



AESIR
STRATEGIC INTELLIGENCE

AESIR™: Strategic AI Platform

Executive Summary

AESIR™ is a sophisticated, enterprise-grade AI platform designed for organizations requiring intelligent document processing, semantic search, and domain-specific AI assistance. Built for both cloud and air-gapped, edge deployments, AESIR™ delivers powerful AI capabilities while maintaining data security and operational flexibility.

Key Value Propositions

- **Cost Flexibility:** Choose between local AI models (cost-effective) or cloud services (maximum performance)
- **Data Security:** On-premises deployment options ensure sensitive information never leaves your infrastructure
- **Domain Expertise:** Specialized AI personas tailored for business and military contexts
- **Cross-Platform Access:** Native desktop applications and browser-based access for maximum user adoption
- **Advanced Document or Telemetry Intelligence:** Transform static documents and telemetry data into searchable, actionable, and conversational knowledge bases

Core Features & Capabilities

 **Dual AI Backend Support**

- **Local Models (Ollama):** Deploy AI models directly on your infrastructure (servers, edge devices, etc) for maximum security and cost control
- **Cloud Services (OpenAI):** Leverage cutting-edge cloud AI for peak performance when connectivity allows
- **Seamless Switching:** Configure and switch between backends based on operational requirements
- **Tested Models:** Pre-validated with high-performance models including Gemma, Qwen, Cogito, and others.

Advanced Document Processing & RAG

- **Intelligent PDF Processing:** Extract text and images from complex documents with OCR capabilities
- **Vision & OCR:** Handle images and image-based documents using AI vision models (like Gemma)
- **Smart Document Chunking:** Semantic Kernel-powered chunking for optimal context preservation
- **Hybrid Search:** Combines keyword matching with semantic vector search for precise information retrieval
- **Citation System:** Automatic source attribution with page-level references for audit trails

Specialized AI Personas

- **Business Professional Mode:** Tailored for corporate environments with business-focused prompts and terminology
- **Military Operations Mode:** OPSEC-compliant assistance for mission-critical information processing
- **Customizable Contexts:** Extensible prompt system for industry-specific adaptations
- **Edge-Optimized:** Designed for performance on edge devices with limited connectivity

Cross-Platform Client Architecture

- **Desktop Application:** Native Avalonia-based client for Windows, macOS, and Linux
- **Browser Access:** WebAssembly deployment for universal web-based access
- **Responsive Design:** Consistent user experience across all platforms
- **Offline Capability:** Continue working even when internet connectivity is limited

Enterprise-Grade Search & Knowledge Management

- **Vector Database Integration:** Qdrant-powered semantic search for intelligent document discovery
- **Global Document Collections:** Organization-wide knowledge bases accessible across teams
- **Conversation-Scoped Documents:** Context-specific document collections for focused discussions
- **Real-Time Citations:** Automatic source linking with file and page references

Technical Architecture

Backend Infrastructure

- **ASP.NET Core API:** Scalable, enterprise-ready web API with health checks and monitoring
- **Microsoft Semantic Kernel:** Advanced AI orchestration and plugin architecture
- **PostgreSQL:** Robust data persistence with including vector search capabilities
- **Qdrant Vector Database:** High-performance semantic search and similarity matching
- **Containerized Deployment:** Docker Compose orchestration with Traefik reverse proxy

AI Integration Layer

- **Modular Service Architecture:** Interface-based design for easy AI provider integration
- **Streaming Responses:** Real-time chat completions with thinking process visibility
- **Token Management:** Comprehensive usage tracking and optimization
- **Error Handling:** Robust retry policies and graceful degradation

Client Technology Stack

- **Avalonia UI Framework:** Modern, cross-platform desktop application framework
- **MVVM Architecture:** Clean separation of concerns with CommunityToolkit.Mvvm
- **WebAssembly Support:** Browser deployment without plugins or extensions
- **Reactive Messaging:** Real-time UI updates and inter-component communication

Target Markets & Use Cases

Enterprise & Business

- **Document Analysis:** Transform contracts, reports, and policies into conversational knowledge
- **Decision Support:** AI-powered insights from corporate documentation
- **Knowledge Management:** Centralized, searchable repository of organizational knowledge
- **Compliance & Audit:** Traceable citations and source attribution for regulatory requirements

Military & Defense

- **Mission Planning:** OPSEC-compliant AI assistance for operational documentation
- **Intelligence Analysis:** Secure processing of classified and sensitive materials
- **Training & Education:** Interactive learning from military manuals and procedures
- **Edge Deployment:** Operational capability in disconnected or low-bandwidth environments

Healthcare & Professional Services

- **Research & Analysis:** AI-powered insights from medical literature and case studies
- **Policy Compliance:** Navigate complex regulatory documentation with AI assistance
- **Training Materials:** Interactive learning from professional documentation
- **Client Consultation:** Informed discussions backed by comprehensive document analysis

Competitive Advantages

Security & Privacy

- **On-Premises Deployment:** Complete data sovereignty with local AI model execution
- **No Data Transmission:** Sensitive information never leaves your infrastructure when using local models

- **Audit Trail:** Complete citation and source tracking for compliance requirements
- **Access Control:** Granular permissions and document collection management

Cost Efficiency

- **Flexible Pricing Models:** Choose between local (one-time cost) and cloud (usage-based) AI backends
- **Resource Optimization:** Efficient processing with configurable batch sizes and delays
- **Scalable Architecture:** Pay only for the resources you need

Technical Excellence

- **Modern Architecture:** Built on proven enterprise technologies (.Net Core, ASP.NET Core, Semantic Kernel, PostgreSQL)
- **Cross-Platform Support:** Single solution for desktop and web deployment
- **Extensible Design:** Plugin, modular architecture for custom integrations and workflows
- **Performance Optimized:** Edge-device optimization with sub-500 token response targets

Implementation & Deployment

Deployment Options

- **Full Container Stack:** Complete Docker Compose deployment with all dependencies
- **Hybrid Deployment:** Can be deployed as containerized services and/or native applications
- **Development Environment:** Rapid setup for evaluation and customization
- **Production Ready:** SSL/TLS termination, health monitoring, and automatic restarts

Integration Capabilities

- **RESTful API:** Standard HTTP endpoints for system integration
- **Plugin Architecture:** Extensible AI kernel plugin system for modular custom functionality
- **Database Flexibility:** Store data of any format to PostgreSQL (including vector extensions)
- **Authentication Ready:** Prepared for enterprise authentication integration

Support & Maintenance

- **Comprehensive Documentation:** Detailed setup and configuration guides
- **Modular Design:** Easy updates and feature additions
- **Monitoring & Logging:** Built-in health checks and diagnostic capabilities
- **Professional Services:** Available for custom implementation and training

Getting Started

Evaluation Process

1. **Quick Setup:** Docker Compose deployment in under 30 minutes
2. **Document Upload:** Test with your organization's sample documents and telemetry
3. **AI Backend Configuration:** Evaluate both local and cloud AI options
4. **User Experience:** Test desktop and browser clients with your team
5. **Performance Assessment:** Measure response times and accuracy with your data

Next Steps

- **Technical Demo:** Schedule a personalized demonstration with your documents
- **Pilot Program:** Limited deployment for evaluation and feedback
- **Custom Configuration:** Tailored setup for your specific requirements
- **Training & Onboarding:** User training and administrator education

Contact Information

For more information about AESIR™ implementation, pricing, or technical specifications, please contact:

Ronin Consulting

- Email: info@ronin.consulting
- Technical Inquiries: Available for detailed architecture discussions
- Demo Requests: Personalized demonstrations with your data

AESIR™: Transforming Data into Intelligent Conversations and Strategic Decisions.

AESIR™ is a trademark of Ronin Consulting.