AI Glossary - 50 Terms Explained in Plain English

Table of Contents

Introduction

AI conversations are full of jargon that gets thrown around as if everyone already knows what it means — tokens, context windows, hallucinations, agents, fine-tuning. Most of these terms have simple explanations once someone actually defines them. This glossary collects 50 of the terms you're most likely to run into, grouped by theme, each explained in one or two plain-English sentences rather than a technical definition written for engineers. Keep it as a reference to come back to whenever a new term comes up in an article, a product page, or a conversation with a colleague.

Part 1: Foundational concepts

Term	Plain-English meaning
Artificial intelligence (AI)	Broadly, software that performs tasks normally associated with human intelligence — understanding language, recognizing images, making judgment calls. Today the term mostly refers to the current wave of chatbots and generators.
Machine learning	The approach behind most modern AI: instead of a programmer writing exact rules, the system learns patterns from large amounts of example data.
Large language model (LLM)	A machine-learning model trained on huge amounts of text to predict and generate language. ChatGPT, Claude, and Gemini are all built on LLMs.
Generative AI	AI that creates new content — text, images, audio, video, or code — rather than just analyzing or classifying existing content.
Neural network	The underlying computing structure most modern AI models are built from, loosely inspired by how neurons connect in a brain. You rarely need to think about this layer directly as a user.
Foundation model	A large, general-purpose model (like GPT, Claude, or Gemini) trained on broad data, which can then be adapted or built upon for many different specific uses.
Training data	The material a model learned from before you ever interacted with it — text, images, and other content used to teach it patterns.
Parameters	The internal values a model adjusts during training to capture patterns in its training data. Loosely, "more parameters" has historically correlated with more capability, though it's not a strict rule anymore.
Knowledge cutoff	The date after which a model has no built-in knowledge, because its training data ends there. Models can still learn about more recent events if they're able to search the web.

Part 2: How models process language

Term	Plain-English meaning
Token	A chunk of text a model processes — often a word or part of a word. Model pricing, speed, and limits are usually measured in tokens rather than words or characters.
Context window	How much text (your conversation, an uploaded document, and the model's own response) a model can "hold in mind" at once. Once a conversation exceeds it, earlier details start getting pushed out.
Prompt	The instruction, question, or input you give an AI model. Prompt quality is the single biggest lever you control over output quality.

System prompt	A set of instructions, usually set by the developer or platform rather than the end user, that shapes how the model behaves throughout a conversation before you type anything.
Temperature	A setting (usually only visible to developers, not casual users) that controls how random or predictable a model's word choices are — low temperature gives more focused, repeatable answers; high temperature gives more varied, creative ones.
Embedding	A way of converting text (or images) into a list of numbers that captures its meaning, so a computer can compare how similar two pieces of content are. Powers most search and recommendation features behind the scenes.
Inference	The act of a trained model actually generating a response to your prompt — as opposed to "training," which is the earlier process of teaching it in the first place.
Fine-tuning	Taking an already-trained foundation model and training it further on a narrower, specific set of data so it performs better on a particular task or in a particular style.

Part 3: Reliability, safety, and limitations

Term	Plain-English meaning
Hallucination	When an AI states something confidently that isn't true. It happens because the model is predicting plausible-sounding text, not looking up verified facts, unless it's explicitly using a search or document-reading tool.
Bias	Skewed or unfair patterns in a model's output, usually inherited from patterns present in its training data rather than intentionally programmed in.
Grounding	Anchoring a model's answer to a specific, verifiable source (a document you provided, a live web search) rather than letting it rely purely on what it learned during training. Grounded answers are generally more trustworthy.
Retrieval-augmented generation (RAG)	A technique where a model looks up relevant information from an external source (a database, a document set, the web) before answering, rather than relying only on what it memorized during training.
Guardrails	Rules, filters, or checks built around an AI system to keep its behavior within intended limits — for example, preventing it from giving certain kinds of harmful advice.
Alignment	The broader effort to make an AI model's goals and behavior match what its developers and users actually intend, rather than technically satisfying an instruction in an unintended way.

Prompt injection	An attack where malicious instructions are hidden inside content an AI is asked to process (a webpage, a document, an email) to try to hijack its behavior. Relevant mainly for AI tools that browse the web or read external content on your behalf.
Jailbreak	An attempt to trick an AI model into ignoring its safety guidelines through cleverly worded prompts. AI providers continually patch known jailbreak techniques.

Part 4: Talking to AI effectively

Term	Plain-English meaning
Zero-shot prompting	Asking a model to perform a task with no examples provided — just a plain instruction. Works well for most everyday tasks with modern models.
Few-shot prompting	Giving a model a small number of examples of what you want before asking it to do the task itself, which often improves consistency for style- or format-sensitive tasks.
Chain-of-thought	Prompting a model to reason through a problem step by step before giving a final answer, which can improve accuracy on multi-step problems for standard models.
Context engineering	The broader discipline of managing everything a model sees when producing an answer — not just the prompt, but the conversation history, any documents provided, and tool results — increasingly seen as more important than wording a single prompt cleverly.
Multimodal	A model that can work with more than just text — understanding or generating images, audio, or video within the same conversation.
Reasoning model	A newer category of model with a built-in "thinking" step before it answers, aimed at improving performance on complex, multi-step problems. These models generally don't need to be told to "think step by step" since they already do it internally.

Part 5: Automation and agents

Term	Plain-English meaning
Agent	Software that pursues a goal by deciding its own intermediate steps — searching, reading, choosing, acting — rather than following a fixed script written in advance.
Agentic AI	A general term for AI systems built around agents: models that can plan, use tools, and take multi-step action toward a goal with limited human input at each step.

Robotic process automation (RPA)	Software that mimics a person's clicks and keystrokes to move data between existing applications, following the exact same steps every time. It automates *what* a person does, not how they think — the older, more rigid cousin of AI agents.
Workflow automation	Tools that model an entire multi-step business process — approvals, handoffs, conditional branching — and run it automatically from a defined trigger to a defined outcome.
Orchestration	Coordinating multiple AI models, tools, or automated steps together so they work as one system, rather than each piece operating in isolation.
Model Context Protocol (MCP)	An open standard that lets an AI assistant connect to outside tools and data (your files, a database, a calendar) in a consistent way, rather than every AI tool needing a custom, one-off integration with every other piece of software.
Tool use / function calling	A model's ability to call an external tool mid-conversation — running a calculation, searching the web, querying a database — rather than answering purely from its own training.
Human-in-the-loop	A design choice where a person reviews or approves an AI's output or action before it takes effect, rather than the system acting fully on its own.
No-code automation platform	Tools like Zapier, Make, or n8n that let non-programmers build automated workflows by connecting apps visually instead of writing code.

Part 6: Business, access, and technical terms

Term	Plain-English meaning
API (application programming interface)	A defined way for one piece of software to talk to another. When a company "offers an API" for its AI model, it means developers can build the model into their own products.
Open-weight model	A model whose underlying trained parameters are published for anyone to download and run themselves, rather than being accessible only through one company's app or paid service.
Closed / proprietary model	A model only accessible through its maker's own app, website, or paid API — the underlying weights aren't published or downloadable.
Rate limit	A cap on how many requests or how much usage an account can make in a given period, used to manage cost and system load.
Latency	The delay between sending a prompt and receiving a response. Lower latency matters most for real-time or conversational use cases.

Benchmark	A standardized test used to compare different AI models' performance on a specific type of task, allowing rough side-by-side comparison.
Model card	A document a company publishes alongside a model describing its capabilities, limitations, intended uses, and known risks.
Data retention policy	A company's stated rules on whether and how long it keeps the conversations, files, or data you send to its AI tool, and whether that data is used for further training.
Free tier / usage limits	Most AI tools offer a capped amount of free use per day or month, with paid plans removing or raising that cap.
Responsible AI / AI governance	The broader set of practices, policies, and oversight a company or organization puts in place to use AI safely, fairly, and in compliance with relevant regulations.

Closing thought

You don't need to memorize any of this to use AI well — most people pick these terms up naturally through use, the same way anyone learns the vocabulary of a new tool over time. Keep this glossary bookmarked for whenever an unfamiliar term shows up in an article, a product announcement, or a conversation with a more technical colleague, and it'll usually take less than a sentence to clear up.

Let us find the right AI for you!

This glossary should help the jargon make sense whenever you run into it, but understanding the vocabulary is only the first step toward putting AI to work well. If you are ready to go further and apply these ideas to your own business, we are here to help.

Contact us today through email or on our website to schedule your free 30-Minute Introduction Call to see what AI can do for you!

Email: Sales@IntelliTechDynamics.com

Website: www.IntelliTechDynamics.com