

Llm Terminology Cheat Sheet

Plain-English AI reference guide | cyberside.ai

Term	Plain-English meaning
A/B testing	A way to compare two versions of a page, message, prompt, or workflow with similar audiences to see which one performs better. Change one important thing at a time when possible.
API key	A secret credential used by software to authenticate to an API. Treat it like a password: do not paste it into public chat tools or commit it to public code.
Attention mechanism	A technique that helps many modern language models weigh which parts of the input are most relevant while producing an output.
Authentication	The process of proving who or what is requesting access, often with passwords, passkeys, tokens, or multi-factor authentication.
Confabulation	Another word for an AI hallucination: an invented or unsupported claim delivered with apparent confidence.
Context	The information an AI receives to help it answer or act, including instructions, conversation history, documents, examples, and tool results.
Context window	The amount of information an AI model can consider at one time in a conversation or task. More context can help, but it does not guarantee the answer is correct.
Embeddings	Numeric representations of text, images, or other data that capture useful similarities. They are often used for semantic search and RAG.
Evaluation	A structured way to test whether an AI system performs well enough for a defined task and risk level.
Hallucination	When an AI produces information that sounds believable but is wrong, unsupported, or invented. Treat high-stakes answers as drafts requiring verification.
Instruction hierarchy	The order of importance among instructions given to an AI system, such as system rules, developer rules, and user requests.
Jailbreak	An attempt to bypass a model's intended safety rules or restrictions through clever instructions or other inputs.
Prompt	The instructions, question, context, and examples you give an AI. Better prompts make the task clearer; they do not turn an unreliable claim into a fact.
Prompt engineering	Designing prompts, examples, constraints, and output formats so an AI can complete a task more reliably.
RAG	Retrieval-Augmented Generation: a setup where an AI looks up relevant documents or data before answering. It can improve grounding, but you still need to check the source material.
Retrieval	Finding relevant information from a collection of documents, records, or data.
Semantic search	Search that tries to match meaning, not only exact words.
System prompt	High-priority instructions that define an AI assistant's overall role, boundaries, and behavior.
Text-to-image	AI that creates images from written instructions.
Token	A small chunk of text that an AI model processes. A token might be a word, part of a word, punctuation, or a short sequence of characters.

Term	Plain-English meaning
Vector database	A database designed to store and search embeddings, often used for semantic search and RAG.
Vector search	Searching for items with similar embeddings, which can help find related meaning even when wording differs.
Zero-shot prompting	Asking a model to perform a task without providing examples in the prompt.

Educational reference. Verify high-stakes decisions with qualified professionals.