GLOSSARY · FOR NEWCOMERS

AI, PLAINLY

Forty words you will hear on the show, each in one plain sentence. Read it once; keep it next to the episodes.

WHAT YOU ARE TALKING TO

The thing that answers, what it reads, and the words for how it was made.

Model	A program that learned patterns from a large amount of data and uses them to produce an answer. The thing you talk to.
Large language model (LLM)	A model trained on text that predicts the next word well enough to write, summarise, answer and reason in language.
Token	The unit a model reads and writes: a word or a piece of one. Prices and limits are counted in tokens.
Context window	How much text a model can hold in mind at once, counted in tokens. Anything outside it is forgotten.
Prompt	What you send the model: the question, the instructions and any material you attach.
System prompt	The standing instructions set by whoever built the product, read before your message. It sets the role and the rules.
Hallucination	A confident answer that is not true. The model fills a gap with something plausible.
Temperature	A setting for how adventurous the model's word choices are. Low is steady, high is varied.
Multimodal	A model that handles more than text: images, audio, video.
Training	The process that builds the model from data. Expensive, done rarely, by the lab.

WHAT IT COSTS, WHERE IT RUNS

The words on the bill, the words on the download page, and how a model finds your documents.

Fine-tuning	Training a finished model a little further on your own examples so it behaves a particular way.
Weights	The numbers inside the model that training produced. "Open weights" means you can download and run them yourself.
Inference	Running the model to get an answer. This is what you pay for, per request.
Latency	How long a request takes to come back.
API	The door a program uses to send a model a prompt and get a reply, instead of a chat window.
API key	The password-like string that lets a program use an API, and bills its owner. Treat it like money.
Local model	A model that runs on your own machine, with no internet needed. BLACKOUT is one.
Quantization	Shrinking a model's numbers so it fits in less memory and runs on smaller hardware, at a small cost in quality.
Embedding	A list of numbers that stands for the meaning of a piece of text, so similar texts land near each other.
Vector database	A store built to search embeddings: find the documents that mean something like this.

HOW IT REACHES THE WORLD

The words for a model that reads your documents, uses tools and takes its own next step.

RAG (retrieval-augmented generation)	Looking up relevant documents first and handing them to the model with the question, so it answers from your material.
Tool use	The ability of a model to call a function: search, send an email, run code, query a database.
Agent	A model given tools and a goal, which decides its own next steps: read this, call that, try again. It acts, not just answers.
Workflow (automation)	A fixed sequence of steps that runs when something happens. Predictable, unlike an agent that chooses.
n8n	A visual tool for building workflows and agents by connecting nodes. The builds on the show use it.
Claude Code	An agent that works inside your code and terminal, run from the command line. Also built with on the show.
Human in the loop	A person who has to approve a step before the agent can take it.
Evaluation (eval)	A test set with known answers, used to measure a model or an agent before and after a change.
Deepfake	Fake video, audio or images made by a model, good enough to pass as real.

WHERE IT GOES WRONG

The security half of the vocabulary. Most episodes of the show live here.

Prompt injection	An attack that hides instructions inside content the model reads (a web page, an email, a file) so it follows the attacker instead of you.
Jailbreak	A prompt designed to talk a model out of its safety rules.
Guardrails	Checks placed around a model to stop certain inputs or outputs. Useful, and never the only layer.
Alignment	Whether a model's behaviour matches what its makers and users actually intend.
Misalignment	When it does not: the model pursues the goal in a way nobody asked for, or a different goal entirely.
Red teaming	Attacking your own system on purpose, to find what breaks before someone else does.
Least privilege	Giving an agent only the access its job needs, and nothing more.
Data exfiltration	Getting data out of a place it should not leave. The most common goal of an attack on an agent.
Supply chain attack	Compromising something you rely on (a package, a tool, a model) rather than attacking you directly.
Zero-day	A flaw nobody has patched yet, because nobody knew it existed.
Kill switch	A way to stop an agent immediately, from outside it. Build it before you need it.