

Generative AI Solutions

Enterprise copilots and private LLMs—on your data, under your control

Overview

We deliver generative AI at enterprise scale: private LLM deployment, RAG systems, internal knowledge assistants, and agentic workflows—with secure data access and model orchestration built in.

Business problems we address

- Need enterprise copilots and internal knowledge assistants without sending data to public APIs.
- Private LLM deployment and RAG systems must run on-prem or in your VPC for compliance.
- Agentic workflows and model orchestration lack a single, secure, governable platform.
- Secure data access—who can see what—is unclear or inconsistent across teams and tools.
- Regulators and procurement require evidence of data sovereignty, guardrails, and auditability.

How we deliver

We deliver end-to-end: use-case and secure data access design, private LLM and RAG architecture (including retrieval pipelines and model orchestration), enterprise copilots and internal knowledge assistants with guardrails and citation, and agentic workflows with human-in-the-loop. Deployment runs in your environment (on-prem, VPC, or sovereign cloud) with documentation, monitoring, and a path to scale.

Delivery steps

1. Use case & data access

Define copilot scope, data sources, and secure access model.

2. Private LLM & RAG design

Model selection, retrieval pipeline, and orchestration architecture.

3. Build, guardrails & deploy

Copilot build, guardrails, and production rollout with monitoring.

Security & governance

Every solution is designed with secure data access and governance by default: data stays in your perimeter with role-based access, PII is redacted or tokenized at the edge, prompt injection and jailbreaking are mitigated via orchestration and guardrails, and outputs are validated for hallucination and policy. We align to EU AI Act, NIST AI RMF, and your internal policies and document controls for auditors.

Who should use this service

Ideal for enterprises deploying enterprise copilots, internal knowledge assistants, private LLM deployment, RAG systems, or agentic workflows—CIOs, CDOs, and product leaders who need production-grade generative AI with secure data access and model orchestration, without compromising compliance or control.

Example use case

A global insurer needed an internal knowledge assistant over 2M policy documents and underwriting guidelines. We deployed a private RAG-based copilot with secure data access, model orchestration for routing and guardrails, and citation-backed answers—reducing research time by 40% while keeping all data on-prem and audit-ready.

Key capabilities

- Enterprise copilots and internal knowledge assistants
- Private LLM deployment (on-prem, VPC, sovereign cloud)
- RAG systems with citation and secure data access
- Agentic workflows and model orchestration
- Fine-tuning and retrieval pipelines

Moaisus Global Solutions — Enterprise AI, securely and at scale.