

PLATFORM DOCUMENTATION

# Architecture Deep Dive

Stack · Security · Infrastructure

*A technical overview of the MIDAS platform — the six components it is built on, how they work together, how client data is secured, and what this means for you as a client or partner.*

**6**

Core  
Components

**100%**

Data  
Sovereignty

**VPS**

Private  
Infrastructure

**Docker**

Fully  
Containerised

## PLATFORM OVERVIEW

# The Architecture at a Glance

MIDAS is a purpose-built marketing intelligence platform deployed on a private Virtual Private Server (VPS) that is owned and controlled by the client. Every component — from data collection to AI inference to dashboard delivery — runs within this client-owned environment. NettScience builds, operates, and maintains the platform as a fully managed service, but the infrastructure and the data within it belong entirely to the client. There are no external software dependencies that hold client data, and no black-box SaaS platforms in the middle.

The platform is built from six proven, production-grade open-source components, orchestrated into a coherent intelligence system. Three architectural principles guide every design decision:

| DATA SOVEREIGNTY                                                                                                                                          | LOCAL AI CAPABILITY                                                                                                                                  | UNIFIED DATA MODEL                                                                                                                    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|
| Client data is never routed through third-party platforms for processing. All storage and analysis occurs within the client's own controlled environment. | AI inference runs within the platform for high-frequency tasks — not through external APIs. Eliminates cost at scale and preserves data sovereignty. | All sources feed one consistent warehouse. One definition of a campaign, a conversion, a cost — across every channel, every platform. |

|                            |                            |                        |                                |
|----------------------------|----------------------------|------------------------|--------------------------------|
| 6                          | 100%                       | VPS                    | Docker                         |
| Core Technology Components | Client Data Stays In-House | Private Infrastructure | Fully Containerised Deployment |

## THE INTELLIGENCE CYCLE — Daily Automation Flow

|    |          |                              |                                                                                                                                                                                |
|----|----------|------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 01 | 06:00 AM | <b>Data Collection</b>       | n8n pulls previous day's data from all connected platforms via official APIs — Google Ads, Meta, LinkedIn, GA4, CRM, email — writing raw records to PostgreSQL landing tables. |
| 02 | 07:00 AM | <b>Transform &amp; Model</b> | n8n triggers KNIME via SSH. KNIME runs anomaly detection, recalculates cross-channel attribution, updates lead scores, and writes modelled results back to PostgreSQL.         |
| 03 | 07:30 AM | <b>AI Intelligence</b>       | n8n reads anomaly flags. Ollama (local LLM) classifies each anomaly by type and severity. Claude API generates plain-language explanations for the highest-severity events.    |
| 04 | 08:00 AM | <b>Dashboard Ready</b>       | Superset dashboards refresh with live data. Account managers begin with intelligence — not raw numbers. Anomaly reports route to the Layer 2 approval queue.                   |

TECHNOLOGY STACK

# The Six Components

Each component in the MIDAS stack has a precise, non-overlapping role. Every one is a proven, production-grade open-source platform — widely deployed in enterprise environments globally. Together they form a coherent intelligence system that is greater than the sum of its parts.

## 01 n8n

### ORCHESTRATION ENGINE

Self-hosted workflow automation managing all data collection, scheduling, AI routing, approval orchestration, and channel execution. 400+ native platform integrations. Runs entirely within client infrastructure via Docker. The connective tissue of the entire platform.

## 02 KNIME

### ANALYTICS ENGINE

Visual data science platform handling ETL transformation, statistical modelling, cross-channel attribution analysis, anomaly detection, lead scoring, and lifetime value prediction. Triggered by n8n after each data collection cycle.

## 03 PostgreSQL

### UNIFIED DATA WAREHOUSE

Central data warehouse and system of record with four logical zones: raw data, modelled data, AI context, and operational logs. Row-level security enforced at the database layer. pgvector extension enables semantic AI memory.

## 04 Ollama

### LOCAL AI RUNTIME

Local AI model runtime serving open-source LLMs within the platform perimeter for classification, embedding generation, and routine analytical tasks. Zero external data routing for high-frequency tasks. Complemented by Claude/GPT-4 APIs for complex reasoning.

## 05 Apache Superset

### DASHBOARDS & VISUALISATION

Open-source BI platform serving client-facing dashboards and internal analytics via secure browser. Role-based access control. Direct connection to PostgreSQL — no intermediate data layer. Originally built at Airbnb; deployed at thousands of enterprises globally.

## 06 Caddy

### SECURE GATEWAY

The sole internet-facing component. Auto-provisions TLS certificates, handles rate limiting, IP access controls, and internal service routing. All other components sit on a private Docker network — invisible to the public internet under any circumstances.

## INFRASTRUCTURE

All six components run as Docker containers on a dedicated private VPS — owned by the client, managed by NetScience. Docker ensures consistent, reproducible environments and enables per-client isolation. Deployable in any geography to meet data residency requirements.

## DATA ARCHITECTURE

## The Unified Data Warehouse

Every piece of data that enters the system — and every model output it produces — is stored in a single structured PostgreSQL warehouse. This unified model is the architectural decision that makes cross-channel intelligence possible. It has four logical zones, each with a distinct function:

| ZONE               | DESCRIPTION                                                                                                                                                                                                                                 | KEY PROPERTY                          |
|--------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------|
| Raw Data Zone      | Data as received from source platforms, preserved exactly as collected. This is the historical record that allows models to be rerun when methodology improves. Never overwritten.                                                          | <i>Immutable historical record</i>    |
| Modelled Data Zone | Unified, transformed, business-ready data. One view of campaign performance across all channels. Attribution results, lead scores, audience segments. This is what powers dashboards and AI prompts.                                        | <i>Powers all dashboards &amp; AI</i> |
| AI Context Zone    | The accumulated institutional knowledge of each client engagement. Client preferences, historical recommendations, outcome measurements, creative performance patterns — making AI recommendations progressively more calibrated over time. | <i>Preference learning engine</i>     |
| Operational Zone   | A complete record of everything the system itself does. Workflow execution, data quality checks, approval decisions, execution logs. This is what makes the service observable and auditable.                                               | <i>Full system observability</i>      |

## SECURITY &amp; DATA PRIVACY

## How Client Data Is Protected

Data security is not a compliance layer applied on top of the MIDAS architecture — it is built into the architecture itself. The decisions that determine how data flows, where AI runs, and what faces the internet were made with security as a primary constraint, not an afterthought.

**01****Data Never Leaves Client Infrastructure**

Client data is stored and processed within the client's own managed environment. Row-level security in PostgreSQL ensures strict isolation — each client's data is accessible only through queries that include the correct identity filter, enforced at the database layer. No application-layer failure can expose one client's data to another.

**02****Local AI — Zero External Exposure for Routine Tasks**

AI inference for high-frequency tasks — content classification, anomaly flagging, embedding generation — runs locally via Ollama. No client data is transmitted to external services for the vast majority of AI workloads. When frontier models are required (Claude, GPT-4), this occurs under enterprise data processing agreements with explicit guarantees that data is not used for model training.

**03****Single External Gateway via Caddy**

Caddy is the only component with an internet-facing presence. It enforces HTTPS on all connections, applies rate limiting, enforces IP access controls, and routes traffic internally. Every other component — PostgreSQL, n8n, KNIME, Ollama, Superset — sits on a private Docker network. They are not accessible from the internet under any circumstances.

**04****Full Audit Trail**

Every workflow execution, data quality event, approval decision, and system action is logged to the Operational Zone of the data warehouse. This creates a complete, immutable audit trail of everything the system does — essential for compliance, debugging, and service accountability.

ARTIFICIAL INTELLIGENCE

## How AI Is Embedded in MIDAS

AI in MIDAS is not a feature added to a traditional analytics system. It is embedded throughout the platform's architecture, performing distinct and well-defined functions at every stage of the intelligence cycle. The distinction between local and cloud AI is a deliberate architectural choice — not a limitation.

### LOCAL AI — Ollama (Runs Inside Client Infrastructure)

|                                                 |                                                                                                                                                                                                                         |
|-------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Continuous Classification &amp; Tagging</b>  | Every ad creative, social post, email, and competitor activity is classified by type, intent, tone, and funnel stage. Runs perpetually across all client accounts at zero marginal API cost with full data sovereignty. |
| <b>Anomaly Detection &amp; Severity Scoring</b> | Statistical anomalies in campaign performance are classified by type and severity. High-volume and high-frequency — unsuitable for cloud API cost structures at scale.                                                  |
| <b>Embedding Generation</b>                     | All content, recommendations, and performance data is encoded into vector representations stored in PostgreSQL (pgvector), enabling semantic similarity search across institutional memory.                             |

### CLOUD AI — Claude / GPT-4 (Complex Reasoning Only)

|                                            |                                                                                                                                                                                                                      |
|--------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Anomaly Explanation &amp; Diagnosis</b> | For the highest-severity anomalies, the system assembles relevant context and generates a plain-language explanation of the most likely cause. Account managers receive a reasoned narrative — not a raw data alert. |
| <b>Strategic Recommendations</b>           | Cross-channel attribution analysis, budget reallocation recommendations, and audience expansion suggestions that require nuanced reasoning across multiple data signals.                                             |
| <b>Preference Learning</b>                 | AI recommendations route through a human-in-the-loop approval system. Approval and rejection signals update the AI Context Zone, making future recommendations progressively more calibrated to each client.         |

COMPONENT SUMMARY

### Quick Reference

| COMPONENT | ROLE          | KEY PROPERTIES                                                  |
|-----------|---------------|-----------------------------------------------------------------|
| n8n       | Orchestration | Self-hosted · 400+ integrations · Docker · API-based collection |

|                        |                  |                                                                            |
|------------------------|------------------|----------------------------------------------------------------------------|
| <b>KNIME</b>           | Analytics Engine | ETL · Attribution modelling · Anomaly detection · Lead scoring             |
| <b>PostgreSQL</b>      | Data Warehouse   | 4-zone schema · Row-level security · pgvector · Multi-tenant               |
| <b>Ollama</b>          | Local AI Runtime | In-perimeter inference · No external routing · Classification & embeddings |
| <b>Apache Superset</b> | Dashboards       | Role-based access · Direct PostgreSQL · Browser-based · Open-source        |
| <b>Caddy</b>           | Secure Gateway   | Auto TLS · Rate limiting · IP controls · Sole internet-facing component    |

### Want a live demonstration?

We run MIDAS on real anonymised data. Request a walkthrough of the platform in action.

[analytics@nettscience.com](mailto:analytics@nettscience.com)