

## Definition

# AI as a Service (AIaaS)

Jorge Guerra Pires<sup>1</sup>

1. Independent researcher

**AI as a Service (AIaaS)** refers to the **modular, externally accessible provisioning of artificial intelligence capabilities**—especially large language models (LLMs) and domain-specific systems—**through APIs or cloud-based platforms**. It enables users to embed advanced cognitive functions (such as reasoning, generation, and classification) into their workflows **without managing or even accessing the underlying architecture**.

In contrast to earlier periods—when AI systems were often **siloe**d within academic labs or industrial R&D groups—AIaaS offers a **scalable, managed alternative**, shifting the locus of innovation from code ownership to API consumption. The term “as a Service” is literal: the AI itself **is** the product. Rather than downloading and deploying an application that uses AI, **the AI is deployed directly**, abstracted as a cloud service.

This model departs sharply from prior paradigms. **Open-source AI** (e.g., TensorFlow) required users to host, maintain, and debug their own deployments. **Proprietary AI**, in turn, was typically expensive, opaque, and limited in scope. In contrast, AIaaS providers (such as OpenAI) expose the same APIs they use internally, offering **real-time feedback logs**, uptime reports, and error transparency—features rarely seen in previous business models.

AIaaS, therefore, represents a **new layer of abstraction in the AI stack**, one that centralizes power and control in platform providers, while dramatically lowering barriers to integration for end users.

## Declarations

**Potential competing interests:** No potential competing interests to declare.