

COURSE GLOSSARY

Generative AI Concepts

- **Generative AI:** A type of AI that creates new content such as text, images, or audio based on patterns in training data.
- **Prompt:** An input instruction that guides a generative AI model in producing an output.
- **Response:** The content generated by a generative AI model in reaction to a prompt.
- **GAN (Generative Adversarial Network):** A machine learning model composed of two neural networks (generator and discriminator) competing to produce realistic outputs.
- **Transformer:** A neural network architecture designed for processing sequential data like text, widely used in modern generative AI.
- **Reinforcement Learning with Human Feedback (RLHF):** A training method where human feedback is used to fine-tune AI model outputs.
- **Artificial General Intelligence (AGI):** A form of AI capable of understanding and performing any intellectual task that a human can do.
- **Fine-tuning:** Adapting a pre-trained model to specialize it for a specific task.
- **Transfer Learning:** Leveraging knowledge from a model trained on one task to improve performance on another task.
- **Custom Embeddings:** Mathematical representations of data optimized for specific AI applications.
- **Preprocessing:** Cleaning and preparing raw data before it is used to train AI models.
- **Bias (Algorithmic / Representation / Temporal):** Systematic errors in AI models caused by skewed data, poor representation, or limitations in training timeframes.
- **Performance Metrics:** Measures such as accuracy, precision, recall, and fairness used to evaluate model performance.
- **Turing Test:** A test that evaluates whether an AI can mimic human intelligence convincingly enough to fool humans.
- **Anonymization:** The removal of personally identifiable information (PII) from datasets.
- **Consent Management:** The process of obtaining and managing user permission to use their data in AI systems.
- **AI Alignment:** Ensuring AI systems' goals and behaviors match human values and safety requirements.
- **Safety Constraints:** Rules or limitations placed on AI systems to prevent harmful outputs.
- **Iterative Development:** Building AI models step by step, testing and refining to ensure effectiveness and safety.
- **Augmentation:** Using AI to enhance human productivity and creativity.
- **Replacement:** Using AI to fully automate tasks traditionally performed by humans.
- **Co-creation:** Humans and AI collaborating together to produce content or solve problems.
- **Deepfake:** AI-generated synthetic media that mimics real content, often in image or video form.
- **Misinformation Campaigns:** The use of AI-generated content to spread false or misleading information.
- **Cross-sector Collaboration:** Cooperation between industries, governments, and academia to advance AI development.
- **Open-source Communities:** Developer groups that create and share AI tools with publicly available source code.
- **Data Governance:** Policies and practices for managing the availability, usability, and security of data used in AI.
- **Ethical AI:** The practice of ensuring AI development and deployment align with fairness, accountability, and transparency principles.