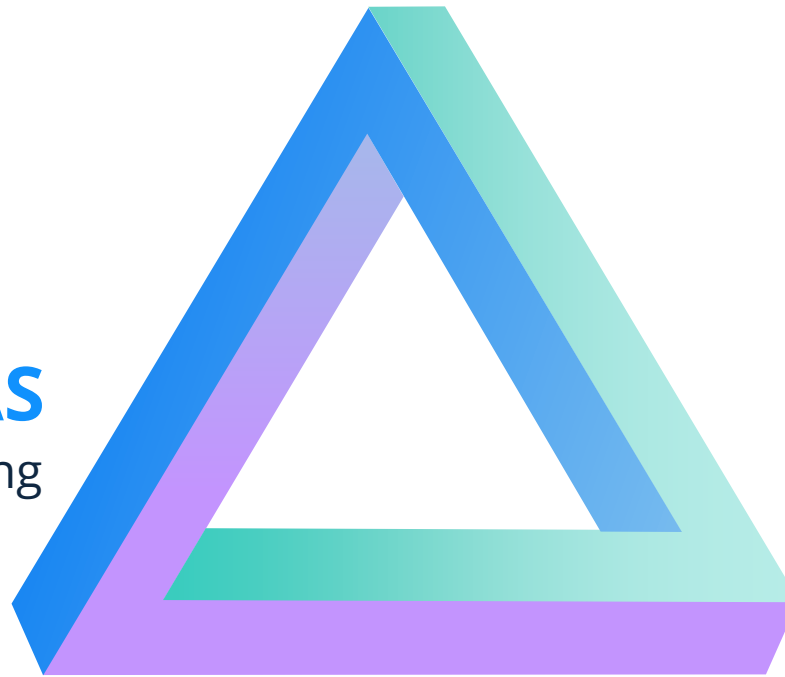
LLMs can hallucinate facts - even when accessing information from websites

## BIAS

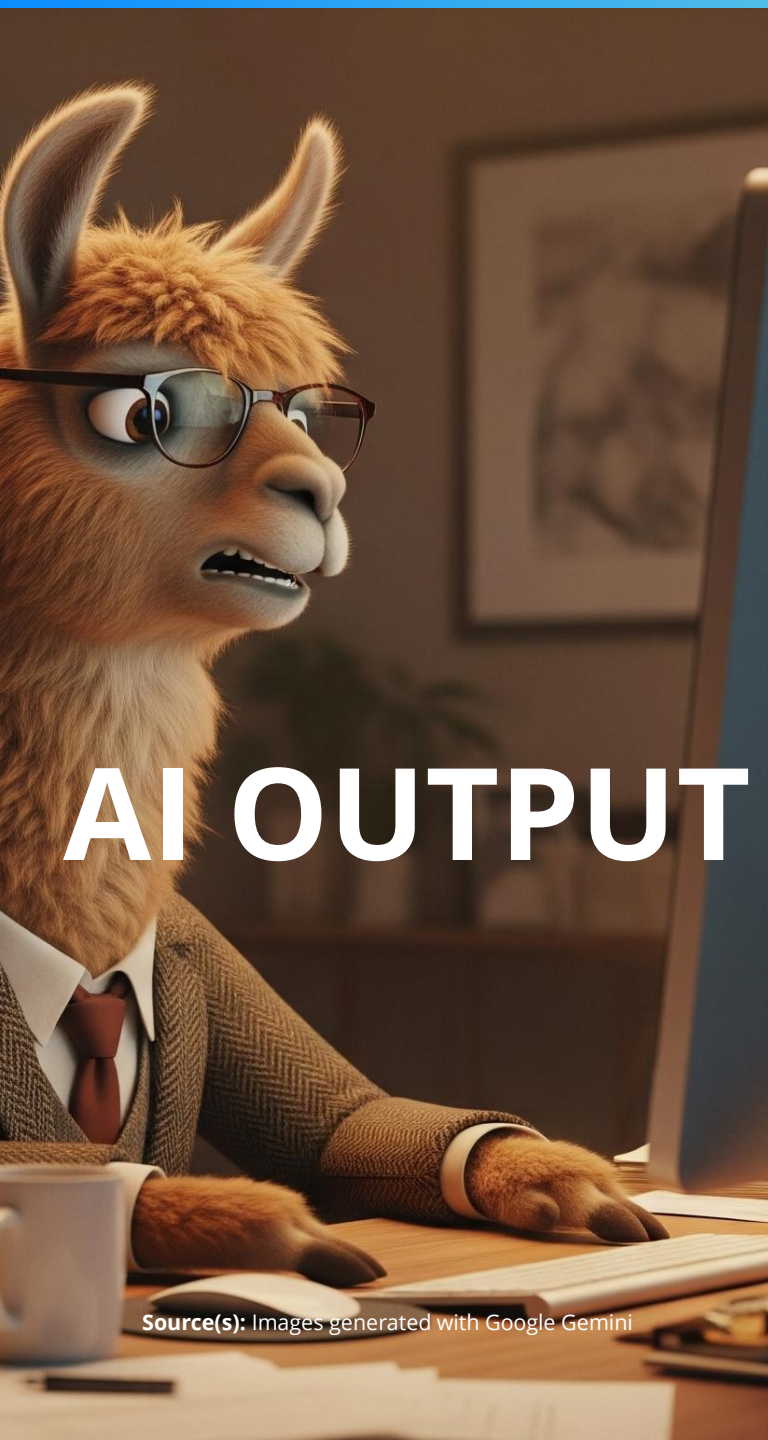
LLMs can reproduce bias, omitting critical facts and/or perspectives

## SOURCE SELECTION

LLMs often select unreliable, out of date, and/or irrelevant sources







AI OUTPUT

Source(s): Images generated with Google Gemini

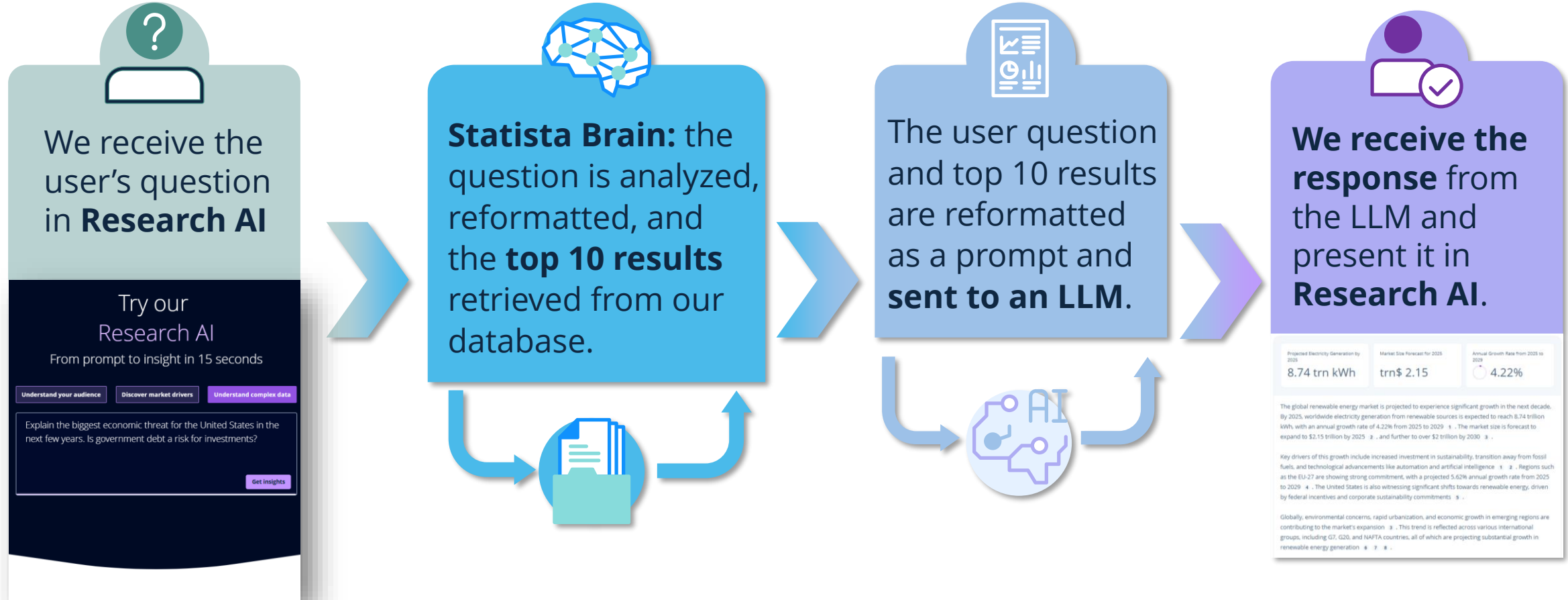


DISINFORMATION

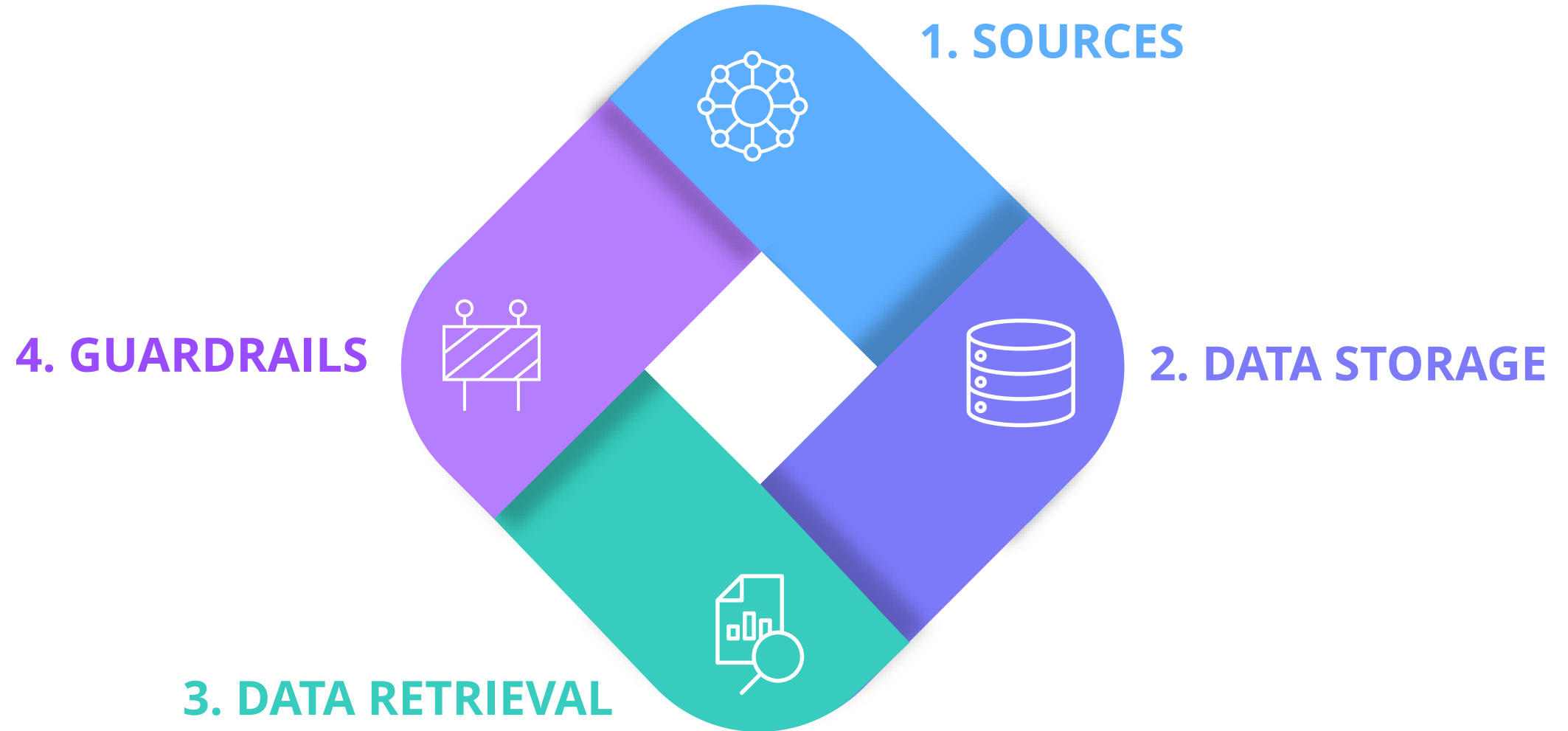


AI ADOPTION

# Retrieval-augmented generation (RAG)


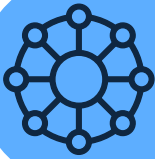





# Elements of Retrieval-augmented generation (**RAG**)





# Elements of RAG

Element		Requirements	statista  solution
	<b>SOURCES</b>	Reliable and up to date	Human verified; curated collection of 22.5k sources covering 80k topics
	<b>DATA STORAGE</b>	Short “chunks” containing facts and context	Data series plus a written description; no breaking up of larger documents / data sets
	<b>DATA RETRIEVAL</b>	Identify relevant data without providing too much information	Custom pipeline with two-stage process combining traditional search and LLM
	<b>GUARDRAILS</b>	Reduce hallucinations and other output issues; use-case specific	Prompt engineering to optimize output for short form research



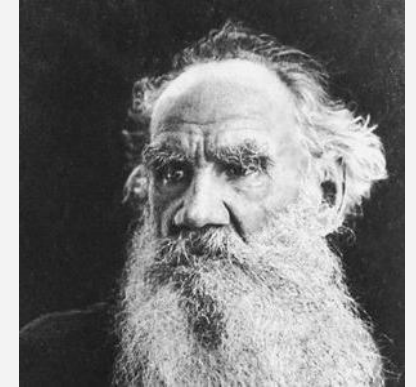
# Chunking can easily lose context

## Chunk 1



### Leo Tolstoy

**Leo Tolstoy** (born August 28 [September 9, New Style], 1828, Yasnaya Polyana, Tula province, Russian Empire—died November 7 [November 20], 1910, Astapovo, Ryazan province) was a Russian author, a master of realistic fiction, and one of the world's greatest novelists. He is best known for his two longest works, *War and Peace* (1865–69) and *Anna Karenina* (1875–77), which are commonly regarded as among the finest novels ever written. *War and Peace* in particular virtually defines this form for many readers and critics. Among **Tolstoy's** shorter works, *The Death of Ivan Ilyich* (1886) is usually classed among the best examples of the novella.



## Chunk 2



Especially during his last three decades **he** also achieved world renown as a moral and religious teacher. His doctrine of nonresistance to evil had an important influence on the Indian independence leader **Mahatma Gandhi**. Although **his** religious ideas no longer command the respect they once did, interest in his life and personality has, if anything, increased over the years.

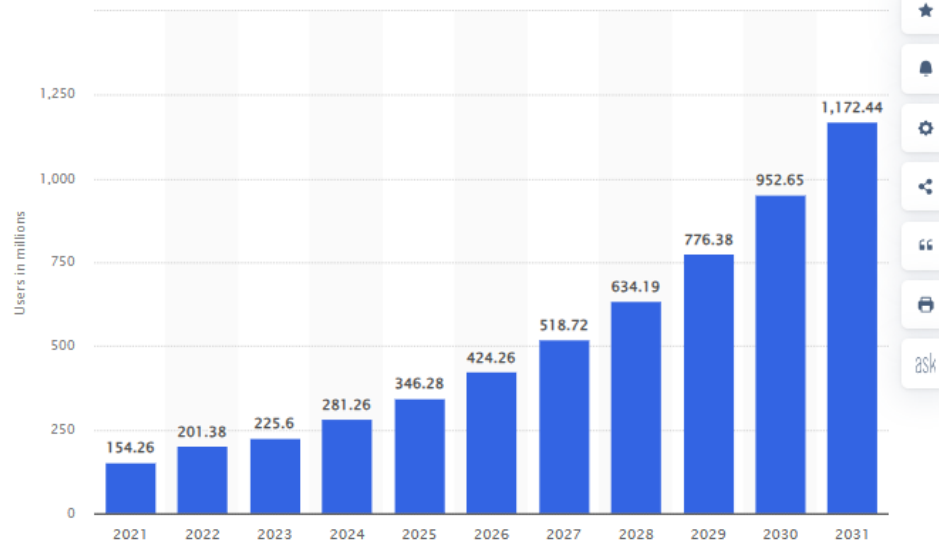
## Chunk 3



Most readers will agree with the assessment of the 19th-century British poet and critic Matthew Arnold that a novel by **Tolstoy** is not a work of art but a piece of life. The Russian author Isaak Babel commented that, if the world could write by itself, it would write like **Tolstoy**. Critics of diverse schools have agreed that somehow **Tolstoy's** works seem to elude all artifice. Most have stressed his ability to observe the smallest changes of consciousness and to record the slightest

# Short, self-contained text is **best practice for RAG** systems

Number of artificial intelligence (AI) tool users globally from 2021 to 2031  
(in millions)



Details: Worldwide; 2021 to 2031

© Statista 2025

## AI tool user numbers worldwide from 2021-2031

Published by [Statista Research Department](#), Apr 7, 2025



The global number of AI tools users in the 'AI Tool Users' segment of the artificial intelligence market was forecast to continuously increase between 2025 and 2031 by in total 826.2 million (+238.59 percent). After the tenth consecutive increasing year, the number of AI tools users is estimated to reach 1.2 billion and therefore a new peak in 2031. Notably, the number of AI tools users of the 'AI Tool Users' segment of the artificial intelligence market was continuously increasing over the past years.

Find more key insights for the number of AI tools users in countries and regions like [the market size in the 'Generative AI' segment of the artificial intelligence market in Australia](#) and [the market size change in the 'Generative AI' segment of the artificial intelligence market in Europe](#).

The [Statista Market Insights](#) cover a broad range of additional markets.

Hide

## Competitor AI **Performance Evaluation** Across Key Metrics – May 2025

		Hallucination (25%)	Context (10%)	Factuality (15%)	Relevance (10%)	Source links (15%)	Recommend- -ations (15%)	Total Score
	Research AI	0.70	0.89	0.85	0.89	1.00	1.00	<b>0.86</b>
	Perplexity.ai Pro	0.57	0.89	0.89	0.91	0.95	1.00	<b>0.83</b>
	Gemini	0.57	0.88	0.85	0.85	0.70	0.00	<b>0.72</b>
	DeepSeek	0.57	0.88	0.84	0.88	0.02	0.00	<b>0.62</b>
	ChatGPT	0.59	0.80	0.67	0.81	0.14	0.00	<b>0.59</b>
	Claude 3.5 Sonnet	0.54	0.87	0.75	0.85	0.00	0.00	<b>0.59</b>

GLOBAL & LOCAL EXPERTISE

# Unmatched coverage of data & markets

**1 M+**

Statistics

**150**

Countries

**170**

Industries

**60 K+**

Reports

**40 K+**

Sources

**15 K+**

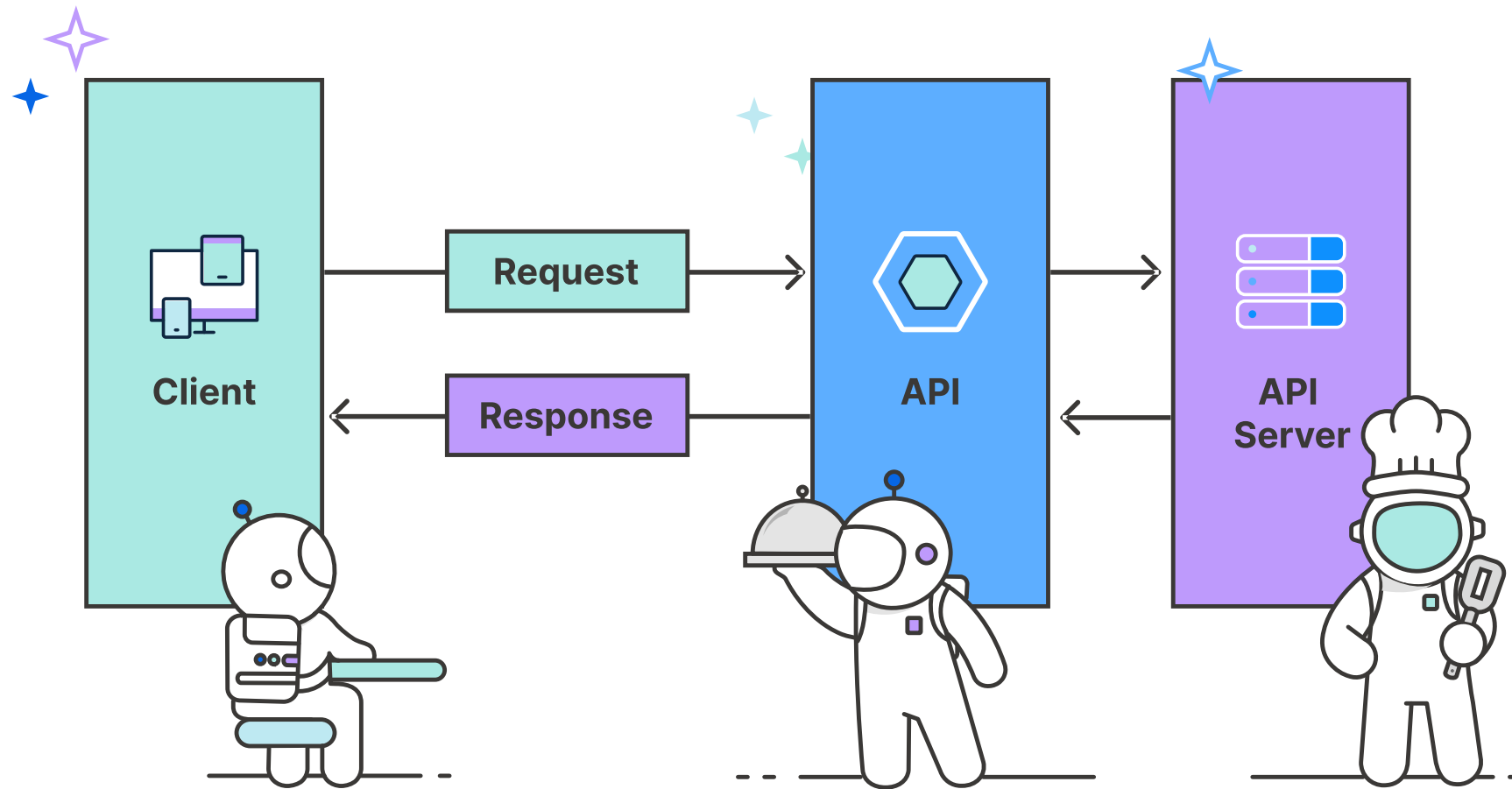
Brands

**1 K+**

Markets



# What is an Application Programming Interface - **API** ?



**Statista Connect** allows users to **leverage trusted data from Statista** wherever they need it and integrate our expert insights directly into their workflows and tools



### Search & Find Data

>>Search through **all your data sources** at once within your working environment to find the **right information quicker**.<<

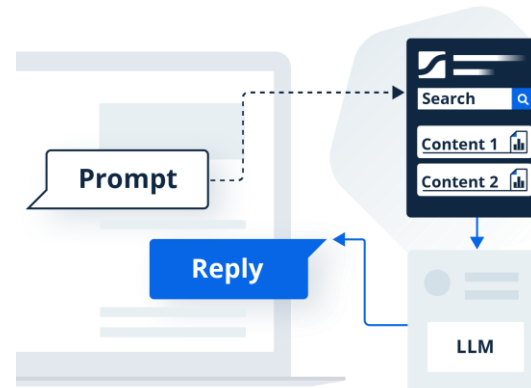


- **Integrate the Statista search** into your federated search or knowledge management system
- Our API lets you **search & discover** our extensive database within the system environment of **your choice**



### Data-powered Chatbot

>>Integrate our **Research AI** or power your **own chatbot** with Statista data to **quickly get answers** to your research questions.<<



- **Integrate trusted Statista data** into an AI-powered Chatbot for fact-based answers to your prompts
- Choose between our ready to use **Research AI** or connect **your custom Chatbot/RAG** to Statista's database



### Automatic Content Creation

>>Ingest **high-quality raw data** to your automated content-creation workflow or application for **faster & better results**.<<



- **Integrate trusted Statista data** into your custom (AI-powered) tool to automatically generate content
- Our API lets you directly **feed our fact checked data into your workflow** to create slides, reports, summaries etc.

# Our new partnerships **bring our trusted data to you**

## Canva

Create polished, data-driven visuals in Canva with Statista's expert-verified insights—perfect for presentations, documents, social posts, and more.



## perplexity

Instantly search, verify, and explore our expert-backed market data—seamlessly integrated into one of the world's leading AI-powered answer engines.



## Microsoft 365 Copilot

From Word and Excel to PowerPoint, Outlook, and Teams—access trusted insights directly in your workflow, without switching apps or losing focus.