

NEXUS AI TECHNOLOGIES

Enterprise AI Solutions & Services Catalog

Formulated Frameworks & Deployment Architectures

Document Version: **2026.1.4 (Updated)**

Classification: **Commercial Confidential**

Target Audience: **CIOs, CTOs, & Enterprise Architecture Teams**

Published Date: **June 2026**

1. AI Strategy & Architecture Consulting

We partner with enterprise organizations to align operational goals with cutting-edge artificial intelligence capabilities, creating scalable, low-risk execution roadmaps.

AI Readiness & Capability Assessment

Evaluation of corporate data infrastructure, engineering talent metrics, and current compliance stances against state-of-the-art Generative AI prerequisites. Delivers a clear gap analysis.

Timeline: 4-6 Weeks

Delivery: Strategic Report

Enterprise LLM Architecture Design

Blueprints for hybrid execution environments. Designing network topologies that mix local private models with frontier hyperscaler API endpoints to maximize throughput and minimize legal footprint.

Timeline: 6-8 Weeks

Delivery: Technical Blueprints

2. Custom Model Engineering & Orchestration

Tailoring underlying open-weights foundational models to match specific domain realities, proprietary corporate lexicons, and strictly guarded operational workflows.

Domain-Specific Fine-Tuning

Supervised Fine-Tuning (SFT) and Direct Preference Optimization (DPO) pipelines applied to open foundational models utilizing internal knowledge vaults safely isolated within private tenant borders.

Timeline: Custom

Delivery: Weights Deployment

Advanced RAG Pipelines (Retrieval-Augmented Generation)

Engineering resilient chunking, embedding generation, semantic caching layers, and cross-encoder re-ranking mechanisms. Minimizes hallucination tracking to $H < 0.01\%$ across target structural databases.

Timeline: 8-12 Weeks

Delivery: Production Engine

3. MLOps, Security & Governance

Ensuring that operational AI pipelines conform to global standard frameworks while guaranteeing continuous performance and optimization metrics.

AI Guardrails & Red-Teaming Integration

Implementation of deterministic runtime safety interceptors monitoring runtime traffic. Prevents data exfiltration vectors, prompt injection exploits, and non-compliant toxic outputs dynamically.

Timeline: 4 Weeks

Delivery: Deployment Policy

Continuous Evaluation & Drift Monitoring

Automated pipeline suites tracking model output shift over time. Measures latent variance to signal model retrain triggers before performance thresholds fall below specified SLAs.

Timeline: Recurring

Delivery: Managed Dashboards