



GLOSSARY

FREE

THE PLAIN-ENGLISH AI GLOSSARY

Fifty terms you will actually meet, defined without the hype.

01

HOW TO USE THIS

Fifty AI terms defined in a sentence or two each, written for the next meeting where someone says RAG and half the room nods anyway. No hype, no doom, no jargon defined with more jargon.

Dated below, reviewed every six months. The field moves, a stale glossary is worse than none.

Agent. An AI set up to complete multi-step tasks on its own, using tools like a browser, files or your business software, rather than just answering a question.

Agentic. Describes AI working that way. An agentic workflow means the AI carries out steps, not just drafts text.

AI slop. Low-quality, mass-produced AI content published without human review. The thing your content should never be mistaken for.

Alignment. The work of making an AI behave the way its makers intend, helpful and within bounds.

API. The plumbing that lets one piece of software talk to another. Most AI features inside business tools are API calls to a model provider.

Assistant. A general-purpose AI you converse with. Distinct from a single-purpose chatbot.

Automation. Rules-based software, when this happens, do that. Not AI on its own, though the two increasingly combine.

Chain of thought. A model working through steps before answering. Often improves results on harder problems.

Chatbot. A scripted or narrow conversational tool, often for support. Many pre-date modern AI and follow decision trees.

Context window. How much text a model can hold in mind at once, measured in tokens. Exceed it and earlier material falls out of scope.

Copilot. Marketing shorthand for an AI assistant embedded inside a product.

Deepfake. AI-generated audio, image or video of a real person doing something they did not do. A fraud and reputation risk, not a novelty.

Deep learning. The family of techniques behind modern AI, built on layered neural networks. You will hear it often and need it rarely.

Embedding. Text turned into numbers that capture meaning, so software can find similar content. The machinery under semantic search.

Few-shot. Giving a model a handful of examples of what good looks like before asking for its own attempt. Examples lift quality more than adjectives do.

Fine-tuning. Training an existing model further on your own examples so it behaves a particular way. Rarely the first tool to reach for, prompting and retrieval get most businesses there first.

Frontier model. One of the most capable current models from the major labs.

Generative AI. AI that produces new content, text, images, audio, video, code, rather than just sorting or ranking existing material.

GPU. The chips AI runs on. The reason AI is expensive to build, and the reason chip supply makes the front page.

Grounding. Tying a model's answer to supplied source material so it draws on given facts rather than memory alone.

Guardrails. Controls around a model that block certain inputs or outputs, from safety filters through to "do not discuss pricing".

Hallucination. A model stating something false with complete confidence, invented citations included. The reason human review exists.

Human in the loop. A process where a person reviews or approves AI output before it counts. The sensible default for anything client-facing.

Inference. Running a trained model to get an answer. Training builds the model, inference uses it. (You will meet the word on pricing pages.)

Knowledge cutoff. The date a model's training data ends. Anything after it, the model cannot know without live retrieval or search.

Latency. How long a model takes to respond.

LLM (large language model). The kind of model behind current AI assistants, trained on vast amounts of text to predict language, which turns out to cover an enormous range of work.

Machine learning. Software that learns patterns from data instead of following hand-written rules. The umbrella over most modern AI.

MCP (Model Context Protocol). An open standard that lets AI assistants connect to other software, your CRM, your files, your calendar, in a consistent way.

Model. The trained system itself. Different models trade capability against speed and cost.

Multimodal. A model that handles more than text, images and audio in, sometimes video, and various formats out.

Neural network. The layered mathematical structure models are built from. Loosely brain-inspired, mostly a term to nod at.

Open weights. A model whose trained parameters are published, so anyone can run or adapt it. Often called open source, not always accurately.

Orchestration. Software that coordinates several AI steps or tools into one workflow.

Parameters. The internal settings a model learns during training, counted in billions. Bigger is not automatically better.

PII (personally identifiable information). Data that identifies a person. The category to keep out of AI tools without a clear, approved reason.

Prompt. What you give the model: instructions, questions, material. Quality in, quality out.

Prompt engineering. Writing prompts deliberately, with structure, examples and constraints, so outputs are reliable. Less mystical than it sounds.

RAG (retrieval augmented generation). The model looks information up, from your documents or a database, then answers using what it found. How assistants answer from your content instead of guessing.

Red teaming. Deliberately attacking an AI setup to find the failures before someone else does.

Semantic search. Search by meaning rather than exact keywords, built on embeddings.

Synthetic data. Data generated by AI to train or test other systems.

System prompt. Standing instructions a model receives before your message, setting its role, tone and rules.

Temperature. A setting for how varied a model's output is. Low is consistent, high is creative.

Token. The unit models read and write, roughly three quarters of a word in English. Pricing and context windows are counted in tokens.

Training. Building a model from data. Enormously expensive, done by the labs, not by you.

Training data. The text and other material a model learned from.

Vector database. A database built to store embeddings and find similar items quickly. The storage layer under RAG and semantic search.

Vibe coding. Building software by describing what you want to an AI and iterating on the result. Powerful for prototypes, still needs proper review before production.

Workflow automation. Chaining tools and triggers so work moves without manual steps. AI increasingly sits inside these chains as one of the steps.

Zero-shot. Asking a model to do a task with no examples at all. Works more often than you would expect, and worse than few-shot.



If you would rather have this done **with** you
than **to** you, this is the work I do.

START A CONVERSATION

PAUL PARSONS

CONSULTANT | DIGITAL MARKETING, AI & MARTECH

0403 127 478

www.paulparsons.com.au

pp@paulparsons.com.au

Perth, Western Australia 6111