

Sprint 1:

2. let AI do the job for you! Using AI to detect AI writing

Prompt: - On a scale of 1-100, give a score for the likelihood that the following text was heavily written by AI?

Result: - selected models correctly scores given text as AI generated. 90/100 | 80 | 90/100 | 80

The screenshot shows the nexos.ai web application. At the top, there's a navigation bar with 'nexos.ai' (BETA), 'Chat', 'Compare Models' (which is highlighted in purple), 'Projects', and 'Assistants'. On the right side of the header are 'Docs' and a user icon. Below the header, there's a row of model selection buttons. The first button for 'Gemini 2.5 Pro' has a checkmark and is highlighted in blue, indicating it's the active model. The other three models—'Llama 4 Scout', 'mistral-medium', and 'o3-mini-2025-01...'—also have checkmarks and are in a standard grey state.

On a scale of 1-100, give a score for the likelihood that the following text was heavily written by AI:

AI is set to change the world because it enables machines to perform cognitive tasks that were once thought to be uniquely human—learning, reasoning, decision-making, and even creativity. This shift dramatically expands the scope of automation from routine physical labor to complex mental work. Unlike previous technological revolutions that primarily affected specific industries (e.g., the steam engine in transport or electricity in manufacturing), AI has the potential to transform virtually every sector—healthcare, education, finance, logistics, law, and beyond. Its ability to process vast amounts of data, identify patterns, and optimize decisions at speeds and scales far beyond human capability makes it a foundational tool for solving problems ranging from drug discovery to climate modeling.

This screenshot shows the same nexos.ai interface as the previous one, but with a different set of four models selected for comparison. From left to right, the selected models are 'Gemini 2.5 Pro', 'Llama 4 Scout', 'mistral-medium', and 'o3-mini-2025-01...'. Each model has a small icon and its name followed by '(gemini-2.5-pro by ...)', '(meta-llama-llama-4...)', '(mistral-m...)', and '(openai) (o3-m...)' respectively.

This screenshot shows the nexos.ai interface with the same four models selected: Gemini 2.5 Pro, Llama 4 Scout, mistral-medium, and o3-mini-2025-01... Each model section contains a brief summary of its assessment of the provided text.

This screenshot shows the nexos.ai interface with the same four models selected. The 'Gemini 2.5 Pro' section includes a 'Reasoning for the Score' section that details the indicators used to score the text as AI-generated.

Prompt: - Analyze and provide a score for text below. How likely this text is written by AI? By what most likely AI model? Provide short answer.

Result: - selected models correctly rate text as AI generated and even precisely can detect which model generated it.

nexos.ai BETA Chat Compare Models Projects Assistants Docs S

Analyze and provide a score for text bellow. How likely this text is written by AI? By what most likely AI model? Provide short answer.

What makes AI revolutionary isn't just its power, but its adaptability and exponential progress. It improves over time, learns from interaction, and can be embedded into virtually any digital system. The convergence of AI with other technologies—robotics, biotech, quantum computing—further amplifies its impact. Just as electricity became a general-purpose technology that quietly powered every device and industry, AI is becoming an invisible infrastructure of intelligence behind modern systems. It doesn't just change how we do things—it changes what is possible. That shift alters the trajectory of scientific progress, the structure of economies, and even the nature of human work and creativity.

G Gemini 2.5 Pro (gemini-2.5-pro by ...)	∞ Llama 4 Scout (meta-llama-llama-4...)	M mistral-medium (mistral) (mistral-m...)	o3-mini-2025-01-31 (openai) (o3-m...
Likelihood Score: 95/100	AI-Generated Text Score: 90	Likelihood Score: 95/100	Score: 85/100
This text is highly characteristic of a state-of-the-art AI. The reasoning is	This text is likely written by AI, specifically a model like GPT-4 or	GPT-4 or similar advanced	Likely AI Model: GPT-4
			Reasoning:

nexos.ai BETA Chat Compare Models Projects Assistants Docs S

G Gemini 2.5 Pro (gemini-2.5-pro by Google)

Likelihood Score: 95/100

This text is highly characteristic of a state-of-the-art AI. The reasoning is based on its conceptual density, flawless structure, and use of a very common and powerful analogy ("AI as electricity"). The prose is polished and authoritative but lacks any unique human voice or perspective, making an AI origin extremely likely.

Most Likely AI Model: A top-tier large language model such as OpenAI's GPT-4 (or a successor), Google's Gemini 1.5 Pro, or Anthropic's Claude 3 Opus. These models excel at synthesizing complex

∞ Llama 4 Scout (meta-llama-llama-4-scout-...)

AI-Generated Text Score: 90

This text is likely written by AI, specifically a model like GPT-4 or similar large language models (LLMs) trained on a vast corpus of text data.

Reasoning:

- The text exhibits fluent, coherent, and generic descriptions of AI's capabilities and impact.
- It uses buzzwords and phrases (e.g., "exponential progress," "invisible infrastructure of intelligence") common in AI discourse.

M mistral-medium (mistral) (mistral-medium-...)

Likelihood Score: 95/100

Most Likely AI Model:

- GPT-4 or similar advanced language model (e.g., Claude 3, Gemini 1.5).
 - The text exhibits high coherence, nuanced phrasing, and a structured flow typical of state-of-the-art models.
 - The use of metaphors (e.g., "invisible infrastructure of intelligence") and the seamless integration of complex ideas suggest a model trained on diverse, high-quality data.

o3-mini-2025-01-31 (openai) (o3-mini-202...

Score: 85/100

Likely AI Model: GPT-4

Reasoning:

- The text is polished, generic, and rich in buzzwords, which is common in AI-generated content.
- It incorporates broad, sweeping statements and analogies that are characteristic of models like GPT-4.
- The style, coherence, and structure suggest high-level pattern-matching typical of modern AI language models.

Task: - Check if 'lower grade' models can detect AI generated text.

Prompt: - Analyze and provide a score for text bellow. How likely this text is written by AI? By what most likely AI model? Provide short answer.

Result: - Some models start to struggle, still most of them rises a flag of AI generated text. Only Llama 4 gives lowest score.

ChatGPT 5 ▾

Generate a small paragraph about what came first egg or chicken?

The classic question of what came first, the chicken or the egg, can be approached through biology and evolution. From a scientific perspective, the egg came first—because long before chickens existed, other animals (like reptiles) were already laying eggs. Over time, genetic mutations inside an egg laid by a bird that wasn't quite a chicken eventually produced the first true chicken. So, in evolutionary terms, the egg predates the chicken.

nexos.ai BETA Chat Compare Models Projects Assistants Docs

G Gemini 2.5 Pro (gemini-2.5-pro by Google)
Likelihood Score: 90/100
This text is a perfect, encyclopedic summary of the scientific answer to a classic paradox. Its structure is highly logical, and the language is clear, factual, and efficient, but it lacks any personal flair or novel phrasing. This style of providing a concise, textbook-correct answer is a hallmark of AI-generated content.
Most Likely AI Model: This is a standard factual retrieval task that almost any modern LLM could handle perfectly. The quality and conciseness are typical of models like OpenAI's GPT-4, Google's

∞ Llama 4 Scout (meta-llama-llama-4-scout-...)
AI-Generated Text Score: 40
This text is likely partially written or influenced by a human, but may have been polished or generated by AI. A likely candidate could be a light editing or content generation model, rather than a fully-fledged LLM like GPT-4.
Reasoning:

- The text provides a clear, scientifically accurate explanation of the chicken-and-egg problem.
- The structure and language are straightforward and informative, typical of AI-generated educational content.

M mistral-medium (mistral) (mistral-medium-...)
Likelihood Score: 85/100
Most Likely AI Model:

- GPT-3.5 or similar mid-tier language model (e.g., earlier versions of Claude or Gemini).
- The explanation is clear and logically structured but lacks the depth or stylistic flair of more advanced models.
- The phrasing is straightforward and informative, typical of AI-generated educational content.

Key Indicators of AI Authorship:

o3-mini-2025-01-31 (openai) (o3-mini-202...)
Score: 60/100
Likely AI Model: GPT-3.5
Reasoning:

- The text provides a clear, logical explanation and uses accessible language, which can be typical of both human and AI writing.
- It adheres to a common format for explaining evolutionary concepts without overly complex or stylized language.
- While the content is coherent and informative, its simplicity leans slightly toward a straightforward AI-

Task: - Check if 'lower grade' models can detect AI generated text that was prompted to be generated not to look like Ai generated..

Prompt: - Analyze and provide a score on a scale from 0 to 100 for text bellow. How likely this text is written by AI? By what most likely AI model? Provide short answer.

Result: - even if superior model tries to generate text that should be AI-like other models still can identify that it is AI generated. This time lowest score is given by by GPT-3o-mini - it defines it as Human-like text.

Generate a small paragraph about what came first egg or chicken? Try to make it doesn't look like Ai generated.

People have argued for centuries about whether the chicken or the egg came first. If you look at it through evolution, the egg has the upper hand: creatures were laying eggs long before chickens showed up. At some point, a not-quite-chicken bird laid an egg, and thanks to a tiny genetic change, the chick that hatched grew into what we'd now call the first chicken. In that sense, the egg came first.

The screenshot shows a user interface for comparing AI detection across four models. Each model's panel includes its name, likelihood score, analysis, most likely AI model, and short analysis.

- G Gemini 2.5 Pro (gemini-2.5-pro by Google)**
 - AI Likelihood Score:** 95/100
 - Analysis:** The text is highly likely to be AI-generated. It presents a flawless, concise, and perfectly structured explanation of a common paradox, which is a hallmark of large language models. The phrasing is efficient and lacks any unique human voice or stylistic quirks.
 - Most Likely Model:** A model from the GPT-4 or Gemini family, known for producing high-quality, logical, and explanatory text.
- ∞ Llama 4 Scout (meta-llama-llama-4-scout-...)**
 - AI Likelihood Score:** 80
 - The text exhibits characteristics common in AI-generated content, such as:
 - A clear and concise structure
 - Use of transitional phrases ("If you look at it through evolution")
 - A logical and coherent argument
 - Avoidance of personal opinions or biases
 - Most Likely AI Model:** The text is likely generated by a model similar to **LLaMA** or **PaLM**, which are known for their ability to produce well-structured and informative
- M mistral-medium (mistral) (mistral-medium-...)**
 - AI Detection Score:** 90/100
 - Most Likely AI Model:** GPT-3.5 or GPT-4
 - Short Analysis:** The text exhibits high coherence, logical structure, and a neutral tone, which are characteristic of AI-generated content. The explanation is clear and concise, typical of models trained for informative and balanced responses. The lack of personal anecdotes or emotional depth further suggests AI authorship.
- o3-mini-2025-01-31 (openai) (o3-mini-202...)**
 - Score:** 20/100
 - Likely AI Model (if AI-generated):** GPT-3.5 or GPT-4
 - The text shows clear, logical reasoning and natural phrasing typical of human writing. If it were AI-generated, it would most likely come from a model like GPT-3.5 or GPT-4, but overall it has a low likelihood of being produced by AI.

Conclusion: - Ai models overall can easily identify which text is AI generated and which is not. They can even identify what model generated those - what is most fascinating finding of this exercise.