

Chapter 1: What is AI: Types, Definitions & Simple Analogies

Domain: I — Understanding the foundations of AI governance

Competency: I.A — Define AI and understand its foundational concepts

Performance Indicators Covered: * Define Artificial Intelligence (AI) and its core components.

- Distinguish between different types of AI (e.g., Narrow AI, General AI, Super AI).
- Explain AI through simple analogies for non-technical stakeholders.
- Identify the fundamental shift from traditional programming to machine learning.

AI Foundations: From Human Logic to Machine Learning

Course Overview: AIGP Body of Knowledge v2.1 | Chapter 1: What is AI: types, definitions & simple analogies. Stripping away the sci-fi to reveal AI as a mathematical system that infers patterns from data.

The Paradigm Shift: Software 1.0 vs. 2.0

Traditional Programming (Logic-First)

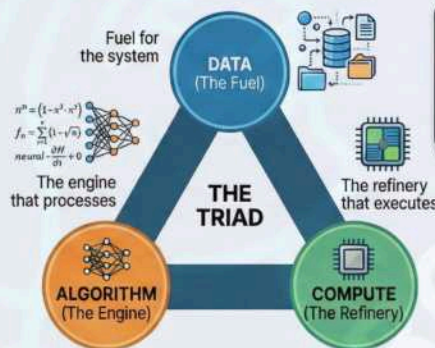


Machine Learning (Data-First)



The OECD Definition of AI:
A machine-based system that infers how to generate outputs to influence environments.

The AI Triad and Current Reality



The "Hyper-Fast Intern" Analogy

Like a brilliant intern with no common sense who follows instructions literally.

Current Reality: Artificial Narrow Intelligence (ANI)

The only AI that currently exists; it is designed for specific, limited tasks.

COMPARING AI TYPES FOR EXAM READINESS

AI Type	Current Status	Description
Narrow (ANI)	Reality	Task-specific systems like facial recognition or chatbots.
General (AGI)	Theoretical	Systems matching human intellect across all intellectual domains.
Super (ASI)	Speculative	Future AI that surpasses all human creative and social wisdom.

1 Chapter Overview

This introductory chapter serves as the bedrock for the entire **Artificial Intelligence Governance Professional (AIGP)** body of knowledge. Before a professional can govern, regulate, or audit a technology, they must first strip away the science-fiction tropes and marketing hyperbole to understand what AI *actually* is.

In this chapter, we define AI not as a singular "brain," but as a constellation of technologies used to perform tasks that typically require human intelligence. We explore the transition from **Rule-Based Systems** (where humans provide the logic) to **Machine Learning** (where the system derives the logic from data). This shift is the "Big Bang" of AI governance; because

humans no longer hand-write every rule, the systems become "black boxes," necessitating the transparency, accountability, and oversight frameworks discussed in later chapters.

Exam Relevance: This chapter maps to the foundational "Understand" level of Bloom's Taxonomy. While the exam focuses heavily on risk and law, you cannot correctly classify an AI system under the **EU AI Act (Chapter 21)** or apply the **NIST AI RMF (Chapter 27)** without first identifying if a system meets the technical definition of AI. Expect 5–8% of the exam to draw directly from these foundational definitions.

2 Learning Objectives

By the end of this chapter, the candidate will be able to:

- **Define** Artificial Intelligence using internationally recognized standards (OECD, ISO/IEC 22989) [Remember/Understand].
- **Differentiate** between Artificial Narrow Intelligence (ANI) and Artificial General Intelligence (AGI) [Analyze].
- **Contrast** traditional software engineering ("If-Then" logic) with Machine Learning (statistical pattern matching) [Understand/Apply].
- **Identify** the three core components of the AI triad: Data, Algorithms, and Compute [Remember].
- **Explain** complex AI concepts to board-level stakeholders using verified analogies [Apply].

3 Core Concepts — Deep Dive

A. Concept Explanation: Defining the Indefinable

The term "Artificial Intelligence" was coined in 1956, but its definition remains a moving target. For the purposes of governance, we move away from "thinking machines" and toward **mathematical optimization**.

Official Definition (OECD/EU AI Act Alignment):

An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

The "Triad" of AI:

1. **Data:** The "fuel." AI requires massive datasets to find patterns.
2. **Algorithms:** The "engine." Mathematical instructions that process data.
3. **Compute:** The "refinery." The raw processing power (GPUs/TPUs) required to run the math.

B. Technical Dimension: The Paradigm Shift

To govern AI, one must understand how it differs from the software that came before it.

1. Traditional Programming (Logic-First):

In traditional coding, a human programmer writes a recipe.

- *Input:* Data + Rules.
- *Output:* Result.
- *Analogy:* A tax calculator. A human writes the tax brackets into the code. If income is \$X\$, then tax is \$Y\$. It is predictable and 100% explainable.

2. Machine Learning (Data-First):

In Machine Learning (ML), we give the computer the answers and ask it to figure out the recipe.

- *Input:* Data + Results (Examples).
- *Output:* Rules (The Model).
- *Governance Implication:* Because the "rules" are statistical weights rather than human-readable code, we face the "**Black Box**" problem. We know what went in and what came out, but the "why" in the middle is mathematically dense.

C. Types of AI: From Narrow to Super

The BOK requires a distinction between the current reality and future possibilities.

Type	Description	Current Status
Artificial Narrow Intelligence (ANI)	Designed for one specific task (e.g., facial recognition, chess, spam filters).	Reality. All current AI is ANI.
Artificial General Intelligence (AGI)	A system that can perform any intellectual task a human can do across multiple domains.	Theoretical. Does not yet exist.
Artificial Super Intelligence (ASI)	AI that surpasses human intellect across all fields, including creativity and social wisdom.	Speculative/Sci-Fi.

Generative AI (GenAI): A subset of ANI that creates *new* content (text, images, audio) rather than just classifying existing data. While it feels "General" because it can write poetry and code, it is still statistically predicting the next token in a sequence.

D. Legal & Regulatory Context

Governments are moving away from vague definitions to "Regulatory Certainty."

- **ISO/IEC 22989:** Provides the standardized vocabulary for AI. This is the "dictionary" for AIGP professionals.
- **EU AI Act:** Uses the OECD definition to ensure that simple "if-then" spreadsheets are not over-regulated as AI.
- **NIST AI RMF:** Focuses on the socio-technical nature—AI isn't just code; it's the people and processes around it.

E. Risk Dimension: Why Governance is Different

Traditional IT governance focuses on **Security** (Is it hacked?) and **Availability** (Is it up?).

AI Governance adds **Probability** and **Alignment**:

- **Stochastic Nature:** AI is probabilistic, not deterministic. It might give a different answer to the same question twice.
- **Brittleness:** An AI trained on "X" data might fail catastrophically when it sees "Y" data (Distribution Shift).
- **Bias:** If the data has historical prejudices, the AI will "learn" and automate those prejudices.

4 Governance in Practice: The "Explainability" Toolkit

Simple Analogies for Stakeholders

When explaining AI to a Board of Directors or a Legal team, avoid "Neural Networks." Use these:

1. **The Intern Analogy:** AI is like a brilliant, hyper-fast intern who has read every book in the library but has zero common sense. They will follow your instructions literally, even if it leads to a cliff, and they will occasionally "hallucinate" facts with total confidence.
2. **The Ultra-High-Speed Mirror:** AI doesn't "know" things; it reflects the patterns in the data we give it. If the mirror shows a distorted image, it's not the mirror's fault—it's the data it is reflecting.

Documentation Checklist: The "Is it AI?" Test

Use this to determine if a project falls under the AI Governance framework:

- [] Does the system derive its own decision-making logic from data?

- [] Is the output probabilistic (a "best guess") rather than a fixed calculation?
- [] Does the system "learn" or improve its performance over time with more data?
- **If Yes to any:** Apply AI Governance controls.

5 Comparative Framework Analysis

Feature	Traditional Software	Machine Learning (AI)
Primary Logic Source	Human Programmer	Statistical Data Patterns
Transparency	High (Code is readable)	Low (Weights/Biases are math)
Failure Mode	Crashes or logic errors	Biased outputs or "Hallucinations"
Maintenance	Bug fixes/Updates	Retraining with new data
Governance Focus	Security & Compliance	Ethics, Bias, & Robustness

6 Case Study: The "Auto-Lend" Implementation

Organization: Global Bank Corp.

The System: A new "Smart Credit" tool designed to replace a manual spreadsheet-based approval process.

The Shift: The old system had 20 fixed rules (e.g., "If Credit Score < 600, Reject"). The new AI system analyzes 5,000 variables, including social media activity and shopping habits.

The Failure: Six months in, the bank realizes the AI is rejecting applicants from specific zip codes at a 400% higher rate, despite those applicants being creditworthy.

The Exposure: The bank cannot explain *why* the AI chose those zip codes as a risk factor. Regulators cite them for "Lack of Explainability" under consumer protection laws.

Scenario Questions:

1. **Which core AI component caused the bias?** (Answer: The Data. Historical lending patterns in the data contained systemic geographic biases).
2. **Is this ANI or AGI?** (Answer: ANI. It is specialized for credit scoring).
3. **How would traditional programming have avoided this?** (Answer: A human would have had to explicitly write a "Reject Zip Code" rule, which would have been caught in legal review).

7 Common Exam Traps

- **The "AGI" Trap:** The exam may ask if a current sophisticated model like GPT-4 is "General AI." The answer is **No**—it is still Artificial Narrow Intelligence (ANI) because it operates within the specific domain of language modeling.
- **The "Algorithm" Misconception:** AI is not just an algorithm. An algorithm is just math. AI is the *system* (Data + Algorithm + Compute + Human Context).
- **Automation vs. AI:** Not all automation is AI. A factory robot following a fixed path is "Automation." A robot navigating a changing warehouse using sensors is "AI."

8 Chapter Summary

- **AI Definition:** Machine-based systems that infer how to generate outputs to influence environments.
- **ML vs. Programming:** Shift from human-defined rules to data-derived patterns.
- **The Triad:** Data, Algorithms, Compute.
- **ANI:** The only AI that currently exists; task-specific.
- **Governance Need:** Arises from the "Black Box" nature and probabilistic outputs of AI.

9 15 Exam-Style Questions

1. According to the OECD definition, what is a primary characteristic of an AI system?

- A. It possesses human-like consciousness.
- B. It operates entirely without human objectives.
- C. It infers how to generate outputs from inputs.
- D. It replaces the need for hardware compute.

Correct: C. AI systems use inference to turn inputs into outputs (predictions/decisions).

2. A company uses a system that follows a strict "If-Then" logic tree written by their legal department. Is this AI?

- A. Yes, it automates a task.
- B. No, it is a traditional rule-based system.
- C. Yes, it is Narrow AI.
- D. No, unless it uses Generative AI.

Correct: B. Traditional programming where humans provide the rules is not AI.

3. In the "AI Triad," which component provides the mathematical instructions used to identify patterns in a dataset?

- A. Compute
- B. Algorithm
- C. Inference engine
- D. Large Language Model (LLM)

- **Correct: B.** The algorithm is the set of mathematical rules or processes.
- **Why others are wrong:** Compute (A) is the hardware/power; Inference engine (C) is a specific part of a runtime system; LLM (D) is a specific type of model, not a foundational component of all AI.

4. A governance professional is explaining the "Black Box" problem to the Board. Which statement best describes this concept?

- A. The physical server housing the AI is inaccessible to auditors.
- B. The AI system has been encrypted by a third-party vendor.
- C. The internal logic of how the model reached a specific output is not easily interpretable by humans.
- D. The data used to train the model is proprietary and cannot be viewed.

- **Correct: C.** This is the core challenge of Machine Learning transparency.
- **Why others are wrong:** A and B refer to physical or security barriers; D refers to data privacy/IP, not the algorithmic "opacity" of the model itself.

5. Which type of AI is currently considered "theoretical" and does not yet exist in any commercial application?

- A. Artificial Narrow Intelligence (ANI)
- B. Machine Learning (ML)
- C. Generative AI (GenAI)
- D. Artificial General Intelligence (AGI)

- **Correct: D.** AGI, which matches human intellectual flexibility across all domains, is currently theoretical.

- **Why others are wrong:** A, B, and C all exist and are currently deployed in the market.

6. What is the primary difference between traditional software and Machine Learning regarding "Rules"?

- A. Traditional software has no rules; ML has many rules.
- B. In traditional software, humans write the rules; in ML, the system derives rules from data.
- C. In ML, rules are permanent; in traditional software, they change every time the code runs.
- D. Traditional software uses "Probabilistic" rules; ML uses "Deterministic" rules.

- **Correct: B.** This is the fundamental paradigm shift in AI development.
- **Why others are wrong:** A is factually incorrect; C is the opposite (ML models are often retrained); D swaps the two—traditional software is deterministic, while ML is probabilistic.

7. Under the ISO/IEC 22989 standard, "Inference" is defined as:

- A. The process of cleaning data before training.
- B. The act of a human overriding an AI decision.
- C. The reasoning process of an AI system to produce an output from given input data.
- D. The legal assumption that an AI system is biased until proven otherwise.

- **Correct: C.** Inference is the "running" phase of an AI model where it applies its learned patterns to new data.
- **Why others are wrong:** A is data engineering; B is human-in-the-loop; D is a legal/policy stance, not a technical definition.

8. Which analogy is most appropriate for explaining the "brittleness" of an AI system to a non-technical stakeholder?

- A. A librarian who knows where every book is but can't read.
- B. A calculator that occasionally rounds up numbers for fun.
- C. A champion swimmer who drowns the moment they are put in a lake instead of a pool.
- D. A mirror that reflects only the person standing directly in front of it.

- **Correct: C.** Brittleness refers to AI failing when the input data (the lake) differs even slightly from the training data (the pool).
- **Why others are wrong:** A refers to a lack of understanding; B implies randomness; D refers to bias/reflection.

9. Why is "Compute" becoming a central pillar of AI Governance and regulation?

- A. High compute costs prevent small businesses from using AI.
- B. Environmental impact and the ability to track large-scale model training via hardware signatures.
- C. Compute is the only part of AI that can be patented.
- D. Modern privacy laws require AI to run on "Zero-Compute" chips.

- **Correct: B.** Regulators (and the EU AI Act) are increasingly looking at compute power (FLOPs) as a threshold for identifying "High-Risk" or "Systemic" models.
- **Why others are wrong:** While A is true, it's an economic issue, not the primary governance driver; C and D are factually incorrect.

10. When an AI model generates an output that is statistically likely but factually incorrect, this is known as:

- A. Overfitting
- B. Underfitting
- C. Hallucination
- D. Data Poisoning

- **Correct: C.** Particularly in GenAI, "hallucination" is a byproduct of the system's probabilistic nature.
- **Why others are wrong:** A and B are training errors related to model generalization; D is a security attack.

Scenario-Based Applied Questions

11. Scenario: A healthcare startup develops a "Diagnostic Assistant." It was trained on 10 million X-rays to identify pneumonia. When deployed in a new hospital using a different brand of X-ray machine, the AI's accuracy drops from 98% to 45%.

Which AI characteristic is being demonstrated?

- A. Artificial General Intelligence
- B. Distribution Shift
- C. Rule-based logic failure
- D. Human-in-the-loop error

- **Correct: B.** The "Distribution Shift" occurred because the new machine produced data slightly different from the training set, causing the "brittle" AI to fail.
- **Why others are wrong:** It is still Narrow AI (A); it isn't rule-based (C); and no human error was mentioned (D).

12. Scenario: An HR department uses a "Resume Screener." The tool was built by a vendor using "Deep Learning" to identify top talent. The HR Director asks for a list of the specific criteria the AI uses to reject candidates, but the vendor says, "The model's neural network is too complex to provide a simple list of rules."

As an AI Governance Professional, how should you classify this system's risk?

- A. Low risk, because it is just a resume tool.
- B. High risk, because it lacks "Explainability" in a high-stakes decision-making context.
- C. No risk, because the vendor owns the IP.
- D. Medium risk, because it can be fixed with more compute.

- **Correct: B.** In HR (employment), the lack of explainability (Black Box) is a major governance and legal risk under both the EU AI Act and US non-discrimination laws.
- **Why others are wrong:** HR is consistently classified as high-stakes (A); IP ownership does not negate regulatory risk (C); compute does not solve explainability (D).

13. Scenario: A retail company implements a chatbot that can answer customer questions, write emails, and even draft simple legal terms of service. The CEO claims they have achieved "AGI."

How should the Governance Lead correct this statement?

- A. Agree, because the system can do three different things.
- B. Disagree; it is still Narrow AI because it is limited to the domain of "Language Prediction" and lacks true cross-domain reasoning.
- C. Disagree; it is only AGI if it has a physical robotic body.
- D. Agree, but only if the chatbot can pass the Bar Exam.

- **Correct: B.** Performing multiple linguistic tasks is "Broad" but still "Narrow AI" (ANI) because it operates within one modality (text) and lacks general autonomous reasoning.
- **Why others are wrong:** A and D are common misconceptions; C is not a requirement for AGI (which can be software-only).

Analytical Comparison Questions

14. Compare "Deterministic" systems and "Probabilistic" systems. Which statement is true for AI Governance?

- A. Deterministic systems require more ethics oversight than probabilistic ones.
- B. Probabilistic systems (AI) are harder to audit because the same input might result in different outputs or "confidence scores."
- C. Probabilistic systems are always more accurate than deterministic systems.

D. Deterministic systems are the foundation of Generative AI.

- **Correct: B.** The lack of a 1:1 consistent output for every input makes auditing and safety testing AI much more complex than traditional code.
- **Why others are wrong:** A is reversed; C is false (calculators are deterministic and 100% accurate); D is false (GenAI is probabilistic).

15. Analyze the shift from "Software 1.0" (Traditional) to "Software 2.0" (AI). What is the primary change in the role of the Human Developer?

- A. The developer no longer needs to understand math.
B. The developer shifts from writing explicit logic to curating high-quality data and designing objective functions.
C. The developer is no longer responsible for the system's errors.
D. The developer's role is replaced by the "Compute" component.

- **Correct: B.** In AI development, "coding" is replaced by "data engineering" and "model architecture."
- **Why others are wrong:** A is false (AI requires more math); C is a major governance fallacy—humans remain accountable; D is false, as compute is a tool, not a role.

■ Summary Checklist for Chapter 1

- I can define AI using the **OECD/ISO** standard.
- I understand that **ANI** is the only existing form of AI.
- I can explain why **Machine Learning** creates a "Black Box" compared to traditional "If-Then" code.
- I can identify the three parts of the **AI Triad**: Data, Algorithm, Compute.
- I can explain the risks of **Stochastic** (probabilistic) systems to a non-technical audience.