

Glossary of Terms

CONTENTS

[A](#)
[B](#)
[C](#)
[D](#)
[E](#)
[F](#)
[G](#)
[H](#)
[I](#)
[J](#)
[K](#)
[L](#)
[M](#)
[N](#)
[O](#)
[P](#)
[Q](#)
[R](#)
[S](#)
[T](#)
[U](#)
[V](#)
[W](#)
[X](#)
[Y](#)
[Z](#)

A

Accuracy

A measure that describes the proportion of correct and relevant results among the total number of tests conducted. In some cases, a response may be partially accurate and any accuracy measures should make it clear how accuracy is determined. AI systems may also have different measures of Accuracy for different Use Cases.

Agents

In the context of “Agentic AI”, agents refer to autonomous software components designed to perform specific tasks and make decisions without constant human intervention. There is a wide spectrum of agents and they are key to Agentic AI systems, which are capable of autonomous action, decision-making, and continuous learning from interactions.

Agentic AI

A type of AI system that uses one or more “agents” and that can make decisions, initiate workflows, and otherwise act autonomously within defined parameters with limited human supervision.

AI Benchmarking

The process of evaluating and comparing the performance of different AI systems or models to identify which performs best for specific tasks.

AI Ethics

The issues that AI stakeholders such as engineers, businesses and governments must consider to ensure the technology is developed and used responsibly. This means adopting and implementing systems that support a safe, secure, unbiased, and environmentally friendly approach to artificial intelligence.

Algorithm

A sequence of rules that a computer uses to complete a task. An algorithm takes an input (e.g., a dataset) and generates an output (e.g., a pattern that it has found in the data).

Algorithmic Bias

Unfairness that can arise from problems with an algorithm’s process or the way the algorithm is implemented, resulting in the algorithm inappropriately privileging or disadvantaging one group of users over another group. Algorithmic biases often result from biases in the data that has been used to train the algorithm, which can lead to the reinforcement of systemic prejudices around race, gender, sexuality, disability or ethnicity.

Application Prompt

Could be considered as a specific type of “System Prompt” tailored for a particular application or use case. It provides additional context or instructions specific to the task at hand, like “System Prompts” but with a narrower focus.

Artificial Intelligence

The design and study of machines that can perform tasks that would previously have required human (or other biological) brainpower to accomplish. AI is a broad field that incorporates many different aspects of intelligence, such as reasoning, making decisions, identifying and classifying things, learning from mistakes, communicating, solving problems, and moving around the physical world.

Auditability

The ability to inspect and review the processes and outcomes of an AI system to ensure compliance with legal standards and ethical considerations. This includes being able to trace decisions back to their source data and logic, which is currently difficult with existing large language models.

Automated decision-making

A process where decisions are made by machines or software algorithms without human intervention.

B

Backpropagation

A method used in training neural networks by adjusting weights based on the error from the output.

Black Box

A system where the internal workings or decision-making processes are not transparent or easily understood.

C

Chatbot

A software application that is designed to imitate human conversation through a text, voice or video interface.

Chain of Thought

A process of breaking down a complex query or task into simple steps, thereby allowing for improved logical reasoning, interpretability/explainability, and accuracy.

Completions

The generated responses or outputs from a language model when given a prompt. For example, in response to the prompt "What is the capital of France?", the completion would be "Paris" or in relation to the prompt "Mary had a little" the completion would be "lamb".

Cross-Modal Learning

A learning approach in AI where the model is trained to understand and relate information across different modalities, such as linking text with corresponding images or audio. This capability enables the model to draw connections and transfer knowledge between modalities, enhancing its ability to perform tasks that require multi-modal understanding and interaction.

D

Data Bias

Data bias refers to bias reflected in the data used to train machine learning (ML) models. Data bias can lead to ML models being trained to generate biased outputs and predictions. There are several potential sources of data bias including incomplete data, and data that reflects societal bias.

Data Cleaning

A step in preparing the data used to train a machine learning (ML) model. Data cleaning involves identifying and correcting errors in the data. For example, fixing typing errors and removing duplicates in text data.

Data Scraping

The automated process of extracting and collecting information from websites or other digital sources for analysis or to train and test models.

Dataset

A collection of data that AI systems use to learn, test, or validate outputs, decisions and predictions.

Deep Learning

A form of machine learning that uses computational structures known as 'neural networks' to automatically recognise patterns in data and provide a suitable output, such as a prediction or evidence for a decision. Deep learning neural networks are loosely inspired by the way neurons in animal brains are organised, being composed of multiple layers of simple computational units ('neurons').

E

Embedding

In the context of AI, embeddings are a way to represent words, phrases, or other pieces of data in a Vector Space. This allows AI systems to understand and process the semantic relationships between different pieces of information. By converting legal sources and documents into embeddings, AI systems can more effectively analyse contracts, monitor regulations, draft legal documents, and perform other tasks that require a deep understanding of legal terminology.

Error Rate

The proportion of all false or irrelevant outputs or predictions out of the total predictions made by the AI system, indicating the frequency of errors.

Ethical AI

The practice of designing, developing, and deploying AI systems in a manner that aligns with widely accepted ethical principles, such as fairness, accountability, and respect for privacy.

Explainability

The degree to which the internal mechanics of an AI or machine learning model can be explained in human terms. This is crucial for understanding how decisions are made.

Evaluations

Also known as customer evaluations (or 'evals'), these are systematic tests designed to measure how well an LLM performs on your specific use case. They serve as a critical bridge between the generalized capabilities of LLMs and the unique demands of your business application

F

F1-Score

A measure that combines Precision and Recall into a single metric by taking their harmonic mean, providing a single score to assess a model's accuracy.

Few-shot Learning

A machine learning approach where a model is trained to recognise patterns or make decisions based on a very limited set of labelled data examples.

Few-shot Prompting

Few-shot prompting is a technique in which an AI model is given a few examples of a task to learn from before generating a response, using those examples to improve its performance on similar tasks.

Fine-tuning

The process of taking a pre-trained model (a model trained on a large dataset) and training it further on a smaller, specific dataset to adapt it for a particular task or to align with certain preferences.

Foundation Model

A large-scale, pre-trained machine learning model that serves as a base and can be fine-tuned or adapted for various specific tasks or applications (eg; using prompt engineering or RAG).

G

Generalisation

The ability of a machine learning model to produce accurate predictions or outputs using new, unseen data.

Guardrails

Restrictions and rules placed on AI systems to make sure they handle data appropriately and don't generate unethical content.

H

Hallucination

hallucination refers to instances where an AI system generates content that is factually incorrect, nonsensical, or entirely fabricated, despite appearing coherent and plausible at first glance. This phenomenon occurs because AI models, particularly large language models, generate responses based on statistical probabilities derived from their training data, rather than true understanding or reasoning capabilities. In legal applications, hallucinations can lead to significant risks, such as the introduction of incorrect legal information, misinterpretation of case law, or the inclusion of nonsensical clauses in legal documents.

Human in The Loop (HITL)

A collaborative approach that combines human expertise and input with artificial intelligence (AI) and machine learning (ML) systems.

Hyperparameter

A parameter, or value, which affects the way an AI model learns or the way an AI model behaves. These tend to be external configurations for a model that are set either before training or before deployment and are not learned from the data.

I

In-context Learning

A process by which generative AI models adapt their outputs based on the specific context provided within a user's prompt, allowing the model to perform tasks without additional training. This capability relies on the model's ability to interpret and utilise the information and instructions embedded in the prompt, enabling it to generate responses that are relevant and coherent to the given context.

Interoperability

The ability of different AI systems and software to work together or exchange and make use of information, which is crucial for integrating AI into existing legal technology infrastructures.

J

Jailbreak

A method by which users attempt to bypass the intended restrictions or limitations of a generative AI model, enabling it to produce outputs that it would typically be restricted from generating. Jailbreaking can exploit vulnerabilities in the model's design or implementation, potentially leading to the generation of inappropriate or harmful content that the model is otherwise programmed to avoid.

K

L

Large Language Model (LLM)

A sub-class of Generative AI trained to model a generalised conception of language and thereby have the capacity to process, interpret, analyse and generate human language.

Latency

The time it takes for an AI system to complete a task or make a decision, important for applications requiring real-time analysis.

M

Model Governance

The framework and processes put in place to ensure the responsible use of AI models, including oversight of development, deployment, and continuous monitoring for compliance and performance.

Multi-Hop Reasoning

A cognitive process that allows for information to be collected and synthesized across contexts or repositories of knowledge. It involves connecting various inference steps to reach a comprehensive answer, enabling enhanced comprehension and problem-solving.

Multi-Modal AI

A type of artificial intelligence that integrates and processes information from multiple data modalities, such as text, images, audio, and video, to perform tasks or generate outputs. This approach allows the AI to leverage diverse sources of information, enhancing its ability to understand and respond to complex inputs by combining insights from different types of data.

N

Natural Language Processing (NLP)

A field of AI that focuses on the interaction between computers and human language. NLP techniques are often used in conjunction with Embeddings to process and analyse texts. NLP enables AI systems to understand, interpret, and generate human language and interact with users in natural language.

Neural Network

A deep learning technique designed to resemble the human brain's structure.

O

Open-Source AI

AI systems whose source code is available publicly, allowing for transparency, scrutiny, and contribution from the global community, which can help in understanding and improving the technology.

Overfitting

Overfitting occurs in machine learning training when the algorithm can only work on specific examples within the training data. A typical functioning AI model should be able to generalise patterns in the data to tackle new tasks.

P

Parameters

The adjustable values in a model that are learned from data to best predict outcomes or representations.

Precision

A measure of how many of the AI system's predictions were correct and relevant.

Predictive Analytics

A type of analytics that uses technology to predict what will happen in a specific time frame based on historical data and patterns.

Prescriptive Analytics

Prescriptive analytics is an advanced form of data analytics that goes beyond predicting future outcomes to recommending specific courses of action. It uses data, advanced algorithms, and mathematical models to analyse various factors such as possible scenarios, past and present performance, and available resources. The goal is to provide actionable insights that help organizations make better strategic decisions.

Prompt

An input that a user feeds into an AI system to get a desired result or output.

Prompt Engineering

The practice of carefully crafting and refining prompts to get more accurate or specific responses from a language model.

Q

R

Retrieval Augmented Generation (RAG)

A technique that combines Embeddings with retrieval mechanisms that connect AI-generated content to verifiable sources of information, to generate more accurate and contextually relevant outputs. RAG can be used to enhance the quality of document analysis and information retrieval. It's also often referred to as Grounding.

Recall

This measures the ability of an AI system to find all relevant instances within a dataset. In legal AI, this could refer to the system's ability to identify all relevant legal precedents or applicable regulations.

Reinforcement Learning

A type of machine learning where agents learn how to behave by receiving rewards or penalties.

Robotic Process Automation (RPA)

Technology that uses software robots to automate repetitive tasks previously done by humans in business processes.

Robustness

The ability of an AI system to maintain its performance when faced with changes in data or conditions, ensuring reliability and trustworthiness in various scenarios.

S

Scalability

The ability of an AI system to handle growing amounts of work or its capability to accommodate growth, such as analysing larger sets of legal documents efficiently.

Semantic Similarity

The measure of how similar two pieces of text are in meaning. Embeddings help AI systems determine semantic similarity by representing words and phrases as Vectors.

Supervised Learning

A machine learning task where the algorithm is trained on labelled data, meaning the input data comes with metadata / the expected response.

Synthetic Data

Data that is artificially generated, rather than collected from real-world events, often used for training or testing purposes without compromising privacy.

System Prompt

A prompt that is built into an AI system or Agent that describes to the system how it is to behave and that provides context about the purpose of the system to reduce the need for users to provide more complex prompts.

T

Temperature

A hyperparameter used in probability scaling, where higher values produce more random (and therefore more creative) outputs, and lower values make model outputs more deterministic or confident.

Throughput

The number of tasks or operations an AI system can handle within a given time frame, indicating its efficiency and capacity.

Tokens

Chunks or segments of text that AI models read. In the context of language models, tokens can be as short as one character or as long as one word (e.g., “a” or “apple”). Tokenisation is the process of converting input text into such tokens, which can then be further analysed or processed. Commercial AI models also charge based on the number of tokens in the input and output.

Top_p (Nucleus Sampling)

A sampling technique used in generative AI to control the diversity of generated outputs by selecting from the most probable set of tokens whose cumulative probability exceeds a specified threshold, p. A high “top_p” value means the model looks at more possible words.

Training

The process of feeding data into an AI system to help it learn and make accurate predictions or decisions.

Training Data

The information or examples given to an AI system to enable it to learn, find patterns, and create new content.

Transfer Learning

Using a pre-trained model on a new, but related task.

Transparency

The clarity and openness with which a model operates, allowing its decision-making process and inner workings to be easily understood, explained, and interpreted by humans.

Transformer Models

A class of neural network architectures designed for processing sequential data, particularly effective in natural language processing tasks. Transformer models utilise mechanisms such as self-attention to capture dependencies and relationships within the data, allowing them to generate coherent and contextually relevant outputs.

U

Unsupervised Learning

A type of machine learning where the algorithm is given data without explicit instructions on what to do with it.

Use Case

A use case refers to a specific scenario or application where AI technology is employed to address a particular problem or enhance a specific process. Use cases articulate how AI can solve specific challenges, streamline workflows, and improve decision-making. They serve as practical examples that bridge the gap between technological potential and real-world utility, ensuring that AI solutions are aligned with the needs and complexities of legal practice. Use cases can be high level (eg; “drafting”) or very specific (eg; “drafting a set of buyer-friendly warranties that are relevant for the purchase of a target company with no subsidiaries”).

User Prompt

A dynamic instruction or query provided by the end-user to interact with the AI. It specifies what the user wants the AI to do or respond to in real-time to meet their requirements.

V

Vector Space/Vectors

A mathematical construct in which Embeddings are represented as Vectors. This allows AI systems to perform operations on the embeddings, such as measuring distances or finding clusters of semantically similar terms.

Verifiability

The extent to which the outputs of an AI system can be traced back to their sources. Verifiability is essential in legal contexts to ensure transparency and accountability.

Validation

The process of testing an AI system with a separate set of data (not used in training) to evaluate its accuracy and effectiveness.

W

Weights

The parameters in a neural network that are adjusted during training to minimise the error in predictions

X

Y

Z

Zero-shot Prompting

A process where a generative AI model is tasked with performing a new task or generating a response without having been explicitly trained on examples of that task. The model relies solely on its pre-existing knowledge and the context provided in the prompt to generate an appropriate output, demonstrating its ability to generalise and adapt to novel situations without prior specific training data.

The [ICO Glossary](#) has a number of definitions which may be useful in this context such as 'Discrimination' 'Misinformation' 'Bias,' 'Fairness' 'federated learning' and 'Perturbation'