AI Glossary
Key Terms and Definitions

Accountability – The obligation of AI system developers and operators to ensure their systems are used responsibly and to address any negative impacts.

Adaptive Learning – The ability of an AI system to improve its performance by learning from new data and experiences over time

AI Governance – The framework of policies, procedures, and practices designed to ensure that AI technologies are developed and used responsibly, ethically, and in compliance with regulations.

Algorithm – A set of rules or steps that a computer follows to complete a task.

Algorithmic Transparency – The extent to which the processes and decisions made by AI systems are visible and understandable to humans.

Artificial Intelligence (AI) – The capability of a machine to imitate intelligent human behavior.

Artificial General Intelligence (AGI) – Artificial General Intelligence (AGI) refers to a type of AI that possesses the ability to understand, learn, and apply knowledge across a wide range of tasks at a level comparable to human intelligence.

Artificial Neural Network (ANN) – Computational models inspired by the human brain, used in certain AI applications.

Chatbots – Software applications designed to simulate human conversation.

Data Mining – The process of discovering patterns and knowledge from large amounts of data.

Data Scrubbing – The process of detecting and correcting (or removing) errors and inconsistencies in data to improve its quality.

Data Training – The phase in which an AI system learns from data to make predictions or decisions without being specifically programmed.

Deepfakes – Digitally crafted images, recordings, or videos that convincingly misrepresent someone's actions or words.

Deep Learning – A subset of machine learning, using neural networks with many layers.

De-skilling – The reduction in the skill level required to perform a job due to technology advancements.

Explainability – The ability to explain how an AI system reaches its decisions in a way that humans can understand.

Generative Adversarial Networks (GAN) – GenAI models that pit two AI models against one other to generate increasingly convincing fake output data, such as audio, images, and videos.

Generative AI (GenAI) – AI models that can generate new content, such as images, sounds, or texts.

Hallucination – When an LLM system responds to a prompt with an inaccurate, irrelevant, or illogical answer.

Large Language Model (LLM) – A kind of neural network designed, essentially, to predict the next word; often used to power chatbots.

Machine Learning (ML) – A type of artificial intelligence where computers learn from data without being explicitly programmed.

Narrow AI – Narrow AI (or weak AI) is a type of artificial intelligence designed to perform a specific task or a limited set of tasks with a high level of efficiency.

Natural Language Processing (NLP) – The ability of computers to understand and generate human language.

Neural Network – A computer system modeled on the human brain and nervous system, used by AI researchers to process data in a brain-like manner.

Prompt Engineering – The art and science of designing specific input structures to guide machine learning models, especially language models, to produce desired outputs.

Reinforcement Learning – One of the three core machine-learning paradigms, it involves algorithms independently using trial-and-error to determine behaviors that maximize a desired reward.

Retrieval Augmented Generation (RAG) – A technique for optimizing LLMs for particular use cases, by feeding them certain targeted information (for instance, by uploading all of a company's contracts).

Robustness – The ability of an AI system to perform reliably under diverse and unforeseen circumstances.

Singularity – A hypothetical point in the future when technological growth becomes uncontrollable and irreversible, leading to unforeseeable changes to human civilization; sometimes described as artificial intelligence surpassing human intelligence.

Superintelligence – An entity, often envisioned as a computer, that exceeds human capability in terms of overall intelligence or specific intellectual measures.

Supervised Learning – A method of AI training where algorithms are taught using labeled input data, such as training an algorithm to recognize cats using images explicitly labeled as cats.

Training Computation – The computational work involved in training a machine learning model on a dataset, typically measured in terms of the number of operations or the energy consumed.

Transformers – Introduced by Google researchers in 2017, transformers are models trained on huge quantities of data to understand context, selectively retaining relevant information similar to human memory; they form the foundation of ChatGPT.

Trustworthy AI – AI that is reliable, fair, transparent, and respects privacy, fostering trust among users and stakeholders.

Unsupervised Learning – A method of AI training where algorithms identify patterns and similarities in unlabeled data, grouping information without prior training or explicit labeling.