The Executive Glossary

Lexicon of the Ensemble Era

```
┌──────────┐          ╭───────╮          ┌──────────┐
│   Data   │ ───────▶ │ Model │ ───────▶ │  Wisdom  │
└──────────┘          ╰───────╯          └──────────┘
```
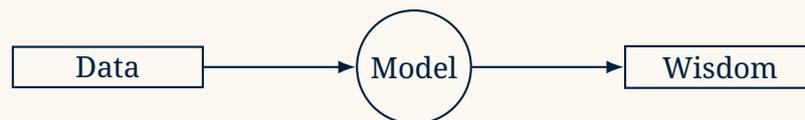
A handy reference for common AI terms you'll encounter on your journey toward computational mastery.

In the 21st century, precision of language is the ultimate hedge against strategic obsolescence. To govern AI is to understand its vocabulary. This glossary provides the definitions required to bridge the gap between technical complexity and executive clarity.

I. MODEL ARCHITECTURE & FUNDAMENTAL SYSTEMS

ARTIFICIAL GENERAL INTELLIGENCE (AGI)
A theoretical state of AI where a system possesses the ability to understand, learn, and apply its intelligence across any intellectual task, equaling or exceeding human capacity.

LARGE LANGUAGE MODEL (LLM)
A sophisticated neural network trained on vast datasets to identify, predict, and generate human-like text. These models function as broad "generalists."

SMALL LANGUAGE MODEL (SLM)
A more compact AI model optimized for specific tasks or domains. SLMs offer higher efficiency, lower operational costs, and the capability for secure, local deployment.

TRANSFORMER
The underlying mathematical architecture that enables AI to process language using "Attention," allowing the system to weigh the relevance of different parts of an input simultaneously.

NEURAL NETWORK
A computational system inspired by the biological structure of the brain, consisting of interconnected layers that process and pass information through weighted nodes.

II. DATA, TRAINING & KNOWLEDGE SYNTHESIS

FINE-TUNING
The process of taking a pre-trained model and further training it on a specialized dataset to adapt its performance to a specific industry or corporate persona.

PARAMETERS
The internal variables of a model that it "learns" during training. In simpler terms, these are the billions of "knobs and dials" that determine the model's behavior and logic.

RETRIEVAL-AUGMENTED GENERATION (RAG)
A framework that allows an AI to look up external, proprietary information (like a company's private documents) before generating an answer, ensuring outputs are grounded in fact.

TOKEN
The atomic unit of text processed by an AI. A token is roughly equivalent to 0.75 of a word; it is how the machine "reads" and "writes" linguistic patterns.

WEIGHTS
Numerical values within the neural network that determine how much importance the model assigns to one data point versus another during a calculation.

Chain-of-Thought (CoT)

A prompting technique that instructs the AI to "think step-by-step," forcing the model to articulate its logic before providing a final answer, which significantly reduces errors.

Few-Shot Prompting

Providing the AI with several examples of a desired output within the instruction to align the machine's probabilistic direction with human expectations.

Hallucination

A statistical phenomenon where an AI generates factually incorrect or nonsensical information but presents it with high confidence and plausible logic.

Inference

The act of an AI model calculating a response based on a prompt. This is the "thinking" phase where the model uses its training to predict the best output.

Prompt Engineering

The professional discipline of refining linguistic inputs to minimize noise and maximize the accuracy and utility of an AI's response.

Temperature

A configuration setting that controls the "randomness" or "creativity" of the output. Low temperature (0.1) results in clinical precision; high temperature (0.8+) results in creative variance.

AI Ensemble Solutions

Knowledge is the ultimate ensemble partner.
www.aiensemblesolutions.com