

AI GLOSSARY

# Stay up to date with industry terms

## A2A (Agent-to-Agent)

A2A is how autonomous AI agents communicate and collaborate directly, delegating tasks, sharing knowledge, and solving problems together, enabling distributed decision-making and multi-agent workflows without constant human oversight.

---

## Action task

An Action Task is a predefined, trigger-based AI task that executes automatically, handling routine actions like notifications, record updates, or workflow launches instantly, without manual effort.

---

## Adaptive RAG

Adaptive RAG (Adaptive Retrieval-Augmented Generation) is a method that dynamically adjusts how and when external information is retrieved based on the complexity of a query. Instead of using a fixed retrieval process, it determines whether retrieval is needed and selects the appropriate strategy, such as no retrieval, single-step retrieval, or multi-step retrieval. This improves efficiency, reduces unnecessary computation, and ensures better response quality across different query types.

[Read more →](#)

---

## Agent embeddings

Agent Embeddings are vector-based representations of an agent's role, skills, and context, enabling intelligent task routing and matchmaking by helping the system identify the best-suited agent for any goal.

---

## Agent orchestrator

Agent Orchestrator dynamically assigns tasks to the right AI agents based on goal and context, coordinating multi-agent collaboration to ensure smooth interaction, conflict-free execution, and efficient workflow completion.

[Read more →](#)

---

## Agent planner

Agent Planner breaks down high-level goals into executable steps, generating multi-step action plans based on intent, context, and available tools, enabling agents to act with foresight and adapt autonomously.

[Read more →](#)

---

## Agent platform (Kore.ai)

Agent Platform is enterprise-grade infrastructure for building, deploying, and orchestrating sophisticated AI agents at scale, supporting autonomy levels from guided assistants to fully independent systems tailored to business needs.

---

## Agent reasoning

Agentic Reasoning enables AI agents to break down complex goals, make context-driven decisions, learn from outcomes, and self-correct, transforming them from reactive tools into proactive, autonomous problem-solvers.

---

## Agent traceability

Agent Traceability tracks and audits how AI agents make decisions, including model calls, tools used, and contextual inputs, providing transparency that supports governance, compliance, and error resolution.

[Read more →](#)

---

## Agent washing

Agent-washing, per Gartner, is branding simple bots or rule-based tools as Agentic AI without true autonomy, reasoning, or orchestration, misleading buyers with the "agent" label on systems that cannot think, plan, or adapt.

[Read more →](#)

---

## Agentic AI

Agentic AI refers to autonomous AI systems that combine reasoning, planning, memory, and adaptability to make decisions, execute complex tasks across tools, and solve problems proactively with minimal human input.

---

## Agentic applications

Agentic Applications are AI-powered software systems where autonomous agents take actions, make decisions, and adapt in real time using LLMs, computer vision, and reinforcement learning with minimal human guidance.

---

## Agentic memory

Agentic Memory enables AI agents to store and recall short-term context and long-term knowledge, maintaining continuity, personalizing responses, and supporting informed decision-making across complex, multi-step workflows.

[Read more →](#)

---

## Agentic mesh

Coined by McKinsey, Agentic Mesh is a flexible, vendor-agnostic AI architecture where multiple autonomous agents collaborate across tools, systems, and models to make decisions and adapt in real time at enterprise scale.

[Read more →](#)

---

## Agentic orchestration

Agentic orchestration is the process of coordinating AI agents, automation systems, and human inputs across workflows to achieve a defined outcome. It manages how tasks are distributed, executed, and monitored, ensuring agents operate within business rules and governance frameworks. This allows enterprises to run complex, multi-step processes with greater control, adaptability, and visibility.

[Read more →](#)

---

## Agentic RAG

Agentic RAG combines Retrieval-Augmented Generation with agent autonomy, enabling agents to actively decide what to retrieve, how to interpret it, and when to act, supporting context-aware, adaptive multi-step workflows.

[Read more →](#)

---

## Agentic reasoning

Agentic Reasoning enables AI agents to break down complex goals, make context-driven decisions, learn from outcomes, and self-correct, transforming them from reactive tools into proactive, autonomous problem-solvers.

[Read more →](#)

---

## Agentic workflows

Agentic Workflows are dynamic task sequences planned and executed by AI agents with minimal human input, adapting in real time based on context and outcomes to achieve goals efficiently and autonomously.

[Read more →](#)

---

## Agentic X

Agentic X brings reasoning, planning, and autonomy into any application or domain, enabling systems to independently manage complex tasks, adapt on the fly, and coordinate actions without constant supervision.

## AI agents

AI Agents are intelligent software entities that independently break down goals into tasks, execute them using tools or APIs, and deliver results with minimal human input, adapting across systems in real time.

## AI analytics

AI Analytics encompasses tools and dashboards that monitor AI performance across interactions, intent detection, and resolution outcomes, helping businesses assess accuracy, identify bottlenecks, and optimize workflows for business goals.

[Read more →](#)

---

## AI augmentation

Augmentation enriches AI models with external context, real-time data, or tools to produce more accurate, relevant outputs, turning general-purpose models into domain-aware, task-specific assistants without retraining.

[Read more →](#)

---

## AI copilot

AI Copilot is a context-aware assistant that works alongside users, offering real-time suggestions, automating repetitive tasks, and surfacing relevant insights to boost productivity and streamline workflows.

---

## AI for process (Kore.ai)

AI for Process uses process mining, cognitive task modeling, and reinforcement learning to optimize workflows, manage exceptions, and enforce compliance, enabling agents to autonomously adapt execution based on real-time data.

---

## AI for service (Kore.ai)

AI for Service integrates agentic AI, multi-modal NLP, and adaptive reasoning to automate customer interactions across voice and digital channels, supporting intent-driven automation, agent augmentation, and scalable omnichannel self-service.

---

## AI for work (Kore.ai)

AI for Work is an enterprise productivity framework using context-aware AI agents for knowledge retrieval, task automation, and workflow optimization, enabling semantic reasoning and decision intelligence across business functions.

---

## AI governance

AI governance refers to the frameworks, policies, and controls used to manage how AI systems are developed, deployed, and monitored. It ensures that AI operates in a secure, compliant, and ethical manner by addressing risks such as bias, data privacy, and accountability. AI governance provides oversight and auditability, which are critical for enterprise adoption.

---

## AI hallucination

A hallucination occurs when AI generates confident but factually incorrect output, leading to misleading answers or flawed actions. Grounding and validation are essential to keeping responses accurate and reliable.

[Read more →](#)

---

## AI safety

AI Safety is the practice of designing AI systems to operate securely, ethically, and aligned with human values, preventing bias, misuse, and unintended actions through governance, monitoring, and human oversight.

[Read more →](#)

---

## AI simulation

AI Simulation uses synthetic environments to train and test AI models safely, allowing agents to learn through trial and error, explore complex scenarios, and refine behaviors without real-world consequences.

---

## AI supercomputing

AI Supercomputing is high-performance infrastructure built to train and run large language models and generative AI workloads at scale, delivering the speed needed for complex reasoning and real-time enterprise inference.

---

## AI TRiSM

AI TRiSM, coined by Gartner, is a framework for managing AI trust, risk, and security, ensuring model governance, fairness, reliability, robustness, and data protection throughout the AI lifecycle.

[Read more →](#)

---

## Alert task

Alert Task is an AI-triggered response to anomalies or threshold breaches, instantly notifying the right people or systems about suspicious activity, errors, or performance drops without manual monitoring.

---

## Anthropomorphism

Anthropomorphism is attributing human traits like emotions or consciousness to AI systems. While making interactions feel natural, it creates false expectations by blurring the line between what AI appears to do and what it truly understands.

[Read more →](#)

---

## API

API is a set of rules and protocols enabling AI systems to communicate with other software, apps, or databases, acting as a bridge that allows data and actions to flow between tools automatically.

---

## Artificial General Intelligence (AGI)

Artificial General Intelligence is theoretical AI capable of understanding, learning, and applying knowledge across any task like a human, unlike today's specialized systems. It remains the long-term ambition of the AI field.

[Read more →](#)

---

## Artificial Intelligence (AI)

Artificial Intelligence is a branch of computer science focused on creating machines that mimic human reasoning, learning, and decision-making, enabling automation that adapts and improves over time without explicit programming.

---

## Auto-NLP

Auto-NLP is a toolkit that automates key natural language processing tasks like text classification, sentiment analysis, and intent detection with minimal manual setup. It's ideal for teams who need fast, reliable NLP results without building custom pipelines from scratch.

[Read more →](#)

---

## Automated Speech Recognition (ASR)

ASR is technology that converts spoken words into written text in real time, powering voice input in apps, IVR systems, and virtual assistants to enable machines to understand and respond to human speech.

---

## Autonomous agents

Autonomous Agents are AI systems that independently plan, act, and learn to achieve goals without human direction, breaking down tasks, making real-time decisions, and adapting to handle complex, multi-step processes.

---

## Autoregressive model

An autoregressive model is a machine learning technique that predicts the next value in a sequence by learning from the values that came before it. The core idea is simple: what happens next is influenced by what has already happened. By identifying patterns in historical data, the model can make informed predictions about what comes next.

[Read more →](#)

---

## Basic RAG

Basic RAG enhances language models by retrieving relevant information from external sources to generate more accurate, grounded responses, supplementing pre-trained knowledge with real-time or domain-specific context.

---

## **Benchmark**

A benchmark is a standardized evaluation that measures AI model performance on tasks like reasoning or language understanding, enabling consistent comparison, progress tracking, and identification of strengths and weaknesses.

---

## **BM25**

BM25 is a keyword-based retrieval algorithm that ranks documents by matching search queries using term frequency and document length, making it fast and effective for classic information retrieval tasks.

---

## **Chain of Thought (CoT) prompting**

Chain of Thought Prompting encourages AI models to reason through problems step by step before answering, rather than jumping to conclusions, producing more accurate and explainable outputs.

---

## **Chunking**

Chunking breaks large documents into smaller, meaningful pieces so AI can understand and retrieve them more efficiently, making RAG and enterprise search smarter and faster.

---

## **Cloud connector**

A Cloud Connector is a plug-and-play integration linking AI systems to third-party cloud apps like CRMs or databases, enabling seamless data exchange and real-time actions without custom code or middleware.

---

## **Cognitive services**

Cognitive Services are pre-built AI capabilities for tasks like speech recognition, image analysis, and language translation, enabling teams to add intelligent, human-like abilities to applications without building models from scratch.

---

## **Component reusability**

Component Reusability means building AI elements like intents, prompts, or connectors that can be reused across agents or applications, speeding up development, ensuring consistency, and eliminating redundancy when scaling AI.

---

## Composable AI

Composable AI builds AI capabilities as modular, reusable blocks, including agents, tools, and workflows, enabling enterprises to scale, customize, and adapt AI for various use cases without starting from scratch.

---

## Confidence score

A Confidence Score indicates how certain an AI is about its prediction or response, typically shown as a percentage, helping determine whether to proceed, seek clarification, or escalate to a human.

---

## Context engineering

Context Engineering is designing AI systems to capture and apply real-world context like user identity, intent, and history, enabling agents to respond more intelligently, personalize actions, and handle complex workflows accurately.

---

## Context router

A Context Router directs requests to the right agent, model, or workflow using contextual signals like user intent, system state, and confidence levels, ensuring accurate, relevant responses without unnecessary handoffs.

---

## Context window

The Context Window is the amount of information a language model can process at once, measured in tokens. Larger windows enable more coherent, context-aware outputs across long conversations, document summarization, or multi-turn reasoning.

---

## Contextual embedding

Contextual Embedding transforms words or data into numerical vectors that capture meaning based on surrounding context, helping AI distinguish nuances like "bank" as a financial institution versus a riverbank for smarter retrieval and reasoning.

---

## Contextual intelligence engine

A Contextual Intelligence Engine collects and analyzes signals like user roles, conversation history, and business rules, giving AI agents memory, situational awareness, and the ability to make smarter, adaptive decisions.

---

## Controllability

Controllability refers to how well AI behavior can be guided or constrained through boundaries on responses, tool usage, and tone, ensuring alignment with business rules, compliance standards, and safety in enterprise settings.

---

## Conversational AI

Conversational AI enables machines to interact with humans using natural language across text, voice, and messaging, combining language understanding, intent recognition, and dialogue management to automate support and execute tasks.

---

## Conversational UI

A Conversational UI enables user interaction through natural language instead of buttons or forms, making experiences more intuitive across chat interfaces and voice assistants, especially for complex or dynamic tasks.

---

## Data augmentation

Data Augmentation expands training data by generating or tweaking existing examples, like rephrasing sentences or adding noise, improving model robustness and performance without collecting large amounts of new data.

---

## Data preprocessing

Data Preprocessing cleans and formats raw data by removing errors and standardizing text, ensuring AI models receive structured, consistent inputs they can effectively learn from.

---

## Data retention

Data Retention defines how long user and system data is stored before deletion, ensuring compliance with privacy laws like GDPR and HIPAA and preventing sensitive information from being kept longer than necessary.

---

## Deliberation engine

A Deliberation Engine enables AI agents to pause and evaluate options before acting, choosing the most effective path rather than defaulting to the first available option, especially useful in complex workflows.

---

## Dense retrieval

Dense Retrieval uses vector embeddings to find semantically similar information rather than matching keywords, enabling more relevant search results in RAG systems and enterprise search applications.

---

## Deployment

Deployment is when an AI system goes live for real users, whether integrated into a chatbot, voice assistant, or internal tool, marking the point where performance and reliability truly matter.

---

## Deterministic model

A Deterministic Model always produces the same output for the same input, making it ideal for consistency and traceability in legal, financial, or safety-critical workflows where predictability matters most.

---

## Dialog builder

A Dialog Builder is a visual workspace for designing conversation flows using drag-and-drop tools, enabling teams to build smart, functional bots without requiring developer involvement at every step.

---

## Dialog task

A Dialog Task is a guided conversation path designed to complete a specific goal, like scheduling a meeting or checking an order, connecting user inputs to backend systems through logical, action-driven steps.

---

## DialogGPT (Kore.ai)

Dialog GPT is an intelligent orchestration engine that powers natural, multi-turn conversations by autonomously managing intent detection, task execution, and ambiguity resolution, combining generative models with domain knowledge at scale.

---

## Domain-Specific Language Model (DSLML)

A DSLML is a language model fine-tuned for a specific industry like healthcare or banking, understanding domain-specific terminology and context to deliver smarter, more relevant responses.

---

## Edge AI

Edge AI runs directly on local devices like phones or kiosks rather than in the cloud, enabling fast, private, real-time decisions even in environments with limited connectivity.

---

## Embedding models

Embedding Models convert language into numerical vectors, enabling AI to make connections based on meaning rather than keywords, powering smarter search, retrieval, and reasoning across applications.

---

## Embeddings

Embeddings are numerical representations of words or data that capture meaning, enabling AI to find similar content, rank search results, and maintain context across multi-step tasks.

---

## Encryption

Encryption scrambles data so only authorized parties can access it, protecting passwords, personal information, and business data as AI systems store and transfer it.

---

## Enterprise AI

Enterprise AI is AI built for business, designed to operate securely and at scale across departments and systems, handling real business logic while maintaining compliance and accountability.

---

## Enterprise RAG

Enterprise RAG combines intelligent retrieval with LLM-generated responses, pulling answers from internal knowledge bases and documents while ensuring accuracy, brand alignment, security, and full traceability.

---

## Enterprise search

Enterprise search is the ability to retrieve and surface relevant information from across an organization's data sources, including documents, databases, and applications. It uses indexing, retrieval models, and often AI to deliver accurate and context-aware results. Enterprise search helps employees access knowledge quickly, improving productivity and decision-making.

---

## Entity

An entity is a specific piece of information the AI is trying to extract like a person's name, a date, or an account number. Think of it as a key detail that makes a vague request actionable.

---

## Entity extraction

Entity Extraction identifies and pulls key details from user inputs, like "March" and "invoice" from a sentence, helping AI accurately route tasks and understand what the user needs.

---

## Ethical AI

Ethical AI means building systems that are fair, responsible, and aligned with human values, avoiding harmful bias, protecting privacy, and ensuring AI is deployed in ways that go beyond just business goals.

---

## Explainable AI (XAI)

Explainable AI reveals why an AI made a decision rather than leaving it opaque, building trust in high-stakes industries like finance and healthcare where transparency and accountability are essential.

---

## FAQ

In AI, an FAQ refers to pre-trained question-answer pairs used by virtual assistants to deliver fast, accurate responses to common queries without requiring full conversations or complex workflows.

---

## Federated learning

Frontier Models are the most advanced AI systems available, pushing the boundaries of reasoning, planning, and autonomy. Typically massive and multimodal, they are often in research or tightly controlled release.

---

## Few-Shot learning

Few-shot learning enables AI models to understand new tasks from just a handful of examples provided in the prompt, eliminating the need for retraining and offering speed and flexibility at scale.

---

## Fine-Tuning

Fine-tuning trains a general AI model on specific data to align it with a particular tone, vocabulary, or industry, making it more accurate and relevant for targeted use cases.

---

## Foundation models

Foundation Models are large, general-purpose AI models trained on massive datasets, adaptable to many tasks like summarization, question answering, or classification through fine-tuning or prompting.

---

## Frontier models

Frontier Models are the most advanced AI systems available, pushing the boundaries of reasoning, planning, and autonomous action. Typically massive and multimodal, they remain in research or tightly controlled release.

---

## Generative AI

Generative AI refers to systems that create content like text, images, or code by learning patterns from data, generating new, dynamic output in real time rather than selecting from pre-set options.

---

## GPT (Generative pre-trained transformer)

GPT is a family of generative language models pre-trained on massive datasets, capable of understanding and generating human-like text for applications ranging from chatbots to summarization and AI agents.

---

## Graph of Thoughts (GoT)

Graph of Thoughts (GoT) is an advanced reasoning framework that extends structured thinking beyond linear or tree-based approaches. It represents reasoning as a graph, allowing multiple interconnected paths and dependencies between ideas. This enables more flexible and efficient exploration of complex problems where relationships between steps are not strictly hierarchical.

---

## Graph-RAG

Graph-RAG combines retrieval-augmented generation with knowledge graphs, understanding relationships between data points rather than pulling isolated chunks, improving reasoning, context, and relevance in generated answers.

---

## Grounding

Grounding ensures AI agent outputs are based on trusted sources like enterprise documents or real-time data rather than guesswork, giving agents a reliable foundation for factual, relevant, and safe responses.

---

## Guardrails framework

Guardrails Framework sets boundaries around what AI can say or do, ensuring outputs are safe, compliant, and on-brand by blocking certain content, guiding tone, or restricting tool access.

---

## Human in the loop

Human in the Loop keeps a person involved in AI decision-making for oversight, approvals, or intervention, balancing automation with control in workflows where accuracy, judgment, or compliance matter.

---

## Hybrid search

Hybrid Search combines keyword-based and semantic search to retrieve both exact matches and meaning-based results, delivering more relevant, complete answers especially for open-ended or complex queries.

---

## Hyperparameter tuning

Hyperparameter tuning optimizes settings that control how an AI model learns, like learning rate or model size, improving accuracy, speed, and reliability without changing the model's core architecture.

---

## In-Context Learning (ICL)

In-Context Learning enables AI models to understand and handle new tasks by reading examples within the prompt, without retraining, making it ideal for custom tasks and dynamic use cases.

---

## Indexing

Indexing organizes and stores data so AI can quickly search and retrieve it, ensuring documents, transcripts, or knowledge articles are structured for fast, accurate information retrieval.

---

## Ingestion

Ingestion is the process of importing external data like documents, PDFs, or knowledge base articles into an AI system, making content searchable, retrievable, and usable in conversations or workflows.

---

## Instruction-Tuning

Instruction-tuning trains AI models to follow human instructions effectively, teaching them to respond in expected ways, whether answering clearly, summarizing concisely, or taking action when asked.

---

## Intent

An intent is what a user wants to accomplish, like resetting a password or checking a balance. Accurately detecting intent enables AI to determine the right response, action, or workflow to trigger.

---

## Joint learning

Joint learning trains multiple AI tasks or models simultaneously, sharing knowledge across tasks like intent detection and entity recognition to improve accuracy in complex systems like virtual assistants.

---

## Knowledge base

A Knowledge Base is a centralized repository of information FAQs, how-to articles, documents, and internal guides that the AI can use to answer questions or support tasks. It's like the AI's internal library, helping it respond with consistent, approved information.

---

## Knowledge graphs

Knowledge Graphs organize information into connected nodes and relationships, like a map of how concepts, entities, and data points relate to each other. This helps the AI reason more intelligently, so it understands how they connect.

---

## Knowledge task

A knowledge task is when AI finds, understands, and delivers information like answering policy questions or summarizing documents. It relies on connecting to the right sources, retrieving relevant content, and presenting it clearly for accurate, helpful outcomes.

---

## LangOps

LangOps (Language Operations) is the practice of managing and optimizing how language models are deployed and used across the enterprise. It includes performance tuning, governance, training data management, and model versioning essentially DevOps for LLMs.

---

## Large Language Model (LLM)

A Large Language Model (LLM) is an advanced AI trained on massive text datasets to understand, process, and generate human language. It analyzes queries, summarizes, completes tasks, and adapts to contexts, delivering coherent, accurate, and scalable language-driven output.

---

## LLM Orchestration

LLM orchestration manages how large language models interact with tools, memory, APIs, and agents. It ensures models don't just generate text but operate as part of a system that can reason, retrieve, act, and adapt across workflows.

---

## LLMOps

LLMOps (large language model operations) refers to the practices and tools used to deploy, manage, and monitor large language models in production. It covers the full lifecycle, including data preparation, model tuning, deployment, evaluation, and ongoing optimization. LLMOps ensures that AI systems remain reliable, scalable, and compliant while maintaining performance over time.

[Read more →](#)

---

## Long-Term memory

Long-term memory allows AI agents to remember information across interactions like user preferences, past actions, or previous answers. It helps make responses more personalized, consistent, and goal-aware over time.

---

## Low-Code

Low-code platforms let users build AI-powered applications or automations using visual interfaces instead of traditional coding. It helps business users and non-engineers create workflows, bots, or integrations quickly and safely.

---

## Low-Rank Adaptation (LoRA)

LoRA is a technique for fine-tuning large models efficiently, without needing to retrain the whole thing. It makes updates lighter, cheaper, and easier to deploy-perfect for customizing foundation models in enterprise settings.

---

## MCP

Model Context Protocol (MCP) is an open standard that enables AI applications to connect with external data sources, tools, and systems through a unified interface. It standardizes how models access context, execute actions, and exchange information, allowing AI systems to operate beyond their training data. MCP simplifies integration, improves interoperability, and supports scalable, context-aware AI applications across enterprise environments.

[Read more →](#)

---

## Memory

Memory allows AI systems to retain and reuse information over time, like past interactions, user preferences, or task history. It helps the AI stay context-aware, make better decisions, and maintain continuity across conversations or workflows.

---

## Model router

A Model Router decides which AI model to use for a specific task. Based on factors like prompt type, confidence score, or domain, it directs requests to the best-fitting model ensuring the system stays efficient, accurate, and scalable.

---

## ModelOps

ModelOps is the practice of managing the full lifecycle of AI models, from training and testing to deployment, monitoring, and retirement. It's essential for keeping models secure, updated, and aligned with business needs over time.

---

## Multi-Agent orchestration

Multi-agent orchestration is the coordination of multiple specialized AI agents working together to complete complex tasks. Each agent focuses on its part retrieving data, executing actions, or reasoning and the orchestration layer ensures everything flows smoothly.

[Read more →](#)

---

## Multi-Agent Systems

A multi-agent system is a setup where several autonomous AI agents collaborate, communicate, and share context to solve a broader goal. It's like a digital team, each agent with its own role working towards the same objective.

---

## Multi-Vector Search

Multi-vector search improves retrieval by using more than one semantic representation to find relevant information. It helps surface better results by capturing different meanings or perspectives behind a single query.

---

## Multimodal AI

Multimodal AI refers to systems that can understand and process more than one type of input, like text, images, audio, or video. It enables richer, more flexible interactions across a wider range of tasks and channels.

---

## Natural Language Generation (NLG)

NLG is the process of turning structured data or internal knowledge into clear, human-sounding language. Whether it's summarizing a report or answering a user question, NLG helps AI systems respond naturally and intelligently in real time.

---

## Natural Language Processing (NLP)

Natural Language Processing is the broad field of AI that helps machines understand, interpret, and work with human language. It covers everything from analyzing text to extracting meaning, enabling systems to handle unstructured input like messages, emails, or voice commands.

---

## Natural Language Understanding (NLU)

NLU is a subset of NLP focused on interpreting meaning and intent behind language. It enables AI systems to understand user goals, even with vague phrasing, driving accurate, context-aware responses and reliable actions.

---

## No-Code

No-code platforms let users build AI applications, workflows, or automations without coding. Using visual interfaces like drag-and-drop tools, they empower business teams to quickly launch and manage solutions, saving time and reducing reliance on deep technical expertise.

---

## Omni-Channel

Omni-channel means delivering a seamless AI experience across chat, voice, email, web, and mobile, maintaining context and continuity. It ensures consistent support, letting users resume interactions smoothly across channels, enhancing convenience, reliability, and customer satisfaction.

---

## Ontology

An ontology is a structured framework for organizing knowledge, defining relationships among concepts, entities, and categories within a domain. In AI, it enables context-aware reasoning by showing how things connect, improving accuracy, consistency, and meaningful responses.

---

## Open-Source LLMs

Open-source LLMs are large language models that are freely available for anyone to use, customize, or deploy. They offer flexibility and transparency, making them a strong option for enterprises that want control over model behavior, cost, or deployment environment.

---

## Parameters

Parameters are internal values a language model learns during training. They control how the model interprets language, forms associations, and generates responses. In simple terms, more parameters generally mean the model can capture more complexity, but also requires more computation.

---

## Pre-Trained Model

A pre-trained model is an AI system trained on large datasets that can be fine-tuned or used directly for specific tasks. It saves time and resources by offering a solid foundation that can be adapted quickly to new use cases.

---

## Probabilistic Model

A probabilistic model makes decisions or predictions based on the likelihood of different outcomes. Instead of producing one "correct" answer, it weighs possibilities and selects the most likely one, making it useful for language, reasoning, and uncertain scenarios.

---

## Prompt Chaining

Prompt chaining links multiple prompts together using the output of one as the input to the next to guide the model through multi-step reasoning or tasks. It helps break down complex problems into manageable steps for more reliable outcomes.

---

## Prompt Engineering

Prompt engineering is the art of crafting instructions that guide an AI model's behavior. How a prompt is written can shape the tone, format, and accuracy of the response, making it a tool for improving results without retraining the model.

---

## Prompt Pipelines

Prompt pipelines are structured sequences of prompts, logic, and decision steps that together drive a larger task. Think of them as reusable flows where each step builds on the last, helping AI systems complete end-to-end actions more reliably.

---

## Query optimization

Query optimization involves refining a query to make it more efficient, precise, or context-aware so the AI retrieves the best possible answers faster. This could include rephrasing, ranking priorities, or eliminating unnecessary noise in the input before processing it.

---

## Reasoning

Reasoning is the AI's ability to think through a problem, break it into steps, and make informed decisions. It's what separates reactive bots from intelligent agents that can handle ambiguity, follow goals, and adapt in real time.

---

## Reinforcement learning

Reinforcement learning is a method where AI learns by trial and error, getting rewarded for good outcomes and penalized for bad ones. It's useful for training agents to improve over time in dynamic or goal-driven environments.

---

## Reinforcement Learning from Human Feedback (RLHF)

RLHF combines reinforcement learning with human guidance. Instead of just learning from rules, the AI improves by watching how humans rate or correct its outputs, leading to responses that better match expectations and values.

---

## Responsible AI

Responsible AI means building and deploying AI systems that are ethical, transparent, fair, and aligned with human values. It covers things like avoiding bias, respecting privacy, and making sure decisions can be explained and trusted.

---

## Retrieval-Augmented Generation (RAG)

RAG is an AI method combining retrieval, augmentation, and generation fetching trusted external data, enriching user queries with context, and producing grounded, accurate responses that reduce hallucinations, enhance transparency, and improve reliability for enterprise use cases requiring precision and traceability.

---

## Robotic Process Automation (RPA)

RPA automates repetitive tasks using bots that mimic human actions, like clicking buttons or copying data between systems. While powerful for rule-based tasks, it lacks the reasoning and flexibility of agentic AI, which can adapt to changing goals and context.

---

## Role-Based Access Control (RBAC)

RBAC restricts access to features or data based on a user's role, like admin, agent, or end user. It's essential in enterprise AI platforms for protecting sensitive information and enforcing security policies across teams.

---

## Scaffolding

Scaffolding is a technique where a complex task is broken into smaller steps that the AI can reason through, often using intermediate prompts, models, or agents. It's helpful for multi-step reasoning, planning, and decision-making.

---

## Search and Data AI (Kore.ai)

Search and Data AI is Kore.ai's intelligent framework for enterprise knowledge discovery. It brings together agentic RAG, semantic understanding, multi-source connectors, and hybrid vector search to turn scattered internal data, like documents, databases, or web content, into context-rich answers.

---

## Self-RAG

Self-RAG (Self-Reflective Retrieval-Augmented Generation) is an approach where an AI system evaluates and refines its own retrieval and generation process during response creation. It retrieves information, assesses the relevance and quality of that information, and iteratively improves the output by validating or correcting its reasoning. This self-evaluation loop helps improve accuracy, reduce hallucinations, and ensure that responses are grounded in reliable data.

---

## Semantic search

Semantic search goes beyond keywords to understand the meaning behind a query. It helps AI systems find relevant content, even if the wording doesn't exactly match, by looking at intent, context, and relationships between concepts.

---

## Sentiment analysis

Sentiment analysis helps AI understand emotions behind text, whether it's positive, negative, or neutral. It's useful in support, marketing, and feedback systems to assess customer tone and urgency.

---

## Sequence modeling

Sequence modeling is the process of analyzing or predicting patterns in ordered data, like sentences, clickstreams, or time-series events. It's essential for tasks where the order of information affects the outcome, such as language processing or behavior prediction.

---

## Short-Term memory

Short-term memory stores recent inputs, decisions, or conversational context that an agent uses during an active session. It helps the system stay coherent and relevant within a task, without mixing it up with long-term data or unrelated past interactions.

---

## Small Language Models (SLMs)

Small Language Models are compact AI models trained for specific tasks or domains. They're faster, more cost-effective, and easier to control than massive models, making them ideal for use cases that require speed, privacy, or domain precision

---

## Software Development Kit (SDK)

An SDK is a collection of tools, libraries, and documentation that helps developers build or extend AI applications. It provides everything needed to integrate with APIs, build custom features, or embed AI into enterprise workflows.

---

## Sparse retrieval

Sparse retrieval relies on traditional keyword matching methods to retrieve content. It's fast and effective for exact matches, but often less flexible than semantic search when queries are vague or varied.

---

## Supervised learning

Supervised learning is when an AI model is trained using labeled data, examples where the input and correct output are known. It's widely used for tasks like classification, prediction, and intent recognition.

---

## Synthetic data generation

Synthetic data generation involves creating artificial data, like text, images, or records, to train or test AI models. It's especially useful when real data is limited, sensitive, or needs to be balanced for fairness.

---

## Temperature

Temperature is a setting that controls how "creative" the AI gets. A low temperature keeps responses focused and predictable. A higher one makes answers more diverse, but sometimes less accurate.

---

## Testing

Testing is how we make sure AI works as expected before it goes live. It includes checking accuracy, behavior, and edge cases, so there are no surprises when customers or teams start using it.

---

## Tokens

Tokens are the chunks of text that an AI model reads or writes, like words or parts of words. The more tokens you give the model, the more context it has to work with.

---

## Tool calling

Tool calling is the ability of an AI model to interact with external tools, APIs, or systems to retrieve data or perform actions. It allows models to go beyond static responses by executing tasks such as fetching real-time information or triggering workflows. This capability enables AI systems to operate as active participants within applications rather than only generating text.

[Read more →](#)

---

## Toxicity

Toxicity is when AI says something harmful, offensive, or inappropriate. It's not always intentional, it's just repeating patterns it has seen. That's why filters and safeguards are used to catch and prevent it from showing up in responses.

---

## Training data

Training data is what an AI system learns from. It could be text, documents, conversations, anything that teaches the model how language works and what to expect. The better the training data, the smarter and more accurate the AI becomes.

---

## Transfer learning

Transfer learning is when an AI takes what it learned from one task and uses it for another. It's like reusing knowledge, saving time, effort, and making the model smarter, faster.

---

## Transformer

The transformer is a type of AI model architecture that made today's powerful language models possible. It helps the AI understand how words relate to each other in a sentence so it can generate responses that make sense.

---

## Transparency

Transparency means you can see and understand how an AI system came to its answers. It helps build trust, especially in business settings where decisions need to be explained, tracked, and improved over time.

---

## Tree of Thoughts (ToT)

Tree of Thoughts (ToT) is a reasoning approach where an AI model explores multiple possible solution paths in a structured, tree-like format before selecting the best outcome. Instead of generating a single response, the model evaluates different intermediate steps and branches. This improves problem-solving for complex tasks that require planning and multi-step reasoning.

---

## Unstructured data

Unstructured data is information that doesn't follow a fixed format, like emails, chats, PDFs, images, or audio files. It's messy but rich with insights, and AI systems are designed to make sense of it by extracting meaning, context, and intent.

---

## Unsupervised learning

Unsupervised learning is when AI is trained on data without labels. It learns by spotting patterns, clusters, or relationships on its own, which is useful for organizing data or discovering hidden insights without manual setup.

---

## Vector databases

A vector database is where those embeddings are stored and searched. Instead of matching text directly, it compares how similar the ideas behind different pieces of content are making search more accurate, especially for open-ended queries.

---

## Vector search

Vector search finds information based on meaning, not just keywords. It compares "embeddings", numerical representations of content, to return the most relevant results, even when the user's words don't exactly match the document.

---

## Zero-Shot learning

Zero-shot learning lets an AI handle tasks it hasn't been explicitly trained on just by understanding the instructions. It's like giving the model a prompt and having it figure things out without needing examples or retraining



**kore.ai**