DATA&
SOCIETY

Understanding AI

Glossary: Frequently Used Terms

Artificial intelligence (AI): A machine-based system that is capable of performing complex tasks, including making predictions, recommendations, or decisions that can influence real or virtual environments.

Agent: An “AI agent” is a kind of AI program that can act on your behalf, such as buying plane tickets and booking a hotel for you, without explicit instructions or supervision. The industry presents AI agents as the “next big thing” in AI, and they are still very much under development.

Automation: Refers to the use of technology, machines, or software to perform tasks without human intervention.

Compute: Refers to the resources — both hardware and software — used to process data, or the processing power needed to run applications or train an AI model.

Cloud computing: Delivers on-demand access to a shared pool of computing services, using the internet rather than local machines. Cloud providers include Amazon Web Services (AWS), Microsoft Azure, and Google Cloud.

Critical minerals: The changing set of resources considered to be critical to national security and technological development, and thus open to interpretation. This includes minerals and rare earth elements that are fundamental to technological advancement, including cobalt, lithium, copper, nickel, and palladium. Critical mineral lists change over time and depend on the region and industry.

Data center: A physical facility that stores servers and related physical computing infrastructures. On-premises data centers are housed on a company’s own property and managed by the company using in-house hardware and software, whereas cloud computing data centers allow companies to rent space and infrastructure.

Generative AI: A type of AI that can create new content, such as text, images, videos, music, and more, based on learned patterns from existing data. Large language models are a type of generative AI.

Graphical processing unit (GPU): A GPU is a specialized electronic circuit for processing digital images. GPUs are used for a specific task, like graphics rendering or machine learning. GPUs are essential to the growth of AI.

Hyperscaler: A large data center that uses enterprise-scale infrastructure to provide cloud services.

Large language model (LLM): The term “large language model” is used to describe a specific type of model for predicting and constructing natural-sounding text. It is created via a specific machine learning method for “deep” pattern-matching applied to an extremely large data set.

Machine learning: The most popular technical approach to achieve AI used today. Before machine learning, computers required specific, step-by-step instructions for how to complete a task. Machine learning is a technique for showing a computer enough examples that specific instructions are not necessary; this is called “training.” Through training, the computer observes patterns to re-create the steps needed implicitly, without stepwise instructions.

Model: In the context of AI, a model refers to a “trained” machine learning model; that is, a carefully-crafted function that converts inputs into outputs. In many cases, an AI system comprises many models; complex systems like your Facebook timeline or X feed rely on hundreds of interconnected models.

Open data: Data that is released in accessible formats for reuse and sharing by anyone.

Prompt: The input or instruction a user gives to an AI model to guide its output and achieve a desired outcome. It's the way you ask the AI to do something.

Public interest technology: Often refers to the development, deployment, and governance of technology in ways that prioritize ethical considerations and the public good.

Scraping: Refers to the act of making a copy of data from an available source (typically an unauthorized copy).

Acknowledgements | Glossary terms were adapted with permission from Data & Society's AI 101 workshop, hosted by Meg Young and Emnet Tafesse (2025); or sourced from Data & Society’s *Keywords of the Datafied State* anthology (2024), edited by Jenna Burrell, Ranjit Singh, and Patrick Davison.