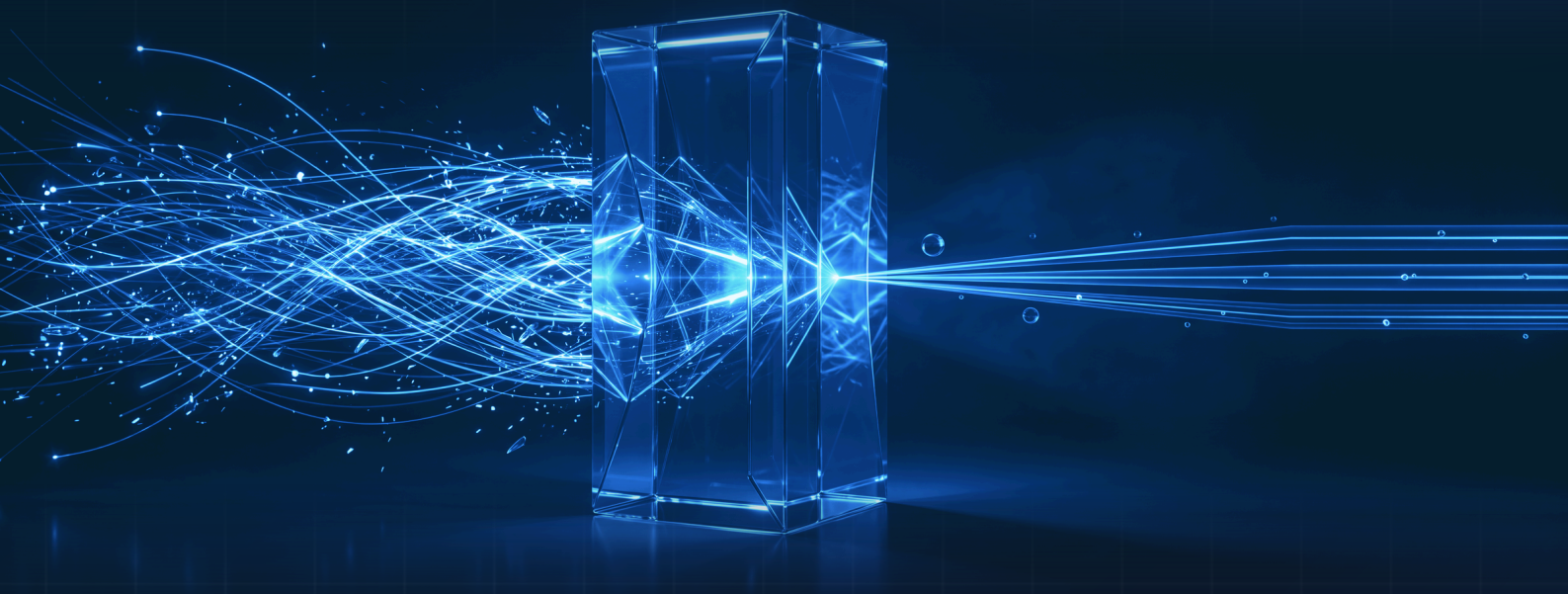
allclouds.pl
AI by design

SAIE PANTERA · Q3 2026

WHITE PAPER



SAIE · CONTROL · SECURITY · AI GOVERNANCE

PROXY:AI

A central layer of control, security and governance for artificial intelligence

PROXY:AI acts as a unified gateway between users, applications and AI agents on one side and language models - cloud and local - on the other. It eliminates Shadow AI, protects sensitive data and keeps costs fully under control - from the cloud, through on-premises, to air-gapped mode

This paper outlines the architecture and operating principles of PROXY:AI: the request lifecycle, platform modules, regulatory compliance mechanisms and deployment options. It is intended for CISOs, IT departments and senior management of organizations adopting AI

01

AI Gateway

An intelligent router that centralizes all traffic to AI models

02

AI Firewall

A real-time firewall: content analysis, policies, data masking

03

Governance

An immutable WORM register and automatic evidence packages for auditors

04

Cost Control

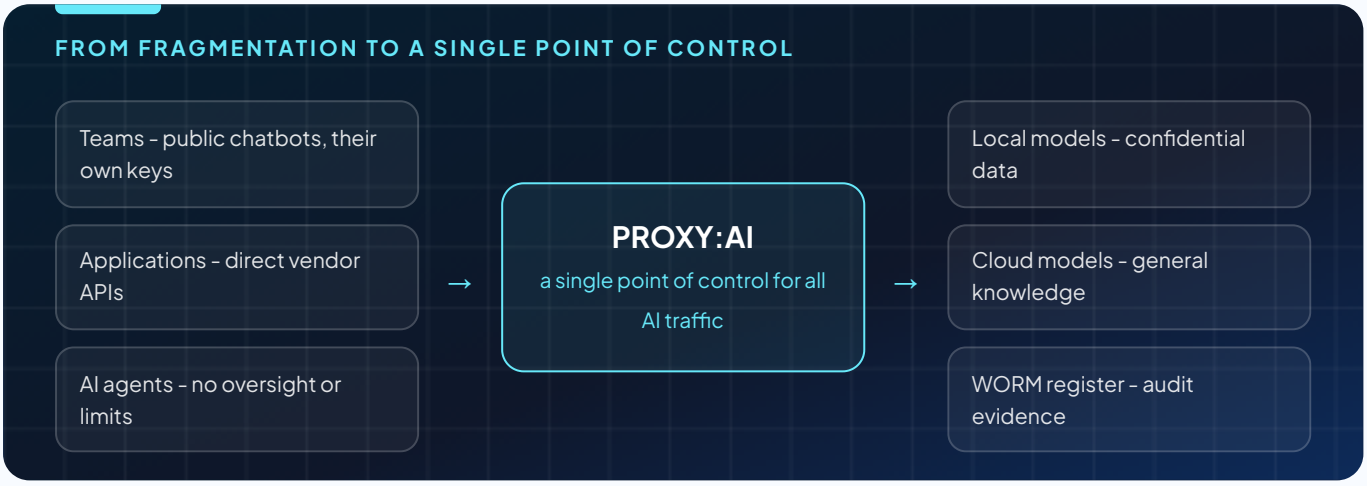
FinOps for AI - AI cost management: budgets, limits, full financial predictability

THE PROBLEM: AI BEYOND THE ORGANIZATION'S CONTROL

In most organizations, access to language models grows bottom-up and without a plan. Teams connect directly to various vendors' APIs, and employees move company data into public chatbots. This phenomenon, Shadow AI, means losing control over who uses AI, with which tool and with what data

The consequences grow along three dimensions: **security** (sensitive data reaches external models outside policy), **costs** (an agent stuck in a loop runs up spiraling bills - Denial-of-Wallet) and **technology dependence** (vendor lock-in)

At the same time, regulatory pressure is rising: the EU AI Act, GDPR, DORA, NIS2/KSC (Polish Cybersecurity Act), KRI (Polish National Interoperability Framework) and ISO/IEC 42001 require demonstrable oversight of AI, backed by evidence. With fragmented access, gathering such evidence is practically impossible



THE CORE OF THE PROBLEM

An organization cannot protect, account for or prove what it cannot see. Control, predictable costs and compliance require a single point through which all AI traffic flows

DATA SECURITY

AI Firewall - masks sensitive data and blocks attacks on the model (prompt injection)

VENDOR LOCK-IN

Zero-Code Switch - switch models without rewriting applications

UNCONTROLLED COSTS

Cost Control - budgets and limits, protection against runaway costs (Denial-of-Wallet)

REGULATORY PRESSURE

Governance - WORM register and evidence packages

WHAT PROXY:AI IS

PROXY:AI is a single entry point between users, applications and AI agents and more than 100 models - cloud and local. Users work through an assigned access profile, and the system decides which model handles each task

MANAGING PROFILES, NOT MODELS - ZERO-CODE SWITCH

01 **User, application or AI agent** sends a query → 02 **Assigned access profile** instead of picking a raw model → 03 **System decision** permissions · cost · data class → 04 **Cloud or local model** over 100 models, incl. Bielik, Llama

HOW IT WORKS – THE REQUEST LIFECYCLE

The value of PROXY:AI is best seen in the path of a single query. Every request - from a person, an application or an agent - goes through the same complete control cycle

01 **Authentication**
RBAC: roles, groups, keys - no direct API keys

02 **Context assessment**
ABAC: who, from where, which model, what risk

03 **Input scan**
AI Firewall: data masking, blocking attacks on the model (prompt injection, jailbreak)

04 **Routing and budget**
model selection based on the scan result, limit control

05 **Model call**
or a semantic cache for recurring questions

06 **Output scan**
analysis of the model's response

07 **Human approval**
high-risk operations (human-in-the-loop)

08 **Audit record**
immutable WORM register

PLATFORM MODULES

The lifecycle stages are handled by ten modules grouped into the platform's four pillars

AI GATEWAY

- **Model Registry** – catalog of models and access profiles
- **ABAC Gateway** – attribute-based request evaluation
- **RBAC and identity** – roles, groups and keys; zero-downtime rotation
- **MCP Gateway** – oversight of autonomous AI agents

AI FIREWALL

- **AI Firewall** – scanning in 12 languages, data masking
- **Guardrails** – prompt injection / jailbreak filters

GOVERNANCE

- **Prompt management** – repository of response policies
- **Immutable audit** – WORM register, Prometheus and Grafana

COST CONTROL

- **FinOps and Anti-DoW – cost protection** – budgets, Twin:Coin, Basic/Advanced classes
- **AI Cost Forecast and Executive AI Dashboard** NEW – cost forecast (subscription or API) and an AI cost dashboard for the board

<50 ms	100+	12	6+ mo.
REAL-TIME ANALYSIS	SUPPORTED MODELS	SCANNING LANGUAGES	AUDIT HISTORY

VENDOR CREDIBILITY

allclouds.pl implemented its own AI management system and obtained ISO/IEC 42001 certification - the same standard that PROXY:AI helps organizations demonstrate compliance with

ISO/IEC 42001

REFERENCE DEPLOYMENT

LW Bogdanka - PROXY:AI in critical infrastructure, fully on-premises

REGULATORY COMPLIANCE AND EVIDENCE

The foundation of compliance is the WORM register: a database that records every model call, its cost and decision context - permanently and with cryptographic protection. The last six months of history are available on demand, and the retention period can be extended

On this basis, PROXY:AI generates evidence packages: documentation per Annex IV of the EU AI Act, DORA risk reports, NIS2/KSC materials and ISO/IEC 42001 audit packs - in one click, from data collected during normal operation. The evidence supports demonstrating compliance - it does not replace legal opinion or certification

DEPLOYMENT OPTIONS AND DATA SOVEREIGNTY

PROXY:AI offers three deployment options for different levels of data sovereignty requirements. Each deployment is a separate, fully isolated environment. In the On-Premises and Air-Gap options, sizing GPU infrastructure for local models is part of the deployment. In the Air-Gap option, the Offline Model Registry selects local models for the hardware you have, prepares a package from Harbor and runs the model without internet access

SaaS

PROFILE
managed service, fastest launch

DATA LOCATION
provider's data center (Poland or EU)

TYPICAL USE
evaluations, no in-house infrastructure

On-Premises

PROFILE
installed on your own infrastructure

DATA LOCATION
the organization's server room

TYPICAL USE
regulated sectors, public administration

Air-Gap

PROFILE
disconnected environment, local models only

DATA LOCATION
isolated infrastructure

TYPICAL USE
defense, critical infrastructure

The platform integrates with DLP and SIEM solutions and lets you build your own data classifiers and add industry-specific extensions. Typical areas: finance, public administration, healthcare, defense, critical infrastructure

SUMMARY

AI must be visible, controlled, accountable and demonstrable to an auditor. PROXY:AI delivers these four properties in a single layer - with real-time policies, budgets and a register that cannot be disputed - while keeping full freedom in choosing models

PROXY:AI

One entry point to the world of AI models.
Control, security and compliance evidence

Book a demo

Pilot in your infrastructure

allclouds.pl
AI by design

allclouds.pl sp. z o.o.
ul. Jutrzenki 139, 02-231 Warsaw, Poland

CONTACT
www.allclouds.pl · office@allclouds.pl
tel. +48 22 100 43 80 · fax +48 22 100 43 84
VAT ID: PL5223052539 · REGON: 363597531
KRS: 0000598708

CERTIFICATION
ISO 9001 · ISO/IEC 27001 · ISO 14001 · ISO 22301 · ISO/IEC 27017 · ISO/IEC 27018 · ISO/IEC 42001